Application of Dempster-Shafer Theory to a Track Association Problem

C. M. Rogers

University College London and QinetiQ Malvern

Abstract: This paper gives an overview of the Demspter-Shafer algorithm, with a brief comparison to a Bayesian methodology, and shows how it may be used for the track association problem encountered in a multispectral seeker application. Three application starting points are considered, incorporating various degrees of ignorance, and compared. The benefits of using Dempster-Shafer are then weighed against the potential drawbacks. In conclusion, the Dempster-Shafer algorithm gives an advantage in the route taken to a decision point, and allows the incorporation of ignorance, something that is not possible with the Bayes approach which offers more rigorous analysis based on specific assumptions.

1 Introduction

The problem of association arises in many applications where multiple sensors, or more generally, sources of data, are used to gather information about an object. Before the information from these separate sources of information can be combined to gain a more complete or more accurate representation of the object in question, it must be ascertained that the information relates to the same object. Usually a probabilistic method is used to determine the probability of the information relating to the same object, and popular candidates are Bayesian probability, Dempster-Shafer, Neural Networks and Fuzzy logic.

In this paper, the particular problem of track association in a multispectral seeker is considered, and the use of the Dempster-Shafer algorithm is compared to a current Bayesian technique. Tracks are formed in each of 3 different sensors, and then associated and fused appropriately, with track fusion being prefered to measurement fusion for robustness.

The debate as to the validity and applicability of the Dempster-Shafer method compared to the Bayesian and more classical probability methods has raged since Dempster first suggested his method in the 1960's. As yet no definitive conclusion has been reached, although certain applications often make better use of one method or the other (an example would be object identification/classification, where Dempster-Shafer has been shown to give significant benefits over simple Bayes[1], although sometimes at the expense of computational resources[2]). The track association problem has as yet shown no tendency to one method or the other.

2 The Track Association Problem

A simple but effective demonstration of the track association problem (and in fact the actual starting situation being researched) is the 2 sensor - 2 track case, where two tracks are formed by each sensor on 2 objects from an unknown number of real objects in the scenario. The tracks contain information on the object position and velocity, depending on the capabilities of the sensors (an IR sensor can measure angular position and derive angular velocity, whilst a radar can measure angular position, range and range rate, and derive angular velocity). The problem is then

> Given that each sensor α and β forms two tracks, α_1, α_2, and β_1, β_2, which pairs $(\alpha_j \beta_k)$ associate, i.e. relate to the same object?

Once this has been decided, the information from each sensor about an object can be fused to gain a better estimate of the object's position and velocity.

3 Applying Dempster Shafer Theory[3]

The first step in applying Dempster-Shafer theory to the problem is to create the Frame of Discernment, Θ. This is a set of mutually exclusive and exhaustive possibilities, i.e. a set of hypotheses. For a 2 sensor - 2 track problem there are 7 hypotheses:

Hypothesis	β_1	β_2	Number of Associations
H_0	-	-	0
H_1	α_1	-	1
H_2	-	α_1	1
H_3	α_2	-	1
H_4	-	α_2	1
H_5	α_1	α_2	2
H_6	α_2	α_1	2

and so the Frame of Discernment is:

$$\Theta = \{H_0, H_1, H_2, H_3, H_4, H_5, H_6\} \tag{1}$$

From this Frame of Discernment a Power Set is derived, 2^Θ, which is a set of all elements in Θ and all possible combinations of elements in Θ, i.e. all possible subsets of Θ, including the empty set.

$$\begin{aligned} 2^\Theta = & \{\{\emptyset\}, \{H_0\}, \{H_1\}, \{H_2\}, \ldots \{H_6\}, \\ & \{H_0, H_1\}, \{H_0, H_2\}, \ldots \{H_5, H_6\} \\ & \vdots \\ & \{H_0, H_1, H_2, H_3, H_4, H_5, H_6\}\} \end{aligned} \tag{2}$$

There are three groups of elements that are of particular interest in this problem. The first of these includes the singleton elements, as they hold the individual hypotheses. The second group of interest includes those elements that represent associations, e.g. $\{H_2, H_6\}$ represents the association $\alpha_1\beta_2$, while $\{H_0, H_1, H_5\}$ represents α_1 associating with neither of the β tracks. The third group is simply the last element in 2^Θ and is the Frame of Discernment.

Each element A in 2^Θ can be assigned a Basic Probability Number (BPN), $m(A)$, which is a measure of the belief assigned exactly to that element, yet implies nothing about belief assigned to proper subsets of that element. Further to this, a Belief Function can be calculated over 2^Θ, where $\mathrm{Bel}(A)$ is the sum of the belief assigned exactly to A and all subsets of A in 2^Θ (this is similar to cumulative probability in classical probability), and is the amount to which that element is supported by the available evidence.

$$\mathrm{Bel}(A) = \sum_{\forall B: B \subseteq A} m(B) \tag{3}$$

There are two rules for the BPN's:

$$m(\emptyset) = 0 \tag{4}$$

$$\sum_{\forall A \in 2^\Theta} m(A) = 1 \tag{5}$$

This in turn implies that $\mathrm{Bel}(\Theta) = 1$

Any belief assigned exactly to Θ, $(m(\Theta) > 0))$, implies a measure of complete ignorance, as this implies belief that any of the hypotheses may be true, and it is through this mechanism that Dempster-Shafer theory allows ignorance to be incorporated.

3.1 The Dempster Combination Rule

The Dempster Combination Rule is a means for calculating the new BPN's of a Power Set when new data is received. The new data is in the form of a second Power Set with it's respective BPN's, and so the combination of 2_1^Θ with 2_2^Θ to give 2_{12}^Θ has BPN's calculated using:

$$m_{12}(B) = \frac{\sum_{\emptyset \subset C \cap D = B} m_1(C) m_2(D)}{1 - \sum_{C \cap D = \emptyset} m_1(C) m_2(D)} \tag{6}$$

It should be noted that there is no mathematical basis for this rule, although it does have the desirable features of commutativity and associativity, which in turn means that the result of combining any number of belief functions is independent of the order of combination. However, other methods exist (such as that due to Fagin and Halpern[4]) that possess these same features but yield different results.

4 Calculation Procedure

The current association probability calculation procedure involves the steps shown in figure 1. To obtain the association probabilities after new data has been gathered, the hypotheses probabilities must first be calculated, and from these the association probabilites may be calculated. Figure 2 shows the equivalent procedure using Dempster-Shafer, whereby it is possible to calculate the element BPN's directly after new data has been gathered.

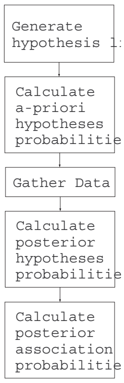

Figure 1: Bayes Process Flow

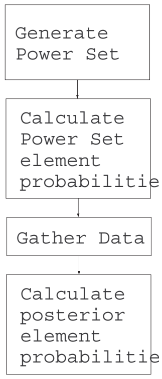

Figure 2: Dempster-Shafer Process Flow

5 Comparison with Bayes

Dempster-Shafer theory differs significantly from Bayes theory, in particular for this application with the handling of ignorance. A distinction is drawn between a lack of supporting evidence and the existence of contradictory evidence. Consider the case of 2 mutually exclusive and exhaustive hypotheses. If there is no knowledge about which hypothesis is true, Bayes would require the assumption that each hypothesis has an equal probability of 0.5. However, if there is equal evidence for both hypotheses, they are still each assigned a probability of 0.5. In contrast, Dempster-Shafer allows each hypothesis to be assigned a belief of 0 when no evidence is available, and equal belief for equal evidence, although not necessarily a belief of 0.5 if some ignorance remains. In this way differing degrees of ignorance can be incorporated.

6 Definition of the Power Set *a priori* Basic Probability Numbers

There are several possible starting points when applying Dempster-Shafer theory to the track association problem. The current Bayesian method makes several assumptions and calculates *a priori* probabilities based on these. However, the availability of a method for incorporating ignorance allows more freedom with Dempster-Shafer. Three methods for defining the *a priori* Power Set suggest themselves:

***a priori* hypothesis probabilities** The first of these would be to take a similar approach to that employed already with Bayes. *a priori* hypothesis probabilities can be calculated with an assumed probability distribution class and mean, and used as the basic probability numbers for the singleton elements within the power set. As these *a priori* probabilities sum to 1 no ignorance is assumed, or indeed may be included in accordance with equation 5. As new data arrives these BPN's will be modified, and decisions can be made accordingly.

This method is certainly valid, but is in contrast to the normal use of Dempster-Shafer, in that it starts with all the belief in the singleton elements. The normal approach is to start with belief in the higher order elements of the power set, and then as data is applied, the information is refined so that only the singleton elements have any belief. The method is also questionable in it's advantages over Bayes in that it still makes use of potentially incorrect assumptions (e.g. the distribution of the number of objects) in the calculation of the *a priori* hypothesis values. The only apparent advantage is the route via which a decision is made, as discussed in section 4.

***a priori* association probabilities** The second method makes use of the *a priori* association probabilities. Whilst these may be derived from the hypotheses probabilities, which assume a distribution class and mean, they may also be calculated without assuming a distribution. Instead, only one number needs to be assumed. This incorporation of ignorance (one less assumption) can be seen in that the BPN's are assigned to higher order elements of 2^{Θ} (the elements that contain all hypotheses that imply a particular association), rather than the singleton elements as in the first method. As the association probabilities are joint probabilities, they sum to more than 1. In order to incorporate them as BPN's in a single Power Set, they are normalised, again meaning that no other ignorance exists.

This method is more akin to the standard use of Dempster-Shafer as the belief is focused on higher elements of 2^{Θ} and then propagates down to the singleton elements as new data arrives. Again there is a potential advantage in the route via which a decision is made and it should be noted that the ignorance may be preserved if insufficient data is available to replace the ignorance. This in effect means that no decisions will be made that are based on poor evidence.

Complete Ignorance The final method considered is to assume complete ignorance until any data is received. The initial power set would therefore have all belief assigned to Θ. New data will result in the propagation of the belief to lower order elements of 2^{Θ}, eventually into the singleton elements. This method is beneficial in that it assumes no prior knowledge, and is the way in which Dempster-Shafer is most often used. Ignorance may be preserved as in the previous method.

7 Utilising New Data

Once the *a priori* BPN's have been defined, the task of updating the BPN's on the arrival of new data using equation 6 remains.Calculating the BPN's from the new data is often the most inexact part of using Dempster-Shafer, as BPN's are not equivalent to probabilities, and are by definition open to subjectivity. The method that is proposed for this association problem utilises the statistical distance between two tracks, $(a-b)/\sigma_c$, $\sigma_c = \sqrt{\sigma_a^2 + \sigma_b^2}$, whereby the BPN for an association is given as:

$$m(A) = \frac{\sigma_c \sigma_{prior}^2}{(a-b)\sigma_c^2} \tag{7}$$

where $\sigma_{prior}^2/\sigma_c^2$ is a scaling factor defining how good the data is. If the sum of association BPN's is greater than 1 they are normalised, and if the sum is less than one (i.e. there is insufficient data), the remaining belief is assigned to Θ to express ignorance. This approach is intuitive, as it implies that the belief in association reduces with increasing track separation $(a-b)/\sigma_c$ (tracks are less likely to be associated if they are measured to be far apart), whilst the belief also reduces with increasing track covariance σ_c (the quality of the track is poor or very noisy, implying a lack of good quality evidence). This is only one suggested method, and others may provide better performance.

8 Conclusions

This paper has shown the basic mathematical basis of the Dempster-Shafer algorithm. It is noted that there is no mathematical basis for the Dempster combination rule, although it does seem logical and fulfils the commutativity and associativity requirements. Dempster Sahfer has been widely applied to many problems where previously Bayesian or classical inference methods have been used, with some success.

The algorithm has then been applied to the problem of track to track association, an area that has apparently not received much attention from Dempster-Shafer proponents in the past. Three starting points have been proposed, each with a varying degree of ignorance or assumption included. An intuitive method for incorporating new data has also been suggested, although its performance has not been tested.

The Dempster-Shafer algorithm seems advantageous in that it arrives at the hypothesis/association probability via a more direct route than the current Bayes method. It also allows tailoring of the algorithm according to confidence in assumptions, whilst complete ignorance (0 confidence) can also be used, something that can not readily be done using the Bayes approach which offers more rigorous analysis based on the specific assumptions.

References

[1] Braun J B; *Dempster-Shafer Theory and Bayesian Reasoning in Multisensor Data Fusion*; Proc. SPIE Vol. 4051; 2000

[2] Buede B M; *A Target Identification Comparison of Bayesian and Dempster-Shafer Multisensor Fusion*; IEEE Transactions on Systems, Man and Cybernetics; Vol 27, no. 5, pp 569-577; 1997

[3] Shafer G; *A Mathematical Theory of Evidence*; Princeton University Press; 1976

[4] Fagin R, Halpern J Y; *A New Approach to Updating Beliefs*; Uncertainty in Artificial Intelligence, Vol. 6, pp347-374; 1991