

Genetic Algorithm Based on Similarity for Probabilistic Transformation of Belief Functions [☆]

Yilin Dong^a, Xinde Li^{a,*}, Jean Dezert^b, Pei Li^a, Xianghui Li^a

^a*Key Laboratory of Measurement and Control of CSE, School of Automation, Southeast University, Nanjing, Jiangsu Province, China*

^b*The French Aerospace Laboratory, Palaiseau F-91761, France*

Abstract

Recent studies of alternative Probabilistic Transformation (PT) in Dempster-Shafer (DS) theory have mainly focused on investigating various schemes for assigning the mass of compound focal elements to each singleton in order to obtain Bayesian belief function for decision making problems. In the process of such transformation, how to precisely evaluate the closeness between original Basic Belief Assignments (BBAs) and transformed BBAs is important. In this paper, a new aggregation measure is proposed by comprehensively considering the interval distance between BBAs and also the rank order of focal elements depending on their positive mass inside BBAs. Relying on this new measure, we propose a novel Multi-Objective Probabilistic Transformation (MOEPT) thanks to global optimizing capabilities inspired from genetic algorithm (GA). From the perspective of mathematical theory, convergence analysis of MOEPT is given to prove the rationality of GA used here. Finally, various scenarios in evidence reasoning are presented to evaluate the effectiveness of MOEPT in PT. Moreover, a simple constraint-handling strategy with MOEPT is developed to tackle target type tracking (TTT) problem. The simulation results of the constrained MOEPT on TTT problem prove MOEPT's extendibility.

Keywords: Probabilistic Transformation (PT), Similarity Measure,

[☆]Fully documented templates are available in the elsarticle package on CTAN.

*Corresponding author

Email address: xinde1i@seu.edu.cn (Xinde Li)

1. Introduction

Since the pioneering work of Dempster and Shafer [1, 2], known as Dempster-Shafer evidence Theory (DST), the belief functions are widely used in information fusion for decision making [3],[4]. However, the computational complexity of reasoning with DST is one of the major points of criticism this formalism has to face.

To overcome this difficulty, various approximating methods have been suggested that aim at reducing the number of focal elements in the Frame of Discernment (FoD) in order to maintain the tractability of computation. One common strategy is to simplify FoD by removing and/or aggregating focal elements for approximating original belief function [5]. Among these methods, Probabilistic Transformations (PTs) seem particularly desirable for reducing such computational complexity by means of assigning the mass of non-singleton elements to some singletons of the FoD [6], [8]. The research on this probabilistic measure has received a lot of attentions [7] and many efficient PTs have been proposed by scholars in recent years. Among them, a classical transformation, denoted as BetP [6], was usually adopted because it offers a compromise between the maximum of credibility (Bel) and the maximum of plausibility (Pl) for decision making. Unfortunately, BetP does not provide the highest Probabilistic Information Content (PIC) [10] and Shenoy really argued against BetP in his publication [11]; Sudano [12] also proposed series of alternatives and principles of these similar to BetP, which were called PrPl, PrBel and PrHyb; CuzzP [13], which was proposed by Cuzzolin in the framework of DST in 2009, showed its ability of probabilistic transformation; Another novel transformation was proposed by Dezert and Smarandache in the framework of Dezert-Smarandache Theory (DSmT), which was called Dezert-Smarandache Probability (DSmP) [10] and comprehensive comparisons have been made in [10] to prove the advantages of DSmP with respect to other PTs.

Simultaneously, various techniques have been proposed to evaluate PTs. Typically, PIC or Shannon's entropy criterion is applied to evaluate PTs which less uncertainty of BBAs obtained from PTs are preferred in order to make decision easily. However, Han et.al [14] illustrated the irrationality of the over-emphasize of Shannon entropy or PIC. Besides, Bucci in [16] also pointed out that PIC may not be sufficient for evaluating the quality of PTs. The distance of evidence [17] is used to characterize the difference between the transformed BBAs and original BBAs to quantify the degree of similarity. But it is not good enough to capture the difference between BBAs in some cases as it has been seen in [18],[19]. Unfortunately, either PIC or distance alone is not an efficient technique and it is actually difficult to quantify all sorts aspects of dissimilarity which inevitably need to be involved. To address un-comprehensive evaluation, several two-dimensional measures [20], [21] have been proposed in order to make sure that results obtained by PTs be consistent in some manner with original BBAs. Han in [22] proposed a 2-D criteria, which joint uses distance and PIC measures, to make a balance between fidelity and clarity ¹. Liu [18] used a two-dimensional measure to effectively detect conflict among evidence. In [23], Liu proposed both a distance and a conflict coefficient based on probabilistic transformations *BetP* to characterize the dissimilarity, which are complementary in a certain sense. Recently, Ma [24] integrated fuzzy closeness and correlation coefficient to generate a new dissimilarity measure to characterize not only the difference between BBAs but also the divergence degree of the hypothesis that two BBAs support.

By analyzing the mentioned existing methods, the relationship between techniques of PTs and their corresponding evaluations are almost independent except [22]. That is to say, the methods of evaluations only assess the existing PTs instead of facilitating the development of novel PTs themselves. In this paper, we present a novel PT method based on Multi-Objective algorithm (MOEPT) using a reasonable and comprehensive two-dimensional criteria in order to cap-

¹a probability with higher clarity and bigger fidelity should be preferred.

ture the similarity in the process of PT. This new method has some connections with the recent algorithm proposed in [22]. However, the main differences lie in the following aspects: 1. 2-D criteria, PIC and Jousselme’s distance, has been pointed out its drawbacks in many reference [19], [28]. Thus, an efficient and different aggregation measure is proposed. Its novelty lies in considering the drawback of the past description of the distance between the evidences. In other words, up to now, most distances were defined according to the corresponding focal elements between two sources of evidence, and the sequence inside the assignments of focal elements itself was not considered. The sequence might also lead to dissimilarity, which is referred to as self-conflict or self-contradiction [25]; 2. More specific steps of evolutionary-based algorithms are given in details; 3. the convergence analysis of MOEPT is illustrated to prove the rationality of using GA; 4. some bugs are detected and fixed when using MOEPT with traditional constraints; 5. Target Type Tracking (TTT) problem has been efficiently solved based on the proposed method with a novel simple constraint. Compared to traditional PTs, global search replaces designing various assigning operator in classical PTs and evaluation criteria is embedded into MOEPT to provide important guidance for searching procedure. Specifically, masses of singletons are randomly generated in evolutionary-based framework, which need to satisfy with the basic constraints for probability distributions in evidence reasoning. Also, a assessment factor is presented to assess the best individual in all populations by a special objective function (desirable evaluation criteria). Simulation results on 4D FoD test cases show that the proposed MOEPT, in these problems, is able to outperform other PTs from the perspective of 2-D criteria. Moreover, we propose a simple constraint-handling strategy within MOEPT that suits well for two target type tracking (2-TTT) problems, which in some extents encourages the applications of MOEPT to more complex and real-world decision making problems.

The reminder of this paper is structured as follows. In Section 2 we briefly summarize the basis of DST. The new aggregation measure is proposed in section 3; In Section 4, a multi-objective Evolutionary Algorithm (EA) based on

two-dimensional objective function is proposed. In Section 5, several examples
 90 and comprehensive comparisons are carried out. A simple pattern recognition
 problem and also target type tracking problem are presented and solved in de-
 tails at the end of this section. The conclusion is drawn in Section 6.

2. Basis of belief functions

In this section, we introduce the belief functions terminology of DST and
 95 the notations used in the sequel of this paper.

2.1. DST basis

In DST [2], the elements θ_i ($i = 1, \dots, N$) of the frame of discernment (FoD)
 $\Theta \triangleq \{\theta_1, \dots, \theta_N\}$ must be mutually exhaustive and exclusive. The power set of
 the FoD is denoted 2^Θ and a basic belief assignment (BBA), also called a mass
 function, is defined by the mapping: $2^\Theta \rightarrow [0, 1]$, which satisfies $m(\emptyset) = 0$ and

$$\sum_{A \subseteq 2^\Theta} m(A) = 1 \quad (1)$$

where $m(A)$ is defined as the BBA of A . The element A is called a focal element
 of $m(\cdot)$ if $m(A) > 0$. The belief and plausibility functions, which are in one-to-
 one mapping with the BBA $m(\cdot)$, are defined for all $A \subseteq \Theta$ by

$$Bel(A) = \sum_{B \in 2^\Theta | B \subseteq A} m(B) \quad (2)$$

$$Pl(A) = 1 - Bel(\bar{A}) = \sum_{A, B \in 2^\Theta | A \cap B \neq \emptyset} m(B) \quad (3)$$

where $\bar{A} \triangleq \Theta \setminus A$ is the complement of A in Θ . The belief interval $[Bel(A), Pl(A)]$
 represents the uncertainty committed to A and the bounds of this interval are
 usually interpreted as lower and upper bounds of the unknown (possibly sub-
 100 jective) probability of A .

In order to fuse n Bodies Of Evidences (BOEs), Dempster's rule of combi-
 nation is usually used in DST framework. The combination of n distinct BOE

is done by:

$$m(A) = \begin{cases} 0, & \text{if } A = \emptyset \\ \frac{\sum_{\cap A_i = A} \prod_{1 \leq i \leq n} m_i(A_i)}{\sum_{\cap A_i \neq \emptyset} \prod_{1 \leq i \leq n} m_i(A_i)}, & \text{if } A \neq \emptyset \end{cases} \quad (4)$$

2.2. Classical Probabilistic Transformations

The efficiency of Probabilistic Transformation (PT) in the field of decision making has been analyzed in deep by Smets [6]. Various PTs have been proposed in the open literature such as BetP[6],[9], CuzzP [13], DSmp [10], PrBP1 and 105 PrBP2 [26] and Cobb-Shenoy's Normalization of Plausibility [11]. And herein, the simple and classical transformation (BetP) is briefly recalled.

Smets in [6],[9] first proposed pignistic (also called *betting*) probability to make decision which aims to transfer the mass of belief of each non-specific element onto the singletons. The classical pignistic probability is defined as $BetP(\emptyset) = 0$, and $\forall A \in 2^\Theta \setminus \{\emptyset\}$:

$$BetP(\theta_i) \triangleq \sum_{A \subseteq 2^\Theta, A \neq \emptyset} \frac{|\theta_i \cap A|}{|A|} \frac{m(A)}{1 - m(\emptyset)} \quad (5)$$

Because in Shafer's framework $m(\emptyset) = 0$, the formula (5) can simply be rewritten for any singleton $\theta_i \in \Theta$ as

$$BetP(\theta_i) = \sum_{B \in 2^\Theta, \theta_i \subseteq B} \frac{1}{|B|} m(B) = m(\theta_i) + \sum_{B \in 2^\Theta, \theta_i \subset B} \frac{1}{|B|} m(B) \quad (6)$$

3. Similarity Measure

3.1. Distances Between Two BBAs

The Jousselme's distance, which was widely denoted as D_J in [17], was 110 applied in many recent references [27], [22], but unfortunately it doesn't provide satisfactory results for all BBAs structures. This has been clearly explained recently in [19], [28], where another strict distance metric between two BBAs, denoted d_{BT}^E , has been proposed which is briefly recalled here.

Assuming that two independent BBAs $m_1(\cdot)$ and $m_2(\cdot)$ are defined on $\Theta =$ 115 $\{\theta_1, \theta_2, \dots, \theta_N\}$. For each focal element $\theta_i \in \Theta$ ($i = 1, 2, \dots, 2^N - 1$), belief

intervals of θ_i for $m_1(\cdot)$ and $m_2(\cdot)$ can be calculated respectively, which are denoted by $[Bel_1(\theta_i), Pl_1(\theta_i)]$ and $[Bel_2(\theta_i), Pl_2(\theta_i)]$. The strict distance between interval numbers $[a, b]$ and $[c, d]$ was defined by [31]

$$d^I([a, b], [c, d]) = \sqrt{\left[\frac{a+b}{2} - \frac{c+d}{2}\right]^2 + \frac{1}{3}\left[\frac{b-a}{2} - \frac{d-c}{2}\right]^2} \quad (7)$$

Therefore, we can calculate the distance between $BI_1(\theta_i) : [Bel_1(\theta_i), Pl_1(\theta_i)]$ and $BI_2(\theta_i) : [Bel_2(\theta_i), Pl_2(\theta_i)]$ according to Eq.(7). Thus, we can obtain totally $2^N - 1$ belief interval distance values for all $\theta_i \in \Theta$. Besides, Euclidean-family Belief Interval-based Distance d_{BI}^E can be rewritten as

$$d_{BI}^E(m_1, m_2) = \sqrt{N_c \cdot \sum_{i=1}^{2^N-1} [d^I(BI_1(\theta_i), BI_2(\theta_i))]^2} \quad (8)$$

Here, $N_c = 1/2^N - 1$ is the normalization factor. In this paper, we regard d_{BI}^E as one criteria for evaluating the degree of similarity² between original BBAs and the transformed ones.

3.2. A New Evidence Similarity Characterization

As mentioned in previous section, those distances, i.e., Jousselme's distance [27] and also other metric like PIC [29] or Entropy [26] were widely applied to measure the degree of "similarity or dissimilarity" between BBAs. But only the corresponding focal elements (or the relevant focal element set) between two sources of evidence are described or characterized. This one-sided view does not consider the order of size of the assignment of each focal element in an evidence, which might lead to "self-conflict or self-contradiction". To consider such "information" produced by evidence itself, here a new evidence similarity measure is defined between two evidential sources according to the order of size of the assignment. Prior to this, to give this new similarity measure, first we define the order correlation coefficient between two sets of data.

²similarity represent the degree of difference between original BBAs and the transformed ones in [32].

Definition 1. [30] Given two sets of data $\{x_1, x_2, \dots, x_n\}, \{y_1, y_2, \dots, y_n\}$, here, x_1, x_2, \dots, x_n and y_1, y_2, \dots, y_n are in an ascending order. After sorting, two
140 sets of data are $x_{p_1}, x_{p_2}, \dots, x_{p_n}$ and $y_{q_1}, y_{q_2}, \dots, y_{q_n}$, respectively, meet $x_{p_1} \leq$
 $x_{p_2} \leq \dots \leq x_{p_n}$ and $y_{q_1} \leq y_{q_2} \leq \dots \leq y_{q_n}$, for each p_i , index its position
from q_1, q_2, \dots, q_n , assuming it is q_j , that is, $q_j = p_i$. Note that $j = f(i)$, the
correlation coefficient is

$$\mu = \frac{\sum_{i=1}^n (i - j)^2}{\sum_{i=1}^n [n - (i - 1) - i]^2} \quad (9)$$

It satisfies $0 \leq \mu \leq 1$. When $\mu = 0$, the convergence of two sets of data is
145 the largest; When $\mu = 1$, it is the smallest.

3.2.1. The consistency of focal elements between two BOEs

Definition 2. For any two sources of evidence, i.e., $S_1, S_2, m_1(\cdot), m_2(\cdot)$ are the
basic belief assignments over the discernment framework Θ of size N . The num-
ber of focal elements and the focal elements of $m_1(\cdot)$ and $m_2(\cdot)$ can be different.
150 We denote X_i and Y_i the index of the focal elements whose mass is sorted by
increasing order. The similarity function of evidence to characterize the order
of the size of the assignments over subsets as follows:

$$Sim_{seq}(m_1, m_2) = 1 - \frac{\sum_{i=1}^N (X_i - Y_i)^2}{\sum_{i=1}^N [N + 1 - 2i]^2} \quad (10)$$

To be useful, a similarity function must be homogeneous to a true distance
metric which must satisfy the following properties:

- 155 • symmetry: $\forall m_i(\cdot), m_j(\cdot), Sim(m_i, m_j) = Sim(m_j, m_i)$;
- consistency: $\forall m(\cdot), Sim(m, m) = 1$ and
- nonnegative: $\forall m_i(\cdot), m_j(\cdot), 0 \leq Sim(m_i, m_j) \leq 1$.
- triangle inequality: $Sim(X, Y) + Sim(Y, Z) \geq Sim(X, Z)$

According to our previous work [30], it is easy to prove that Sim_{seq} satisfy
160 symmetry, consistency and nonnegativity but the last important condition is
lost. So we prove the property of triangle inequality of Sim_{seq} here.

Proof:

Based on Eq.(10), the triangle inequality can be rewritten as follows:

$$1 - \frac{\sum_{i=1}^n (X_i - Z_i)^2}{\sum_{i=1}^n [n+1-2i]^2} \leq 1 - \frac{\sum_{i=1}^n (X_i - Y_i)^2}{\sum_{i=1}^n [n+1-2i]^2} + 1 - \frac{\sum_{i=1}^n (Y_i - Z_i)^2}{\sum_{i=1}^n [n+1-2i]^2}$$

$$\Rightarrow \sum_{i=1}^n [n+1-2i]^2 \geq \sum_{i=1}^n (X_i - Y_i)^2 + \sum_{i=1}^n (Y_i - Z_i)^2 - \sum_{i=1}^n (X_i - Z_i)^2$$

Using vector notations $\mathbf{X} = [X_1, \dots, X_N]^T$, $\mathbf{Y} = [Y_1, \dots, Y_N]^T$ and $\mathbf{Z} = [Z_1, \dots, Z_N]^T$,

one must prove that

$$\Rightarrow \sum_{i=1}^n [n+1-2i]^2 \geq (\mathbf{X} - \mathbf{Y})^T (\mathbf{X} - \mathbf{Y}) + (\mathbf{Y} - \mathbf{Z})^T (\mathbf{Y} - \mathbf{Z}) - (\mathbf{X} - \mathbf{Z})^T (\mathbf{X} - \mathbf{Z})$$

According to the Squared Sum Formula (SSF)³,

$$\begin{aligned} \sum_{i=1}^n [n+1-2i]^2 &= \sum_{i=1}^n [(n+1)^2 + 4i^2 - 4(n+1)i] = \\ &= n(n+1)^2 + 4 * \frac{1}{6} n(n+1)(2n+1) - 4(n+1) \frac{n(n+1)}{2} \\ &= \frac{2}{3} n(n+1)(2n+1) - n(n+1)^2 \\ &= n(n+1) \cdot [\frac{4}{3}n + \frac{2}{3} - n - 1] \\ &= \frac{1}{3} \cdot n \cdot (n^2 - 1) \end{aligned}$$

Because $1 \leq X_i \leq n, 1 \leq Y_i \leq n, 1 \leq Z_i \leq n$, so we obtain:

$$(\mathbf{X} - \mathbf{Y})^T (\mathbf{X} - \mathbf{Y}) + (\mathbf{Y} - \mathbf{Z})^T (\mathbf{Y} - \mathbf{Z}) - (\mathbf{X} - \mathbf{Z})^T (\mathbf{X} - \mathbf{Z}) \leq 1 + (k-1)^2 + (k-3)^2 + \dots$$

- When $n-1 = 2k \Rightarrow k = \frac{n-1}{2}, 1 + (k-1)^2 + (k-3)^2 + \dots = \frac{1}{6} \cdot k(k+1)(k+2)$,

³squared sum formula of Natural Number: $1^2 + 2^2 + 3^2 + \dots + n^2 = \frac{n(n+1)(2n+1)}{6}$; squared sum formula of odd number: $1^2 + 3^2 + 5^2 + \dots + (2n-1)^2 = \frac{1}{3}n(4n^2-1)$; squared sum formula of even: $2^2 + 4^2 + \dots + (2n)^2 = \frac{2}{3}n(n+1)(2n+1)$.

thus

$$\begin{aligned}
& \sum_{i=1}^n [n+1-2i]^2 - [(\mathbf{X}-\mathbf{Y})^T(\mathbf{X}-\mathbf{Y}) + (\mathbf{Y}-\mathbf{Z})^T(\mathbf{Y}-\mathbf{Z}) - (\mathbf{X}-\mathbf{Z})^T(\mathbf{X}-\mathbf{Z})] \\
& \geq \frac{1}{3} \cdot n \cdot (n^2 - 1) - \frac{1}{6} \cdot k(k+1)(k+2) \\
& = \frac{1}{3} \cdot n \cdot (n^2 - 1) - \frac{1}{6} \cdot \frac{n-1}{2} \left(\frac{n-1}{2} + 1 \right) \left(\frac{n-1}{2} + 2 \right) \\
& = \frac{1}{3} \cdot n \cdot (n^2 - 1) - \frac{1}{6} \cdot \frac{n-1}{2} \cdot \frac{n+1}{2} \cdot \frac{n+3}{2} \\
& = (n^2 - 1) \cdot \left(\frac{1}{4}n + \frac{1}{6} \right)
\end{aligned}$$

Because $n \geq 2$, thus

$$\sum_{i=1}^n [n+1-2i]^2 - [(\mathbf{X}-\mathbf{Y})^T(\mathbf{X}-\mathbf{Y}) + (\mathbf{Y}-\mathbf{Z})^T(\mathbf{Y}-\mathbf{Z}) - (\mathbf{X}-\mathbf{Z})^T(\mathbf{X}-\mathbf{Z})] \geq 0$$

175

$\Rightarrow Sim_{seq}$ satisfies the triangle inequality;

- When $n-1 = 2k-1 \Rightarrow k = \frac{n}{2}$, $1 + (k-1)^2 + (k-3)^2 + \dots = \frac{1}{3} \cdot k \cdot (4k^2 - 1)$,

thus

$$\begin{aligned}
& \sum_{i=1}^n [n+1-2i]^2 - [(\mathbf{X}-\mathbf{Y})^T(\mathbf{X}-\mathbf{Y}) + (\mathbf{Y}-\mathbf{Z})^T(\mathbf{Y}-\mathbf{Z}) - (\mathbf{X}-\mathbf{Z})^T(\mathbf{X}-\mathbf{Z})] \\
& \geq \frac{1}{3} \cdot n \cdot (n^2 - 1) - \frac{1}{3} \cdot k \cdot (4k^2 - 1) \\
& = \frac{1}{3} \cdot n \cdot (n^2 - 1) - \frac{1}{3} \cdot \frac{n}{2} \cdot \left(4 \left(\frac{n}{2} \right)^2 - 1 \right) \\
& = \frac{1}{3} \cdot n \cdot (n^2 - 1) - \frac{n}{6} \cdot (n^2 - 1) \\
& = \frac{n}{6} \cdot (n^2 - 1)
\end{aligned}$$

Because $n \geq 2$, thus

$$\sum_{i=1}^n [n+1-2i]^2 - [(\mathbf{X}-\mathbf{Y})^T(\mathbf{X}-\mathbf{Y}) + (\mathbf{Y}-\mathbf{Z})^T(\mathbf{Y}-\mathbf{Z}) - (\mathbf{X}-\mathbf{Z})^T(\mathbf{X}-\mathbf{Z})] \geq 0$$

180

$\Rightarrow Sim_{seq}$ satisfies the triangle inequality;

End Proof.

Definition 3. For any two sources of evidence, i.e., S_1, S_2 , $m_1(\cdot)$ and $m_2(\cdot)$ are the BBAs defined over the power-set of the same FoD Θ . We denote by s_1 (with $1 \leq s_1 \leq 2^N - 1$) the number of focal elements of $m_1(\cdot)$, and by s_2 ((with $1 \leq s_2 \leq 2^N - 1$) the number of focal elements of the BBA $m_2(\cdot)$). Assuming that s_1 subsets' BBAs are same in $m_1(\cdot)$, and s_2 subsets' BBAs are same in $m_2(\cdot)$. Wherein, X_i, Y_i are the serial number according to the order of the size of subset' BBAs; the subscript i indicates the i th subset. Due to the BBAs of some sub-propositions are same. For the evidence S_1 , there might be s_1 kinds of sorts. For S_2 , there might be s_2 kinds of sorts. Therefore, there are $s_1 \times s_2$ kinds of sorts for S_1 and S_2 . The similarity measure functions is redefined in this case as follows:

$$Sim'_{seq}(m_1, m_2) = 1 - \frac{\sum_{t=1}^{s_1 s_2} \sum_{i=1}^n (X_i^t - Y_i^t)^2}{s_1 s_2 (\sum_{i=1}^n (n+1-2i)^2)} \quad (11)$$

Similarly, it is easy to prove that $Sim'_{seq}(m_X, m_Y)$ is a true distance metric.

Example 1 (Bayesian BBAs): Assuming two evidences: $m_1 = \{\theta_1, \theta_2, \theta_3\} = \{0.1, 0.2, 0.7\}$; $m_2 = \{\theta_1, \theta_2, \theta_3\} = \{0.7, 0.2, 0.1\}$, then m_1 and m_2 are sorted respectively from small to large so that $X(m_1) = \{\theta_1, \theta_2, \theta_3\} = [1, 2, 3]$ and $Y(m_2) = \{\theta_3, \theta_2, \theta_1\} = [3, 2, 1]$. Thus we can calculate the similarity measure based on Eq.(10):

$$Sim_{seq}(m_1, m_2) = 1 - \frac{(1-3)^2 + (2-2)^2 + (3-1)^2}{(3+1-2*1)^2 + (3+1-2*2)^2 + (3+1-2*3)^2} = 0$$

according to $Sim_{seq}(m_1, m_2)$, we see that m_1 and m_2 are completely different and dissimilar.

Remark: When two evidences: m_1 and m_2 have the property of ambiguity (the equal same value), that is $m_1 = \{\theta_1, \theta_2, \theta_3, \theta_4\} = \{0.1, 0.1, 0.2, 0.6\}$; $m_2 = \{\theta_1, \theta_2, \theta_3, \theta_4\} = \{0.6, 0.2, 0.1, 0.1\}$, thus the corresponding sequences are obtained: $X_1 = \{1, 1, 2, 3\}$; $Y_1 = \{3, 2, 1, 1\}$. So we can calculate the similarity based on Eq.(10).

3.2.2. The inconsistency of focal elements between two BOEs

How to calculate Sim_{seq} when the focal elements in BBAs are different? Let
 205 us consider the following example and put forward the different way compared
 to [30]:

Example 2 (General BBAs): Assuming two evidences: $m_1 = \{\theta_1, \theta_2 \cup \theta_3, \theta_2 \cup$
 $\theta_4, \theta_1 \cup \theta_3 \cup \theta_4\} = \{0.3, 0.2, 0.2, 0.3\}$; $m_2 = \{\theta_1, \theta_1 \cup \theta_3, \theta_3, \theta_2 \cup \theta_3 \cup \theta_4, \theta_1 \cup \theta_2 \cup$
 $\theta_3 \cup \theta_4\} = \{0.4, 0.1, 0.1, 0.2, 0.2\}$,

210 Borrowing ideas from Dezert d_{BI}^E in [28], for each element x of the power-set
 $2^\Theta - \{\emptyset\}$ of the FoD Θ , the belief intervals of x denoted $[Bel_1(x), Pl_1(x)]$ and
 $[Bel_2(x), Pl_2(x)]$ can be calculated from BBAs: $m_1(\cdot)$ and $m_2(\cdot)$ respectively.
 According to the theory of evidence, the width of such interval $[Bel(x), Pl(x)]$
 represents the degree of uncertainty for the corresponding focal element x .
 215 Therefore, X_i and Y_i in Eq.(10) are obtained which refer to the index of the
 width of the interval for each focal elements whose value is sorted by increasing
 order. Steps of this mechanism are illustrated as follows:

- Step 1:

$$\text{For } x = \theta_1, [Bel_1(\theta_1), Pl_1(\theta_1)] = [0.3, 0.6], [Bel_2(\theta_1), Pl_2(\theta_1)] = [0.4, 0.7];$$

220 $\text{For } x = \theta_2, [Bel_1(\theta_2), Pl_1(\theta_2)] = [0, 0.4], [Bel_2(\theta_2), Pl_2(\theta_2)] = [0, 0.4];$

$$\text{For } x = \theta_3, [Bel_1(\theta_3), Pl_1(\theta_3)] = [0, 0.5], [Bel_2(\theta_3), Pl_2(\theta_3)] = [0.1, 0.6];$$

$$\text{For } x = \theta_4, [Bel_1(\theta_4), Pl_1(\theta_4)] = [0, 0.5], [Bel_2(\theta_4), Pl_2(\theta_4)] = [0, 0.4];$$

- Step 2: The parameter l denotes the width of belief interval:

$$l_1(\theta_1) = Pl_1(\theta_1) - Bel_1(\theta_1) = 0.3, l_1(\theta_2) = 0.4, l_1(\theta_3) = 0.5, l_1(\theta_4) = 0.5;$$

225 $l_2(\theta_1) = 0.3, l_2(\theta_2) = 0.4, l_2(\theta_3) = 0.5, l_2(\theta_4) = 0.4;$

- Step 3: X_1 and Y_1 is the index of focal element whose ς is sorted by
 increasing order: $X_1 = \{1, 2, 3, 3\}$ and $Y_1 = \{1, 2, 3, 2\}$;

- Step 4: Sim_{seq} is calculated based on Eq.(10).

3.3. The hybrid similarity measure

230 To consider the influence of the distance of evidence, here based on the d_{BI}^E Eq.(8), we propose a new hybrid similarity measure which is presented as follows:

$$Sim_H(m_1, m_2) = w_1 \cdot d_{BI}^E(m_1, m_2) + w_2 \cdot F(Sim_{seq}(m_1, m_2)) \quad (12)$$

here, $w_1 = w_2 = 0.5$. and $F(\cdot)$ is the decreasing function within the interval $[0, 1]$ which in this paper $F(\cdot) = 1 - x^2$. As a convex combination of true
235 distance metrics, Sim_H is still a true distance metric that we use as a measure of similarity of two BBAs..

Additionally, to consider normalization of equation (12), it can be re-written as follows:

$$Sim_H^*(m_t, m_a) = w_1 \cdot (d_{BI}^{E*}(m_t, m_a)) + w_2 \cdot (F(Sim_{seq}^*(m_t, m_a))) \quad (13)$$

where t is the **TARGET** evidence and a is the **APPROXIMATE** evidence;

$$240 \quad d_{BI}^{E*}(m_t, m_a) = \frac{d_{BI}^E(m_t, m_a) - \min(\mathbf{vector}_1)}{\max(\mathbf{vector}_1) - \min(\mathbf{vector}_1)},$$

$$Sim_{seq}^*(m_t, m_a) = \frac{Sim_{seq}(m_t, m_a) - \min(\mathbf{vector}_2)}{\max(\mathbf{vector}_2) - \min(\mathbf{vector}_2)},$$

$$\mathbf{vector}_1 = (d_{BI}^E(m_t, m_1), d_{BI}^E(m_t, m_2), \dots, d_{BI}^E(m_t, m_j)),$$

$\mathbf{vector}_2 = (Sim_{seq}(m_t, m_1), Sim_{seq}(m_t, m_2), \dots, Sim_j(m_t, m_j))$ and the parameter j is defined in the Sec.4 which represents the number of approximate

245 BBAs.

4. Multi-Objective Evolutionary Algorithm Based on two-dimensional criteria

In this section, we regard PT as a general multi-objective problem consisting of two objectives and is also involved in a number of inequality and equality
250 constraints. Then a corresponding optimization model is proposed for selecting the *best* Bayesian BBA in the set of candidates.

4.1. Multiple-Objective Evolutionary-Based Probabilistic transformation

The idea to approximate any BBA into a Bayesian BBA (i.e. a subjective probability measure) using the minimization of the Shannon entropy under compatibility constraints has been proposed recently by Han et al. in [14, 22] using *on the shelf* optimization techniques. In this paper, we present in details a new optimization method to achieve this PT based on a random evolutionary algorithm to acquire minimization of the new aggregation criteria. And this new comprehensive criteria represents different aspects of information in BBAs, for example, *conflict coefficient* represent the degree of similarity in *conflict* between transformed BBAs and original BBAs (in other words, the more conflicts exist between two BBAs, the less similarity they have); Also, d_{BI}^E represent the interval distance between original BBAs and transformed ones.

Let's assume that the FoD of the original BBA $m(\cdot)$ to approximate by a Bayesian BBA is $\Theta \triangleq \{\theta_1, \theta_2, \dots, \theta_N\}$. The MOEPT method consists of the following steps which are derived from GA:

- Step 0 (setting parameters): t_{\max} is the max number of iterations; n_{\max} is the population size in each iteration; P_s is the selection probability, P_c is the crossover probability, and P_m is the mutation probability.
- Step 1 (population generation/encoding mechanism): A set \mathbf{P}_t of $j = 1, 2, \dots, n_{\max}$ random probability values $P_t^j = \{P^j(\theta_1), \dots, P^j(\theta_N)\}$ is generated ⁴ such that the constraints Eq.(14)-(16) for $j = 1, 2, \dots, n_{\max}$ are satisfied in order to make each random set of probabilities P_t^j compat-

⁴The lower (Bel) and upper (Pl) limits of each focal element are calculated using Eq.(2) and (3) based on the value of $m(\cdot)$.

ible with the original or target BBA $m(\cdot)$ to approximate, that is

$$P^j(\theta_i) \in [0, 1], \quad i = 1, 2, \dots, N \quad (14)$$

$$\sum_{i=1}^N P^j(\theta_i) = 1 \quad (15)$$

$$Bel(\theta_i) \leq P^j(\theta_i) \leq Pl(\theta_i), \quad i = 1, 2, \dots, N \quad (16)$$

- 270 • Step 2 (fitness assignment): For each probability set P_t^j , ($j = 1, 2, \dots, n_{\max}$), we compute its fitness value F based on Eq.(13). More precisely, one takes $F(P_t^j) = Sim_H(m(\cdot), P_t^j)$.
- Step 3 (best approximation of $m(\cdot)$): the best set of probability $P_t^{j_{\text{best}}}$ with minimum value of fitness is sought, and its associated index j_{best} is stored respectively in *Best Individual* and *Index of BestIndividual*.
275
- Step 4 (selection, crossover and mutation): The tournament selection, crossover and mutation operators drawn from evolutionary theory framework [33] are implemented to create the associated offspring population \mathbf{P}'_t based on the parent population \mathbf{P}_t . If $F(P_t^{j_{\text{best}}}) \leq F(P_t'^{j_{\text{best}}})$, then the *Best Individual* remains unchanged; otherwise, $Best\ Individual = P_t'^{j_{\text{best}}}$.
280
- Crossover operator: The crossover operator is one of the most important operator in the genetic algorithm. The crossover operation is conducted for the selected pairs of individuals (which are randomly chosen by the roulette wheel method). The feasibility condition of each individual is described as follows: the value of each subsegment must be between 0 and 1; the summation of the individuals should be 1. Although the initial population is made in a way that all individuals are feasible and correct, using the standard crossover operators leads to defective sub-segments which a normalization procedure is needed for such situation. Consider the following two individuals as parents: $X = (0.1, 0.2, 0.3, |0.4)$ and $Y = (0.2, 0.2, 0.1, |0.5)$ ⁵. With
285

⁵Here, this vertical bar represents the intersection point in crossover operator.

the single-point classic crossover operator, the following offspring will be produced: $X' = (0.1, 0.2, 0.3, 0.5)$ and $Y' = (0.2, 0.2, 0.1, 0.4)$, where $\sum_{j=1}^4 X'_j$ is equal to 1.1, which is greater than one and $\sum_{j=1}^4 Y'_j$ is equal to 0.9, which is less than one. Therefore, X', Y' have defective values which a normalization factor is needed as follows:

$$X'' = \frac{X'}{\sum_{j=1}^4 X'_j} = (0.1/1.1, 0.2/1.1, 0.3/1.1, 0.5/1.1),$$

$$Y'' = \frac{Y'}{\sum_{j=1}^4 Y'_j} = (0.2/0.9, 0.2/0.9, 0.1/0.9, 0.4/0.9).$$

– Mutation operator: The mutation operator randomly alters the value of a sub-segment. After applying the mutation operator, the normalization of the changed individuals is required. The normalization will be done in a similar way as the crossover operator.

- Step 5 (Stopping MOEPT) The steps 1–4 illustrate the t -th iteration of MOEPT method. If $t \geq t_{\max}$ then MOEPT method is completed, otherwise another iteration must be done by taking $t + 1 = t$ and going back to step 1.

4.2. Convergence Analysis

In order to mathematically prove the feasibility of MOEPT, convergence analysis of our algorithm is given. Firstly, we give a simplified description of the algorithm and also symbolic representation for simplicity.

- Encoding Mechanism: the size of population is n_{max} ; the length of individual (chromosome) is N and the initial population is \mathbf{P}_1 .
- Retain the best individual directly for the next generation;
- Randomly select the other non-optimal individuals in \mathbf{P}_t to crossover so as to form the intermediate population \mathbf{Y}_t ;
- The population \mathbf{Y}_t is mutated to form population \mathbf{V}_t ;
- The better individuals in the population \mathbf{V}_t are selected as the new generation population \mathbf{P}_t .

Algorithm 1 Multi-Objective Evolutionary-Based PT (MOEPT)

- 1: Define Stopping Criteria, ($t \leq t_{\max}$); population Size n_{\max} for each iteration; crossover probability P_c , mutation probability P_m and selection probability P_s .
 - 2: Generate an initial random population \mathbf{P}_t of consistent probabilities P_t^j with $m(\cdot)$.
 - 3: For each individual P_t^j in \mathbf{P}_t do
 - 4: Calculate Fitness $F(P_t^j) = Sim_H(m(\cdot), P_t^j)$ of P_t^j
 - 5: Store the best individual $P_t^{j_{\text{best}}}$
 - 6: End
 - 7: Repeat:
 - 8: Crossover: exchange parts of individuals with probability P_c
 - 9: Mutation: mutate the child individuals with probability P_m
 - 10: Selection: Select individuals based on fitness according to P_s
 - 11: After these three sub-steps, the updated population \mathbf{P}'_t is obtained
 - 12: Calculate the fitness of individuals of \mathbf{P}'_t , and store the best individual $P_t'^{j_{\text{best}}}$
 - 13: If $F(P_t'^{j_{\text{best}}}) \leq F(P_t^{j_{\text{best}}})$
 - 14: Best-Individual remains unchanged
 - 15: else
 - 16: Best-Individual = $P_t'^{j_{\text{best}}}$
 - 17: If $t \geq t_{\max}$ then stops, otherwise $t + 1 \rightarrow t$ and go back to line 7
-

Specifically, three operators (crossover operator, mutation operator and selection operator) can be described by transition probability as follows:

- crossover operator: For the single point crossover, new individual k is produced based on their parents: individual i and j :

$$P_C^t(i \times j, k) = \begin{cases} |k|p_c/N, k \neq i, j \\ (1 - p_c) + |k|p_c/N, k = i \end{cases} \quad (17)$$

where $|k|$ is the number of individual k ; $0 \leq p_c \leq 1$ is the crossover probability; a is the minimum probability for individual $|k|$.

$$a = 1 - p_c + p_c/N. \quad (18)$$

- mutation operator:

$$P_M^t(i, j) = p_m^{d(i,j)}(1 - p_m)^{N-d(i,j)} \quad (19)$$

where $0 \leq p_m \leq 1$ is the mutation probability; $d(i, j)$ is the Hamming distance between i and j ; b is the minimum probability.

$$b = (1 - p_m)^N. \quad (20)$$

- selection operator: MOEPT uses the strategy of retaining the elite and the best individual is retained for the next generation which do not participate in the competition. Assuming that m individuals are selected based on the following equations:

$$P_S^t(\mathbf{P}_t, P_t^j) = \frac{\sigma_n(F(P_t^j))}{\sum_{k=1}^{n_{max}} (F(P_t^k))}, j \in \mathbf{P}_t, n = 1, 2, \dots \quad (21)$$

where σ_n represents a increasing scale function. Besides, the probability of selecting the first individual in the next generation population is

$$P_S^{t*}(\mathbf{P}_t, P_t^j) = \frac{|P_t^j|}{|B(\mathbf{P}_t)|}, P_t^j \in \mathbf{P}_t. \quad (22)$$

where $|P_t^j|$ is the number of individual P_t^j in \mathbf{P}_t and $B(\mathbf{P}_t)$ is the cardinality of optimal set of \mathbf{P}_t .

In order to facilitate the convergence analysis, the changing process of fitness value $F(P_t^j)$ is regarded as a Markov chain. If MOEPT obtains the best individual $P_t^{j_{best}}$ in t generation, we can denote this as $\{\hat{F}(P_t)\} = P_t^{j_{best}}$. Then all the other populations in $t + 1$ generations will also reach the best fitness value due to the Elite strategy [34]. Therefore, markov chain $\{\hat{F}(P_t)\}$ constitute the lower martingale. According to the properties of the lower martingale and the convergence theorem of the lower martingale [35], convergence analysis of MOEPT is converted into the convergence of $\{\hat{F}(P_t)\}$. The following three theorems are given, in which theorem 4 is to prove that $\{\hat{F}(P_t)\}$ satisfy the conditions of martingale theorem; Theorem 5 proves the global convergence of MOEPT and theorem 6 constructs three conditions of the convergence of the lower martingale so that the optimal solution can be almost obtained everywhere.

Theorem 4. *The process of describing values of fitness functions in MOEPT is a non-bounded martingale:*

$$E\{\hat{F}(\mathbf{P}_{t+1})/\mathbf{P}_t\} \geq \hat{F}(\mathbf{P}_t) \quad (23)$$

Proof: Because the algorithm retains the maximum fitness value of the previous generation to the next generation, and does not participate in the genetic operation, the best individual mode is not destroyed, so the maximum fitness value of the next generation population will not be less than the maximum fitness value of previous generation.

$$E\{\hat{F}(\mathbf{P}_{t+1})/\mathbf{P}_t\} \geq \hat{F}(\mathbf{P}_t) > 0 \quad (24)$$

Theorem 5. *MOEPT converges to the global optimal solution based on probability, which is mathematically expressed by the condition [Give reference]*

Proof: When population \mathbf{P}_t is updated to t generation and the minimum or best fitness is recorded as $P_t^{j_{best}}$ and the global optimal solution is noted as F^* , assuming that MOEPT can converge to global optimal solution at t generation, that is

$$\{\hat{F}(\mathbf{P}_t)\} = F^* \quad (25)$$

Based on Theorem 4:

$$E\{\hat{F}(\mathbf{P}_{t+1})/\mathbf{P}_t\} = F^* \quad (26)$$

Defined by conditional expectation:

$$\begin{aligned} E\{\hat{F}(\mathbf{P}_{t+1})/\mathbf{P}_t\} &= \\ &\sum_{i,j \in \mathbf{P}_t} P_C^t(i \times j, y) \sum_v P_M^t(y, v) \sum_k P_S^t(v, k) \hat{F}(k) \geq \\ &\sum_{i,j \in \mathbf{P}_t} P_C^t(i \times j, i) \sum_v P_M^t(y, v) \sum_k P_S^t(v, k) \hat{F}(k) \geq \\ &a \sum_v P_M^t(y, v) \sum_k P_S^t(v, k) \hat{F}(k) \geq \\ &a \sum_v P_M^t(y, y) \sum_k P_S^t(v, k) \hat{F}(k) \geq \\ &ab^m \left\{ \sum_{k \in B(\mathbf{P}_t)} [P_S^t(v, k) - P_S^{t*}(v, k)] \hat{F}(k) + \sum_{k \in B(\mathbf{P}_t)} P_S^{t*}(v, k) \hat{F}(k) \right\} \end{aligned}$$

When $k \notin B(\mathbf{P}_t)$, $P_S^{t*}(v, k) = 0$ and when $k \in B(\mathbf{P}_t)$, $\hat{F}(k) = F^*$. So that $E\{\hat{F}(\mathbf{P}_{t+1})/\mathbf{P}_t\}$ can be rewritten as

$$\begin{aligned} E\{\hat{F}(\mathbf{P}_{t+1})/\mathbf{P}_t\} &\geq \\ &ab^m \sum_{k \in B(\mathbf{P}_t)} [P_S^t(v, k) \hat{F}(k) + F^*] \geq ab^m F^*. \end{aligned}$$

So we obtain

$$ab^m F^* \leq F^*. \quad (27)$$

Because $F^* > 0$, one gets

$$ab^m \leq 1. \quad (28)$$

Based on the above formula derivation, MOEPT converges to the global
340 optimal solution.

Theorem 6. $\forall n \geq 1$, the following conditions are satisfied:

- $E[\hat{F}(\mathbf{P}_1)] < \infty, F^* < \infty$,

- $E[\hat{F}(\mathbf{P}_t)/\mathbf{P}_{t-1}] = \hat{F}(\mathbf{P}_{t-1}) + c_{t-1}F^*$,
- $c_t \in [0, 1], \lim_{t \rightarrow \infty} \sum_{k=0}^{t-1} c_k = 1 - \frac{\hat{F}(\mathbf{P}_1)}{F^*}$,

345 Then that random sequence $\hat{F}(\mathbf{P}_t) \xrightarrow{a,s} F^*$

Proof: By taking the mathematical expectation on both sides of the condition (2) one has:

$$\begin{aligned} E[\hat{F}(\mathbf{P}_t)] &= E[\hat{F}(\mathbf{P}_{t-1})] + c_{t-1}F^* = \\ E[\hat{F}(\mathbf{P}_{t-2})] &+ c_{t-1}F^* + c_{t-2}F^* = \dots = \\ E[\hat{F}(\mathbf{P}_1)] &+ F^* \sum_{k=0}^{t-1} c_k. \end{aligned}$$

According to condition (1) and (3),

$$\begin{aligned} E[\hat{F}(\mathbf{P}_t)] &< E[\hat{F}(\mathbf{P}_1)] + F^* < \infty \\ \text{Sup}_t E[\hat{F}(\mathbf{P}_t)] &< \text{sup}_t E[\hat{F}(\mathbf{P}_1)] + \text{sup}_t F^* < \infty \end{aligned}$$

Because $\hat{F}(\mathbf{P}_t)$ is a non-bounded martingale, we have:

$$\hat{F}(\mathbf{P}_t) \xrightarrow{a,s} \hat{F}(\mathbf{P}_\infty) = \lim_{t \rightarrow \infty} \hat{F}(\mathbf{P}_t) \quad (29)$$

$$\begin{aligned} \lim_{t \rightarrow \infty} E[\hat{F}(\mathbf{P}_t)] &= \\ \lim_{t \rightarrow \infty} E[\hat{F}(\mathbf{P}_1)] &+ F^* \lim_{t \rightarrow \infty} \sum_{k=0}^{t-1} c_k = \\ E[\hat{F}(\mathbf{P}_1)] &+ F^* \left(1 - \frac{\hat{F}(\mathbf{P}_1)}{F^*}\right) = F^*, \end{aligned}$$

$$\hat{F}(\mathbf{P}_t) \xrightarrow{a,s} F^*. \quad (30)$$

5. Simulation Results

According to the first step of MOEPT, we initially set the related parameters
 350 as follows: $t_{max} = 50$, $n_{max} = 1000$, $P_s = 0.3$, $P_c = 0.5$, $P_m = 0.1$.

5.1. Simple Examples

Example 3: Let's consider the frame $\Theta = \{\theta_1, \theta_2, \theta_3, \theta_4\}$ and the corresponding BBA's is illustrated as follows:

$$\begin{aligned}
m(\theta_1) &= 0.16, m(\theta_2) = 0.14, m(\theta_3) = 0.01, m(\theta_4) = 0.02 \\
m(\theta_1 \cup \theta_2) &= 0.20, m(\theta_1 \cup \theta_3) = 0.09, m(\theta_1 \cup \theta_4) = 0.04 \\
m(\theta_2 \cup \theta_3) &= 0.04, m(\theta_2 \cup \theta_4) = 0.02, m(\theta_3 \cup \theta_4) = 0.01 \\
m(\theta_1 \cup \theta_2 \cup \theta_3) &= 0.10, m(\theta_1 \cup \theta_2 \cup \theta_4) = 0.03 \\
m(\theta_1 \cup \theta_3 \cup \theta_4) &= 0.03, m(\theta_2 \cup \theta_3 \cup \theta_4) = 0.03 \\
m(\Theta) &= 0.08
\end{aligned}$$

Based on the classical PTs respectively, the original BBA is transformed into their corresponding probabilities as illustrated in Table 1. Their corresponding Sim_H can be calculated using Eq.(13), which has been already listed in Table

355 1. Clearly, several interesting characteristics which are presented in Table 1 are deserved to be mentioned: (1) $MOEPT_{d_{BI}^E + Sim_{seq}}$ has the minimum value from the perspective of Sim_H criteria, which considers both aspects of d_{BI}^E and Sim_{seq} rather than concentrating on single aspect; (2) Comparing to other PTs, especially $MOEPT_{D_J + Sim_{seq}}$ ⁶, our method performs better than mentioned
360 methods.

Example 4: Let's consider another situation in the frame $\Theta = \{\theta_1, \theta_2, \theta_3, \theta_4\}$ and the corresponding BBA's is illustrated as follows:

$$\begin{aligned}
m(\theta_1) &= 0.16, m(\theta_2) = 0.16, m(\theta_3) = 0.16, m(\theta_4) = 0.16 \\
m(\theta_1 \cup \theta_2) &= 0.04, m(\theta_1 \cup \theta_3) = 0.04, m(\theta_1 \cup \theta_4) = 0.04 \\
m(\theta_2 \cup \theta_3) &= 0.04, m(\theta_2 \cup \theta_4) = 0.04, m(\theta_3 \cup \theta_4) = 0.04 \\
m(\theta_1 \cup \theta_2 \cup \theta_3) &= 0.03, m(\theta_1 \cup \theta_2 \cup \theta_4) = 0.03 \\
m(\theta_1 \cup \theta_3 \cup \theta_4) &= 0.03, m(\theta_2 \cup \theta_3 \cup \theta_4) = 0.03
\end{aligned}$$

⁶Here, to show the property of d_{BI}^E , we replace d_{BI}^E by D_J in MOEPT to make comparisons.

Table 1: Results of Different PTs in Example 3 ($w_1=w_2=0.5$).

	θ_1	θ_2	θ_3	θ_4	Sim_H
CuzzP	0.3860	0.3382	0.1607	0.1151	0.2800
BetP	0.3983	0.3433	0.1533	0.1050	0.2799
DSmP₀	0.5176	0.4051	0.0303	0.0470	0.1897
DSmP_{0.001}	0.5162	0.4043	0.0319	0.0477	0.1896
PrBP1	0.5419	0.3998	0.0243	0.0340	0.1918
PrBP2	0.5578	0.3842	0.0226	0.0353	0.1933
MOEPT_{D_J+Sim_{seq}}	0.3980	0.3322	0.1156	0.1541	0.1849
MOEPT_{d_{BI}^E+Sim_{seq}}	0.3985	0.3983	0.0623	0.1409	0.0733

Actually, the Example 4 is the extension of the case studied by Han in [14] which assumes a special scenario where no difference exists between $m(\theta_1)$, $m(\theta_2)$, $m(\theta_3)$, $m(\theta_4)$ and where the traditional PTs become invalid and give unreasonable results which can be seen in Table 2. The property of original BBA which no difference exists between $m(\theta_1)$, $m(\theta_2)$, $m(\theta_3)$, $m(\theta_4)$ is almost lost when classical PTs have been applied. When sequence is not considered in MOEPT, which is denoted as $MOEPT_{Distance}$, the feature of equal mass in original BBAs is also missing as other classical PTs. Fortunately, when information of “sequence” is added into objective function, MOEPT performs better in keeping the original information as expected.

Example 5: $\Theta = \{\theta_1, \theta_2, \theta_3, \theta_4\}$

To investigate the robustness of MOEPT from a statistical point of view, in this example, we randomly generate BBAs and compare MOEPT with classical PTs (BetP[6],[9], CuzzP [13], DSmP [10], PrBP1 and PrBP2 [26]). The original BBAs to approximate are generated according to Algorithm 2 of [36].

In our test, we have set the cardinality of the FoD to 4 and fixed the number of focal elements to $l = N_{max} = 15$. We randomly generate $L = 100$ BBA’s. Six PT methods are tested and Sim_H is used to evaluate the quality of their corresponding results are shown in Figure 1. As we can see and as naturally

Table 2: Results of Different PTs in Example 4.

	θ_1	θ_2	θ_3	θ_4	Sim_H
<i>BetP</i>	0.3983	0.3433	0.1533	0.1050	0.3974
<i>DSmP</i> ₀	0.2500	0.2500	0.2500	0.2500	0.5458
<i>PrBP</i> ₁	0.5419	0.3998	0.0243	0.0340	0.6368
<i>PrBP</i> ₂	0.5578	0.3842	0.0226	0.0353	0.6412
<i>MOEPT</i> _{Distance}	0.2500	0.1597	0.3578	0.2325	0.3415
MOEPT _{D_J+Sim_{seq}}	0.2483	0.2485	0.2496	0.2536	0.1708
MOEPT _{d_{BI}^E+Sim_{seq}}	0.2484	0.2488	0.2489	0.2539	0.0450

380 expected, the MOEPT outperforms significantly other methods based on minimum of Sim_H criterion, which is absolutely normal because the method has been developed to this aim.

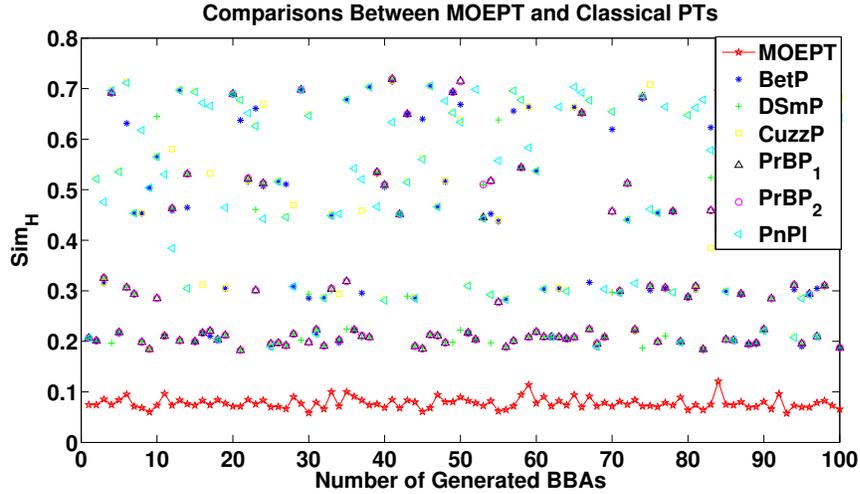


Figure 1: Comparisons Between MOEPT and state of art PTs.

5.2. Example of Pattern Classification Using MOEPT

In this example, we use the evaluation of decision making under the evidence theory framework to indirectly evaluate MOEPT. We consider seven classes of aircrafts which are illustrated in Figure 2. And the classifier used in this example

385

Algorithm 2 Random generation of BBA

- 1: Input: Frame of Discernment $\Theta = \{\theta_1, \theta_2, \theta_3, \theta_4\}$
 - 2: N_{max} :Maximum number of focal element
 - 3: Output : BBA-m
 - 4: Generate $K(\Theta)$, which is the power set of Θ
 - 5: Generate a random permutation of $K(\Theta) \rightarrow R(\Theta)$
 - 6: Generate an integer between 1 and $N_{max} \rightarrow l$
 - 7: For each First k elements of $R(\Theta)$ do
 - 8: Generate a value within $[0, 1] \rightarrow m_i, i = 1, \dots, l$
 - 9: End
 - 10: Normalize the vector $m = [m_1, m_2, \dots, m_l] \rightarrow m'$
 - 11: $m(\theta_i) = m'_i$
-

is the Probabilistic neural networks (PNNs). For each test example, the output of the classifier is represented by a BBA. The corresponding BBA for each test sample is generated according to our previous work [38]:



Figure 2: The binary image of seven kinds of airplanes.

- 390 1. Firstly, the image is preprocessed with binarization and then multiple features are extracted, such as Hu moments, normalized moment of inertia, affine invariant moments, discrete outline parameters and singular values. Secondly, five BBAs can be assigned to the evidence sources by each PNN
- 395 ⁷. Thirdly, all these five BBAs are fused by PCR6 [8] to form a single BBA $m(\cdot)$.
2. For the two classes t_1 and t_2 ($t_1, t_2 \in 1, 2, 3, \dots, 7, t_1 \neq t_2$), with the top two values of $m(i), i = 1, 2, 3, \dots, 7$, the corresponding updated mass

⁷Specifically, transfer functions in five PNNs are set to Gaussian function; the weighting function is set to the Euclidean distance; the input function is netprod and also the output function is compet.

assignments are generated according to [39]

$$m'(i) = m(i), \forall i = t_1, t_2 \quad (31)$$

The remaining mass is assigned to the total set Θ

$$m'(\Theta) = 1 - m'(t_1) - m'(t_2). \quad (32)$$

For example, for a test sample *target-1*, we obtain the corresponding BBA from PNNs: $m(1) = 0.7, m(2) = 0.05, m(3) = 0.2, m(4) = 0.01, m(5) = 0, m(6) = 0.02, m(7) = 0.02$. The dominant class is class 1 and class 3 is at the second place. The updated corresponding BBA is $m'(1) = 0.7, m'(3) = 0.2, m'(2, 4, 5, 6, 7) = 0.1$.

There are 100 samples for each one class with a total of 700 samples. For each class, 50 samples are randomly selected for training PNNs and the remaining samples are used for testing. For MOEPT, the decision result will be class t_{final} if

$$t_{final} = \operatorname{argmax}(MOEPT) \quad (33)$$

As we can see from Figure.3, MOEPT performs well in this task of pattern classification.

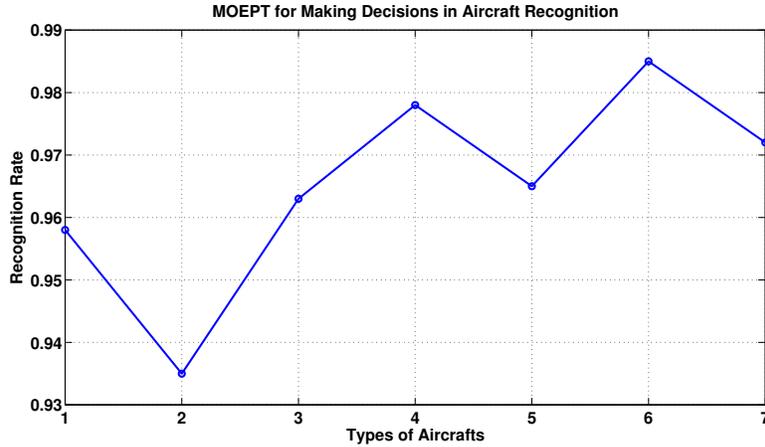


Figure 3: Recognition Rate of MOEPT.

5.3. Example of Target Type Tracking Using MOEPT

To further discuss the interest of the proposed MOEPT, Target Type Tracking (TTT) problem in the area of decision making has been used which is briefly described below [40]:

410 5.3.1. Target Type Tracking Problem (TTT)

1. Considering $\zeta = 1, 2, \dots, \zeta_{max}$ be the time index and let N possible target types $Tar_{\zeta} \in \Theta = \{\theta_1, \theta_2, \dots, \theta_N\}$ in the surveillance area; For instance, in the normal air target surveillance systems the FoD could be $\Theta = \{Fighter, Cargo\}$. That is, $Tar_1 = \theta_1 \triangleq Fighter$, $Tar_2 = \theta_2 \triangleq$
415 $Cargo$. Similarly, the FoD in a ground target surveillance systems could be $\Theta_{ground} = \{Tank, Truck, Car, Bus\}$. In this paper, we just consider the air target surveillance systems to prove the practicability of EPT.
2. At every time ζ , the true type of the target $Tar(\zeta) \in \Theta$ is immediately observed by an attribute-sensor (here, we assume a possible target prob-
420 ability).
3. A defined classifier is applied to process the attribute measurement of the sensor which provides the probability $Tar_d(\zeta)$ on the type of the observed target at each instant ζ .
4. The sensor is in general not totally reliable and is characterized by an $N \times N$ confusion matrix:

$$\mathbf{M} = [M_{ij} = P(Tar_d = Tar_j | TrueType = Tar_i)] \quad (34)$$

where $0 \leq i \leq N; 0 \leq j \leq N$.

425 Here, we briefly summarize the main steps of TTT using MOEPT.

1. Initialization. Determine the target type frame $\Theta = \{\theta_1, \theta_2, \dots, \theta_N\}$ and set the initial BBA $m^{initial}(\theta_1 \cup \theta_2 \cup \dots \cup \theta_N) = 1$ since there is no information about the first target type that will be observed;
2. Updating BBA. An observed BBA $m_{obs}(\cdot)$ on types of unknown observed
430 target is defined from current target type declaration and confusion matrix \mathbf{M} ;

3. Combination. We combine the current BBA $m_{obs}(\cdot)$ with initial BBA $m^{initial}(\cdot)$ according to PCR6 combination rule [8]: $m_{PCR6}(\cdot) = m_{obs}(\cdot) \oplus m^{initial}(\cdot)$;
- 435 4. Approximation. Using $MOEPT(\cdot)$ to approximate $m_{PCR6}(\cdot)$ into a Bayesian BBA;
5. Decision Making. Taking a final decision about the type of the target at current observation time based on the obtained Bayesian BBA;
6. Updating BBA. Set $m^{initial}(\cdot) = m_{PCR6}(\cdot)$, and increase time index $\zeta =$
- 440 $\zeta + 1$ and go back to step 2.

5.3.2. Raw Dataset of TTT

We have tested our MOEPT-based TTT on a very simple scenario for a 2D TTT, namely $\Theta = \{Fighter, Cargo\}$ for two types of classifiers. The matrix \mathbf{M}_1 corresponds to the confusion matrix of the good classifier, and \mathbf{M}_2 corresponds to the confusion matrix of the poor classifier.

$$\mathbf{M}_1 = \begin{bmatrix} 0.95 & 0.05 \\ 0.05 & 0.95 \end{bmatrix}; \mathbf{M}_2 = \begin{bmatrix} 0.75 & 0.25 \\ 0.25 & 0.75 \end{bmatrix} \quad (35)$$

In our scenario, a true Target Type sequence over 120 scans is generated according to Figure 4. We can observe clearly from Figure 4 that Cargo (which is denoted as Type 2) appears at first in the sequence, and then the observation

445 of the Target Type switches three times onto Fighter Type (Type 1) during different time durations (namely, 20s, 10s, 5s).

A pathological case for TTT: Our analysis has shown that MOEPT can nevertheless be in troubles for tracking two target types as proved in this simple particular example (when $0 \leq m(\theta_1 \cup \theta_2) \leq 0.1$). Let's consider the following BBA

$$m_{target}(\cdot) = [\theta_1, \theta_2, \theta_1 \cup \theta_2] = [0, 1, 0]$$

According to the compatibility constraints Eq.(14)–Eq.(16), the population \mathbf{P}'_t is obtained from \mathbf{P}_t through a selection procedure. Next, individual $P'_t{}^j$ in

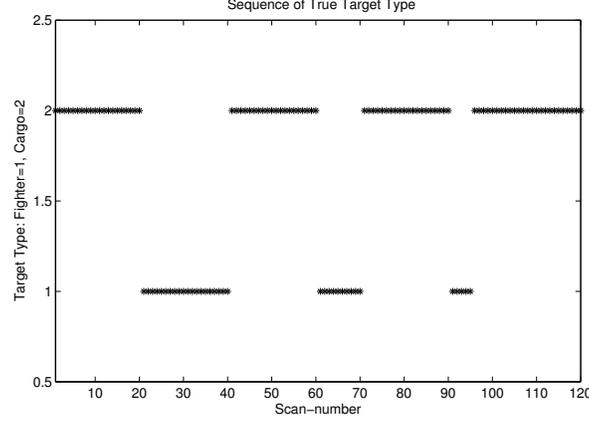


Figure 4: Raw Sequence of True Target Type.

\mathbf{P}'_t which is denoted as $P_t^{ij} = [m'(\theta_1), m'(\theta_2)]$ is subject to initial constraint Eq.(1) and Eq.(36):

$$\begin{aligned}
 m'(\theta_1) &\geq (Bel(\theta_1) = m(\theta_1) = 0) \\
 m'(\theta_1) &\leq (Pl(\theta_1) = m(\theta_1) + m(\theta_1 \cup \theta_2) = 0 + 0 = 0); \\
 m'(\theta_2) &\geq (Bel(\theta_2) = m(\theta_2) = 1) \\
 m'(\theta_2) &\leq (Pl(\theta_2) = m(\theta_2) + m(\theta_2 \cup \theta_1) = 1 + 0 = 1);
 \end{aligned} \tag{36}$$

From the above inequalities, one sees that only one probability measure $P_t^S = [m(\theta_1), m(\theta_2)] = [0, 1]$ (where the superscript index S means *Single*) satisfies this constraint⁸. However because of mechanism of MOEPT Eq.(14)-(16), P_t^j in population \mathbf{P}_t which are randomly generated in the interval $[Bel(\theta_i), Pl(\theta_i)]$, $i = 1, 2, \dots, N$, will be unable to generate enough candidates for evolutionary computation⁹. That is why MOEPT becomes inefficient in this case which occurs

⁸the constraint is $m(\theta_1) \in [Bel(\theta_1), Pl(\theta_1)] = [0, 0], m(\theta_2) \in [Bel(\theta_2), Pl(\theta_2)] = [1, 1]$.

⁹A sufficient number of candidate sets are prerequisites for ensuring the global optimization performance of evolutionary algorithms.

with a probability of $1/n_{\max}$, where n_{\max} denotes the size of population¹⁰ \mathbf{P}_t . Unfortunately, in TTT decision making problems, such case cannot be avoided
455 because it can really happens.

To circumvent this problem and make MOEPT approach working in most circumstances, we need modify a bit MOEPT method to generate enough individuals for making selection steps efficient when the bounds of belief interval
460 $[Bel, Pl]$ take their min and max values ($[0.9, 0.05, 0.05]$, $[0.05, 0.9, 0.05]$). For achieving this, we propose to enlarge this particular interval through a parameter λ , and maintain the property of original interval in some degree at the same time. More precisely, the modified belief interval, denoted $[Bel', Pl']$, is heuristically computed by a simple thresholding technique as follows:
465 First, we assume that the original BBA we consider here for FoD $\Theta = \{\theta_1, \theta_2\}$ is $[\theta_1, \theta_2, \theta_1 \cup \theta_2] = [a, b, c]$, with $(a + b + c) = 1$ and with $0 \leq c \leq 0.1$

Step 1: Let $m'(\theta_1 \cup \theta_2) = c + \lambda$;

Step 2: if $a > b$

$$m'(\theta_1) = a - \lambda; m'(\theta_2) = b; m'(\theta_1 \cup \theta_2) = c + \lambda; \quad (37)$$

Step 3: if $a \leq b$

$$m'(\theta_1) = a; m'(\theta_2) = b - \lambda; m'(\theta_1 \cup \theta_2) = c + \lambda; \quad (38)$$

So the value of $[Bel'(\theta_1), Pl'(\theta_1)]$ and $[Bel'(\theta_2), Pl'(\theta_2)]$ can be calculated based on Eq.(37),Eq.(38), which are presented as follows:

When $a > b$:

$$\begin{cases} Pl'(\theta_1) = m(\theta_1) + m'(\theta_1 \cup \theta_2) = a - \lambda + c + \lambda = a + c; \\ Bel'(\theta_1) = 1 - Pl'(\bar{\theta}_1) = 1 - (b + c + \lambda) = a - \lambda. \end{cases} \quad (39)$$

¹⁰In our simulation, we did take $n_{\max} = 1000$.

$$\begin{cases} Pl'(\theta_2) = m(\theta_2) + m'(\theta_1 \cup \theta_2) = b + c + \lambda = b + c + \lambda; \\ Bel'(\theta_2) = 1 - Pl'(\bar{\theta}_2) \\ \quad = 1 - (a - \lambda + c + \lambda) = 1 - (a + c) = b. \end{cases} \quad (40)$$

When $a \leq b$:

$$\begin{cases} Pl'(\theta_1) = m(\theta_1) + m'(\theta_1 \cup \theta_2) = a + c + \lambda; \\ Bel'(\theta_1) = 1 - Pl'(\bar{\theta}_1) \\ \quad = 1 - (b - \lambda + c + \lambda) = 1 - (b + c) = a. \end{cases} \quad (41)$$

$$\begin{cases} Pl'(\theta_2) = m(\theta_2) + m'(\theta_1 \cup \theta_2) = b - \lambda + c + \lambda = b + c; \\ Bel'(\theta_2) = 1 - Pl'(\bar{\theta}_2) = 1 - (a + c + \lambda) = b - \lambda. \end{cases} \quad (42)$$

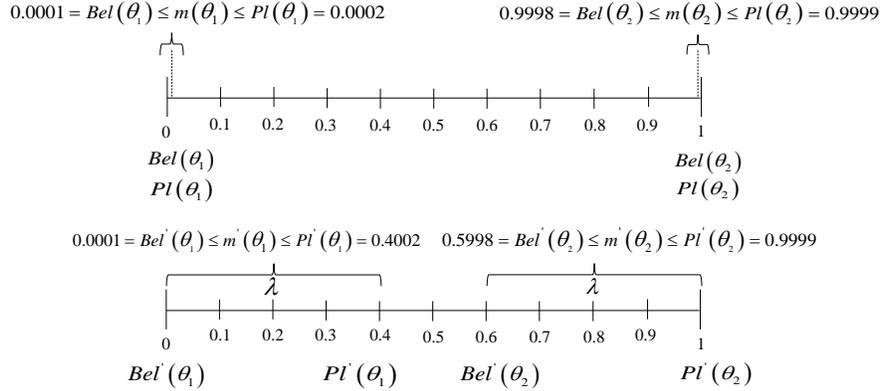


Figure 5: The principle of modified-constraint MOEPT ($\lambda = 0.4$).

Explanation: Through step 1, one computes the total singleton mass one has in the entire BBA and the threshold value 0.9 allows to evaluate if the percentage of singleton mass is big enough or not. Here, we not only consider the
470 unique extreme case $m_{target}(\cdot) = [\theta_1, \theta_2, \theta_1 \cup \theta_2] = [0, 1, 0]$, but also other possible cases such as $m_{target}(\cdot) = [\theta_1, \theta_2, \theta_1 \cup \theta_2] = [0.0001, 0.9998, 0.0001]$. Why

do we consider the concept of percentage? Actually, the higher percentage of singleton mass, the smaller interval for P_t^j , in other words, the higher value of $m(\theta_1 \cup \theta_2)$, the bigger interval for P_t^j which can be shown in Eq.(36); The step 2 and step 3 give the way of calculating the updated upper bound of belief interval $[Bel', Pl']$ and Eq.(39)–Eq.(42) prove that the parameter λ determines the range of the interval; Next, we give two examples to show how the above method works:

480

The pathological case one for TTT (revisited with modified MOEPT)

$$m_{target}(\cdot) = [\theta_1, \theta_2, \theta_1 \cup \theta_2] = [0.0001, 0.9998, 0.0001].$$

Here, the parameter λ is arbitrarily¹¹ set to 0.4. Then one computes in step 2 the modified plausibility bounds $Bel'(\theta_1) = 0.0001$, $Pl'(\theta_1) = 0.0001 + 0.0001 + \lambda = 0.4002$ and $Bel'(\theta_2) = 0.9998 - 0.4 = 0.5998$, $Pl'(\theta_2) = 0.9999$. So we get $[Bel'(\theta_1), Pl'(\theta_1)] = [0.0001, 0.4002]$ and $[Bel'(\theta_2), Pl'(\theta_2)] = [0.5998, 0.9999]$. The relationship between original interval $[Bel, Pl]$ and the updated interval $[Bel', Pl']$ is illustrated in Figure 5.

485

Consequently, any Bayesian BBA $P_t^j = [m'(\theta_1), m'(\theta_2)]$ must be generated according the (modified) compatibility constraints

$$\begin{aligned} m'(\theta_1) &\in [Bel'(\theta_1), Pl'(\theta_1)] = [0.0001, 0.4002] \\ m'(\theta_2) &\in [Bel'(\theta_2), Pl'(\theta_2)] = [0.5998, 0.9999] \end{aligned}$$

The pathological case two for TTT (revisited with modified MOEPT)

$$m_{target}(\cdot) = [\theta_1, \theta_2, \theta_1 \cup \theta_2] = [0.45, 0.48, 0.07].$$

Here, the parameter λ is set to 0.2. Then any Bayesian BBA $P_t^j = [m'(\theta_1), m'(\theta_2)]$

¹¹The value of the parameter λ can be chosen to any value in $[0, 1]$ by the designer for his/her own reason to ensure the alternative interval effectively in modified MOEPT version.

must be generated according the (modified) compatibility constraints

$$m'(\theta_1) \in [Bel'(\theta_1), Pl'(\theta_1)] = [0.45, 0.72]$$

$$m'(\theta_2) \in [Bel'(\theta_2), Pl'(\theta_2)] = [0.28, 0.55]$$

In order to evaluate the influence of the parameter λ , we have reexamined all the pathological cases based on the following procedure:

1. The value of parameter λ is taken to five possible values: 0, 0.1, 0.2, 0.3,
490 0.4, 0.5;
2. We randomly generate initial population \mathbf{P}_t based on λ , which is also subject to the constraints (14)–(16).

With this simulation, we can observe in Figure 6 and Figure 7 the impact of λ value on the number of P_t^j in \mathbf{P}_t . When we set $\lambda = 0$ ¹², there exists
495 no suitable P_t^j for case one which demonstrates the necessity to circumvent the pathological case problem. Obviously, the number of P_t^j increases with the increase of λ value, which efficiently proves the advantage of using the modified MOEPT approach to make selection step of the evolutionary algorithm more efficient. One point we need to clarify is that the intervals i.e. $[Bel'(\theta_1), Pl'(\theta_1)]$,
500 $[Bel'(\theta_2), Pl'(\theta_2)]$ induced from parameter λ above aims at guaranteeing enough number of P_t^j in \mathbf{P}_t in the implementation of MOEPT. Another point we also need to mention is that number of P_t^j in \mathbf{P}_t is not influenced by the weight¹³ which in some degree guarantee the implementation of MOEPT.

5.3.3. Simulation Results of TTT Based on Modified MOEPT

505 Our simulation consists in 100 Monte-Carlo runs and we show in the sequel the averaged performances of MOEPT. The figures Figure 8 and 9 illustrate the Bayesian BBA's obtained by our new MOEPT method for solving TTT using PCR6 fusion rule. One sees that regardless of the good classifier \mathbf{M}_1

¹²which actually the original MOEPT is applied.

¹³Here, Weight equals w_2 in Eq.(13). And thus, $w_1 = 1 - Weight$.

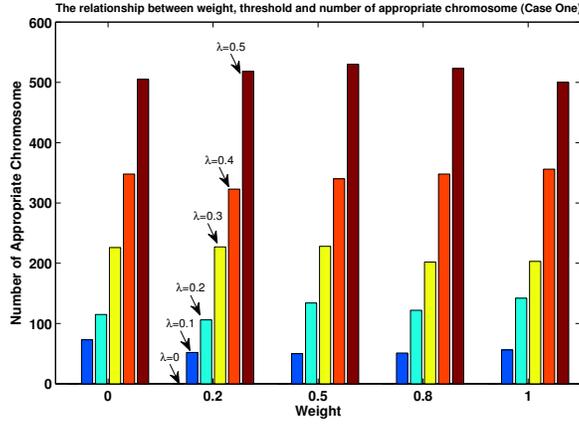


Figure 6: Impact of λ (x-axis) on individuals in \mathbf{P}_t (y-axis).

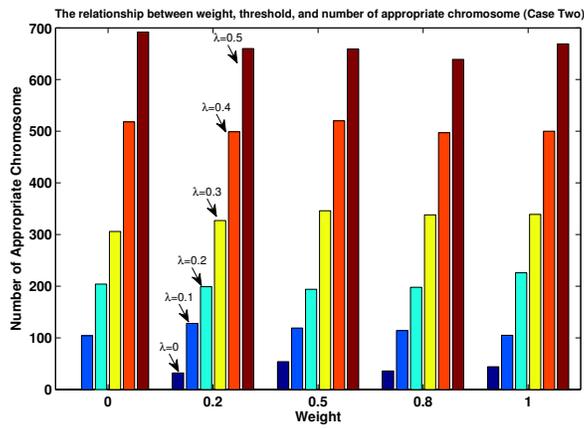


Figure 7: Impact of λ (x-axis) on individuals in \mathbf{P}_t (y-axis).

(Recognition rate: 90.83 %) and poor classifier \mathbf{M}_2 (Recognition rate: 80.83%),

510 MOEPT is able to track properly the quick changes of target type.

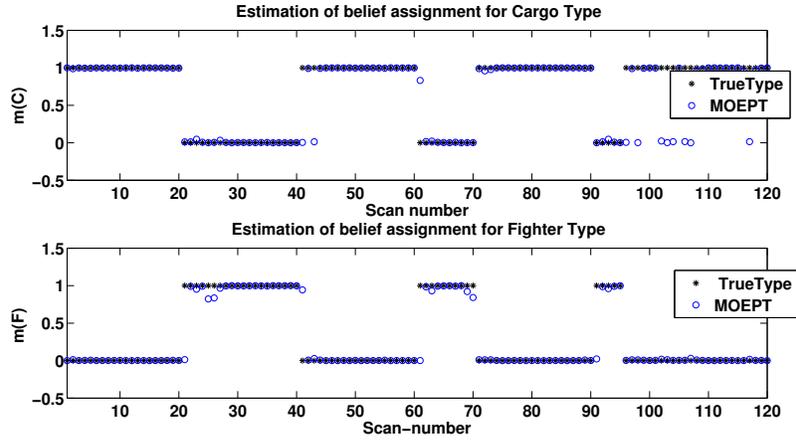


Figure 8: Result of MOEPT for Cargo and Fighter types Using M_1 .

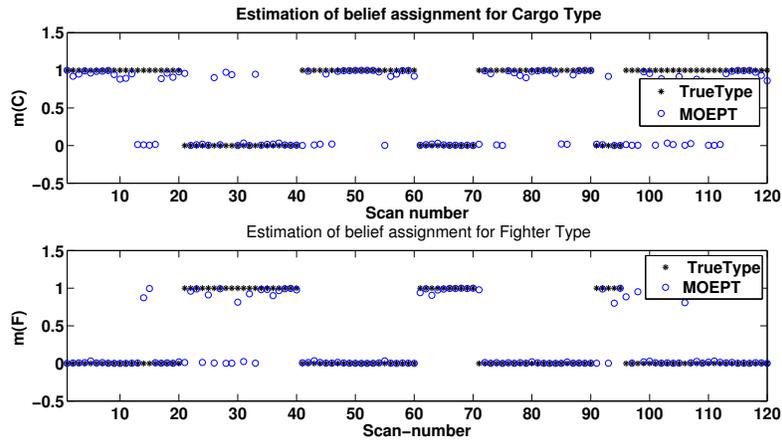


Figure 9: Result of MOEPT for Cargo and Fighter types Using M_2 .

6. Conclusion

A multi-objective evolutionary-based algorithm for probabilistic transformation (MOEPT) has been proposed in this paper. It uses genetic algorithm to obtain Bayesian belief function to offer a comprehensive consideration concerning distance closeness between original BBA and the Bayesian approximate one.

Also, a new aggregation measure is proposed in this paper so as to be combined into a more accurate "distance closeness" measure for MOEPT. More importantly, the convergence analysis of MOEPT is given to prove the rationality of our proposed method. The effectiveness of MOEPT has been compared with respect to several probabilistic transformations proposed in the literature. Furthermore, the shortcomings of original MOEPT version have been clearly identified on two targets type tracking problem, and they have been overcome thanks to a modification of belief interval constraints. As future works, we would like to establish adaptive scheme on the selection of weight in MOEPT and make more comparisons between performance of this MOEPT approach with other recent proposed evolutionary algorithms. We would also make more investigations on MOEPT to extend it to work with more than two targets.

Acknowledgment

This work was supported by NNSF of China (No.61175091), Qing Lan Project of Jiangsu Province, Aeronautical Science Foundation of China (20140169002), and Six Major Top-talent Plan of Jiangsu Province.

References

- [1] A.P., Dempster, Upper and lower probabilities induced by a multivalued mapping, *Ann. Math. Stat.*, **1967**, *38 (4)*,325-339.
- [2] G, Shafer, A mathematical theory of evidence, *Princeton Univ. Press*, **1976**.
- [3] F.J., Li, Y.H., Qian, J.T., Wang, J.Y., Liang, Multigranulation information fusion: A Dempster-Shafer evidence theory-based clustering ensemble method, *Information Sciences*, **2016**, *000*, 1-21
- [4] X.Y., Su, S., Mahadevan, W.H., Han, Y., Deng, Combining dependent bodies of evidence, *Applied Intelligence*, **2015**, *3 (44)*,634-644

- [5] T., Dencœux, Inner and outer approximation of belief structures using a hierarchical clustering approach, *International Journal of Uncertainty, Fuzziness, and Knowledge-based Systems*, **2001**, 9 (4), 437-460.
- 545 [6] P., Smets, The Combination of Evidence in the Transferable Belief Model, *IEEE Trans. on PAMI*, **1990**, 5, 29–39.
- [7] M. Daniel, Probabilistic Transformations of Belief Functions, in *Proc. 8th Rur. Conf. Symbol. Quant. Approach. Reason. Uncertainty (ECSQARU) Barcelona, Spain*, Jul. **2005**, 539-551.
- 550 [8] F., Smarandache, J., Dezert (Editors), Advances and applications of DS_mT for information fusion, American Research Press, Rehoboth, NM, U.S.A., Vol. 1–4, 2004–2015.
- [9] P., Smets, Decision making in the TBM: the necessity of the pignistic transformation, *Int. J. of Approximate Reasoning*, **2005**, 38.
- 555 [10] J., Dezert, F., Smarandache, A new probabilistic transformation of belief mass assignment, *In Proc. of 11th Int. Conf. on Information Fusion*, Cologne, Germany, June-July **2008**, 1-8.
- [11] B.R., Cobb, P.P., Shenoy, On the plausibility transformation method for translating belief function models to probability models. *International*
560 *Journal of Approximate Reasoning*, **2006**, 41 (3), 314–330.
- [12] J., Sudano, Equivalence Between Belief Theories and Naïve Bayesian Fusion for Systems with Independent Evidential Data-Part I, The theory, *In Proc. of Fusion 2003*, Cairns, Australia, July 2003.
- [13] F., Cuzzolin, The Intersection probability and its properties, in *Proc. of*
565 *ECSQARU*, Verona, Italy, July **2009**, 1-3.
- [14] D., H., Han, J., Dezert, C., Han and Y., Yang, Is entropy enough to evaluate the probability transformation approach of belief function? *In Proc. 13th Conf. Inf. Fusion*, Edinburgh, U.K., Jul. **2010**, 286-293.

- [15] D.Q., Han, J., Dezert, C.Z., Han, New Basic belief assignment approxima-
570 tions based on optimization, *In Proc. of 15th Int. Conf. on Information
Fusion*, July **2012**, 286–293.
- [16] D.J., Bucci, S., Acharya, T.J., Pleskac and T.J., Kam, Performance of prob-
ability transformations using simulated human opinions, *In Proceedings of
17th Int. Conf. On Information Fusion*, **2014**, 1-8.
- [17] Y.,Yang, D.,Q., Han, A new distance -based total uncertainty measure in
575 the theory of belief function, *Knowledge-Based System*, **2016**, *94*, 114-123.
- [18] W.R., Liu, Analyzing the degree of conflict among belief functions, *Artifi-
cial Intelligence*, **2006**, *170*, 909-924.
- [19] D.,H., Han, J., Dezert and Y., Yang, New Distance Measures of Evidence
580 based on Belief Intervals. *In: Proc. of Belief 2014*, Oxford, UK, **2014**.
- [20] A., Frikha, on the use of a multi-criteria approach for reliability estimation
in belief function theory, *Information Fusion*,**2014**,*18*, 20-32.
- [21] L.G.O., Silva, A.T., Almeida-Filho, A multicriteria approach for analysis of
conflicts in evidence theory, *Information Sciences*,**2016**,*346-347*, 275-285.
- [22] D.Q., Han, J., Dezert, Z.S., Duan Evaluation of Probability Transforma-
585 tions of Belief Functions for Decision Making. *IEEE Transactions on Sys-
tems, Man, And Cybernetics: Systems*, **2015**,*99*.
- [23] Z.G., Liu, J., Dezert, Q., Pan, G., Mercier, Combination of sources of
evidence with different discounting factors based on a new dissimilarity
590 measure, *Decision Support Systems*, **2011**, *52*, 133-141.
- [24] M.M., Ma, J.Y., An, Combination of evidence with different weighting fac-
tors: A novel probabilistic-based dissimilarity measure approach, *Journal
of Sensors*, **2015**.

- [25] D.Q. Han, Y., Deng, C.Z, Han, Sequential weighted combination for un-
595 reliable evidence based on evidence variance. *Decision Support System*,
2013,*56*, 387-393.
- [26] W., Pan, J.Y., Hong, New methods of transforming belief functions to pig-
nistic probability functions in evidence theory, *Int. Workshop on Intelligent
System and Applications*, **2009**, (In Chinese).
- 600 [27] A.L., Jusselme, C.S., Liu, D., Grenier and E., Bossé, Measuring Ambi-
guity in the Evidence Theory, *IEEE Transactions on Systems, Man, And
Cybernetics-Part A: Systems and Humans*, **2009**, *36* (5), 890-903.
- [28] J. Dezert, D. Han, J.-M. Tacnet, S. Carladou, Y. Yang, Decision-Making
with Belief Interval Distance, in Proc. of Belief, **2016**, Int. Conf., Prague,
605 CZ,
- [29] J., Sudano, The system probability information content (PIC) relationship
to contributing components, combining independent multisource beliefs,
hybrid and pedigree pignistic probabilities, in *Proc. of Int. Conf. on Infor-
mation Fusion*, Annapolis, Maryland, U.S.A, July **2002**, 1277–1283.
- 610 [30] X.D., Li, F.Y., W, A Clustering-Based Evidence Reasoning Method. *Inter-
national Journal of Intelligent System*, **2016**, *31*, 698-721.
- [31] A., Irpino, R., Verde, Dynamic clustering of interval data using a
wasserstein-based distance. *Pattern Recognition Letters*, **2008**, *29*, 1648-
1658.
- 615 [32] T., Burger, Geometric views on conflicting mass functions: From distances
to angles. *International Journal of Approximate Reasoning*, **2016**, *70*, 36-
50.
- [33] N., Srinivas, K., Deb, Multiobjective function optimization using nondom-
inated sorting genetic algorithms, *Evolutionary Computation*, **1995**, *2* (3),
620 221-248, .

- [34] K., Deb, A., Pratap, S., Agarwal, T., Meyarivan, A Fast and Elitist Multi-objective Genetic Algorithm: NSGA-II, *IEEE Transactions on Evolutionary Computation*,**2002**,*6 (2)*, 182-197.
- [35] Q.B., Zhu, Analysis of Convergence of Ant Colony Optimization Algorithms, *Control and Decision*,**2006**, *21 (7)* 764-770.
625
- [36] A.-L., Jusselme, P., Maupin, On some properties of distances in evidence theory, in *Workshop on Theory of Belief Functions*, Brest, France, March 31st–April 2nd,**2010**, 1-6.
- [37] W., Jiang, J., Zhan, A Modified Combination Rule in Generalized Evidence Theory, *Applied Intelligence*, **2016**, 1-11.
630
- [38] X.,d., Li, W.,D., Yang, J., Dezert, An Airplane Image Target’s Multi-feature Fusion Recognition Method, *ACTA AUTOMATICA SINICA*, **2012**, *38*, 1299-1307.
- [39] Y. Bi,D. Bell,J. Guan, Combining evidence from classifiers in text categorization, in *Proc.8th Int. Conf. Knowl.-Based Intell. Inf. Eng. Syst.*, Wellington, New Zealand, **2004**, 521-528.
635
- [40] J., Dezert, A., Tchamova, F., Smarandache, P., Konstantinova, Target Type Tracking with PCR5 and Dempster’s rules: A Comparative Analysis, in *Proc. of Int. Conf. on Information Fusion*, Florence, Italy, July
640 **2006**.