A Pragmatic Approach for the use of Dempster-Shafer Theory in Fusing Realistic Sensor Data

PIERRE VALIN
PASCAL DJIKNAVORIAN
ÉLOI BOSSÉ

This article addresses the performance of Dempster-Shafer (DS) theory, when it is slightly modified to prevent it from becoming too certain of its decision upon accumulation of supporting evidence. Since this is done by requiring that the ignorance never becomes too small, one can refer to this variant of DS theory as Thresholded-DS. In doing so, one ensures that DS can respond quickly to a consistent change in the evidence that it fuses. Only realistic data is fused, where realism is discussed in terms of data certainty and data accuracy, thereby avoiding Zadeh's paradox. Performance measures of Thresholded-DS are provided for various thresholds in terms of sensor data certainty and fusion accuracy to help designers assess beforehand, by varying the threshold appropriately, the achievable performance in terms of the estimated certainty, and accuracy of the data that must be fused. The performance measures are twofold, first in terms of stability when fused data are consistent, and second in terms of the latency in the response time when an abrupt change occurs in the data to be fused. These two performance measures must be traded off against each other, which is the reason why the performance curves will be very helpful for designers of multi-source information fusion systems using Thresholded-DS.

Manuscript received August 4, 2009; revised December 2, 2009; released for publication December 15, 2009.

Refereeing of this contribution was handled by Jean Dezert.

Authors' address: P. Valin and El. Bossé, Decision Support Systems section, Defence R&D Canada Valcartier, 2459 Blvd Pie XI Nord, Québec, G3J 1X5, Canada; P. Djiknavorian, Département de Génie Électrique et de Génie Informatique, Université Laval, Québec, G1K 7P4, Canada.

1557-6418/10/$17.00 © 2010 JAIF

1. INTRODUCTION

Potential users of Dempster-Shafer (DS) theory [5, 10] are often faced at the outset with a list of its pitfalls, which they must somehow solve or at least live with:

1. When confronted with Bayesian reasoning over N identities, DS theory seems at a disadvantage. Indeed, since DS theory reasons over the power set, which has $2^N - 1$ elements, excluding the null set, the storage of all of the intermediate fusion results and the processing of them quickly can become overwhelming, when compared to Bayesian reasoning. However, many solutions were developed from 1993 until 1997, such as those of Simard et al. [4, 11], Tessem [14], and Bauer [1]. They all involve approximation (or truncation) schemes with 3 tunable parameters, and some have been researched extensively [2, 3] as to which values are appropriate for a given situation. One therefore takes the view that this problem can be solved, and we will then focus on cases with small values of N.
2. When the evidence to be fused is too consistent, DS theory will become certain of it after a sufficient number of steps, and will have an extremely hard time to react to a sudden real change in the evidence to be fused. This was solved by Simard et al. [4, 11] by preventing the ignorance from falling below a certain threshold, hereafter called $I_{\min}$, after each fusion step, one of the three tunable parameters mentioned previously. After setting the ignorance to $I_{\min}$, all the other masses are rescaled proportionately, so that these rescaled masses now sum up to $(1 - I_{\min})$. This is the approach we will follow here.
3. When evidence is too conflicting, the normalization step in DS theory can cause wild behaviours from one extreme to another. This is partially a problem in modeling the uncertainty of the data to be fused. We take the approach that the data must correctly be modeled by specifying its accuracy and certainty in a reasonable and realistic manner.

At this point, one should make more precise what is meant by data certainty and accuracy:

1. Certainty is a feature of the sensor that declares that a certain proposition is true with a given mass value m. With little loss of generality, one can assume for simplicity that the sensor declares only one proposition with mass m, and that the rest is assigned to the ignorance. This is likely the case, when the time allowed for decisions is critical, since it provides at each time step only one likely candidate for the declaration. In the example scenario described later, an Electronic Support Measures (ESM) sensor is likely to provide such a behaviour. In order to stress this point, the article will always mention in the text "sensor certainty."

2. Accuracy refers here to how often the data is likely to be wrong. For example, the association mechanism that is necessary to select which sensor data is to be associated to which track can sometimes be erroneous, particularly if it is single scan in nature. Accuracy is therefore a characteristic of the fusion process, not the sensor itself. In the case of the ESM sensor, miss-associations can occur for the bearing-only reports when the targets are densely found in that bearing angle. In order to stress this point, the article will always mention in the text "declaration accuracy."

One should point out at this time that any sensor will have a value for the uncertainty (or certainty) of its declaration(s), and that, however complex the association mechanism, the association mechanism will occasionally err in its contact-to-track (or track-to-track) correlations, which will provide an inaccuracy in the fusion results. In this sense, the performance characteristics that will be provided later below for Thresholded-DS can be applied to a wide range of sensors and positional fusion algorithms, with only very minor modifications.

2. STATEMENT OF THE PROBLEM AND SCENARIO

The selected problem was already used in publications [6–8] that addressed the use of Dezert-Smarandache (DSm) theory [12, 13] and compared it to Thresholded-DS. When the two approaches were compared in these publications, the focus was on DSm performance, while neglecting Thresholded-DS performance. It became quickly clear that, if one did not insist on conformance to STANAG 1241 [9] (which only DSm can provide), Thresholded-DS theory performed quite well. This article aims to fill this gap by exploring at much greater length the stability and response time of the theory for various threshold levels $I_{\min}$ in terms of sensor data accuracy and declaration certainty.

A possible illustration of the problem chosen is through the fusion of three types of ESM reports: Friend ($*gq_1$), Neutral (θ_2), or Hostile (θ_1). Since $N = 3$, the first pitfall of DS theory mentioned in the introduction is avoided, and no approximation schemes are necessary.

The approach followed in this article will be to study the ESM problem using a Modeling and Simulation (M&S) approach, first on specific representative scenarios, followed by a thousand Monte-Carlo runs to confirm the conclusions that can be reached.

The list of the prerequisites that any scenario must address are:

- Should have a clearly defined ground truth, which is sufficiently complex to test stability and latency in the response time.
- Should contain sufficient miss-associations, leading to values of average fusion accuracy that are in a realistic range.
- Should only provide partial knowledge about the ESM sensor declaration and to varying degrees, which therefore leads to sensor uncertainty (or sensor certainty) values that are in a realistic range.

The following scenario parameters have therefore been chosen accordingly:

1. The known ground truth is Friend (θ_1) for the first 50 time stamps of the scenario, and Hostile (θ_3) for the last 50 time stamps.
2. The percentage of correct associations is approximately Acc%, corresponding to countermeasures appearing (100 – Acc)% of the time. Acc% will be explored over a realistic range between 60% and 90%. If the accurate allegiance is Friend (as is the case for the first 50 time stamps), then the declarations which correspond to miss-associations are equally distributed between Neutral and Hostile. Similarly, for the last 50 time stamps when Hostile is the correct allegiance, the miss-associations are distributed evenly between Friend and Neutral.
3. The ESM declaration has a mass of m, with the rest $(1 - m)$ being assigned to the ignorance, reflecting a certainty percentage Cer% in the declaration. Cer% will be explored over a realistic range between 60% and 90%.

This section will show a representative example of such a scenario, but the rest of the paper addresses the general trends that can be established from 1000 Monte-Carlo runs, where a different random seed is chosen for each member of the sequence in each Monte-Carlo run.

Thresholded-DS should be able to adequately represent the main features of the ground truth (which is known in an M&S approach), namely

1. Show stability under occasional miss-associations, namely show stability when fused data are generally consistent, specifically for the first 50 time stamps (after a short ramp-up time) and the last 50 time stamps (after the ramp-up time, or latency, due to the allegiance change).
2. Switch allegiance when the ground truth does so, namely have a reasonable measured latency in the response time (or delay, hereafter denoted Δ) when an abrupt change occurs in the data to be fused.

A typical scenario, with the random number generator set to produce *on average* (for a set of Monte-Carlo runs) an Acc% = 80%, is shown in Fig. 1, with the x-axis representing the time index.

For this scenario, Thresholded-DS achieves the results shown in Fig. 2, given a typical value of $I_{\min} = 0.02$. In Fig. 2, the x-axis represents the time index, and the y-axis represents the value of basic belief assignment (or mass) associated with the given hypothesis. Note that Thresholded-DS therefore never becomes more than 98% sure of its fused result (as mentioned in the introduction).

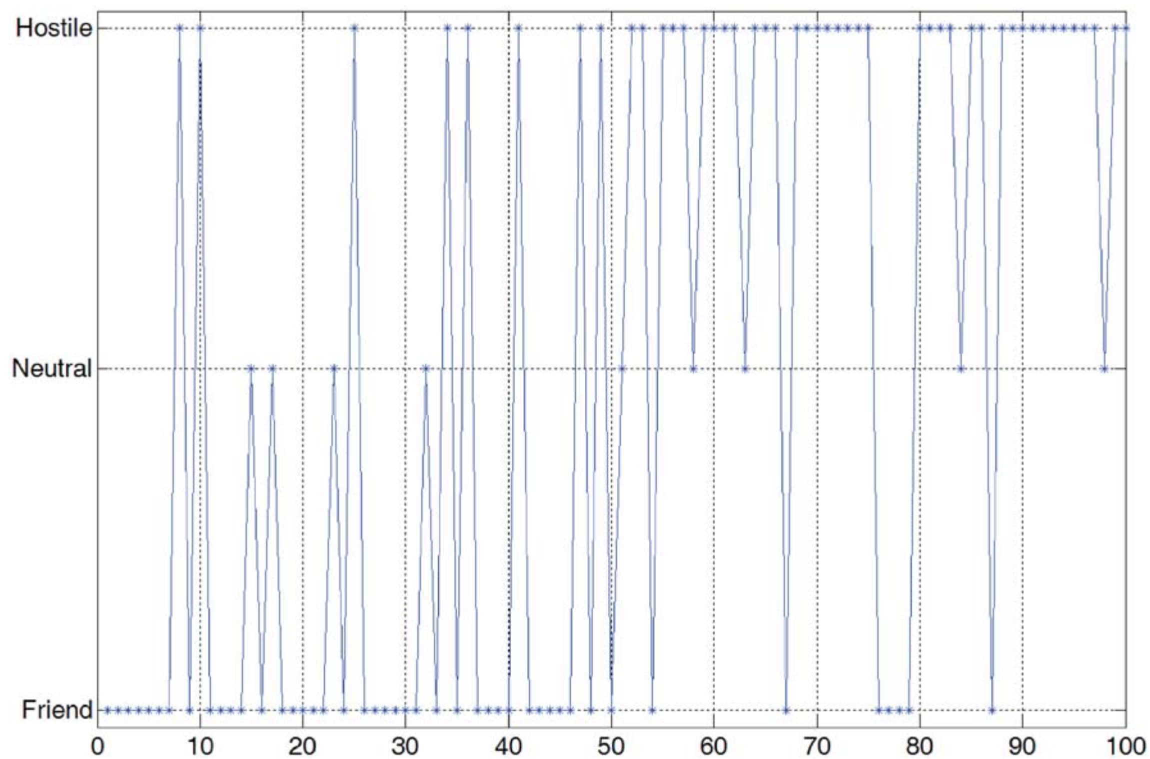

Fig. 1. Typical scenario with Acc% = 80%.

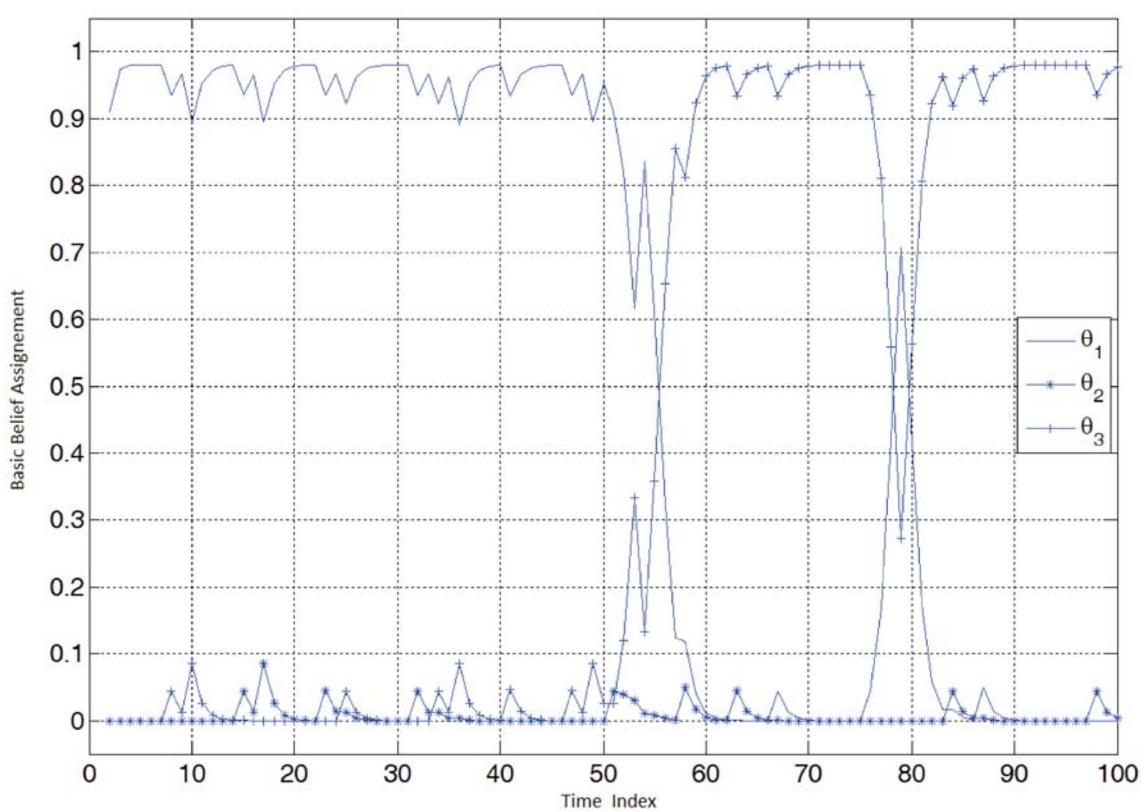

Fig. 2. Thresholded-DS for the typical scenario with Acc% = 80% and Cer% = 70%.

DS never becomes confused, shows good stability when miss-associations arrive randomly spaced out, which is the case until iteration 50. It then reacts reasonably quickly and takes about 8–10 reports before switching allegiance as it should. Furthermore, after being confused for an iteration around the sequence of

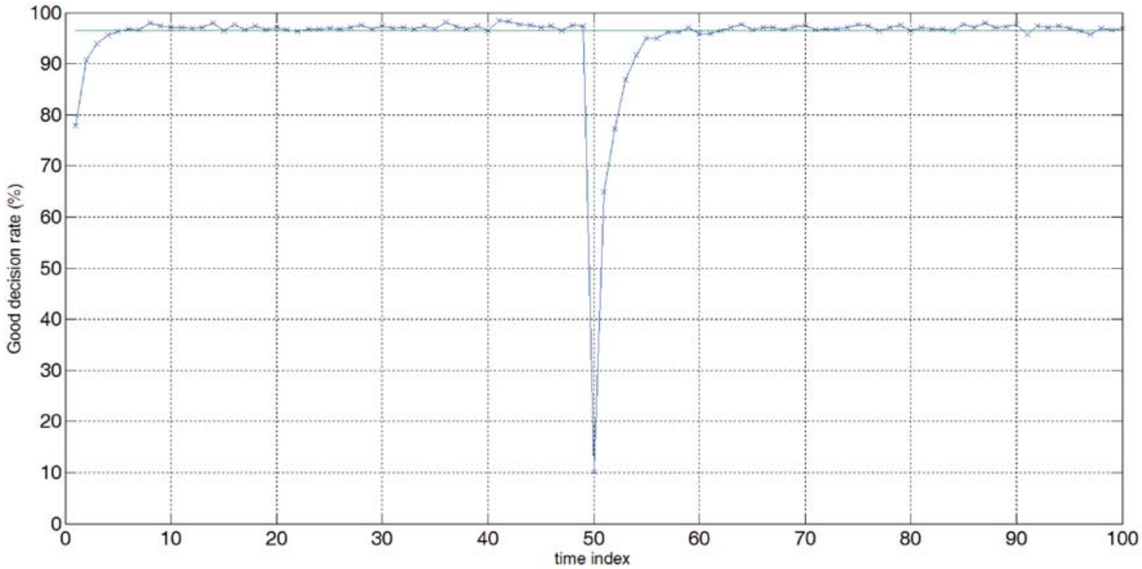

Fig. 3. Good decision rate for the scenario with Acc% = Cer% = 80% and 1000 runs.

four Friend reports starting at iteration 76, it quickly reverts to the correct Hostile status.

Fig. 3 shows a sample of a good decision rate of the target identification for Thresholded-DS using an input case such as the one from Fig. 1 generated randomly 1000 times. More specifically, it is the result of a Monte-Carlo simulation run of 1000 with an ESM sensor having values of accuracy and certainty both at 80% with the DS threshold at $I_{\min} = 0.05$ at every fusion step.

In order to evaluate the latency in the reaction time around iteration 50, we first determine the empirical mean averaged over time index 15 to 45 and 65 to 95, and then we subtract three times the value of the empirical standard deviation (3σ) averaged over the same interval. This interval has been chosen arbitrarily to exclude most of the instability that is mostly due to the initialization instability and the change of allegiance instability. So it will only include the instability of the decision system and the input data. The measure of latency then starts at time index 50, and ends at the time index at which the good decision rate reaches the threshold for reaction time performance shown as a horizontal line in Fig. 3. This horizontal line corresponds to the mean determined by the method above minus three standard deviations σ, which indicate the stability in the above mentioned time periods, according to the formulae for σ:

$$\sigma^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \mu)^2, \qquad \mu = \frac{1}{n}\sum_{i=1}^{n} x_i.$$

The standard deviation σ tends to a fixed value as a function of increasing n, as shown in Figs. 4(a) for 100 Monte-Carlo iterations and 4(b) for 1000 Monte-Carlo iterations (0.16% in this case on the y-axis, with the x-axis being again the time index), but show less noise as n increases. This shows that σ is a dynamical feature of

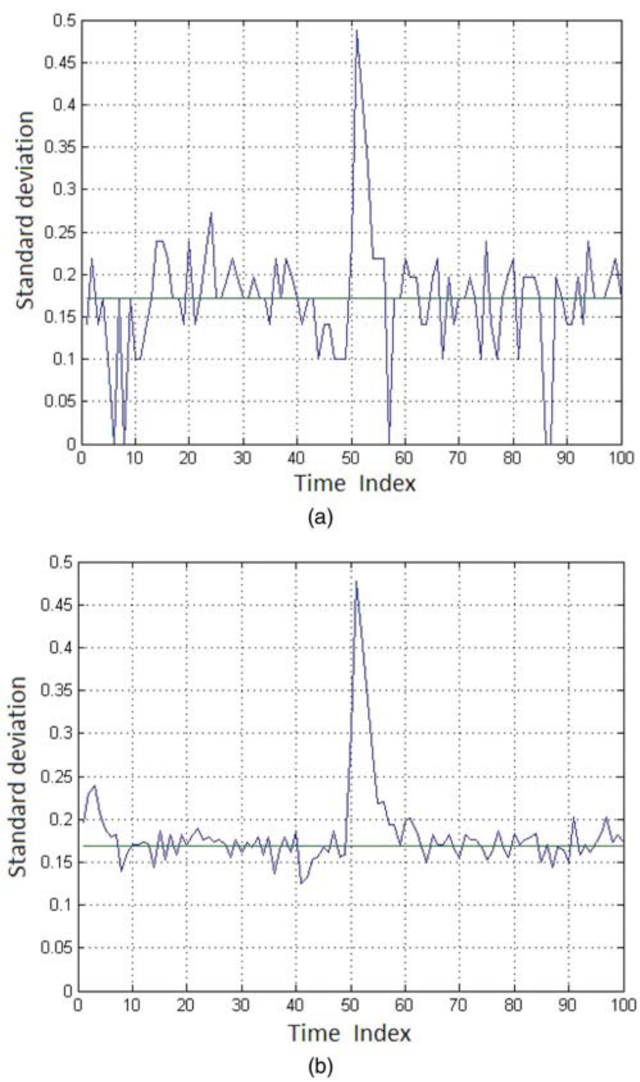

Fig. 4. (a) (top) and (b) (bottom). Standard deviations σ for stability (in %).

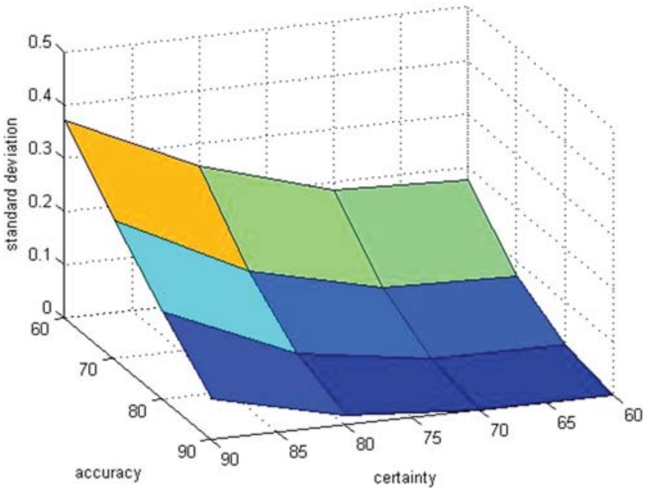

Fig. 5. Measure of stability σ for $I_{\min} = 0.01$.

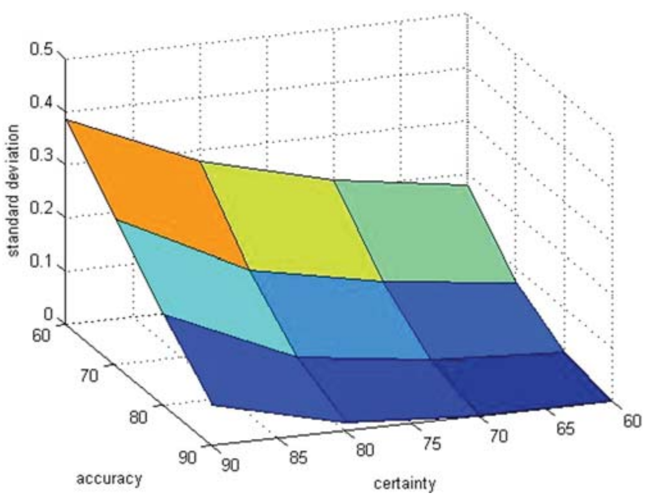

Fig. 6. Measure of stability σ for $I_{\min} = 0.02$.

the process, rather than being dependent on the number of Monte-Carlo runs.

Please note that this is just a practical definition of the latency, in order to show the trends in latency, when the parameters are varied, particularly $I_{\min}$. Other definitions may be more appropriate for other applications.

3. NUMERICAL GRAPHICAL RESULTS FROM MONTE-CARLO RUNS

This section shows the graphs for stability in the first subsection and reaction time latency (or delay) in the second subsection, for 1000 Monte-Carlo runs, for various values of the threshold in Thresholded-DS. Since one has three parameters to vary (certainty, accuracy, and $I_{\min}$), the presentation in this section focuses on showing the stability (in Subsection 3.1), and the reaction time latency (in Subsection 3.2) as a function of certainty and accuracy, with different figures corresponding to different choices for values of $I_{\min}$.

3.1. Stability

For an increase in the threshold of the minimum ignorance of 0.01 for each different figure, we have the following results for the standard deviation σ indicative of stability, for $I_{\min} = 0.01$ (Fig. 5), $I_{\min} = 0.02$ (Fig. 6), $I_{\min} = 0.03$ (Fig. 7), $I_{\min} = 0.04$ (Fig. 8), and $I_{\min} = 0.05$ (Fig. 9).

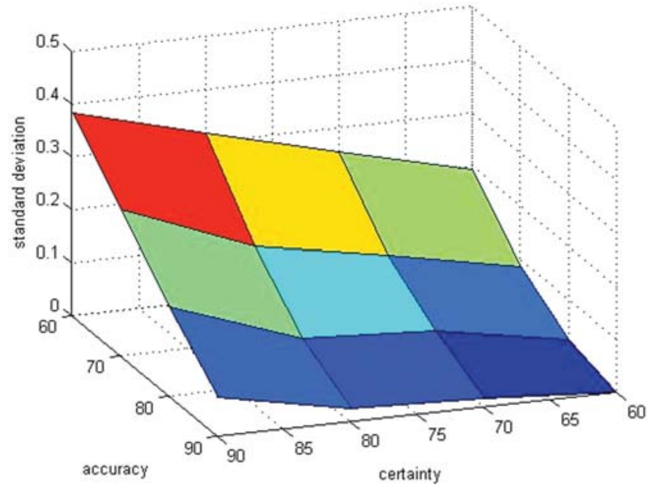

Fig. 7. Measure of stability σ for $I_{\min} = 0.03$.

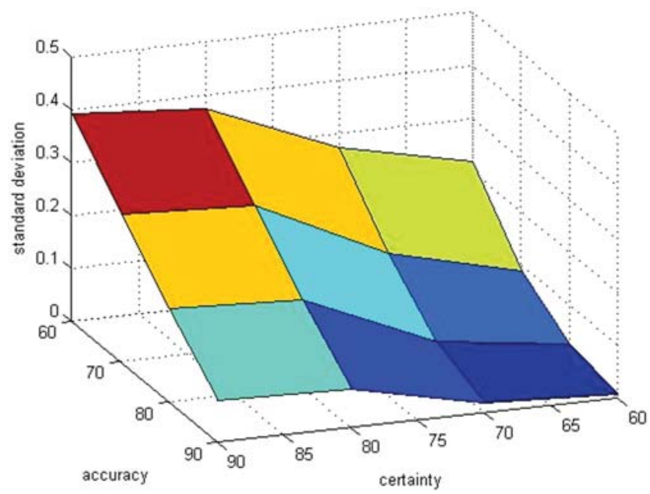

Fig. 8. Measure of stability σ for $I_{\min} = 0.04$.

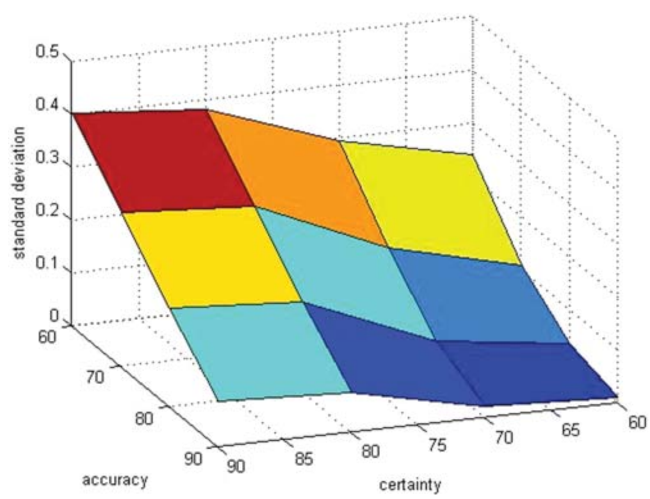

Fig. 9. Measure of stability σ for $I_{\min} = 0.05$.

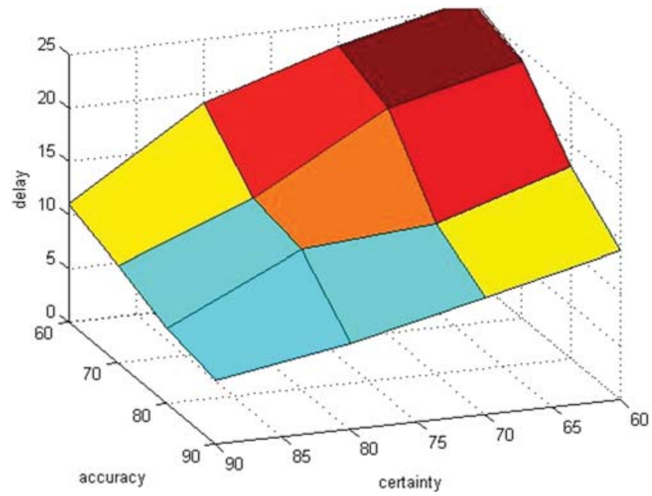

Fig. 10. Reaction time latency or delay for $I_{\min} = 0.01$.

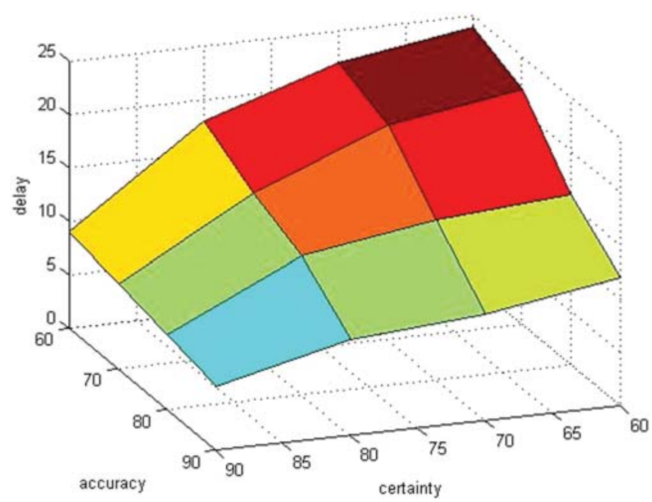

Fig. 11. Reaction time latency or delay for $I_{\min} = 0.02$.

Any much smaller value than 0.01 would result in too much rigidity when an allegiance changes, resulting in longer reaction time latency or delay (as will be shown in the next subsection). These figures show that any much larger result than 0.05 adversely affects stability, as can be seen when comparing Fig. 9, which becomes concave and has higher σ over all of the values of certainty and accuracy, with Fig. 5, which is convex and has lower σ over all of the values of certainty and accuracy. The intermediate figures show the slow deterioration in stability as $I_{\min}$ increases.

3.2. Reaction time latency

For an increase in the threshold of the minimum ignorance of 0.01 for each different figure, we have the following results for the reaction time latency (or delay Δ) in time units of the simulation scenario, with $I_{\min} = 0.01$ (Fig. 10), $I_{\min} = 0.02$ (Fig. 11), $I_{\min} = 0.03$ (Fig. 12), $I_{\min} = 0.04$ (Fig. 13), and $I_{\min} = 0.05$ (Fig. 14). Again this corresponds to 1000 Monte-Carlo runs.

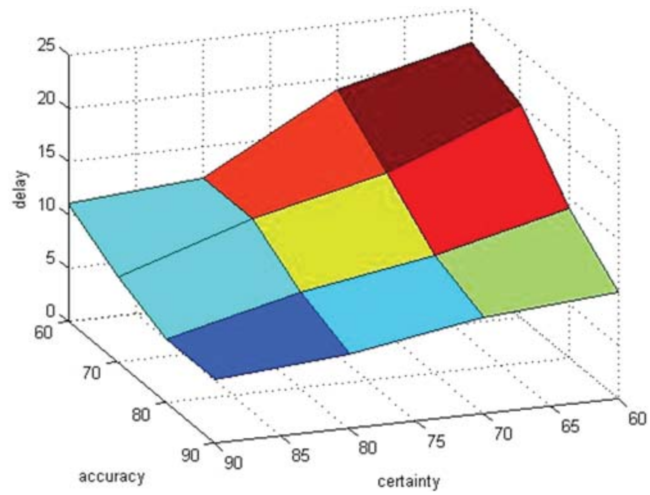

Fig. 12. Reaction time latency or delay for $I_{\min} = 0.03$.

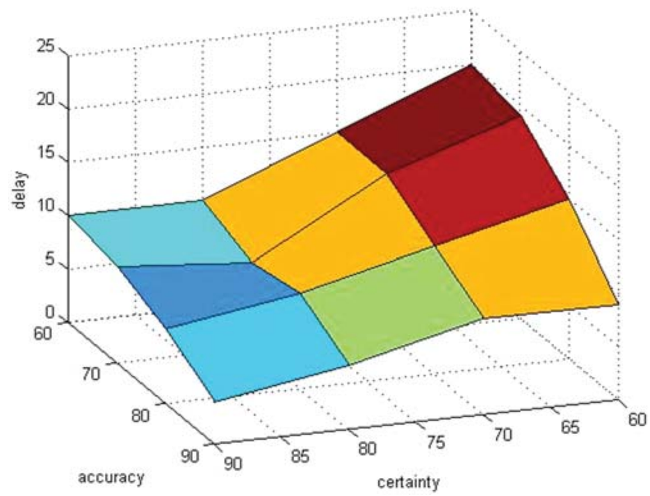

Fig. 13. Reaction time latency or delay for $I_{\min} = 0.04$.

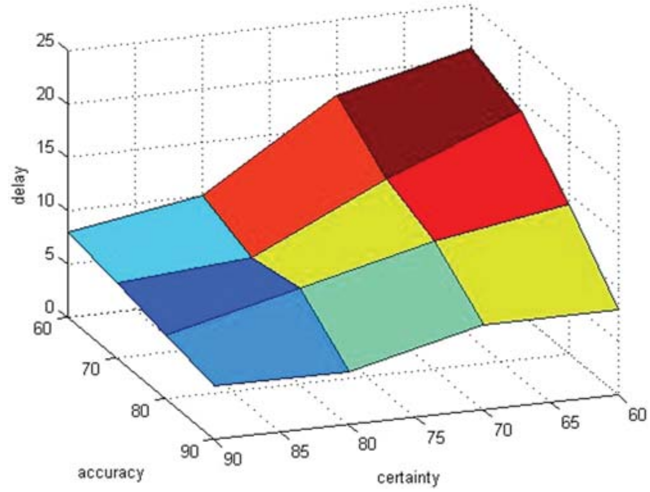

Fig. 14. Reaction time latency or delay for $I_{\min} = 0.05$.

These figures show that much smaller values of $I_{\min}$ than 0.01 result in too much rigidity when an allegiance changes, resulting in longer reaction time latency or delay. This is clearly seen by the much higher values for the delays in the surface of Fig. 10 when compared to Fig. 14, over all of the values of certainty and accuracy.

This is particularly notable for low values of certainty and accuracy: the delay exceeds 25 time units (or more than half the total time to recover from an allegiance change) when compared to Fig. 14, where it is about 20 time intervals. The effect is also very noticeable for high accuracy values (towards the reader). The intermediate figures show the slow improvement in the reaction time latency as $I_{\min}$ increases.

4. ANALYSIS OF THE GRAPHICAL RESULTS IN ORDER TO IDENTIFY TRENDS

The large amount of graphical data shown in the previous section can be interpreted rather simply for the instability (in Subsection 4.1 for Figs. 5–9) and reaction time latency Δ (in Subsection 4.2 for Figs. 10–14). Although the trends discussed in the following subsections can be phrased rather straight-forwardly, the trends themselves are non-linear, as can be seen by close inspection of the figures in the previous section.

4.1. Instability

Analysis of the performance measure of stability (or instability) of the Thresholded-DS system can identify the following trends from our various simulations shown in the last section.

1. For a fixed value of certainty, the value of instability increases when the accuracy decreases.
2. For a fixed value of accuracy, the value of instability increases when the certainty increases.
3. For fixed values of certainty and accuracy, the value of the **instability increases** when the value of the total ignorance threshold $I_{\min}$ **is increased**.
4. A change in accuracy affects more the instability than the certainty does.
5. Lower values of instability (good) are achieved with higher accuracy and lower certainty, and vice versa.

4.2. Reaction time latency

Analysis of the performance measure of reaction time latency (or delay Δ) of the Thresholded-DS system can identify the following trends from our various simulations shown in the last section.

1. For a fixed value of certainty, the value of the delay increases when the accuracy decreases.
2. For a fixed value of accuracy, the value of the delay increases when the certainty decreases.
3. For fixed values of certainty and accuracy, the value of the **delay increases** when the value of the total ignorance threshold $I_{\min}$ **is decreased**.

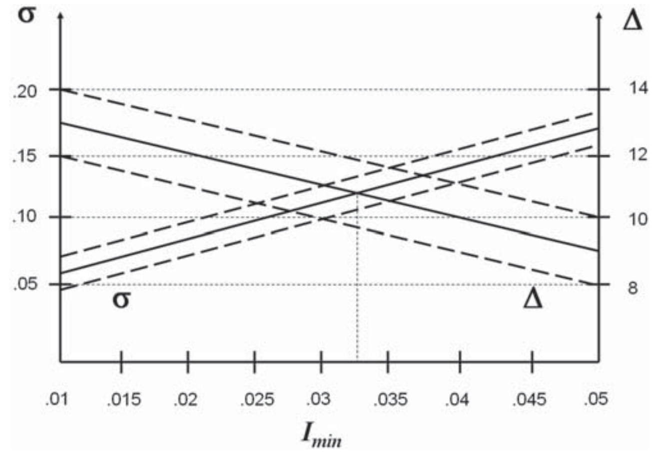

Fig. 15. Reaching a compromise for low σ and low Δ.

4. A change in accuracy affects more the delay than the certainty does.
5. Lower values of delay (good) are achieved with higher accuracy and higher certainty, and vice versa.

Points 3 in the above two lists clearly show that a compromise must be achieved when using Thresholded-DS between being responsive to any real change in the data, yet not being too responsive to fluctuations in the data, due to either poor sensor data certainty or fusion accuracy. In general, a high value for $I_{\min}$ will tend to respond to a stream of false reports rather quickly (bad) but will be very responsive to a real change in the data (good). A low value for $I_{\min}$ will provide excellent stability (good), but will react slowly to a real change in the data (bad).

The trends shown above are correct over the vast majority of the 16 points shown in the preceding Figs. 5–14. Only the exact values are shown in those figures, without the estimated errors from the Monte-Carlo runs.

The following Fig. 15 shows such a compromise as a function of $I_{\min}$, for a value of %Acc = %Cer = 80% with an estimate of errors, which cannot easily be portrayed in Figs. 5–14. The vertical axes represent σ (in % on the left) and Δ (in time units of the simulation) on the right, with the dashed lines showing approximate error bars given the limited number of Monte-Carlo runs (about 1000 runs). The figure shows that the interval $I_{\min} \in [0.025, 0.04]$ with a best value around 0.0325 can be selected.

5. CONCLUSIONS

This paper has provided performance measures of Thresholded-DS for various thresholds in terms of sensor data certainty and fusion accuracy to help designers assess beforehand, by varying the threshold appropriately, the achievable performance in terms of the estimated certainty and accuracy of the data that must be fused, i.e., an operating point for the application.

The threshold that the designers can choose according to figures similar to Fig. 15 depends on appropriate definitions for sensor certainty and latency (or delay) for their given application. Reasonable values were chosen here for an ESM application. In real applications, one should have an independent way of assessing the sensor certainty and the fusion accuracy in real-time. The Monte-Carlo runs provide the operating points, but it has to be assumed that the user can assess these operating points by monitoring the performance of the sensor as the mission develops (for example on well-isolated targets), and has calibrated the performance of the association mechanism in various conditions, which any manufacturer of such software should have done.

The performance measures are twofold, first in terms of stability when fused data are consistent, and second in terms of the latency in the response time when an abrupt change occurs in the data to be fused. These two performance measures must be traded off against each other, which is the reason why the performance curves will be very helpful for designers of identification fusion using Thresholded-DS.

REFERENCES

[1] M. Bauer
Approximation algorithms and decision making in the Dempster-Shafer theory of evidence—An empirical study.
International Journal of Approximate Reasoning, **17**, 2–3 (1997), 217–237.

[2] D. Boily and P. Valin
Truncated Dempster-Shafer optimization and benchmarking.
In *Sensor Fusion: Architectures, Algorithms, and Applications IV*, SPIE Aerosense 2000, vol. 4051, 237–246.

[3] D. Boily and P. Valin
Optimization and benchmarking of truncated Dempster-Shafer for airborne surveillance.
In NATO Advanced Study Institute on Multisensor and Sensor Data Fusion, Kluwer Academic Publishers, NATO Science Series, II, Mathematics Physics and Chemistry, vol. 70, 617–624.

[4] É. Bossé and M-A. Simard
Managing evidential reasoning for identity information fusion.
Optical Engineering, special section on sensor fusion, **37**, 2 (Feb. 1998), 391–400.

[5] A. Dempster
Upper and lower probabilities induced by multivalued mapping.
Ann. Math. Statist., **38** (1967), 325–339.

[6] P. Djiknavorian, D. Grenier and P. Valin
Analysis of information fusion combining rules under the DSm theory using ESM inputs.
In 10th International Conference on Information Fusion, 2007.

[7] P. Djiknavorian, P. Valin and D. Grenier
Dezert-Smarandache theory applied to highly conflicting reports for identification and recognition—Illustrative example of ESM associations in dense Environments.
DRDC Valcartier TR 2008-537, 34 p.

[8] P. Djiknavorian, P. Valin and D. Grenier
Fusion of ESM allegiance reports using DSmT.
In Advances and Applications of DSmT for Information Fusion, Collected Works edited by F. Smarandache and J. Dezert, vol. 3, American Research Press, 2008, 503–518.

[9] NATO Standardization Agreement (STANAG)
NATO standard identity description structure for tactical use.
STANAG 1241 (ed. 4), October 16, 1996.

[10] G. Shafer
A Mathematical Theory of Evidence.
Princeton University Press, 1976.

[11] M-A. Simard, P. Valin and E. Shahbazian
Fusion of ESM, radar, IFF and other attribute information for target identity estimation and a potential application to the Canadian Patrol Frigate.
In AGARD 66th Symposium on Challenge of Future EW System Design, AGARD-CP-546, May 1994, 14.1–14.18.

[12] F. Smarandache and J. Dezert (Eds.)
Advances and applications of DSmT for information fusion.
vol. 1, American Research Press, 2004.

[13] F. Smarandache and J. Dezert (Eds.)
Advances and Applications of DSmT for Information Fusion.
vol. 2, American Research Press, 2006.

[14] B. Tessem
Approximations for efficient computation in the theory of evidence.
Artificial Intelligence, **61** (June 1993), 315–329.

Pierre Valin received a B.Sc. honours physics (1972) and a M.Sc. degree (1974) from McGill University, then a Ph.D. in theoretical high energy physics from Harvard University (1980), under the supervision of the 1979 Nobel Laureate Dr. Sheldon Glashow.

He was faculty lecturer at McGill and Associate Professor of Physics in New Brunswick at Moncton and Fredericton before joining Lockheed Martin Canada in 1993 (then called Paramax), as a Principal Member of R&D. In 2004, he became a defence scientist at Defence R&D Canada (DRDC) at Valcartier, where he currently leads a research group in Future C2 Concepts & Structures. He is thrust leader for Air Command at DRDC since 2007.

He has been particularly active in the International Society of Information Fusion (ISIF) though the organization of FUSION 2001 and 2007. He has been ISIF board member since 2003, VP-membership since 2004, and was president in 2006. He is also an associate editor for the *Journal of Advances in Information Fusion* (JAIF).

Dr. Valin's interests focus mainly on the following topics: Multi-Sensor Data Fusion (MSDF) requirements and design, C2 systems, algorithmic benchmarking, use of a priori information databases, imagery classifiers (EO/IR and SAR) and their fusion, neural networks, fuzzy logic, information processing and uncertainty representation, reasoning techniques for recognition and identification (Bayes, Dempster-Shafer and Dezert-Smarandache), SAR image processing, Network Centric Warfare, distributed information fusion, dynamic resource management, as well as theoretical and mathematical physics.

Pascal Djiknavorian obtained a B.Eng. in computer engineering and a certificate in business administration in 2005 from Laval University. He also completed in 2008 an M.Sc. in electrical engineering on information fusion within the Dezert-Samarandache theory framework applied to ESM reports under STANAG 1241.

He is currently a Ph.D. student in information fusion at Laval University and is supervised by Professor Dominic Grenier and Professor Pierre Valin. He has ten publications in books, journals, and conference proceedings. His research interests include evidential theory, DSm theory, approximations algorithms, optimization methods.

Éloi Bossé received the B.A.Sc., M.Sc. from Laval University in 1981 in electrical engineering and Ph.D. degrees in 1990 from Ottawa/Carleton and Laval Universities.

In 1981 he joined the Communications Research Centre, Ottawa, Canada, where he worked on high resolution spectral analysis and radar tracking in multipath. In 1988 he was transferred to the Defence Research Establishment Ottawa (DREO) and in 1992 to the Defence Research and Development Canada Valcartier (DRDC Valcartier) where he worked on information fusion and decision support. He is an adjunct professor at Laval University (Québec), University of Calgary and McMaster University (Hamilton). He now heads the C2 Decision Support Systems Section at DRDC Valcartier. He was the executive chair of the 10th International Conference on Information Fusion (FUSION 2007), held in July 2007 in Québec City, Canada.

Dr. Bossé has published over 160 papers in journals, conference proceedings and technical reports. He is coauthor of the following book on information fusion: *Concepts, Models, and Tools for Information Fusion*, Artech House, Norwood, Dec. 2006, 352 p. (NAMRAD Principals 2006 achievement AWARD).