14th International Conference on Information Fusion
Chicago, Illinois, USA, July 5-8, 2011

GRP1. A recursive fusion operator for the transferable belief model

Gavin Powell
Innovation Works
EADS UK
Newport, Wales, UK
Gavin.Powell@eads.com

Matthew Roberts
Innovation Works
EADS UK
Newport, Wales, UK
Matthew.Roberts@eads.com

Abstract – Generally, there are problems with any form of recursive fusion based on belief functions. An open world is often required but the empty set can become greedy where, over time, all of the mass will become assigned to it where conjunctive combination is used. With disjunctive combination, all of the mass will move toward the ignorant set over time. In real world applications often the problem will require an open world but due to these limitations it is forced into a closed world solution. GRP1 works iteratively in an open world in a temporally conscious fashion, allowing more recent measurements to have more of an impact. This approach makes it ideal for fusing and classifying streaming data.

Keywords: recursive, iterative, fusion, transferable belief model.

1 Introduction

Sensing devices can classify an enemy target from a selection of 'known' objects. This may be from human intelligence, radar, LADAR etc. This object classification can occur iteratively over time providing a new measurement, or classification, at regular time intervals. To obtain a more informed overall classification, all sensors' measurements need to be fused at each time interval and recursively over time, as shown in Equation 1. From this fused data a better classification of the object can be made.

$$S_{1..t} = S_{1..t-1} + s_t \tag{1}$$

Where S is the fused sensor measurements and s is the sensor measurement at time t

A classification can then be made from $S_{1..t}$. Current set based methods are inadequate in this scenario due to the fusion method used. GRP1 overcomes these problems by providing a much more intelligent form of fusion method, designed specifically for iterative situations.

1.1 Scenario

Within Dempster-Shafer Theory (DST) based methods, such as the Transferable Belief Model (TBM), we are given a world of possible outcomes to choose from. This generally contains everything that we know. For instance when fusing data from sensors to classify a target we will have all of the possible targets in our world that we can identify

$$\Omega = \{Pedestrian, Car, Amphibious\ Light\ Tank\ (ALT), Light\ Tank\ (LT), Main\ Battle\ Tank\ (MBT)\} \tag{2}$$

As well as accounting for all outcomes that we are aware of there is also the empty set, $\varnothing$. Generally this is used to signify conflict in the data being passed to the system, which in itself is often a marker that the problem has been poorly modelled, a sensor is malfunctioning or that there we are in fact looking at something outside of our model space.

Our world of all possible outcomes, Θ, is made up of $\varnothing$ and all the possible combinations of Ω. There will be 2 elements such that

$$\Theta = \{\{Pedestrian\}, \{Car\},\ldots,\{ALT, LT\},\ldots,\{Pedestrian, Car, ALT, LT, MBT\}, \varnothing\} \tag{3}$$

The sensors can assign a mass (which shows their belief in that being the actual truth) to each of these sets. Assigning mass to a set with one element *{Car}* is a definite decision, where as assigning mass to *{Pedestrian, Car, ALT, LT, MBT}* is showing complete vagueness. The set *{Pedestrian, Car, ALT, LT, MBT}* is saying that the truth is within that set, but we have no idea which element it is.

If we have two sensors providing data about what they think the target is then we will have two belief assignments for our world. It is these belief assignments that we wish to fuse together using a combination rule.

It is this fusion operator that we are going to describe and discuss.

1.2 Open and Closed Worlds

Within set theoretic approaches, there exists a notion of open and closed worlds. If we accept that we know everything about our world then we must also accept that the empty set has no meaning. There is nothing that we haven't accounted for, there is no 'anything else'. Contrary to this is the open world, which is that other things do, or could, exist that we haven't accounted for, and these are signified by the empty set.

2 Set theoretic approaches

Early work by both A. P. Dempster [1] and Glenn Shafer [2] later became known as Dempster Shafer Theory (DST). DST is a generalisation of Bayesian Theory and it states:

i) beliefs are created from subjective probabilities
ii) information is fused using Dempster's rule of combination

The basis of the work in the DST has been taken and extended by other parties [3,4] and so this basic framework has evolved into various mutations, but they share the same core idea. One such framework is the TBM [3] and we shall base the rest of the discussion on this to highlight the empty set issues; where appropriate we will compare and contrast to other approaches.

2.1 Basic Elements of the Transferable Belief Model

The TBM is a set theoretic approach, such as the ones discussed in previous sections. The TBM splits the set theory into two stages. Firstly, the *credal* level where beliefs are entertained and quantified by belief functions. Secondly, the *pignistic* level where those beliefs are used to make decisions and are quantified by probability functions.

If we take a weather example then all of the possible weather types that we are going to account for in our world are *wind*, *rain*, and *sun*. The set of all the possible types of weather we know about is given by $\Omega = \{wind, rain, sun\}$. To this we apply a basic belief assignment (bba) $m : 2^{\Omega} \rightarrow [0,1]$ with $\sum_{A \subseteq \Omega} m(A) = 1$ where *m(A)* is the basic belief mass (bbm) given to *A*. This is a method of assigning masses to each of the subsets of Ω to signify our belief, or that there is evidence showing that the truth is somewhere within that set, not necessarily equal to but, within that set. Each of these sets is in fact a hypothesis. Any set that has a *bbm > 0* is called a focal set, and any focal set that only has one member is called a singleton set. The more members within a focal set the more the ignorance, or uncertainty about which single state is true.

If we recall one of the possible subsets of Θ is the empty set $\varnothing$. Values assigned to it carry various meanings, these meanings depend on if we are working within an open or closed world. If the world is closed then we should not be assigning bbm to the empty set as we claim that **Ω** is exhaustive and covers all possible outcomes. Within an open world we can assign bbm to the empty set and it shows how much we believe that the truth is **not** in **Ω**. These are two very important concepts. It can be a very dangerous assumption to make that **Ω** is exhaustive; it can cause an unfair normalisation that is poor for decision making and limits our means of analysing the data and our fusion process [5].

DST has always worked within the closed world. When applied to fusion tasks it can appear that all is well within this model, allowing data to conveniently be fused and decisions to be made. The closed world assumption manages to mask any issues that are occurring due to improper problem modelling, which in turn, forces the user to make an impaired and naïve decision. The TBM and Dezert Smarandache Theory (DSmT) [9] allow for the possibility that **Ω** is not exhaustive.

2.2 Set theoretic combination of data, or Fusion.

If we have more than one piece of data from either the same source over time or multiple sources, or even multiple sources over time then it is normal that we will want to combine all of this data. This combination will enable us to make a more informed decision using all available information, as opposed to just looking at a single piece of evidence. An example of such a task would be classifying an object where we receive data continuously over time from a variety of sensors, and thus want to recursively combine, or fuse, this data, before we make a decision on the classification.

Each piece of evidence will give us a bba over Θ such that bbm's are applied to the various sets within Θ. This will show where we think the truth is. If this evidence is very certain then the singleton sets within Θ will get more mass, if we are uncertain then the set **Ω** will get more mass. Coincidently the set that shows complete ignorance to which is the correct outcome is **Ω**, assigning a mass of value 1 to this set is showing complete ignorance. A completely naïve state would apply all of the bbm to **Ω**. To fuse data we must combine these bba's to create a new bba. There are a multitude of methods to accomplish this [6,8]. The original work on the DST used Dempster's (conjunctive) rule of combination

$$m_{1\otimes 2}(A) = \frac{\sum_{B\cap C} m_1(B)m_2(C)}{1-k}$$

$$\text{where} \quad K = \sum_{B\cap C=\varnothing} m_1(B)m_2(C) \tag{4}$$

.

and *A, B*, and *C* are sets within Θ . It can be seen that *K* is in fact a normalisation factor that redistributes any mass assigned to the empty set after combination. If we are to remove the normalisation we get Dempster's un-normalised rule of combination as used in the TBM [3] as shown below

$$m_{1\otimes 2}(A) = \sum_{B\cap C} m_1(B)m_2(C) \tag{5}$$

As well as the conjunctive rule of combination there is also a disjunctive rule of combination which is a lot more cautious with its approach as it uses the disjunction of two sets rather than the conjunction, this allows for [3]

$$m_{1\oplus 2}(A) = \sum_{B\cup C} m_1(B)m_2(C) \tag{6}$$

2.3 Convergence

Issues that have been highlighted previously [5] show that unnormalised conjunctive, and disjunctive combination, is ineffective at any form of iterative and recursive combination. Convergence happens very quickly where all the mass will tend toward either the empty set, for conjunctive combination, or the ignorant set, for disjunctive combination. This convergence can make the system become unresponsive to any new, or differing inputs, and ultimately fail in its task. Previously it has been suggested that convergence protection [7] is a means to prevent this. This is quite a crude method that simply caps the limit that any set can reach enabling the system to retain a degree of flexibility, and thus not converge completely.

2.4 Convergence Avoidance

If we take both the conjunctive and disjunctive combination rules and find the average [6,8] of the two then we have a fusion operator that wont instinctively, or unnecessarily, converge to the ignorant or empty set when used in a recursive manner. Unlike the conjunctive rule of combination, shown in Equation 5, it can recover from a situation where it incorrectly thinks it knows the truth, while at the same time not being too vague, as can happen with the disjunctive rule of combination as shown in Equation 6.

The arithmetic mean is calculated using

$$m_{mean}(A) = \frac{1}{M}\sum_{j=1}^{M} m_j(A) \tag{7}$$

where *M* is the number of sets you are finding the mean across

3 Vagueness in the system

Our world, Θ, will receive beliefs either from a sensor (of some sort), or by fusing two other worlds. If we look at how that belief is distributed throughout the world, Θ, we can ascertain how vague our world is. If values are assigned to the singleton sets, which have only one element, then the world is precise and any decisions are well educated. If beliefs are given to the uncertain set, **Ω**, then the world is vague and any decisions made from this are uneducated.

This notion of precision (educatedness) is quite important, and can be used to determine how we fuse incoming information. If the world is showing a high degree of precision then we are certain in our beliefs and thus it should take a lot more effort to alter our belief. Alternatively, if we are completely vague about our knowledge and beliefs then we should be more accepting of new information. This concept needs to be accounted for when we are fusing information.

We have the ability to 'discount' incoming information [9]. This discounting process will weight the incoming data and is a measure of how much it is to be trusted, or how much impact it will have.

Smets [10] describes the discounting of data through Equation 8.

$$\begin{aligned} m^{\alpha}(A \mid x) &= (1-\alpha)\cdot m(A) \quad && \forall A \subseteq \Omega, A \neq \Omega \\ m^{\alpha}(A \mid x) &= [(1-\alpha)\cdot m(A)] + \alpha \quad && A = \Omega \end{aligned} \tag{8}$$

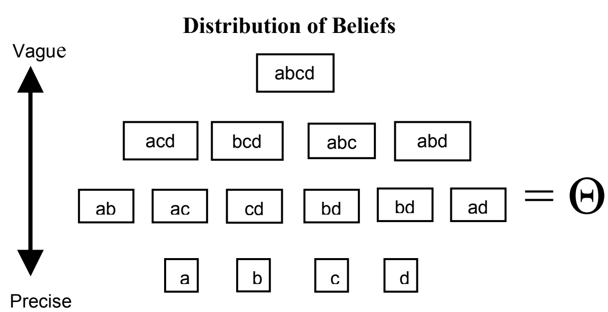
where *x* is some prior knowledge about the validity of the information that that entity is providing, and is accounted for by α. This works perfectly well when we are dealing with the conjunctive combination rule as the masses are passed toward the uncertain set, **Ω**. For the disjunctive rule this will only force it to be vaguer and encourage convergence toward the uncertain set, Ω. When using the disjunctive combination rule you must discount as with Equation 9. This will allow the discounted mass to be passed to the empty set, which when fused with the cautious disjunctive combination rule allows for the mass to be redistributed evenly across the system.

$$
\begin{aligned}
m^{\alpha}(A \mid x) &= (1-\alpha)\cdot m(A) \qquad \forall A \subseteq \Omega, A \neq \varnothing \\
m^{\alpha}(A \mid x) &= [(1-\alpha)\cdot m(A)] + \alpha \qquad A = \varnothing
\end{aligned} \tag{9}
$$

4 Dynamic Discounting

The degree that we choose to discount by is of course related to the degree of precision (educatedness) in our world Θ, and shows how much we can be influenced by incoming data.

If Ω = *{a, b, c, d}* then *Θ* is shown by the following diagram



We can measure the precision, *p*, using Equation 10 [11]

$$
p(\Theta) = \sum \frac{|\Omega| - |A|}{|\Omega| - 1} \times m(A) \qquad \forall A \neq \varnothing, A \subseteq \Theta \tag{10}
$$

Any value added to the empty set is treated as adding to the vagueness. This is a point of argument as to whether the empty set is making the system vaguer or is adding precision, or if in fact it should be ignored. If the empty set is adding precision then

$$
p(\Theta) = \left(\sum \frac{|\Omega| - |A|}{|\Omega| - 1} \times m(A) \right) + m(\varnothing) \qquad \forall A \neq \varnothing, A \subseteq \Theta \tag{11}
$$

If we want to ignore any belief given to the empty set and normalise then we can use Equation 12

$$
p(\Theta) = \sum \frac{|\Omega| - |A|}{|\Omega| - 1} \times \frac{m(A)}{(1 - m(\varnothing))} \qquad \forall A \neq \varnothing, A \subseteq \Theta \tag{12}
$$

4.1 GRP1 process

The GRP1 algorithm can be utilised using the following steps

1. Set up our knowledge as ignorant for fused measurement, Θ_f or use initial prior information
2. Receive measurement from sensor
3. Put measurement into world Θ_m
4. Calculate precision, $p(\Theta_f)$, using Equation 10, 11, or 12
5. Discount Θ_m by $p(\Theta_f)$ using Equation 8 to get Θ_{mc}
6. Discount Θ_m by $p(\Theta_f)$ using Equation 9 to get Θ_{md}
7. Disjunctively combine Θ_f with Θ_{md} to get Θ_{fd} using Equation 6
8. Conjunctively combine Θ_f with Θ_{mc} to get Θ_{fc} using Equation 5
9. Combine Θ_{fd} and Θ_{fc} with the arithmetic mean operator from Equation 7 to get a new Θ_f
10. Return to 2

Steps 1,2,3,7,8,9 and 10 constitute to algorithm GRP1
Steps 4, 5 and 6 are Dynamic Discounting

5 Results

Data was acquired from previous work involving a classification task over time [12]. The data was produced from a simulation of a light amphibious tank travelling over road, grass, and water. Classification at each time step takes place within the TBM. Belief assignments are conditional upon the current estimate of the speed of the target and terrain that the target is travelling over.

From the start of the simulation to approximately time step 225, the target travels quickly along a road. It then slows down, turns at a junction, travels along a short stretch of road and some grass before initiating a river crossing at time step 375. The river crossing finishes at approximately time step 450, where the target travels over more grass.

Figure 1 shows the output from the sensor classification where the possible target types that the sensor understands are given in Equation 2. Other objects will exist outside of this model and so an assumption of a closed world is incorrectly modeling the problem. For the sake of argument we will include results that show a closed world where the empty set is normalised, in either a simplistic manner or a more intelligent one.

Figure 2 shows the results of recursively fusing the data given in Figure 1 using the conjunctive combination rule. It shows how the system very quickly converges to a state where a set is dominant, i.e. approaches a *mass* of 1, as indicated by the high *BetP*. In the open world this is even more of an issue as the empty set becomes dominant. Any decisions that are made from this result will be incorrect. Through normalization we are forcing what is really an open world problem into a closed world solution. Results for this appear to be more promising and the system is suggesting a very high confidence in a particular target. This target classification may not only be incorrect, but the normalization artificially raises the confidence levels and gives a false sense of security.

Figure 3 shows the results of combining the sensor outputs recursively using PCR5 [13], which is a much more intelligent combination rule, with respect to how the empty set can be redistributed. We can see that this works well with the recursive nature of the fusion process and the noise of the data. It also does not converge and become unresponsive over time. The implementation of PCR5 within the TBM framework means that there is no facility to cater for the notion of the target being something outside of the world that we know, so there are still the original issues of normalization occurring, even if it does take place in a more intelligent manner than the conjunctive rule of combination. This is not the case if it is implemented within the DSmT framework.

Figure 4 and Figure 5 show the fused results of using GRP1 with dynamic discounting turned on and off. The stability of these operators working with the open world and with noisy data is an improvement over previous attempts at open world fusion within the TBM. Previous attempts, as shown in Figure 2, were poor to say the least, and were not able to cope with iterative and recursive fusion. The dynamic discounting has the effect of making a much more stable output. Table 1 shows the percentage of correct classification (PT76 Tank), where the classification is taken as the *max BetP*. The rates are very similar for both GRP1 methods and PCR5. The mean squared error, where error is *BetP-(max BetP)*, is also very similar for both GRP1 methods, but with a noticeable difference to PCR5. Due to the normalization and lack of an empty set a useful comparison cannot be made, but the results are included for completeness.

Figure 1. Object classification data received from sensor over time

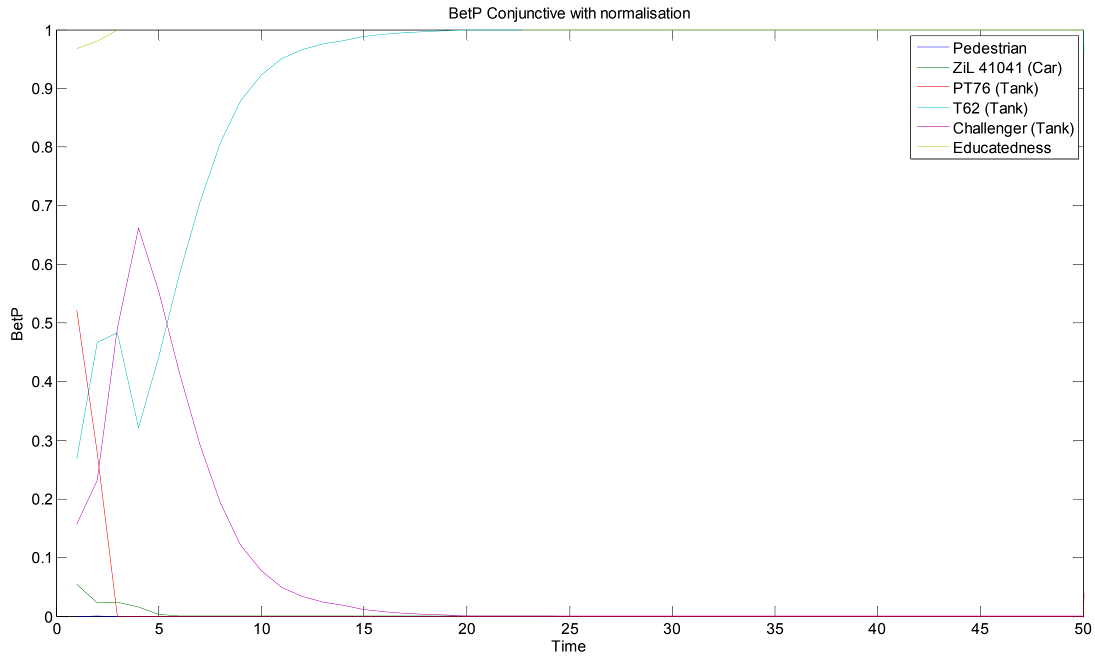


Figure 2. Normalised conjunctive combination

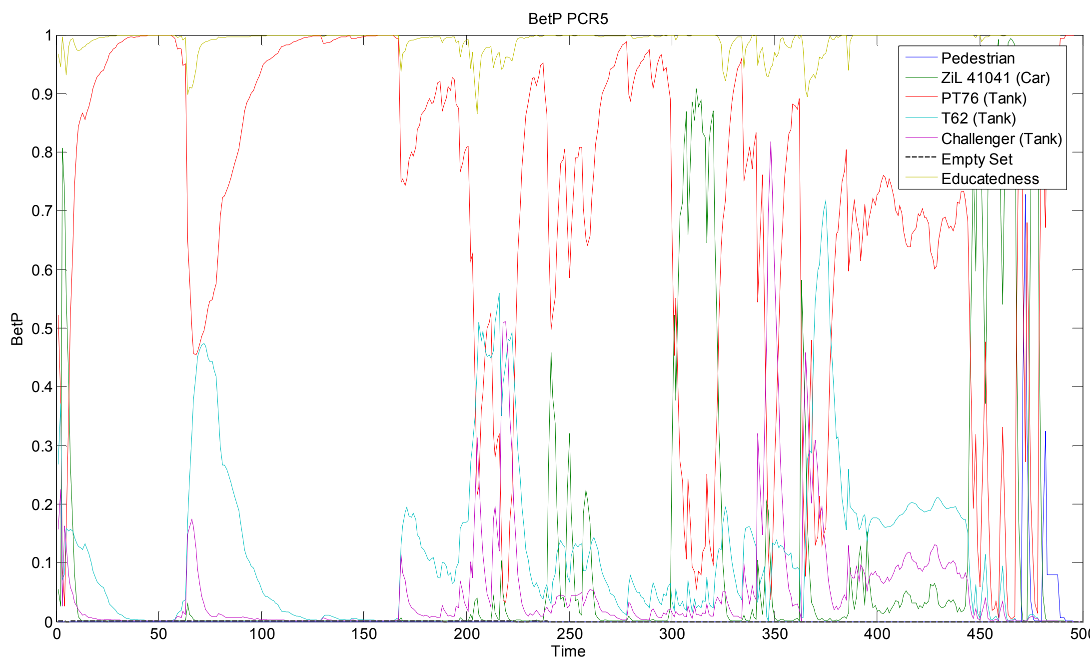


Figure 3. BetP after PCR5, Intelligent empty set redistribution

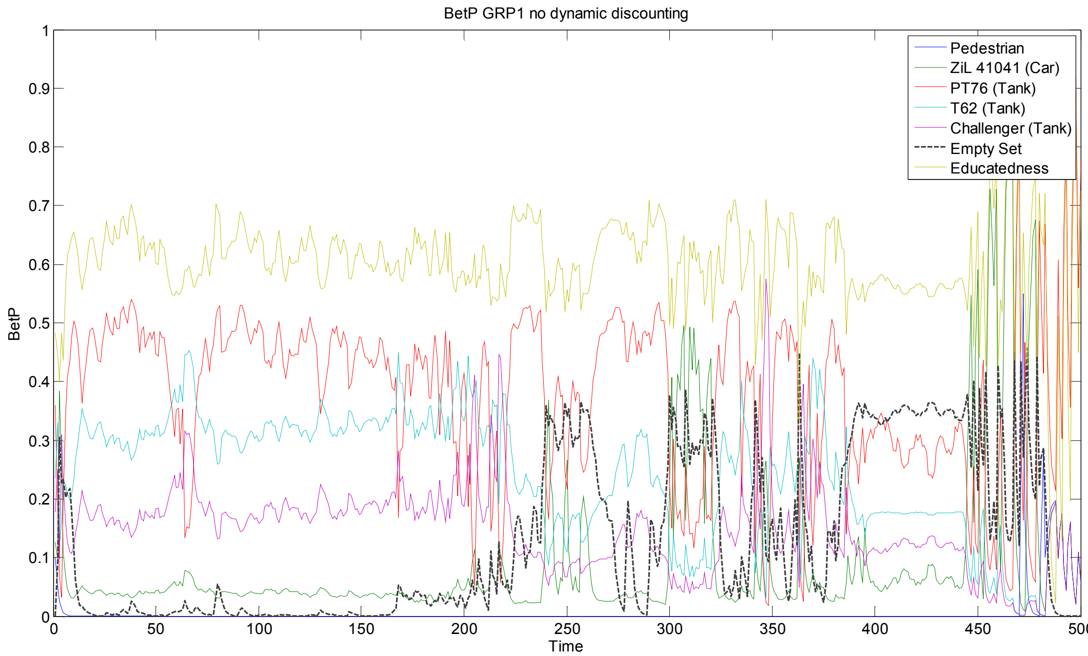


Figure 4. BetP after GRP1 without dynamic discounting

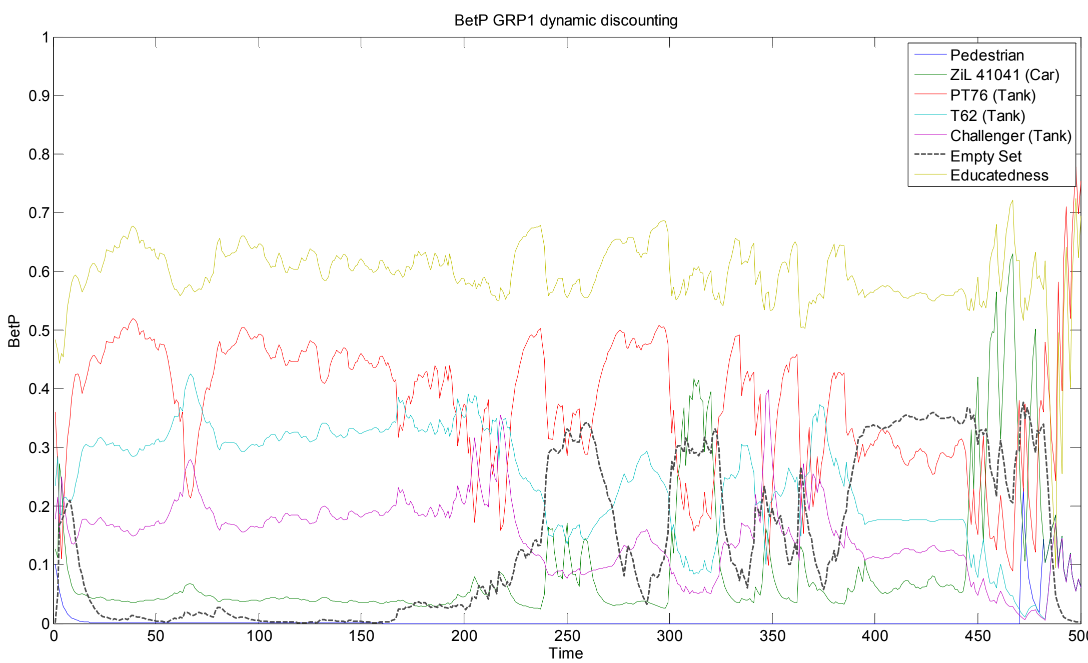


Figure 5. BetP after GRP1 with dynamic discounting

	Correct Classification	*Mean Squared Error*
PCR5	80.3%	0.16
GRP1	79.1%	0.41
GRP1 with dynamic discounting	80.8%	0.40

Table 1. Comparison of fusion operators

6 Conclusions

It has been shown that GRP1 will effectively fuse together information in a recursive manner over time retaining the empty set and the open world. This was not previously possible with the conjunctive or disjunctive combination rules unless normalisation took place. Dynamic discounting allows the algorithm to determine how much bias should be placed on the new and incoming information . This allows for a much more stable output from which decisions can be made.

References

[1] A. P. Dempster, *A Generalisation of Bayesian Inference*, Journal of the Royal Statistical Society, Series B30. pp205-247. 1968.

[2] G. Shafer, *A mathematical theory of evidence*, Princeton University Press, Princeton, NJ 1976.

[3] P. Smets, R. Kennes, *The transferable Belief Model*, Artificial Intelligence, v66 1994, pp191-234.

[4] J. Dezert, *Combination of Paradoxical Sources of Information within the Neutrosophyic Framework*, Proceedings of the First Int. Conf. on Neutrosophics, Univ. of New Mexico, Gallup Campus, December 1-3, 2001.

[5] G. Powell, M. Roberts, D. Marshall, *Pitfalls for recursive iteration in set based fusion*, Theory of Belief Functions, Brest, France, April 2010.

[6] C Osswald, A Martin, *Understanding the large family of Dempster-Shafer theory's fusion operators – a decision-based measure*, 9th International Conference on Information Fusion, Florence, 2006.

[7] G. Powell, D. Marshall, P. Smets, B. Ristic, S. Maskell, *Joint Tracking and Classification of Airbourne Objects using Particle Filters and the Continuous Transferable Belief Model*, International Conference on Information Fusion, Florence, 2006.

[8] C. Osswald, A. Martin , *Understanding the large family of Dempster-Shafer theory's fusion operators – a decision-based measure*, 9th International Conference on Information Fusion, Florence, 2006.

[9] P. Smets, *Belief Functions: the disjunctive rule of combination and the generalised Bayesian theorem*, International Journal of Approximate Reasoning, 9, pp1-35, 1993.

[10] P. Smets and R. Kruse (1997) *The Transferable Belief Model for Belief Representation*. A. Motro and Smets Ph. (eds.) (1993) Uncertainty Management in information systems: from needs to solutions. Kluwer, Boston, (1997) 343-368.

[11] H. Stephanou and S. Lu *Measuring consensus effectiveness by a generalised entropy criterion*. Multisensor integration and fusion for intelligent machines and systems. Ablex publishing ISBN:0-89391-863-6. 1995

[12] M. Roberts, D. Marshall, and G. Powell, *Improving Joint Tracking and Classification with the Transferable Belief Model and Terrain Information,* International Conference on Information Fusion, Edinburgh, 2010.

[13] F. Smarandache and J. Dezert, *Proportional conflict redistribution rules* Proceedings of the 8th international conference on information fusion, 25-29 July 2005.