Defence Research and Development Canada
Recherche et développement pour la défense Canada

DEFENCE DÉFENSE

Information quality effects on information fusion

Galina Rogova
Calspan-UB Research Center, Inc.

Éloi Bossé
DRDC Valcartier

Defence R&D Canada – Valcartier
Technical Report
DRDC Valcartier TR 2005-270
May 2008

Canada

Information quality effects on information fusion

Galina Rogova
Calspan-UB Research Center, Inc.

Éloi Bossé
DRDC Valcartier

Defence R&D Canada – Valcartier

Technical Report

DRDC Valcartier TR 2005-270

May 2008

Principal Author

Éloi Bossé

Approved by

Éloi Bossé
Section Head/C2 Decision Support Systems

Approved for release by

Christian Carrier
Chief scientist

Abstract

In this report, we present results from simulation and analysis that show how to take into account data/information reliability into fusion processes. The report gives the results of an analysis of the main approaches used to address information/data/source reliability. The question of reliability and existence of prior knowledge that impact the data fusion processes, is discussed in the context of probability, possibility, evidential, or fuzzy sets theories. We consider reliability as a second level of uncertainty (uncertainty of evaluation of uncertainty) representing a measure of the adequacy of the model used and the state of the environment observed; so then, reliability represents adequacy of each belief model to the reality. The report reviews, as well, the existing methods of building reliability functions based on probabilistic, possibilistic, evidential frameworks and their effectiveness.

Résumé

Des résultats d'analyse complémentés par des simulations sont présentés dans ce rapport pour montrer comment prendre en compte la notion de fiabilité des sources et de l'information. Le rapport nous donne les résultats d'analyse des principales approches pour aborder la fiabilité des sources d'information. La question de fiabilité et d'existence de la connaissance a priori des processus de fusion nous est présentée par les représentations des théories des probabilités, possibilités, de l'évidence et des ensembles flous. Nous considérons la fiabilité comme un second niveau d'incertitude, c'est-à-dire une évaluation de l'incertitude représentant une mesure de la justesse du modèle utilisé et l'état de l'environnement observé. La fiabilité représente la justesse de chaque modèle de croyance par rapport à la réalité. Le rapport passe aussi en revue les méthodes existantes pour construire des fonctions de fiabilité dans le cadre des probabilités, des possibilités, de la théorie de l'évidence et de leur efficacité.

This page intentionally left blank.

Executive summary

Information quality effects on information fusion:

Galina Rogova, Éloi Bossé; DRDC Valcartier TR 2005-270; Defence R&D Canada – Valcartier; May 2008.

Data and/or information fusion is a key enabler to meeting the demanding requirements of military command decision support systems, mainly regarding situation awareness. The success of information fusion is defined by the quality of knowledge produced by fusion processes, which in turn depends on how well data are represented, how good and adequate are the data and uncertainty model being used, and how accurate and appropriate or applicable a priori knowledge is.

The subject of information/data/source reliability has been receiving significant attention in the last couple of years in areas such as communication, personal computing, and databases. At the same time, there is a relatively small body of literature in the data fusion community that addresses this topic; moreover, there is no clear understanding of what defines information reliability from the perspective of information fusion processes designers and how reliability of sources affect fusion results.

This report provides a comprehensive survey of the main approaches used to address information/data/source reliability in the fusion literature as well as results from simulations to compare approaches. Generally, all the approaches to incorporation of reliability into fusion processes fall into three categories: methods measuring reliability of data input to fusion processes and eliminating data of poor reliability, methods modifying the data and information by considering their reliability before fusion, methods modifying the fusion process to account for the reliability of the input. The choice of one of the strategies is defined by many factors and is directly related to the question of reliability and the existence of prior knowledge, which is modeled in the context of probability, possibility, evidential, or fuzzy sets theory, and the degree to which any given data is representative of that prior knowledge.

We assume that we have a model, which utilizes measurements, observations, intelligence reports, and prior knowledge to provide us with a degree of belief in the occurrence of an event. These degrees of belief take values in a real interval and are modeled within a framework of a particular uncertainty theory (probabilities in probability and Bayesian theory, possibility distributions, membership degrees in fuzzy set theory, certainty factor, basic probability assignments, beliefs, plausibility, or pignistic probabilities in evidence theory).

In this report, we presented results from simulations that show that incorporation of reliability into fusion processes gives "richer behavior" to the fusion system and

can improve performance of fusion systems while producing many theoretical and practical problems not very often addressed in the data fusion literature. Among thm are the problem of estimation of reliability of sources and their temporal analysis; the problem of interrelationship between reliability of information sources and the reliability of fusion results; the problem of incorporating reliability into fusion of heterogeneous information.

Sommaire

Information quality effects on information fusion:

Galina Rogova, Éloi Bossé ; DRDC Valcartier TR 2005-270 ; R & D pour la défense Canada – Valcartier ; May 2008.

La fusion d'information et/ou de données est une fonction habilitante pour satisfaire les besoins des systèmes militaires d'aide à la décision, particulièrement pour augmenter l'éveil situationnel. Pour réussir, la fusion nécessite que la qualité de l'information soit évaluée, qui à son tour dépend de la représentation des données ou de l'information. La justesse du modèle de représentation des connaissances et de la connaissance elle-même ont un impact important sur la performance du processus de fusion.

La fiabilité des sources et de l'information a reçu une attention significative ces dernières années dans le domaine des communications, de l'informatique et des bases de données. D'un autre côté, il y a peu de littérature dans le monde de la fusion qui aborde ce sujet. De plus, il n'y a pas de compréhension claire de la signification de la fiabilité de l'information par rapport aux processus de fusion et de l'impact sur les résultats.

Ce rapport nous donne une analyse des principales approches pour prendre en compte la notion de fiabilité des sources et de l'information par rapport à la fusion. Des simulations supportent l'analyse dans un but de comparaison des différentes approches. On retrouve généralement trois catégories d'approches pour traiter de ce problème : 1) les méthodes qui estiment la fiabilité des entrées au processus de fusion et éliminent les données qui ont une faible fiabilité ; 2) les méthodes qui modifient les données ou l'information en tenant compte de leur fiabilité avant le processus de fusion ; 3) les méthodes qui modifient les processus de fusion pour compenser le degré de fiabilité des entrées. Le choix d'une stratégie est motivé par plusieurs facteurs et est sujet à la question de fiabilité et de l'existence des connaissances a priori qui sont modélisées dans le cadre des probabilités, des possibilités, de la théorie de l'évidence ou des ensembles flous, et à quel degré la donnée est représentative de cette connaissance a priori.

Nous posons comme hypothèse qu'on a un modèle qui utilise des mesures, des rapports du renseignement et de la connaissance a priori nous donnant un certain degré de croyance qu'un événement se produise. Ce degré de croyance prend des valeurs numériques qui sont modélisées selon un cadre particulier de représentation de l'incertitude (probabilité, possibilité, etc).

Dans ce rapport, nous montrons qu'il y a intérêt à incorporer la fiabilité dans le processus de fusion et cela peut améliorer la performance. Par contre, cela entraîne

des problèmes pratiques et théoriques qui ne sont pas relatés dans la littérature. Parmi ces problèmes, on retrouve l'estimation et l'analyse temporelle de la fiabilité des sources. On trouve également l'interrelation entre la fiabilité des sources et la fiabilité du résultat de la fusion. Finalement, on parlera du problème de fiabilité dans la fusion de sources hétérogènes.

Table of contents

List of figures

List of tables

This page intentionally left blank.

1 Introduction

Information Fusion is a field of study utilizing data coming from various input sources, techniques exploiting this data, a priori knowledge, and "known" models of uncertainty in various input sources to produce estimates and knowledge about objects and situations (see e.g. [2–5]). The success of information fusion is defined by the quality of knowledge produced by fusion processes, which in turn depends on how well data are represented, how good and adequate the data and uncertainty models are used, and how accurate and appropriate or applicable a priori knowledge is.

As explained in Meriam-Webster Unabridged Dictionary, data means factual information given or admitted as measurements or statistics to be used as a basis for reasoning, inferencing, discussion, or calculation. Some problems related to the data representation issues are: leaving out significant variables and associated data, not having enough relevant data, using inappropriate data or an inappropriate model of data, and using data beyond its applicable range. Information is defined as the judgments that given data resolve questions related to a data uncertainty model, disclose or reveal distinctions, or enable new action.

Knowledge is defined as beliefs (a hypothesis about some unobservable situation with or without confidence supporting it), facts (something, which makes a belief true or false), and heuristic rules. Knowledge is the capacity for effective action in a domain of human actions. Among the important knowledge quality related issues are interrelationship between the level of confidence of the judgment, hypotheses, or facts and their reliability [6–8].

The subject of information and data reliability has been receiving significant attention in the last couple of years in areas such as communication, personal computing, and databases. At the same time, there is a relatively small body of literature in the data fusion community that addresses this topic; moreover, there is no clear understanding of what defines information reliability from the perspective of information fusion processes designers and how reliability of sources affect fusion results. The main body of the literature on information fusion concerns with building an adequate uncertainty model without paying much attention to the related problems of reliability beliefs produced by this model. The majority of fusion operators is based on optimistic assumptions about reliability of the models producing beliefs and assumes that they are equally reliable and play a symmetrical role. At the same time, different models may have different reliability and it is necessary to account for this fact in order to avoid decreasing in performance of fusion results. Although the concept of reliability has been introduced, it allows for various interpretations and representations and is not yet well established [9]. The focus of the research presented in this report is to provide a discussion of definitions, issues, and a review of methodologies related to incorporation of belief reliability into data fusion.

2 Using reliability in fusion

In general, the problem of incorporating information reliability in information fusion can be represented in the following way. Let $s_i, i = 1, \ldots, I$ be data produced by I sources (e.g. sensors or intelligence reports). Let $\Theta = \{\theta_1, \ldots, \theta_N\}$ be a set of hypotheses under consideration (in belief theory, Θ is called a frame of discernment, in possibility theory it is universe of discourse). Such an hypothesis may be a particular value for a parameter measured by a sensor (e.g. distance to a certain target or the presence or absence of a certain structure, or a certain hypothesis about particular state of the world in the case of a classification problem or the higher level fusion).

We assume that we have a model which utilizes measurements, observations, intelligence reports, and prior knowledge to provide us with a degree of belief in event $A \subseteq \Theta$: $x_i(A), i = 1, \ldots, I$. These degrees of belief take values in a real interval and are modeled within a framework of a particular uncertainty theory (probabilities in probability and Bayesian theory, possibility distributions, membership degrees in fuzzy set theory, certainty factor, basic probability assignments, beliefs, plausibility, or pignistic probabilities in evidence theory). The degree of belief based on fused information is defined by operator $F(x_1, x_2, \ldots, x_I)$ built in the framework of a mathematical theory used for belief representation (see e.g., a classification of different fusion operators in [10]). Most processes of parallel combination of beliefs described in the literature are concerned with adequate modeling of uncertainty, which is the result of noisy, imprecise, erroneous or ill-suited to the problem data, ambiguous observations, and incomplete and poorly defined prior knowledge [11]. The difficulty in modeling uncertainty stems from the difficulty in finding an adequate belief model. Most of the belief models are derived from neighborhood information according to a distance (see, e.g., [12, 13]) or to likelihood functions (see e.g., [3, 14, 15]). An inappropriate choice of metrics or poor estimation of the likelihood functions can provide an inadequate belief model and can lead to conflicting and unreliable beliefs to be combined. Moreover, beliefs can be modeled within different uncertainty frameworks and dealing with different sources may mean also dealing with different framework theories. In this report we assume that beliefs to be combined are represented within the same uncertainty framework.

A natural way to deal with this problem is to establish reliability of the beliefs computed within the framework of the model selected and take it into account while building fusion operators. We consider reliability as a second level of uncertainty (uncertainty of evaluation of uncertainty) representing a measure of the adequacy of the model used and the state of the environment observed. There are at least two approaches used for defining reliability as a higher order uncertainty [16]. In one of them reliability is understood as relative stability of the first order uncertainty. In this case reliability is often measured by the performance of each source (e.g., by the recognition or false alarm rates. Another approach to representing a higher level of uncertainty is to measure accuracy of predicted beliefs). Here reliability represents adequacy of each belief model to the reality. The discussion in this report concentrates on this second notion of reliability.

In the case when we take into consideration source reliability, different situations can be considered [17]:

1. It is possible to assign a numerical degree of reliability to each source.

2. A subset of sources is reliable but we do not know which one.

3. Only an order of the reliabilities of the sources is known but no precise values.

Dealing with these situations calls for one or all of the following strategies of incorporating reliability into fusion processes:

- Strategies for identifying reliability of sources to fusion processes and elimination of data of poor reliability.

- Strategies for modifying beliefs by considering their reliability before fusion.

- Strategies for modifying fusion processes to account for the reliability of the input.

The choice of one of the abovementioned strategies is defined by many factors and is directly related to the question of reliability and existence of prior knowledge, which is modeled in the context of probability, possibility, evidential, or fuzzy sets theory (see e.g. [10,18–24]), and the degree to which any given data is representative of that prior knowledge. If such prior knowledge exists and is reliable, the quality of data can be measured and then adjusted by modifying the data or even eliminating it from consideration for fusion in a filtering-type approach. For example, it is possible to introduce a measure of reliability (reliability coefficients) and use it in a corresponding combination rule. When such a priori information does not exist or cannot be considered reliable, other strategies in order to account for sources reliability can be considered.

Let us assume that we are able to establish a measure of reliability or "trust" for each belief: $R_i \in [0,1], i = 1, \dots, I$. R_i is close to zero if source i is unreliable and is closer to 1 if it is more reliable. Each value of reliability may be "relative" or "absolute" and they may or may not be linked by an equation such as $\sum_i R_i = 1$. Reliability of source i can also be a vector $\overline{R}_i = \{R_{i1}, \dots, R_{iN}\}$, where R_{in} is reliability of belief of source i into a particular hypothesis. In general, $R_{in} \neq R_{im}$, if $n \neq m$ since belief of the same source in one hypothesis can be more reliable than in the other. Reliability coefficients depend not only on a model selected but also on characteristics of the environment and a particular domain of the input and, therefore, we can have $R_i = R(M_i, \overline{\gamma}, \overline{Y})$, where M_i is a model chosen for source i, $\overline{\gamma}$ is a vector parameters characterizing external environment of the source (e.g., meteorological conditions, or the deception ability of the observed object). Vector $\overline{Y}$ represents parameters characterizing the internal environment of the source (tuning parameters).

The values of reliability are used in fusion $F_R = F(x_1, \ldots, x_I, R_1, \ldots, R_I)$ to have a smooth transition between reliability (conjunctive rule) and unreliability (disjunctive rule). The operator F_R is "context dependent" [10] since it depends not only on the values of x_i but also on global knowledge. The form of this operator also depends on the strategy of incorporating reliability factors and the mathematical theory utilized:

- When strategies identifying the quality of data input to fusion processes and elimination data of poor quality are used, then $F_R = F(\overline{X}_j)$ where : $\overline{X}_j = (x_{i_1}, \ldots, x_{i_j}), i_j \leq I$ and $\overline{X}_j$ satisfies a certain criterion.
- When strategies utilizing reliability coefficients are used we can consider two cases:
  - It is possible to include reliability into modeling belief for each source before fusion to compensate for their different reliability and make them totally or, at least, equally reliable before fusion and then fuse transformed beliefs (separable case):

$$F_R = F(g(x_1, R_1), \ldots, g(x_I, R_I)) \tag{1}$$

  - Each source cannot be transformed independently and reliability coefficients modify the fusion operator considered (non-separable case):

$$F_R = F(g(x_1, \ldots, x_I, R_1, \ldots, R_I)) \tag{2}$$

- In many cases, a new fusion function represents a combination of all the abovementioned strategies.

The next chapters will review the existing methods of building function based on probabilistic, possibilistic, evidence methods and their effectiveness.

3 Probabilistic fusion operators

In the probabilistic and Bayesian framework the degrees of belief are represented by a priori, conditional, and a posteriori probabilities. Usually, decisions are made on a posteriori probabilities $P(\theta_n|y_i), \forall i, n$, where y_i is a measurement or feature coming form source i, $P(\theta_n|y_i)$ represents statistics of each source to be combined (data sources, outputs of classifiers to be combined). The fusion is usually performed by the Bayesian rule, which under the condition of source independence is reduced to a product:

$$F_n(x_1,\ldots,x_I)|_{y_i} = F_n(\overline{P})|_{y_i} = P(\theta_n)\prod_i^I \frac{P(\theta_n|y_i)}{P(\theta_n)} \quad ,\forall n \tag{3}$$

This fusion operator is conjunctive and assumes equal reliability of the sources. If the sources are not equally reliable, several fusion rules within the framework of the probability theory have been proposed in the literature. The majority of such methods (see, e.g., [25–29]) are represented by so-called trade off rules involving general procedures of combining single source probability distributions with the assumption that the decisions are based on Bayesian decision theory. One most commonly used rule is the linear opinion pool:

$$F_n(x_1,\ldots,x_I,R_1,\ldots,R_I)|_{y_i} = F_n(\overline{P},\overline{R})|_{y_i} = \sum_i P(\theta_n|y_i)R_i, \tag{4}$$

where R_i is reliability associated with the sources in the global membership function to express quantitatively the goodness of each source.

Among other rules presented in the literature are logarithmic opinion pools:

$$F_n(\overline{P},\overline{R})|_{y_i} = P(\theta_n)\prod_i^I \left[\frac{P(\theta_n|y_i)}{P(\theta_n)}\right]^{R_i} \tag{5}$$

Equations (3) - (5) assume that reliability R_i of each source i is the same for beliefs in each hypothesis θ_i. However, a similar formalism can be used with an assumption that the source reliabilities are different for different hypotheses. In this case classwise reliabilities R_{in} can be used in many fusion models and equations (3) - (5) can be changed in an obvious way. Equation (5) can be reduced to the following formulation when $\sum_{i=1}^I R_i = 1$:

$$\log F_n(\overline{P},\overline{R})|_{y_i} = \sum_i^I R_i \log P(\theta_n|y_i) \tag{6}$$

Several publications addressed theoretical and practical issues of comparison of the performance of weighted average classifiers ("trained classifiers") with simple averaging ("fixed classifiers") [30–35]. The theoretical model designed showed that "trained classifiers" could significantly outperform "fixed classifiers" only for a classifier with a large difference between the error rates of the best and the worst classifier. Moreover, the differences in correlation also play an important role in the performance improvement achievable by weighted averaging. Experiments conducted with remote-sensing images and biometrics supported these theoretic results.

The designed method introduced in [36,37] is based on the idea that sensor reliability depends on the context of sensor acquisition. This method integrates contextual information into the target tracking domain. The contextual analysis supervising tracking is able to detect the sensors which are reliable and those which are not, in distributed operational situations. The developed algorithm automatically increases the importance of measurements of reliable sensors and decreases the importance of unreliable ones. Fuzzy logic is used to represent the expert knowledge to describe validity of the sensors.

The method utilizes the fact that, in any given context, only a subset J of a set I of all sources to be combined is valid or reliable (i.e. their belief model adequately represents reality).

$$F_n(x_1, \ldots, x_I, R_1, \ldots, R_I)|\overline{y} = \sum_{J \subseteq I} P(\theta_n | y_1, \ldots, y_I, A_J) P(A_J), \tag{7}$$

where $P(A_J)$ is the probability of validity of the subset J of sensors. This probability is calculated thanks to the reliability R_i of the individual sensors. It is clear that this strategy is not separable and then falls into the second substrategy.

The optimal Bayesian fusion rule (OBFL) for fusion of unreliable classifiers is introduced in [38]. In this model, each source i can discriminate between a set $\Theta^i \subseteq \Theta$ and its complement $\neg\Theta^i \subseteq \Theta$ with reliability $R_i = (1-\alpha_i, 1-\beta_i)$, where $\alpha_i = P(\theta_n \in \Theta^i | \theta_n \notin \Theta^i)$ and $\beta_i = P(\theta_n \in \neg\Theta^i | \theta_n \in \Theta^i)$ correspond to false alarm and miss probabilities. This modeling assumes that each classifier can discriminate only between several subsets of Θ and is different from the classical Bayesian model, which usually provides probability for each atomic hypothesis. The OBFL computes a posteriori probabilities for each source and each hypothesis based on produced beliefs, reliability factor and prior probabilities and then fuses them with the Bayesian rule. This model requires the full knowledge of prior probabilities and reliability factors R_i. The model has been intensively validated through comparison with Monte Carlo simulation and proved to be of very good accuracy and without a high computational cost.

4 Reliability and evidence theory

Evidence theory is recognized as the "broadest and best-suited to the interpretation of the data considered and also the most federative in terms of synergy with the related processes (especially as concerns the stochastic filtering)" [11]. In the framework of the evidence theory [19–21, 24], information obtained from source i is represented by the basic probability assignments. Let Θ be a set of all possible events under consideration and 2^Θ be the set of all subsets of Θ. A function m is called a basic probability assignment if:

$$m : 2^\Theta \rightarrow [0,1], \qquad m(\varnothing) = 0, \qquad \sum_{A \subseteq \Theta} m(A) = 1 \tag{8}$$

$m(A) \in [0,1]$, where $A \in 2^\Theta$ and Θ is a set of all possible events under consideration. Basic probability assignment $m_i(A)$ characterizes belief of source i that observed data belongs to subset $A \in 2^\Theta$. Other functions, e.g., belief ($\mathrm{Bel}(A)$) and plausibility ($\mathrm{Pl}(A)$), are derived from the basic probability assignment:

$$\begin{aligned} \mathrm{Bel}(A) : 2^\Theta \rightarrow [0,1], \quad & \mathrm{Bel}(A) = \sum_{B \subseteq A} m(B), \\ \mathrm{Pl}(A) : 2^\Theta \rightarrow [0,1], \quad & \mathrm{Pl}(A) = \sum_{B \cap A \neq \varnothing} m(B). \end{aligned} \tag{9}$$

m is said to be a simple support function with focus $F \subseteq \Theta$ if $m(F) = s \neq 0$, $m(\Theta) = 1 - s$ and $m(A) = 0, \forall A \neq F$. A simple support function invests all its belief into its focal element. Combinations of simple support functions are called separable support functions.

Fusion of independent and equally reliable basic probability assignments is performed by the Dempster rule of combination:

$$\begin{aligned} m(\varnothing) =& 0 \\ m(A) =& C^{-1} \sum_{\cap A_j = A} \prod_i m_i(A_j), \qquad A \neq \varnothing \\ C =& \sum_{\cap A_j \neq \varnothing} \prod_i m_i(A_j) \end{aligned} \tag{10}$$

Normalization coefficient C is used in the case of the exhaustive frame of discernment and $1 - C$ measures the conflict between sources.

Let us assume that the beliefs $m_i(A)$ are not equally reliable and $R_i, i = 1, \ldots, I$ are reliability factors. Then one of the combination rules to be considered is the trade off rule, derived from probability, where reliability factors serve as weights assigned to each sources [1]:

$$m(A) = \sum_{i=1}^{I} R_i m_i(A) \tag{11}$$

The trade off rule is similar to a weighted average rule based on probabilities.

There are also several approaches to include belief reliability into the Dempster rule of combination. One is to "discount" beliefs provided by each source by R_i according to certain rules.

In the rule introduced in [19, 24] for each source i we will have:

$$\begin{aligned} m_i^{disc}(A) =& g(m_i(A), R_i) = R_i m_i(A), \qquad\qquad \forall A \subset \Theta, \\ m_i^{disc}(\Theta) =& g(m_i(A)) = 1 - R_i + R_i m_i(\Theta). \end{aligned} \tag{12}$$

The discounting model defined by equations (11) and (12) is used in [39, 40], in which basic probability assignments based on Gaussian distribution learned from a training set are produced. The method has been applied to diagnosis in dermatology.

In [11,41] a different model for incorporating reliability into fusion of different sources has been considered. In this model, a "quality factor" $R_{in} \in [0, 1]$ is associated with a likelihood of each hypothesis $C_{in} \in [0, 1]$. C_{in} is based on sensor measurements and a priori knowledge. It is assumed that R_{in} includes the quality of a priori knowledge used for building C_{in} and is known. It is also assumed that $C_{in} \geq 0$ only for θ_n, $\neg\theta_n$, Θ, i.e. Θ can be considered as a simple support function, and when $C_{in} = 0$ is valid ($R_{in} = 1$), θ_n is not verified.

Two models of $g(m_i(A))$ are considered for each source $i = 1, 2, \ldots, I$ and for each hypothesis θ_n, $n = 1, 2, \ldots N$:

1.

$$\begin{aligned} m_{in}(\theta_n) &= g(m_{in}(\theta_n), R_{in}) = 0 \\ m_{in}(\neg\theta_n) &= g(m_{in}(\neg\theta_n), R_{in}) = R_{in}(1 - C_{in}) \\ m_{in}(\Theta) &= 1 - R_{in}(1 - C_{in}) \end{aligned} \tag{13}$$

2.

$$\begin{aligned} m_{in}(\theta_n) &= g(m_{in}(\theta_n), R_{in}) = R_{in} C_{in} \\ m_{in}(\neg\theta_n) &= g(m_{in}(\neg\theta_n), R_{in}) = R_{in}(1 - C_{in}) \\ m_{in}(\Theta) &= 1 - R_{in} \end{aligned} \tag{14}$$

The basic probability assignments constructed in each case can be combined in the Dempster rule with close attention being paid to the most common case where data originated from statistical processes. The values of reliability factors are assumed to be estimated from any available information; however no estimation method is presented. The developed model has been utilized by other authors (see, e.g. [15, 42, 43]) who successfully applied it to remote sensing, target classification and tracking, and creating a multi-model decision support system.

In [13] a model for incorporating reliability of agents into decision fusion is introduced but beliefs in this model were assumed to be not simple support functions with foci θ_n and $\neg\theta_n$ but separable support functions with all singletons $\theta_n, \forall n$ as foci. Two different rules for building a function $g(x, R)$ for each agent i are designed.

Rule 1 is based on an assumption that when $R_i(A)$, a reliability factor for belief of agent i in hypothesis A, is 0, the value of $m_i(A)$ contributes to our ignorance. In this case m_i^{new} is a basic probability assignment with focal elements $\theta_n, n = 1, 2, \ldots, N$ and Θ and for each sensor i, m_i is transformed into m_i^{new} as follows:

$$\begin{aligned} m_i^{new}(\theta_n) =& R_i(\theta_n) m_i(\theta_n) && \forall n \\ m_i^{new}(\Theta) =& 1 - \sum_n R_i(\theta_n) m_i(\theta_n) \end{aligned} \tag{15}$$

This rule is similar to discounting rule from equation (12) but with reliability coefficients different not only for different agents but also for different hypotheses. This rule was also used in [44].

Rule 2 is based on the assumption that when for $R_i(A) = 0$ the value $m_i(A)$ contributes to our belief in hypothesis $\neg A$. In this case m_i^{new} is a basic probability assignment with focal elements θ_n, $\neg\theta_n$, $n = 1, 2, \ldots, N$ and Θ and for each sensor i, m_i is transformed into m_i^{new} as follows:

$$\begin{aligned} m_i^{new}(\theta_n) &= R_i(\theta_n) m_i(\theta_n) && \forall n \\ m_i^{new}(\neg\theta_n) &= (1 - R_i(\theta_n)) m_i(\theta_n) && \forall n \\ m_i^{new}(\Theta) &= m_i(\Theta) \end{aligned} \tag{16}$$

The feasibility of the model described in [13] was demonstrated in a case study of recognition of natural scene images.

A modification of the Dempster's rule incorporating fuzzy operators applied to the reliability coefficients $\overline{R} = \{R_1, R_2, \ldots, R_I\}$ was suggested first in [1] and later in [45]:

$$\begin{aligned} &m(\Theta) = 1 && \text{if } R_i = 0\ \forall i \\ &m(A) = \frac{1}{F}\left[And(\overline{R}) \sum_{\cup A_j = A} \prod_i m_i(A_j) + Xor(\overline{R}) \sum_{\cap A_j = A} \prod_i m_i(A_j) \right] && \text{otherwise} \end{aligned} \tag{17}$$

where Xor and And are fuzzy operators which can be defined as

$$And(\overline{R}) = \prod_i R_i \tag{18}$$

$$Xor(\overline{R}) = \left[1 - \prod_i (1 - R_i)\right] \left(1 - \prod_i R_i\right) \tag{19}$$

and $F = Xor(\overline{R}) + C\ And(\overline{R})$ is a normalization factor with C being the measure of conflict between the sources as defined by equation (10).

In a more recent work [45,46], the authors propose to see the conflict between sources as an indicator about the reliability of the sources: (1) if the conflict is low, the sources should be considered reliable and the fusion result should be close to the result provided by a conjunctive operator and (2) if the conflict is high, at least one of the sources should be considered as unreliable and the fusion result should be close to the result provided by a disjunctive rule. From this point of view they developed a combination rule as a weighted sum of the conjunctive and disjunctive rules.

$$m(A) = \alpha(C) \sum_{\cup A_j = A} \prod_i m_i(A_j) + \beta(C) \sum_{\cap A_j = A} \prod_i m_i(A_j) \tag{20}$$

Equation (20) is close to equation (17), with the difference that the weighting coefficients α and β are not functions of the reliability coefficients R_i (as in the case of the Xor and And operators) but directly of the conflict C. This fusion rule has the advantage to address the problem of reliability of the sources even when the reliability coefficients R_i are unknown or are not well quantifiable.

The strategy, in which sources of poor reliability are identified and excluded from the fusion process, is implemented, (e.g. in [47]). In this method designed for target tracking, a reference sensor is selected and the sensor, which gives the best fusion result in combination with the reference one is used in fusion. The fusion is performed with the Dempster rule of combination.

In [42,43] two methods of integrating "degree of trust" into fusion are proposed. The first method, local contextual method (LCCM) is based on utilization of reliability coefficients before fusion in the way proposed in [11] and the second one, global contextual processing method (GCPM), inspired by the method developed in [36,37]. Both methods utilize the introduced contextual combination rule (CC) rule, which allows for combination of the contextual information with the prior information. The contextual information is subjective and modeled by the theory of fuzzy events used in association with the probability theory. The LCCM corresponds to the case when the CC rule is applied before the fusion operation. Contextual information is taken into account in the form of basic probability assignments evaluated for each possible contextual variable, which plays the role of reliability. In GCPM method, a basic probability assignment integrates the validity of the different combinations of the sources and the CC rule is applied after fusion. The performance of LCCM and GCPM has been compared in a case study. It has been concluded that the LCCM method is most efficient when the validity degrees of sources are close while the GCPM method performs better when the validity degrees of sources are different. When the most efficient sensor is reliable, the results obtained by both methods and the probabilistic approach are similar. The developed general methodology has been successfully applied by the authors to the problem of pixel fusion in order to increase scene perception.

5 Fusion by possibility and fuzzy operators

In the framework of possibility theory (see, e.g., [9, 17, 18]) information obtained from sensor i is represented by possibility distribution π. The notion of possibility distribution is equivalent to the notion of the basic probability assignment in belief theory with a different constraint:

$$\pi : \Theta \rightarrow [0,1] : \max_{\theta_n \in \Theta} \pi(\theta_n) = 1 \tag{21}$$

A possibility distribution can also be viewed as the membership function of the fuzzy set of possible values of variable x. As with belief functions, two other functions can be derived from possibility distribution:

- Necessity N (counterpart of a Belief function):

$$N : \Theta \rightarrow [0,1] : N(A) = \min_{\theta_n \in A} \pi(\theta_n) \tag{22}$$

- Possibility Π (counterpart of a Plausibility function):

$$\Pi : \Theta \rightarrow [0,1] : \Pi(A) = \max_{\theta_n \in A} \pi(\theta_n) \tag{23}$$

Possibility and necessity measures can be thought of in terms of a consonant belief function or as bounds of a family of probabilities [17, 19].

Possibilistic approach offers several fusion rules. Most of the rules are based on t-norm and or t-conorm, the fuzzy translation of the intersection and union (Table 1).

t-norm ($\cap$)	Dual t-norm ($\cup$)	Name
$\min(x,y)$	$\max(x,y)$	Zadeh
xy	$x+y-xy$	Probabilistic
x, if $y=1$ y, if $x=1$ 0, otherwise	x, if $y=0$ y, if $x=0$ 1, otherwise	Discontinuous
$\max(x+y-1,0)$	$\min(x+y,1)$	Lukasiewicz

Table 1: *Few t-norms and t-conorms [1].*

Disjunctive rule is applied when it is known for sure that at least one source of data is absolutely reliable but it is not known which one:

$$\pi^*_{\cup}(\theta_n) = \bigcup_i \pi_i(\theta_n) = \max_i \pi_i(\theta_n) \qquad \forall \theta_n \in \Theta \tag{24}$$

When combining absolutely reliable sources, a conjunctive rule is applied:

$$\pi^*_{\cap}(\theta_n) = \bigcap_i \pi_i(\theta_n) = \min_i \pi_i(\theta_n) \qquad \forall \theta_n \in \Theta \tag{25}$$

The quality $h = \max_{\theta_n} \pi^*_{\cap}(\theta_n) \in [0, 1]$ corresponds to the level of contradictory between sources:

$$1 - h = \begin{cases} 1, \text{if sources have no intersection} \\ 0, \text{if sources agree} \end{cases}$$

If $h \neq 1$ a normalization step is required.

$$\pi_{\cap}(\theta_n) = \min_i \pi_i(\theta_n)/h \qquad \forall \theta_n \in \Theta \tag{26}$$

A different fusion method was proposed in [48]:

$$\pi_Y(\theta_n) = \min_i \pi_i(\theta_n) + 1 - h \qquad \forall \theta_n \in \Theta \tag{27}$$

As in the Dempster-Shafer's evidence theory, the level of contradiction is linked to the reliability of sources. The fusion methods given by equations (24) - (27) are either the conjunctive or the disjunctive and sources are either reliable or not reliable. Several classes of rule with intermediate behavior have been proposed in the literature, e.g., the rule, in which it is assumed that J sources among I are reliable but the reliable sources are unknown, symmetrical sum, median operators, or rules based on fuzzy integrals (see e.g. [10, 17, 49]) but they still attached the same weight to levels of reliability of the sources.

A number of fusion methods defined in the framework of possibility theory, which take information reliability into account are reviewed in [10, 17, 18, 50].

One of these approaches is a trade off rule, which is similar to the trade-off rule used in evidence theories and probabilistic methods based on consensus theory:

Let $R_i > 0$ be a level of reliability of source i. Then

$$\pi_P(\theta_n) = \sum_{i=1}^{I} R_i \pi_i(\theta_n) \qquad \forall \theta_n \in \Theta \tag{28}$$

If we consider relative reliability we have the following constraint:

$$\sum_{i=1}^{I} R_i = 1$$

In [51] a similar operator is used with logical ones:

$$\pi_{\cup P}(\theta_n) = \sigma \sum_{i=1}^{I} R_i \pi_i(\theta_n) + (1-\sigma)\pi^*_{\cup}(\theta_n), \qquad \forall \theta_n \in \Theta \tag{29}$$

where $\sigma \in [0,1]$.

The reliability of sources can be taken into account by discounting, which modifies the possibility distribution according to their level of reliability before fusion [52]. With this approach, R_i is the degree of certainty that the source is reliable and the possibility distribution can be changed in the following manner:

$$\pi'_i(\theta_n) = \max(\pi_i(\theta_n), 1 - R_i) \qquad \forall \theta_n \in \Theta \tag{30}$$

As it follows from (30), if $R_i = 1$ (source i is fully reliable) then $\pi'_i = \pi$, when $R_i = 0$ (source is absolutely unreliable) then $\pi'_i(\theta_n) = 1, \forall \theta_n \in \Theta$, which means total ignorance.

A discounting method similar to one used in belief functions [53] is

$$\pi'_i(\theta_n) = R_i \star \pi_i(\theta_n) + 1 - R_i \qquad \forall \theta_n \in \Theta \tag{31}$$

where operator $\star$ is t-norm (conjunction). For example, if $\star = \min$, then (31) corresponds to $\pi'_i(\theta_n) = \min(R_i, \pi(\theta_n)) + 1 - R_i, \forall \theta_n \in \Theta$, and is also equivalent to $\pi'_i(\theta_n) = \min(1, 1 - R_i + \pi(\theta_n)), \forall \theta_n \in \Theta$ and if $\star$ =product, then (31) corresponds to $\pi'_i(\theta_n) = R_i \pi_i(\theta_n) + 1 - R_i, \forall \theta_n \in \Theta$.

When sources are discounted according to their reliability levels, they are considered absolutely reliable and then fused by a conjunctive rule. The drawback of these rules is that when $R_i = 0$ then the support of discounted distribution becomes the whole universe. Moreover, the values of reliability levels are supposed to be given.

Two rules are introduced in [1,50], in which "absolute" reliability levels are considered (constraint $\sum_{i=1}^{I} R_i = 1$ is not required):

$$\pi_0(\theta_n) = Xor(\overline{R}) \max_i (R_i \pi_i(\theta_n)) + And(\overline{R}) \min_i (R_i \pi_i(\theta_n)) \qquad \forall \theta_n \in \Theta \tag{32}$$

where Xor and And are equivalents of the logical operators and can be defined as in equations (18) - (19).

The rule is not normalized and, in the case in which the assumption of a closed world is made, this rule requires normalization, i.e. we have to divide (32) by the maximum of the distribution π_0. This rule has the following drawback : when all the sources are absolutely unreliable, it provides a null possibility distribution instead of the complete ignorance ($\pi_0(\theta_n) = 1, \forall \theta_n \in \Theta$). To overcome this problem, the authors [50] modified equation (32) as follows :

$$\pi_0(\theta_n) = Xor(\overline{R}) \max_i (R_i \pi_i(\theta_n)) + And(\overline{R}) \min_i (R_i \pi_i(\theta_n)) + Nor(\overline{R}) \qquad \forall \theta_n \in \Theta \quad (33)$$

where

$$Nor(\overline{R}) = \prod_i (1 - R_i) \quad (34)$$

In a more recent study [54] the authors proposed a new combination rule to improve the one from equation (33). Thus, this new rule is defined $\forall \theta_n \in \Theta$ as :

$$\begin{aligned} \pi_0(\theta_n) &= (1 - \underline{R}^q) \max_i (R_i \pi_i(\theta_n)) + \underline{R}^p \min_i (1 - R_i + R_i \pi_i(\theta_n)) + \underline{R}^p \min_i (1 - R_i) && \text{if } \underline{R} \neq 0 \\ \pi_0(\theta_n) &= \max_i \pi_i(\theta_n) && \text{if } \underline{R} = 0 \end{aligned} \quad (35)$$

where $\underline{R}$ is the average of the reliability vector $\overline{R}$ of all sources and $p > 1$ and $p > q > 0$. The authors proposed to set $p = 2$ and to vary q.

Another idea presented in [54] is that the source' reliability can be time-varying and in this situation the fusion method should adapt to this situation. The authors proposed to study the conflicts between sources dynamically in order to continuously adapt the reliabilities of the sources and to find the defective sources. Thus, two notions of reliability are introduced :

- a scalar reliability $R_{i,S}$ representing the past reliability and
- a dynamic reliability $R_{i,D}$ obtained from the actual conflict between the sources.

These two reliabilities are combined to give the final reliability of a source R_i.

The methods utilizing possibilistic and fuzzy approaches described above form a foundation for building fusion processes applied in many fields. Among them are medical image processing [10], remote sensing [55], face recognition [56], recognition of crises [57], fusion of databases [58].

6 Reliability of fusion results

A very important question related to designing a fusion system is a question of reliability of the fusion results. Although all the publications report improved fusion systems performance as the result of incorporating reliability into fusion processes, the reliability of the results obtained with such fusion processes has seldom been investigated. It is usually assumed that incorporation of source reliability makes fusion result perfectly reliable and the number of publications addressing this problem is very limited.

The authors of [9, 50] are studying reliability of the possibility based fusion operator (see 3.4.2) and propose to measure reliability of the fusion result as a combination of its "direct reliability" and an index of its quality. "Direct reliability" is modeled as a trade off between a reinforcement function of reliability of the sources and a function of the number of sources. Reinforcement refers to the situation, in which the fused result is more reliable than reliability of each participating source. A parameter defining a trade off between these two factors represents users' attitude towards them. The quality index measures internal contradiction of the result and depends on the shape of the result, namely on the extend to which a resulting possibility distribution contains several different modes and the extend to which they overlap.

The authors of [59] assume that reliability of the fusion result obtained by propagating sensor measurements through the layers of a recurrent neural network can be assessed by propagating sensor reliability in the same way. Thus, the reliability of the result is computed as a function of reliability of the sensors and the architecture of the neural network used.

In [60] reliability assessment of fusion results is conducted by treating the aggregated result as a "virtual expert", which allows for comparing different pooling methods considered in the document.

7 Reliability coefficients

One of the major problems of incorporating reliability into fusion is the problem of obtaining reliability coefficients. The value of reliability coefficients may be provided by external sources, modeled by utilizing contextual information (see, e.g. [25, 36, 37, 43]), learned by using training data, as e.g., in a neural network (see e.g., [13, 44, 61]), or constructed as a function of agreement between different sources or sources and fusion results [9, 62]. A discussion on these methods in the literature is presented below.

7.1 Methods utilizing domain knowledge and contextual information

In [36, 37] the context is modeled by a set of parameters influencing reliability of each sensor and expert knowledge is used to represent validity of the sensor domain as a fuzzy membership function of the context. Then reliability coefficients are modeled as the probability of fuzzy events.

Two other methods of defining sensor reliability based on contextual information are introduced in [42, 43]. Both methods utilize subjective contextual information modeled by the theory of fuzzy events and used in connection with probability theory. In one (LCCM method), reliability of each source is defined by associating of each sensor and each hypothesis to the context and computed as a Bayesian probability mass on the frame of discernment defined by validity domain of each source. Then obtained reliability values are used in a discount rule introduced in [11]. In another method, GCPM, inspired by the method suggested in [36, 37], reliability of one or a combination of several sensors is computed as the probability of conjunction of the fuzzy subsets corresponding to each source for each contextual variable. This probability of validity is further used for the construction of reliability of one or several associations of sources and hypothesis.

In [62] expert knowledge is used to represent reliability by a possibility distribution defined on the sensor domain. The computed reliability coefficients are then used in a production system to determine the sources of satisfactory reliability to be used in combination.

The methods utilizing contextual information either in the form of a priori distributions or represented by validity domain for modeling reliability are very successful however this information not always available.

7.2 Obtaining reliability coefficients from training data

A general approach to learning reliability coefficients from training data consists in integrating reliability assessment into the fusion process and explicitly training fusion rules. Reliability coefficients in such methods are computed by minimizing the distance between a vector of beliefs obtained as the result of fusion within the framework of uncertainty theory considered and a target vector from a training set (see, e.g., [13, 44, 61]).

Another method of defining reliability is based on separability of different hypotheses: a source is called reliable if separability is high and not reliable if separability is low [39, 40, 59]. Reliability is measured by average or hypothesis specific distance (e.g., Bhattacharyya distance or Hellinder's distance) between probability distributions estimated from a training set.

The methods of learning reliability factors require a training set and the advantages of incorporating reliability into fusion processes may be diminished by a small number of training patterns available. However, this type of methods can be very useful for establishing relative reliability of legacy classifiers.

7.3 Reliability based on consensus

Another approach to modeling source reliability employs a degree of consensus among various sources or a degree of consensus among sources and fusion results.

One of such methods designed for target tracking [59] adaptively computes a deviation between measurements of each sensor and the fusion result and uses this deviation measure as a degree of reliability of this sensor.

A different consensus based method utilizes the notion of inner trust introduced in [9]. Evaluation of inner trust is performed in two steps. At the first step a pairwise likeliness of the sources is computed and then the inner trust is defined in such a way that a source is considered absolutely reliable if and only if there is no contradiction with other sources while a source in absolute contradiction have a very small reliability.

Consensus based methods have the advantage of not using any supplementary information for modeling reliability. However, this method is not always good since if sources are unreliable they may be conflicting or not. One of the ways to improve the inner trust is to complement it with any external information, which can become available [9].

7.4 Reliability of expert judgments

Reliability weights in the consensus rules for combining experts' subjective probabilities are discussed e.g. in [60, 63]. The author of [63] introduced two measures of performance: calibration and information scores. The information score is defined by the degree to which an expert's distribution is concentrated relative to some background measure (usually of uniform or loguniform distribution over an intrinsic range for each variable). The calibration of expert e, $C(e)$, is the statistical likelihood that an expert's quantile assessment corresponds to a set of experimental results.

This approach inspired development of the methodology introduced in [63], in which expert opinions are modeled by possibility distributions, which represent expert judgments more naturally than probabilities. Reliability of an expert in [60] is measured by membership grades (calibration) and by fuzzy cardinality indexes (level of precision).

8 Case study

The purpose of the case study presented below is to compare two different ways of learning and incorporating reliability coefficients into the fusion process within the framework of the Dempster-Shafer theory of evidence.

We consider here a set of agents observing different features of the environment and producing beliefs into each hypothesis based on the distance to a class representative vector (see Figure 1).

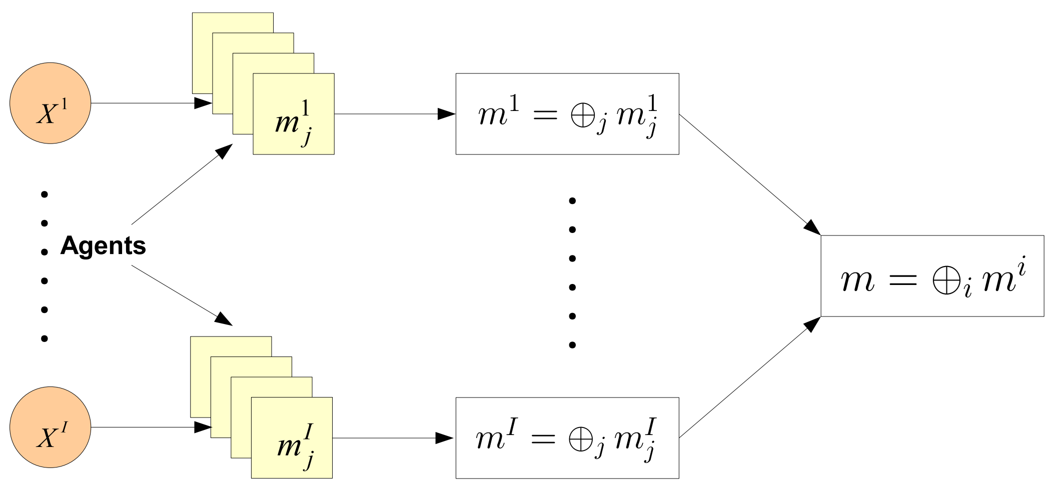

Figure 1: *Architecture of the system considered in the case study*

Let each individual agent $i, (1 \leq i \leq I)$ observe states of environment and extracts a particular type of information represented by a feature vector $\overline{X}^i = (x_1^i, x_2^i, \ldots, x_{M_i}^i)$, where M_i is a feature vector dimension. Let $\Theta = \{\theta_1, \ldots, \theta_N\}$ be a frame of discernment, where θ_k is the hypothesis that a pattern under consideration belongs to class k. For each agent i and each class k, we can define a proximity measure between a feature vector $\overline{X}^i$ and a class representative vector $\overline{W}_k^i$ characterizing a strength of support for hypotheses θ_k :

$$\Phi(\overline{X}^i, \overline{W}_k^i) = \Phi(d(\overline{X}^i, \overline{W}_k^i)) \tag{36}$$

where Φ is a decreasing function of distance $d(\overline{X}^i, \overline{W}_k^i)$ between $\overline{X}^i$ and $\overline{W}_k^i$ such as

$$0 \leq \Phi(d(\overline{X}^i, \overline{W}_k^i)) \leq 1 \quad \text{and} \quad \Phi(d(\overline{X}^i, \overline{W}_k^i)) = 1 \quad \text{if} \quad d(\overline{X}^i, \overline{W}_k^i) = 0 \tag{37}$$

In our case study we consider

$$\Phi(\overline{X}^i, \overline{W}^i_k) = \cos\Big(||\overline{X}^i - \overline{W}^i_k||^2\Big)^2 \tag{38}$$

$\Phi(\overline{X}^i, \overline{W}^i_k)$ yields a simple support function m^i_k :

$$\begin{aligned}
&m^i_k(\theta_k) = \Phi(\overline{X}^i, \overline{W}^i_k) && \forall \theta_k \in \Theta \\
&m^i_k(\Theta) = 1 - \Phi(\overline{X}^i, \overline{W}^i_k) && \\
&m^i_k(A) = 0 && \forall A \subset \Theta, A \neq \theta_k
\end{aligned} \tag{39}$$

The combination of all K simple support functions, corresponding to each hypothesis k provided by each sensor i, with the Dempster rule [21] from equation (10) leads to the basic probability assignment for hypotheses $A \subseteq \Theta$:

$$\begin{aligned}
&m^i(\theta_k) = C^{-1} m^i_k(\theta_k) \prod_{\substack{\ell=1 \\ \ell \neq k}}^{K} m^i_\ell(\Theta) && \forall \theta_k \in \Theta \\
&m^i(\Theta) = C^{-1} \prod_{\ell=1}^{K} m^i_\ell(\Theta) && \\
&m^i(A) = 0 && \forall A \subset \Theta, A \neq \theta_k
\end{aligned} \tag{40}$$

where

$$C = \sum_k m^i_k(\theta_k) \prod_{\substack{\ell=1 \\ \ell \neq k}}^{K} m^i_\ell(\Theta) + \prod_{k=1}^{K} m^i_k(\Theta)$$

and $1 - C$ is the conflict between sources.

These bpas obtained for each sensor i (m_i) are combined by the Dempster rule of combination to produce fused bpas, which are used to make a decision on the hypothesis about the state of the environment observed.

Now we want to incorporate agents' reliability into our system and learn reliability coefficients from a training set. There are two ways to include reliability into fusion of beliefs of these agents. First we can consider reliability of simple support functions of each agent computed for each hypothesis based on the distance $d(\overline{X}^i, \overline{W}^i_k)$ between $\overline{X}^i$ and $\overline{W}^i_k$ which can be transformed into (see equation (12)):

$$\begin{aligned}
&(m^i_k)_{disc}(\theta_k) = R^i_k \Phi(\overline{X}^i, \overline{W}^i_k) && \forall \theta_k \in \Theta \\
&(m^i_k)_{disc}(\Theta) = 1 - R^i_k + R^i_k \Phi(\overline{X}^i, \overline{W}^i_k) && \\
&(m^i_k)_{disc}(A) = 0 && \forall A \subset \Theta, A \neq \theta_k
\end{aligned} \tag{41}$$

The architecture corresponding to this method of incorporation of the reliabilities into the system is shown in Figure 2.

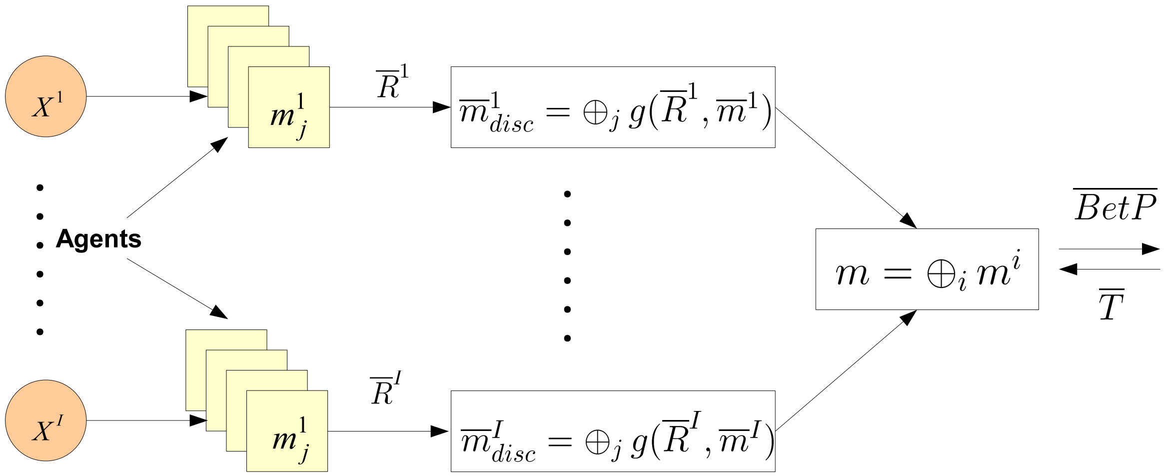

Figure 2: *Incorporating reliability coefficients at the simple support functions level*

Then by combining all the discounted bpas we can obtain systems's bpas

$$m(A) = \bigoplus_{i,k} (m_k^i)_{disc}$$

that are further used for computing $\mathrm{BetP}(\theta_k)$:

$$\mathrm{BetP}(\theta_k) = \sum_{\substack{\theta_k \in A \\ A \subseteq \Theta}} \frac{m(A)}{|A|} \tag{42}$$

The values of reliability coefficients are trained to minimize the following objective function:

$$E(\overline{\mathrm{BetP}}, \overline{T}, \overline{R}) = (\overline{\mathrm{BetP}} - \overline{T})^2 \to \min$$

Minimization is performed by the gradient descent methods while utilizing a training set.

In the second case, reliability coefficients were incorporated at the decision fusion levels (see Figure 3). In this case, $\overline{m}^i$ are computed as in equations (40) but $m(A) = \oplus_i\, g_1(\overline{R}^i, \overline{m}^i) = \oplus_i\, \overline{R}^i \overline{m}^i$ as in equation (15). As in the previous case, the decisions are based on $\overline{\mathrm{BetP}}$ and the values of reliability coefficients are selected to train the following objective function:

$$E(\overline{\mathrm{BetP}}, \overline{T}, \overline{R}) = (\overline{\mathrm{BetP}} - \overline{T})^2 \to \min$$

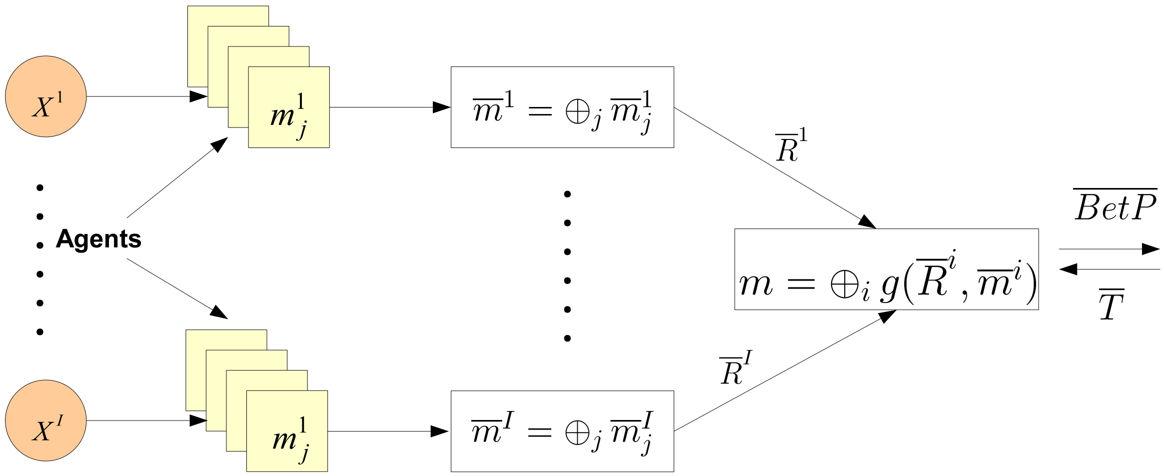

Figure 3: *Incorporating reliability at the decision fusion level*

In order to evaluate the performance of both scheme, a proprietary database containing a set of 2430 images of size 256×256 was used. The images were obtained from the VisTex image database from MIT Media Labs. They contained five corrupted variants of each of 486 different natural scene images from four classes: metal (144 patterns), sand (96 patterns), vegetation (133 patterns), and water (113 patterns).

We have conducted experiments with four agents, which have different expertise and are able to extract different features from the images: one texture agent and three color agents. Testing has been performed using cross-validation in which the group reserved for testing was fixed at 15% of the total available dataset. The result was averaged over 10 runs. Results of the experiments are presented in Table 2. As we can see from Table 2 the accuracy rate of the fused classifiers can be significantly improved by introducing reliability coefficients. We can also see that incorporation of reliability into decision fusion is not that beneficial as into fusion of simple support functions. At the same time, in certain cases, for example, when we deal with legacy classifiers and can fuse their decisions only utilization of reliability coefficients at the decision fusion level can considerably decrease misclassification error.

	Accuracy	Reduction of misclassification
Initial fusion result ($R_k^i = 1$)	81.4%	
Case 1: reliability of beliefs of each agent in each class separately	90.3%	52%
Case 2: Decision fusion (classwise agents' reliability)	85.3%	20.9%

Table 2: *Improvement of accuracy with trained reliability coefficients*

A series of additional experiments were conducted with utilization of reliability coefficients at the decision fusion level. Reliability coefficients representing trust into belief of each agent into each hypothesis obtained as the result of training were averaged over all 10 runs with a very small variance. We define then reliability of each agent as a weighted average of reliability of belief of each agent into each hypothesis:

$$R^i = \sum_k \frac{n_k}{N} R^i_k \tag{43}$$

where n_k is the number of patterns of class k in the training set, and N is the total number of patterns in the training set. Agents' performance and classwise agents' reliability of coefficients are presented in Table 3. The results presented in Table 3 show that reliability of agents are related to their performance. At the same time while performance of red and green agents are very close (53% and 50%, respectively) their reliabilities differ significantly (0.42 and 0.24, respectively).

	TEXTURE AGENT	RED AGENT	GREEN AGENT	BLUE AGENT
Accuracy rate	59.1%	53.0%	50.0%	63.8%
Agent reliability				
Class 1	0.92	0.08	0.05	0.77
Class 2	0.10	0.75	0.13	0.25
Class 3	0.72	0.78	0.70	0.07
Class 4	0.32	0.16	0.08	0.97
Averaged agents' reliability	0.54	0.42	0.24	0.53

Table 3: *Agents' individual accuracy and averaged reliability coefficients*

In another series of experiments we compared the accuracy of the fused results of the three most reliable agents with and without taking into account reliability coefficients.

	Recognition rate without reliability coefficients	Recognition rate with reliability coefficients
4 agents	81.4%	85.3%
3 agents	77%	84%

Table 4: *Recognition rate and accuracy of the fused results with and without reliabilities.*

The results are presented in Table 4. As we can see from Table 4, the incorporation of the recognition rate of the fusion system comprising three agents is notably lower than the recognition rate of the system comprising 4 agents. The incorporation of reliability into the fusion process can significantly improve accuracy of the system of the three most reliable agents, making it very close to the accuracy of the system comprising four classifiers.

9 Conclusion

The subject of information and data reliability has been receiving significant attention in the last couple of years in areas such as communication, personal computing, and databases. At the same time, there is a relatively small body of literature in the data fusion community that addresses this topic; moreover, there is no clear understanding of what defines information reliability from the perspective of information fusion processes designers and how reliability of sources affect fusion results. The main body of the literature on information fusion concerns with building an adequate uncertainty model without paying much attention to the related problems of reliability beliefs produced by this model. The majority of fusion operators is based on optimistic assumptions about reliability of the models producing beliefs and assumes that they are equally reliable and play a symmetrical role. However, different models may have different reliability and it is necessary to account for this fact in order to avoid decreasing in performance of fusion results.

This report provides a comprehensive survey of the main approaches used in the fusion literature as well as simulation results.

Generally, all the approaches to incorporation of reliability into fusion processes fall into three categories:

- Methods measuring reliability of data input to fusion processes and eliminating data of poor reliability.
- Methods modifying the data and information by considering their reliability before fusion.
- Methods modifying the fusion process to account for the reliability of the input.

Selection of one of these methods and a type of uncertainty theory to be utilized depend strictly on the problem to be addressed, on the quality and existence of prior knowledge modeled, and the degree to which any given data is representative of that prior knowledge.

The publications reviewed and our own simulation results show that incorporation of reliability into fusion processes gives “richer behavior” to the fusion system and can improve performance of fusion systems while producing many theoretical and practical problems not very often addressed in the data fusion literature. Among these are the problem of estimation of reliability of sources and their temporal analysis; the problem of interrelationship between reliability of information sources and the reliability of fusion results; the problem of incorporating reliability into fusion of heterogeneous information. Future research addressing these issues will definitely improve performance and effectiveness of fusion systems.

This page intentionally left blank.

References

[1] Delmotte, F., Dubois, L., Desodt, A.-M., and Borne, P. (1995), Using trust in uncertainty theories, *Information and Systems Engineering*, 1, 303–314.

[2] White, Franklin E. (1987), Data Fusion Lexicon, In *Technical Panel for C3, Data Fusion Sub-Panel, Naval Ocean Systems.*

[3] Waltz, E. and Llinas, J. (1990), Multisensor Data Fusion, Artech House Inc. New-York, NY.

[4] Abidi, M. A. and Gonzalez, R. C., (Eds.) (1992), Data Fusion in Robotics and Machine intelligence, Academic Press.

[5] Hall, David and Llinas, James, (Eds.) (2001), Handbook of Multisensor Data Fusion, CRC Press.

[6] Ören, T. I. (2001), Impact of data on simulation: From early practices to federated and agent-directed simulation, In *Proc. of the EUROSIM.*

[7] Denning, P. J. (1997), How we will learn, In Denning, P.J. and Metcalfe, R.M., (Eds.), *Beyond Calculation - The Next Fifty Years of Computation.*, pp. 267–286, Copernicus (Springer Verlag). New York.

[8] Leemis, L. (2000), Input Modeling, In et al., J.A. Jones, (Ed.), *Proc. of the 2000 Winter Simulation Conference*, pp. 17–25, Orlando, FL.

[9] Delmotte, F. and Borne, P. (1998), Context-dependent trust in data fusion within the possibility theory, In *Proc of the IEEE Intl Conf on Systems, Man, and Cybernetics*, Vol. 28, pp. 78–88.

[10] Bloch, Isabelle (1996), Information combination operators for data fusion: a comparativereview with classification, *IEEE Transactions on Systems, Man and Cybernetics, Part A*, 26(1), 52–67.

[11] Appriou, A. (2001), Situation assessment based on spatially ambiguous multisensor measurements, *International Journal of Intelligent Systems*, 16(10), 1135–1166.

[12] Denoeux, T. (1995), A k-Nearest Neighbor Classification Rule Based on Dempster-Shafer Theory, *IEEE Transactions on Systems, Man and Cybernetics*, 25(5), 804–813.

[13] Rogova, Galina (2003), Adaptive decision fusion by reinforcement learning neural network, in Distributed Reinforcement Learning For Decision-making In Cooperative Multi-agent Systems, Part 1., Technical Report CUBRC Technical report prepared for AFRL, Buffalo, NY.

[14] Grandin, J-F, Detexis, T., and Marques, M. (2000), Robust Data Fusion, In *Proc. of the Third Int.Conf. On Information Fusion*, pp. MoC3–3– MoC3–9.

[15] Delmotte, F. and Dubois, L. (1996), Multi-models and belief functions for decision support systems, In *IEEE International Conference on Systems, Man, and Cybernetics*, Vol. 1, pp. 181–186.

[16] Wang, Pei (2001), Confidence as higher level of uncertainty, In *Proc. of the Int. Symp. on Imprecise Probabilities and Their Applications*, pp. 352–361.

[17] Dubois, D. and Prade, H. (1992), Data Fusion in Robotics and Machine Intelligence, Ch. Combination of information in the frame of possibility theory.

[18] Dubois, D. and Prade, H. (2000), Possibility Theory in Information Fusion, In *Fusion 2000*, p. Plenary Speech, Paris.

[19] Shafer, G. (1976), A Mathematical Theory of Evidence, Princeton University Press.

[20] Hutchinson, S.A and Kak, A. (1992), Multisensor Strategies Using Belief Accumulation, In Abidi, M. A. and Gonzalez, R. C., (Eds.), *Data Fusion in Robotics and Machine Intelligence*.

[21] Smets, P. and Kennes, R. (1994), The Transferable Belief Model, *Artificial Intelligence*, 66, 191–234.

[22] Klir, G. J. and Folger, T. A. (1988), Fuzzy Sets, Uncertainty and Information, Englewood Cliffs, New Jersey: Prentice Hall.

[23] Dubois, D. and Prade, H. (1980), Fuzzy Sets Systems: Theory and Application, New York, Academic.

[24] Smets, Ph. (2000), Data fusion in Transferable Belief Model, Possibility theory in Information Fusion, In *Proc. of the 3rd. Int. Conf. On Information Fusion*, pp. PS–6 – PS–19.

[25] Benediktsson, J. A., Sveinsson, J. R., and Swain, P. H. (1997), Hybrid consensus theoretic classification, *IEEE Trans. Geosci. Remote Sensing*, 35(4), 833–843.

[26] Benediktsson, J. A. and Swain, P. H. (1992), Consensus theoretic classification methods, *IEEE Trans. Syst., Man, Cybern.*, 22(4), 688–704.

[27] Florian, R. and Yarowsky, D. (2002), Modeling Consensus: Classifier Combination for Word Sense Disambiguation, In *Proceedings of EMNLP'02"*, pp. 25–32.

[28] Lee, T., Richards, J., and Swain, P. (1987), Probabilistic and evidential approaches for multisource data analysis, *IEEE Trans. Geosci. Remote Sensing*, 25(3), 283–293.

[29] Benediktsson, J. A and Kanellopoulos, I. (1999), Classification of Multisource and Hyperspectral Data Based on Decision Fusion, *IEEE Trans. Geosci. Remote Sensing*, 36(3), 283–293.

[30] Roli, F., Fumera, G., and Kittler, J. (2002), Fixed and Trained Combiners for Fusion of Imbalanced Pattern Classifiers, In *Proc. of the FUSION'2002-Fifth Conference on Multisource- Multisensor Information Fusion*.

[31] Kittler, J. and Roli, F., (Eds.) (2000), Multiple Classifiers Systems, Vol. 1857 of *Lecture Notes in Computer Science*, Springer-Verlag, Berlin.

[32] Tumer, K. and Ghosh, J. (1996), Analysis Of Decision Boundaries In Linearly Combined Neural Classifiers, *Pattern Recognition*, 29(2), 341–348.

[33] Fumera, G. and Roli, F. (2002), Performance Analysis and Comparison of Linear Combiners for Classifier Fusion, In *IAOR Int Workshop on Statistical Pattern Recognition*, Springer-Verlag.

[34] Roli, F. and Fumera, G. (2002), Analysis Of Linear And Order Statistics Combined For Fusion Of Imbalanced Classifiers, In *3rd Workshop on Multiple Classifier Systems*.

[35] Kuncheva, L. I. (1997), An application of OWA operators to the aggregation of multiple classification decisions, In *The Ordered Weighted Averaging operators. Theory and Applciations*.

[36] Nimier, V. (1998), Supervised Multisensor Tracking Algorithm By Context Analysis, In *Fusion'98*, Las Vegas.

[37] Nimier, V. (1994), Introducing Contextual Information in Multisensor Tracking Algorithms, In *5th International Conference on Processing and Management of Uncertainty in Knowledge-Based Systems*.

[38] Dezert, J. (2001), Optimal Bayesian Fusion of Multiple Unreliable Classifiers, In *Fusion 2001*, Number WeB2-11-WeB2-11.

[39] Lefevre, E., Colot, O., and Vannoorenberghe, P. (1999), A Classification Method Based On The Dempster-Shafer's Theory And Information Criteria, In *FUSION'99 - Second International Conference on Information Fusion*, Silicon Valley-San Francisco Bay Area, California.

[40] Lefevre, E., Colot, O., and Vannoorenberghe, P. (1999), Using Information Criteria In Dempster-Shafer's Belief Assignment, In *FUZZ-IEEE'99 - 8th International Conference on Fuzzy Systems*.

[41] Appriou, A. (1998), Uncertain Data Aggregation in Classification and Tracking Processes, In *Aggregation and Fusion of Imperfect Information*, pp. 231–260, Physica-Verlag.

[42] Fabre, S., Appirou, A., and Briottet, X. (2001), Presentation And Description Of Two Classification Methods Using Data Fusion Based On Sensor Management, *Information Fusion*, 2, 47–71.

[43] Fabre, S., Appirou, A., and Briottet, X. (1998), Introduction Of Contextual Information In Pixel Fusion, In *Fusion'98*, pp. 601–609.

[44] Elouedi, Zied, Mellouli, Khaled, and Smets, Philippe (2004), Assessing sensor reliability for multisensor data fusion within the transferable belief model, *IEEE Trans. on SMC-Part B*, 34(1), 782–787.

[45] Florea, M. C. (2007), Combinaison d'informations hétérogènes dans le cadre unificateur des ensembles aléatoires : approximations et robustesse, Ph.D. thesis, Université Laval, Québec, Canada.

[46] Florea, M. C., Jousselme, A.-L., and Grenier, D. (2005), Dempster-Shafer evidence theory through the years : limitations, practical examples, variants under conflict and a new adaptive combination rule, In *Multisensor Data and Information Processing for Rapid and Robust Situation and Threat Assessment*, NATO ASI, Albena, Bulgaria, 16-27 May.

[47] Chen, V. and Alikhani, M. (1996), Evidential Approach To Select Image Sensors For Data Fusion, In *Symposium on Robotics and Cybernetics CESA '96*, pp. 529–537, Lille, France.

[48] Yager, R. R. (1985), Aggregating evidence using quantified statements, *Information Sciences*, 36, 179–206.

[49] Zadeh, L.A. (1978), Fuzzy Sets as a Basis for a Theory of Possibility, *Fuzzy Sets and Systems*, 1, 3–28.

[50] Delmotte, F. and Borne, P. (1998), Modeling Reliability with Possibility Theory, *IEEE Trans. on Systems, Man, and Cybernetics-Part A: Systems and Humans*, 28(1), 78–88.

[51] Yager, R., Filev, D. P., and Sadeghi, T. (1994), Analysis of Flexible Structured Fuzzy Logic controller, *IEEE Trans. on Systems, Man, and Cybernetics*, 24, 1035–1043.

[52] Dubois, D. and Prade, H. (1988), Possibility Theory - An approach to the computerized Processing of Uncertainty, Plenum press, NY.

[53] Yager, R. (1984), Approximate reasoning as a basis for rule-based expert systems, *IEEE Trans. on Systems, Man and Cybernetics*, 14, 636–643.

[54] Delmotte, F. (2007), Detection of defective sources in the setting of possibility theory, *Fuzzy Sets and Systems*, 158, 555–571.

[55] Jaeger, G. and Benz, U. (1999), Supervised fuzzy classification of SAR data using multiple sources, In *Proc. of IEEE 1999 International Geoscience and Remote Sensing Symposium*, Vol. 3, pp. 1603–1605.

[56] Mirhosseini, A. Reza, Yan, A., Lam, Kin-Man, and Pham, T. (1998), Human Face Image Recognition: An Evidence Aggregation Approach, *Computer Vision And Image Understanding*, 71(2), 213–230.

[57] Larsen, H. and Yager, R. (2000), A Framework for Fuzzy Recognition Technology, *IEEE Transactions On Systems, Man, And Cybernetics-Part C: Applications And Reviews*, 30(1), 65–76.

[58] Benferhat, S., Dubois, D., and Prade, H. (1997), Syntactic combination of uncertain information: A possibilistic approach, In *Qualitative and Quantitative Practical Reasoning (Proceedings of the 1st Inter. Joint Conf. ECSQARU/FAPR'97)*, Vol. 1244 of *Lecture Notes in Artificial Intelligence*, pp. 30–42.

[59] Kobayashi, F., Avai, F., and Fucuda, T. (1999), Sensor selection by reliability based on possibility measure, In *Proc. of the Int. Conf. on Robotics and Automation*.

[60] Sandri, S., Dubois, D., and Kalfsbeek, H. (1995), Elicitation, assessment and pooling of expert judgments using possibility theory, *IEEE Trans. on Fuzzy Systems*, 3(3), 322–335.

[61] Rogova, G. and Kasturi, J. (2001), Reinforcement Learning Neural Network For Distributed Decision Making, In *Proc. of the Forth Conf. on Information Fusion*, pp. TuA23 – TuA28, Montreal, Canada.

[62] Parra-Loera, Ramon, Thompson, Wiley, and Salvi, Ajit (1991), Adaptive selection of sensors based on individual performance in a multisensor environment, In *Data structures and target classification; Proceedings of SPIE*, Vol. 1470, pp. 30–36.

[63] Cooke, R. M. (1991), Experts in Uncertainty: Opinion and Subjective Probability in Science, New York: Oxford University Press,.

This page intentionally left blank.

Distribution list

DRDC Valcartier TR 2005-270

Internal distribution

1 Director General

3 Document Library

1 Head/C2 Decision Support Systems(author)

1 Dr A. Guitouni

1 M. Bélanger

1 P. Maupin

1 Dr P. Valin

1 Head/Intelligence and Information

1 R/Visualization and Geo-spatial Systems

1 R/Knowledge Management

1 R/Intelligence Production and Exploitation

1 R/Environments and Systems

1 Head/System of Systems

1 R/Characterization and Exploration

1 R/Engineering & Architecture

1 R/Computing Solutions and Experimentation

1 Maj. A. Lamontagne

1 Ltv. L. St-Pierre

1 Dr M. Allouche

1 Dr A. Benaskeur

1 J. Berger

1 M. Blanchette

1 Dr A. Boukhtouta

1 Dr R. Breton

1 C. Daigle

1 E. Dorion

1 Dr H. Irandoust

1 Dr A.-L. Jousselme

1 N. Pageau

1 S. Paradis

1 L. Pigeon

1 F. Rhéaume

1 A. Sahi

1 J. Roy

Total internal copies: 36

External distribution

1 Director Research and Development Knowledge and Information Management (PDF file)

1 Director Science & Technology Maritime(DSTM)
Constitution Building, 305 Rideau St., Ottawa, ON, K1N 9E5

1 Director Science & Technology Land (DSTL)
Constitution Building, 305 Rideau St., Ottawa, ON, K1N 9E5

1 Director Science & Technology Air (DSTA)
Constitution Building, 305 Rideau St., Ottawa, ON, K1N 9E5

1 Director Science & Technology C4ISR (DSTC4ISR)
Constitution Building, 305 Rideau St., Ottawa, ON, K1N 9E5

1 Director Maritime Requirements Sea (DMRS) 4
Louis St.Laurent Bldg, 555 Boul. de la Carrière, Gatineau, QC, J8Y 6R5

1 Director Maritime Requirements Sea (DMRS) 6
Louis St.Laurent Bldg, 555 Boul. de la Carrière, Gatineau, QC, J8Y 6R5

1 Director Aerospace Requirements (DAR) 4
101 Colonel By Drive, Ottawa, ON, K1A 0K2

1 Director Aerospace Requirements (DAR) 4-2
101 Colonel By Drive, Ottawa, ON, K1A 0K2

1 Director Maritime Ship Support (DMSS) 6
Louis St.Laurent Bldg, 555 Boul. de la Carrière, Gatineau, QC, J8Y 6R5

1 Director Maritime Ship Support (DMSS) 8
Louis St.Laurent Bldg, 555 Boul. de la Carrière, Gatineau, QC, J8Y 6R5

2 DRDC - Atlantic:
9 Grove Street, Dartmouth, NS, B2Y 3Z7
Attn: Dr. Bruce MacArthur
Attn: Dr. Jim S. Kennedy

2 DRDC - Ottawa:
3701, Carling Avenue, Ottawa, ON, K1A 0Z4
Attn: Barbara Ford
Attn: Dan Brookes

2 CF Maritime Warfare School CFB Halifax
PO Box 99000
Stn Forces
Halifax, Nova Scotia, B3K 5X5
Attn: TAC AAW - OIC Modeling and Simulation

2 Canadian Forces Naval Operations School CFB Halifax
PO Box 99000
Stn Forces
Halifax, Nova Scotia, B3K 5X5
Attn: Tactic - CT AWW

1 Canadian Forces Naval Engineering School CFB Halifax
PO Box 99000
Stn Forces
Halifax, Nova Scotia, B3K 5X5
Attn: CSTC

1 Operational Requirements Analysis Cell CFB Halifax
Stn Forces
Halifax, Nova Scotia, B3K 5X5
Attn: Commanding Officer

1 Canadian Forces Fleet School CFB Esquimalt
P.O. Box 17000
STN Forces
Victoria, British Columbia, V9A 7N2
Attn: Commanding Officer/WTD

1 Operational Requirements Analysis Cell CFB Esquimalt
STN Forces
Victoria, British Columbia, V9A 7N2
Attn: Commanding Officer

1 G. Rogova (author)
Calspan-UB Research Center, Inc.
PO Box 400
4455 Genesee Street
Buffalo, NY 14225, USA

Total external copies: 24

Total copies: 60

DOCUMENT CONTROL DATA

(Security classification of title, body of abstract and indexing annotation must be entered when document is classified)

1. ORIGINATOR (The name and address of the organization preparing the document. Organizations for whom the document was prepared, e.g. Centre sponsoring a contractor's report, or tasking agency, are entered in section 8.) Defence R&D Canada – Valcartier 2459 Pie-XI Blvd. North Québec, Quebec, Canada G3J 1X5	2. SECURITY CLASSIFICATION (Overall security classification of the document including special warning terms if applicable.) UNCLASSIFIED

3. TITLE (The complete document title as indicated on the title page. Its classification should be indicated by the appropriate abbreviation (S, C or U) in parentheses after the title.)

Information quality effects on information fusion:

4. AUTHORS (Last name, followed by initials – ranks, titles, etc. not to be used.)

Rogova, G.; Bossé, É.

5. DATE OF PUBLICATION (Month and year of publication of document.)	6a. NO. OF PAGES (Total containing information. Include Annexes, Appendices, etc.)	6b. NO. OF REFS (Total cited in document.)
May 2008	48	63

7. DESCRIPTIVE NOTES (The category of the document, e.g. technical report, technical note or memorandum. If appropriate, enter the type of report, e.g. interim, progress, summary, annual or final. Give the inclusive dates when a specific reporting period is covered.)

Technical Report

8. SPONSORING ACTIVITY (The name of the department project office or laboratory sponsoring the research and development – include address.)

Defence R&D Canada – Valcartier
2459 Pie-XI Blvd. North Québec, Quebec, Canada G3J 1X5

9a. PROJECT NO. (The applicable research and development project number under which the document was written. Please specify whether project or grant.)	9b. GRANT OR CONTRACT NO. (If appropriate, the applicable number under which the document was written.)
11bs	

10a. ORIGINATOR'S DOCUMENT NUMBER (The official document number by which the document is identified by the originating activity. This number must be unique to this document.)	10b. OTHER DOCUMENT NO(s). (Any other numbers which may be assigned this document either by the originator or by the sponsor.)
DRDC Valcartier TR 2005-270	

11. DOCUMENT AVAILABILITY (Any limitations on further dissemination of the document, other than those imposed by security classification.)

(X) Unlimited distribution
( ) Defence departments and defence contractors; further distribution only as approved
( ) Defence departments and Canadian defence contractors; further distribution only as approved
( ) Government departments and agencies; further distribution only as approved
( ) Defence departments; further distribution only as approved
( ) Other (please specify):

12. DOCUMENT ANNOUNCEMENT (Any limitation to the bibliographic announcement of this document. This will normally correspond to the Document Availability (11). However, where further distribution (beyond the audience specified in (11)) is possible, a wider announcement audience may be selected.)

Unlimited

13. ABSTRACT (A brief and factual summary of the document. It may also appear elsewhere in the body of the document itself. It is highly desirable that the abstract of classified documents be unclassified. Each paragraph of the abstract shall begin with an indication of the security classification of the information in the paragraph (unless the document itself is unclassified) represented as (S), (C), (R), or (U). It is not necessary to include here abstracts in both official languages unless the text is bilingual.)

In this report, we present results from simulation and analysis that show how to take into account data/information reliability into fusion processes. The report gives the results of an analysis of the main approaches used to address information/data/source reliability. The question of reliability and existence of prior knowledge that impact the data fusion processes, is discussed in the context of probability, possibility, evidential, or fuzzy sets theories. We consider reliability as a second level of uncertainty (uncertainty of evaluation of uncertainty) representing a measure of the adequacy of the model used and the state of the environment observed; so then, reliability represents adequacy of each belief model to the reality. The report reviews, as well, the existing methods of building reliability functions based on probabilistic, possibilistic, evidential frameworks and their effectiveness.

14. KEYWORDS, DESCRIPTORS or IDENTIFIERS (Technically meaningful terms or short phrases that characterize a document and could be helpful in cataloguing the document. They should be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location may also be included. If possible keywords should be selected from a published thesaurus. e.g. Thesaurus of Engineering and Scientific Terms (TEST) and that thesaurus identified. If it is not possible to select indexing terms which are Unclassified, the classification of each should be indicated as with the title.)

Source reliability, Information fusion, Uncertainty representation.

Defence R&D Canada

Canada's Leader in Defence
and National Security
Science and Technology

R & D pour la défense Canada

Chef de file au Canada en matière
de science et de technologie pour
la défense et la sécurité nationale

DEFENCE DÉFENSE

www.drdc-rddc.gc.ca