December 2012 Volume 7 Number 2 ISSN 1557-6418

JAIF 12

Journal of Advances in Information Fusion

A semi-annual archival publication of the International Society of Information Fusion

Regular Papers Page

Possibilistic Medical Knowledge Representation Model 101
Mohammad Alsun, Institut TELECOM, France
Laurent Lecornu, Institut TELECOM, France
Basel Solaiman, Institut TELECOM, France

Distributed Tracking with a PHD Filter using Efficient Measurement Encoding 114
Biruk K. Habtemariam, McMaster University, Canada
Ampikaithasan Aravinthan, McMaster University, Canada
Ratnasingham Tharmarasa, McMaster University, Canada
Kumaradevan Punithakumar, GE Healthcare, London
Thomas Lang, General Dynamics Canada, Canada
Thia Kirubarajan, McMaster University, Canada

Association Performance Enhancement Through Classification 131
Quirin Hamp, University of Freiburg, Germany
Leonhard Reindl, University of Freiburg, Germany

A Fusion Analysis and Evaluation Tool for Multi-Sensor Classification Systems 141
Rommel Novaes Carvalho, George Mason University, USA
Kuochu Chang, George Mason University, USA

Modified Scoring in Multiple-Hypothesis Tracking 153
Stefano Coraluppi, Compunetix Inc., USA
Craig Carthel, Compunetix Inc., USA

ISIF

INTERNATIONAL SOCIETY OF INFORMATION FUSION

The International Society of Information Fusion (ISIF) is the premier professional society and global information resource for multidisciplinary approaches for theoretical and applied INFORMATION FUSION technologies. Technical areas of interest include target tracking, detection theory, applications for information fusion methods, image fusion, fusion systems architectures and management issues, classification, learning, data mining, Bayesian and reasoning methods.

Journal of Advances in Information Fusion (ISSN 1557-6418) is published semi-annually by the International Society of Information Fusion. The responsibility for the contents rests upon the authors and not upon ISIF, the Society, or its members. ISIF is a California Nonprofit Public Benefit Corporation at P.O. Box 4631, Mountain View, California 94040.

Possibilistic Medical Knowledge Representation Model

MOHAMMAD ALSUN
LAURENT LECORNU
BASEL SOLAIMAN

Medical Decision Support Systems involve two main issues: medical knowledge representation and reasoning mechanisms adapted to the considered representation model. This paper proposes an approach to construct a new medical knowledge representation model, based on the use of possibility theory. The major interest of using the possibility theory comes from its capacity to represent different types of information (quantitative, qualitative, binary, etc.), as well as different forms of information imperfections such as uncertainty, imprecision, ambiguity and incompleteness. Starting from the description, realized by an expert of the medical knowledge, describing the relationship between symptoms and diagnoses, the proposed approach consists on building a possibilistic model including the Medical Knowledge Base. Moreover, the proposed approach integrates several possibilistic reasoning mechanisms based on the considered knowledge. The validation of the proposed approach is then conducted using an Endoscopic Knowledge Base. The proposed representation, reasoning model and the obtained validation results show a real interest in order to realize various goals of Medical Decision Support Systems such as classification, similarity estimation, etc.

Manuscript received May 30, 2011; released for publication September 6, 2011.

Refereeing of this contribution was handled by Jean Dezert.

Author's address: Institut Télécom, Télécom Bretagne, ITI Dpt, Technopôle Brest-Iroise, CS 83818, 29238 Brest Cedex 3, France.

1557-6418/12/$17.00

1. INTRODUCTION

Physician is the direct responsible for health and life of his patients. Therefore, diagnosis delivering is an extremely critical, although difficult task. Furthermore, diagnosis delivering is an error-prone task [3]. Medical Decision Support Systems such as Knowledge Based Systems, Case Based Reasoning Systems, Machine Learning Systems and Medical Data Mining Systems, have been constructed in order to reduce diagnosis error risks, as well as to help physicians making high quality and reliable medical decisions [4]. These systems involve two main issues: the medical knowledge representation and adapted reasoning mechanisms. The medical knowledge, in general terms, has to be considered from two points of view: Expert Knowledge related to the physician's description of different relationships between symptoms and diagnoses, symptoms and symptoms, and diagnoses and diagnoses. Patient Information is collected from each patient (patient data collecting and structuring). The first is crucial in order to establish Medical Knowledge Base, while the second leads to establish the patient database (i.e., Medical Case Base). Experts use their own experience of the medical cases as well as references knowledge sources to define the structure of the medical knowledge base.

Medical knowledge often suffers from different forms of information imperfections (i.e., uncertainty, imprecision, ambiguity, etc.). In addition to the different types of information imperfections, the information can be quantitative (numerical or binary) or qualitative (nominal and ordinal) [17, 29]. Thus, the heterogeneity and imperfection of medical knowledge must be taken into consideration while the construction of a Medical Decision Support System. In other words, Medical Decision Support System has to be able to deal with heterogeneous and imperfect knowledge.

In [27] R. Seising et al. defined the Medical knowledge as follows:

> "*The certain information about relationships that exist between symptoms and symptoms, symptoms and diagnoses, diagnoses and diagnoses and more complex relationships of combinations of symptoms and diagnoses to a symptom or diagnosis are formalizations of what is called medical knowledge*."

The term "symptom" is used for any information about the patient's state health (anamnesis, signs, laboratory test results, etc.).

According to the previous definition, the term "medical knowledge" will be considered in this study to represent the relationship between symptoms and diagnoses, (Symptom)–(Diagnosis). This relation is generally expressed in a probabilistic way based on the use of a linguistic term, referring to the expert's assessment of the occurrence of a given symptom related to a given diagnosis.

In order to be exploited in Medical Decision Support Systems, this Medical Knowledge has to be modeled (translated into a model understandable by the system)

using one of representation's approaches such as probabilistic, fuzzy, possibilistic, etc.

In [4] and [18], theory about clinical decision supports system is presented. The probabilistic approach is one of the first model that can be proposed, regarding the probabilistic nature of the linguistic term. The following probability, $\text{pr}(\mathbf{D} \mid B)$ (where $\mathbf{D}$ and B represent respectively a diagnosis and a symptom), should be computed for each diagnosis. This value is obtained by the Bayes' formula:

$$\text{pr}(\mathbf{D} \mid B) = \frac{\text{pr}(B \mid \mathbf{D}).\text{pr}(\mathbf{D})}{\text{pr}(B)}. \quad (1)$$

In this formula we need two types of information:

—$\text{pr}(B \mid \mathbf{D})$ which is available

—$\text{pr}(\mathbf{D})$ which is difficult to be known, but it can be estimated by a statistic approach.

In our case, we suppose that the only available information is $\text{pr}(B \mid \mathbf{D})$. For this reason the probabilistic representation approach is not adequate in our context.

Fuzzy sets theory, introduced by Zadeh [33] has several interesting properties that make it suitable for formalizing the imperfect information upon which medical diagnosis is usually based on. Firstly, it allows the definition of inexact and/or ambiguous medical entities as fuzzy sets. Secondly, it offers the possibility of using linguistic variables in addition to crisp numerical variables. Finally, fuzzy logic (i.e. mathematical logic allowing the manipulation of fuzzy sets) offers reasoning methods adequate for approximate inferences drawing. Fuzzy sets as a framework representation and fuzzy logic as a reasoning mechanism have been successfully applied to different Medical Decision Support Systems [1, 5, 21, 22, 26].

Progress in this field was characterized by the introduction of the possibility theory as an alternative approach of the inexact reasoning. Although the possibility theory is an extension of the fuzzy sets theory, it has many advantages over to make it more suitable as well as more efficient [24, 35]. In fact, possibility theory provides an approach to formalize subjective uncertainties of events, that is to say means of assessing to what extent the occurrence (realization) of an event is possible and to what extent we are certain of its occurrence, without having the possibility to measure the exact probability of this realization because we lake similar events to be referred to, or because the uncertainty is the consequence of observation instruments reliability absence. It also offers the advantage of decision making based on two set-based measures called the possibility and the necessity measures. At the level of information fusion, the possibility theory uses simple mathematical operations (min, max, etc.). Several studies proved the successful using of possibility theory as a representation framework and as a reasoning mechanism in Medical Decision Support Systems [9].

In this document, we propose the use of the possibility theory [10] as a global framework in our Medical Decision Support System. After studying the existing possibilistic approaches, we can note that these works neglect the issue of the medical knowledge representation, and concentrate their contribution only on the issue of possibilistic reasoning (for instance see [10]). In other words, there is no algorithm describing the phase concerning the transition from the medical description (i.e. the linguistic term expressing the medical knowledge (Symptom)–(Diagnosis)) into a possibilistic description (i.e. numerical value in the interval $[0, 1]$ expressing the occurrence possibility degree of Symptom with a given Diagnosis).

The important contribution of this work is to answer the question concerning the issue of medical knowledge possibilistic representation. Furthermore, this work proposes an algorithm describing, in details, the construction of Possibilistic Knowledge Base (in which the relation (Symptom)–(Diagnosis) is represented by possibilistic value belonging to the interval $[0, 1]$) from Medical Knowledge Base (in which the relation (Symptom)–(Diagnosis) is represented by linguistic term).

This document is organized as follows: Section 2 details a knowledge representation model allowing physicians to express their medical knowledge. Main aspects of possibility theory are briefly introduced in Section 3. Section 4 is devoted to the detailed description of the proposed approach to construct a new possibilistic model of medical knowledge and to the use of this model in order to build Possibilistic Knowledge Base. In Section 5, the evaluation of the reliability of the constructed model will be conducted by realizing several tasks accomplished by Medical Decision Support Systems. The particular Endoscopic Knowledge Base allowing the validation of proposed possibilistic model, obtained results and the comparison with prior ones are detailed in Section 6. Finally, Section 7 presents conclusions concerning the proposed model as well as some propositions for further developments.

2. MEDICAL KNOWLEDGE AND REPRESENTATION (EXPERT VISION)

The objective of the medical knowledge base construction is to perform a reliable information modeling of the medical knowledge description, expressed by physicians, according to a predefined knowledge representation scheme. The knowledge representation schemes had been classified by Carter [6] into four categories: logical, procedural, graph/network and structural. In this paper, we adopt the structural model that has been used by Cauvin [4] in order to construct the Endoscopic Knowledge Base, which represents our medical application. According to this structural model, this section is devoted to present the description of:

—Diagnoses in the medical knowledge base;

—Patient-cases in the medical case base.

From here, we will use the term "feature" to represent the name of symptom, and the term "modality" to represent the value of symptom. For example, the

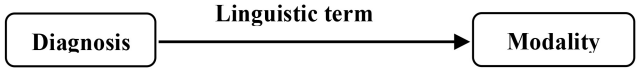


Fig. 1. Qualitative description using linguistic term.

Fig. 2. Example of qualitative description using linguistic term.

TABLE I
Example of Physician Description in a Medical Knowledge Base

	$\mathbf{P}_1$			$\mathbf{P}_2$	
	v_1^1	v_2^1	v_3^1	v_1^2	v_2^2
$\mathbf{D}_1$	always	never	never	rare	usual
$\mathbf{D}_2$	rare	usual	usual	never	always
$\mathbf{D}_3$	exceptional	usual	usual	never	always

feature "Temperature" can take one of three modalities *low*, *normal*, *high*.

2.1. Medical Knowledge Base

The Medical Knowledge Base is assumed to encapsulate the expert knowledge related to the different considered diagnoses. A diagnosis is represented by physicians, using all potential modalities of the predefined features, through describing the relationship between modalities and diagnoses. This relationship expresses the occurrence, assessed by physicians, of a given modality for a given diagnosis.

2.1.1. (Modality)–(Diagnosis) Relationship

From a probabilistic point of view, the ideal representation of this relation is to attribute to each couple (Modality)–(Diagnosis), it's exact occurrence probability value. Nevertheless, these values are rarely known by physicians in terms of exact values. For this reason and in order to express this imprecise/ambiguous knowledge of the probabilistic values, physicians use a qualitative description by means of natural language [2]. This description mode offers physicians the opportunity to express their uncertainty by using linguistic terms more indicative than numerical ones used in possibility or probability theories. The form of the qualitative description using these linguistic terms is shown in Fig. 1. In this form, the linguistic term belongs to the scale $\{never,\ldots,always\}$. For instance, if the relation between a given diagnosis *Flu* and a given modality *fever* is described by the linguistic term *habitual* as shown in Fig. 2, then we can read: the modality *fever* occurs *habitually* with the diagnosis *rheum*.

2.1.2. Medical Diagnosis Representation

Let $\mathbf{D} = \{\mathbf{D}_1,\ldots,\mathbf{D}_M\}$ denote the set of M diagnoses, $\mathbf{P} = \{\mathbf{P}_1,\ldots,\mathbf{P}_G\}$ denote the set of G features used for the description of diagnoses. In this description, each feature is considered independently from the others. Each feature $\mathbf{P}_g$ can assume one of K_g potential modalities defined by the set $\mathbf{V}_g = \{v_1^g, v_2^g,\ldots,v_{K_g}^g\}$. The diagnosis $\mathbf{D}_m$, $m = 1,\ldots,M$, is thus represented in the medical knowledge base by the following model:

$$\mathbf{D}_m = \{(\mathbf{P}_g, v_j^g, R(v_j^g, \mathbf{D}_m));\ g = 1,\ldots G;\ j = 1,\ldots,K_g\} \tag{2}$$

where

- $\mathbf{P}_g$ denotes the feature "g";
- v_j^g is the jth modality ($j = 1,\ldots,K_g$) of the feature "g";
- $R(v_j^g, \mathbf{D}_m)$ represents the linguistic term (defined by physicians) that expresses the occurrence of jth modality related with the given diagnosis $\mathbf{D}_m$;
- $\mathbf{Q} = \{q_1,\ldots,q_L\}$ represents the predefined set of linguistic terms.

Table I shows an example of an expert description in a medical knowledge base [4]. In this example, the physician describes a set of three diagnoses (diseases) $\mathbf{D} = \{\mathbf{D}_1,\mathbf{D}_2,\mathbf{D}_3\}$ using two features: $\mathbf{P}_1$ (with three modalities: $\mathbf{V}_1 = \{v_1^1, v_2^1, v_3^1\}$) and $\mathbf{P}_2$ (with two modalities: $\mathbf{V}_2 = \{v_1^2, v_2^2\}$). Five linguistic terms are used: $\mathbf{Q} = \{q_1 = never,\ q_2 = exceptional,\ q_3 = rare,\ q_4 = usual,\ q_5 = always\}$.

2.2. Patient-Case Representation

The Medical Case Base is assumed to encapsulate the recorded data collected from different patients. An expert standardizes the description such that a case has a unique description and is structured to be used by a computer-aided system [4].

A patient-case is described by physicians using the same set of G features ($\mathbf{P}_g$, $g = 1,\ldots,G$) used in the description of diagnoses. Each feature $\mathbf{P}_g$ can assume one and only one of its potential modalities included in its corresponding feature modalities set $\mathbf{V}_g$, or it can assume the value "0" in the case where this feature is not evaluated (i.e., a missing data) or if the feature is impossible to be observed or to be evaluated.

Let $B = \{B_1,\ldots,B_N\}$ denote a medical case base containing a set of N patient-cases. A patient-case B_n, $n = 1,2,\ldots,N$, is thus represented in the medical case base by the following medical model:

$$B_n = \{(\mathbf{P}_g, x^{g,n}), \mathbf{D}_n\} \tag{3}$$

where

- $x^{g,n}$ is the value of the feature $\mathbf{P}_g$ such that $x^{g,n} \in \mathbf{V}_g \cup \{0\}$, $g = 1,\ldots,G$;
- $\mathbf{D}_n$ is the diagnosis associated with the case B_n, $\mathbf{D}_n \in \mathbf{D} = \{\mathbf{D}_1,\ldots,\mathbf{D}_M\}$, ($\mathbf{D}$ contains all possible diagnoses).

In this model, only a discrete set of modalities is involved, it means that an expert divides each continuous modality in intervals.

An illustrative example of an Endoscopic Medical Case Base is shown in Table II. In this example,

TABLE II
Example of Physician Description in an Endoscopic Case Base

	$\mathbf{P}_1$ = Object Type	$\mathbf{P}_2$ = Origin	Diagnosis
$\mathbf{B}_1$	Not Homogenous Simple	Parietal	Tumor
$\mathbf{B}_2$	Homogenous	Parietal	Spot
$\mathbf{B}_3$	Not Homogenous Multiple	Luminal	Food

three cases B_1, B_2, B_3, are described using two features: $\mathbf{P}_1$ = "Object Type" with three modalities {*Homogenous*, *Not Homogenous Simple*, *Not Homogenous Multiple*} and $\mathbf{P}_2$ = "Origin" with two modalities {*Parietal*, *Luminal*}. The associated diagnosis is respectively given as {$\mathbf{D}_1$ = *Tumor*, $\mathbf{D}_2$ = *Spot*, $\mathbf{D}_3$ = *Food*}.

3. POSSIBILITY THEORY

3.1. Possibility and Necessity Measures

Possibility Theory, introduced by L. Zadeh in 1978 [34] and then developed by Dubois and Prade in 1988 [13], offers an interesting tool allowing to deal with different forms of information imperfections (ambiguity, imprecision, incompleteness, etc.).

The possibility theory constitutes the basis of several recent studies in medicine [11]. The obtained results in these studies confirmed the efficacy of the use of possibility theory as a tool for medical knowledge representation, as well as for the medical diagnostic decision.

Let $\mathbf{\Omega} = \{x_1,\ldots,x_N\}$ denote an exhaustive and exclusive Universe of discourse that means the list of the possible alternatives. At the semantic level, the basic function in possibility theory is the *possibility distribution* denoted as $\pi : \Omega \rightarrow [0,1]$ which assigns to each possible alternative x_n from Ω a value ranging within the interval $[0,1]$. This possibility distribution represents the possibility occurrence degree of x_n, the basic alternative, decision, diagnosis, etc. If, for some x_n, $\pi(x_n) = 1$, then x_n is said to be a *totally possible* alternative; and if $\pi(x_n) = 0$, then x_n is said to be an *impossible* alternative. Based on a possibility distribution, the information concerning the occurrence of an event $A \in \mathcal{P}(\Omega)$ ($\mathcal{P}(\Omega)$ is the power set of Ω) is represented by means of two set functions: a *Possibility Measure* denoted as $\mathbf{\Pi}(.)$ and a *Necessity Measure* denoted as $\mathbf{N}(.)$.

The *possibility measure* $\mathbf{\Pi}(.)$ is defined as follows [11]:

$$\begin{aligned} &\Pi : \mathcal{P}(\Omega) \rightarrow [0,1] \\ &A \rightarrow \Pi(A) = \max_{x_n \in A}(\pi(x_n)) \end{aligned} \tag{4}$$

and satisfying the following requirements:

$$\Pi(\Phi) = 0 \qquad \text{and} \qquad \Pi(\Omega) = 1 \tag{5}$$

$$\Pi\left(\bigcup_{j\in J} A_j\right) = \max_{j\in J} \Pi(A_j) \qquad \forall A_j, \quad j \in [1,J] \tag{6}$$

where J represents the number of elements of the set $\mathcal{P}(\Omega)$.

If the possibility measure of an event $A \in \mathcal{P}(\Omega)$ is equal to the unity (i.e., $\Pi(A) = 1$, then A is said to be *totally possible* event. If $\Pi(A) = 0$ then, A is said to be totally *impossible* event.

Reciprocally, the possibility distribution can be defined from the possibility measure, by affecting the possibility measure of the subset $A = \{x_n\}$ to the alternative x_n: $\pi(x_n) = \Pi(\{x_n\})$.

The second measure, called the *necessity measure* $\mathbf{N}(.)$, is defined as follows [11]:

$$\begin{aligned} &N : \mathcal{P}(\Omega) \rightarrow [0,1] \\ &A \rightarrow N(A) = 1 - \max_{x_n \in A}(1 - \pi(x_n)) \qquad \forall A_j, \quad j \in [1,J] \end{aligned} \tag{7}$$

and satisfying the following requirements:

$$N(\Phi) = 0 \qquad \text{and} \qquad N(\Omega) = 1 \tag{8}$$

$$\Pi\left(\bigcap_{j\in J} A_j\right) = \min_{j\in J} N(A_j) \qquad \forall A_j, \quad j \in [1,J]. \tag{9}$$

If the necessity measure of an event $A \in \mathcal{P}(\Omega)$ is equal to the unity (i.e., $N(A) = 1$, then A is said to be *totally certain*; and if $N(A) = 0$, then A is said to be *totally uncertain*.

Several duality relations link the possibility measure and the necessity measure:

$$A \in \mathcal{P}(\Omega) : 0 \leq N(A) \leq \Pi(A) \tag{10}$$

$$\text{If} \quad N(A) > 0, \quad \text{then} \quad \Pi(A) = 1 \tag{11}$$

$$\text{If} \quad \Pi(A) < 1, \quad \text{then} \quad N(A) = 0 \tag{12}$$

$$N(A) = 1 - \Pi(A^c). \tag{13}$$

3.2. Joint and Conditional Possibility Distribution

Within the application studied here, the expert expresses medical knowledge as the possibility of modality occurrence given a diagnosis. This type of knowledge can be modeled using the conditional possibility concept.

Given two reference sets X and Y where $X = \{x_1,\ldots,x_M\}$ and $Y = \{y_1,\ldots,y_N\}$, a joint possibility distribution $\pi(x_m, y_n)$ where $x_m \in X$ ($m = 1,\ldots,M$) and $y_n \in Y$ ($n = 1,\ldots,N$) can be defined on the Cartesian product $X \times Y$ in order to express the joint occurrence possibility of the singletons $x_m \in X$ ($m = 1,\ldots,M$) and $y_n \in Y$ ($n = 1,\ldots,N$) [34]. The joint possibility distribution provides information on each reference set X and Y individually as two marginal possibility distributions, obtained by retaining the largest value of joint possibility distributions relative to the reference set as it is explicated in the following definitions.

DEFINITION 1 Starting from a given joint distribution π, the *marginal possibility distributions* are defined on the two reference sets X and Y as follows:

$$\forall x_m \in X : \pi(x_m) = \sup_{y \in Y} \pi(x_m, y) \tag{14}$$

$$\forall y_n \in Y : \pi(y_n) = \sup_{x \in X} \pi(x, y_n). \tag{15}$$

The reciprocal influence among the reference sets can be studied through the degree to which an element y_n of Y is possible, knowing that the element x_m of X is considered. In other words, the conditional possibility which is defined as follows:

DEFINITION 2 There is not a unique definition of the *conditional possibility distribution* $\pi(y_n \mid x_m)$ measuring the occurrence for an element y_n from Y knowing that the element x_m from X has occurred [6], but all proposed definitions in the literature [31] are based on the following general formula linking the conditional possibility with the joint and marginal possibilities:

$$\pi(x_m, y_n) = \pi(y_n \mid x_m) * \pi(x_m) \qquad \forall x_m \in X, \quad \forall y_n \in Y \tag{16}$$

where, "*" is a combination operator which can be considered as the *minimum* or the *product* fusion operator.

The decision made by humans, is usually taken based on information fusion of different types and assigned various forms of imperfection: uncertain information, possibilistic information, binary information, ambiguous information, etc. To address these different types of information into a single framework, a transformation from one type to another is fundamental. An important facet of the theory of possibilities lies in the ability to transform probabilistic information in possibilistic information in carrying out the projection of probability distributions to possibility ones. Indeed, this transformation is a useful operation when dealing with heterogeneous information. Several transformations of a probability distribution into a possibility distribution and conversely have been proposed in this direction. In this study, we will adopt Dubois-Prade transformation [13, 15]:

Dubois-Prade transformation procedure:

Given a reference set $X = \{x_1, \ldots, x_M\}$, in which each element x_i is associated with its probability $\mathrm{pr}_i = \Pr(\{x_i\})$, $i = 1, \ldots, M$. In order to perform the transformation from the given probability into a possibility distribution, first, the probability values are arranged in a decreasing order so that $\mathrm{pr}_1 \geq \mathrm{pr}_2 \geq \ldots \geq \mathrm{pr}_M$; then, the following possibility degrees are computed $\forall i$, $i = 1, \ldots, M$:

$$\begin{aligned} \pi_1 &= 1 \\ \pi_i &= \Pi(\{x_i\}) = \sum_{j=1}^{M} (\mathrm{pr}_j) \qquad \text{if} \quad \mathrm{pr}_{i-1} > \mathrm{pr}_i \\ &= \pi_{i-1} \qquad\qquad\qquad\quad \text{otherwise.} \end{aligned} \tag{17}$$

4. POSSIBILISTIC MEDICAL KNOWLEDGE REPRESENTATION MODEL

In Section 2, we have shown the Medical Knowledge Base, and how physicians qualitatively describe, using linguistic terms, the medical knowledge considered mainly as a relationship (Modality)–(Diagnosis).

In order to be exploited in Medical Decision Support Systems, this Medical Knowledge Base has to be modeled using one of representation approaches. Furthermore, the linguistic term, expressing the relationship (Modality)–(Diagnosis), has to be translated into a model understandable by the system.

This section is devoted to present our proposed approach in order to represent this kind of relationships by means of possibilistic model.

4.1. Possibilistic Knowledge Base Construction

Assume that a Medical Knowledge Base (as described in Section 2), containing a set $\mathbf{D}$ of M diagnoses, is available. Each diagnosis in this base $\mathbf{D}_m$, $m = 1, \ldots, M$, is characterized using a set $\mathbf{P}$ of G features. Each feature $\mathbf{P}_g$, $g = 1, \ldots, G$, can assume one of K_g possible modalities grouped in a set $\mathbf{V}_g$. The diagnosis D_m is thus expressed using the model given in (2). The expert will indicated the modality frequency for each diagnosis in using a qualitative scale $\mathbf{Q}$ of L linguistic terms $\mathbf{Q} = \{q_1, \ldots, q_L\}$ running from "*never*" to "*always*" as follows: $q_1 = never, \ldots, q_L = always$. The expert doesn't know the exact probability but only an approximation.

The objective here is to translate the Medical Knowledge Base established by the Expert into a possibilistic model exploitable by medical decision support systems. In other words, we want to build the following possibilistic model of diagnosis $\mathbf{D}_m$, $m = 1, 2, \ldots, M$, in which the relationship (Modality)–(Diagnosis) is represented as a possibility value:

$$\mathbf{D}_m = \{(\mathbf{P}_g, v_j^g, \pi(v_j^g \mid \mathbf{D}_m));\ g = 1 \ldots, G;\ j = 1, \ldots, K_g\}. \tag{18}$$

The proposed approach to realize reach this target consists on performing the following steps:

Step 1 Transforming the qualitative scale of linguistic terms $\mathbf{Q} = \{q_1, \ldots, q_L\}$ into a quantitative one of numerical values $\boldsymbol{\alpha} = \{\alpha_1, \ldots, \alpha_L\}$ where $\alpha_i \in [0, 1]$, $\alpha_1 = 0, \ldots, \alpha_L = 1$, and $\forall j \in [0, L-1] : \alpha_j < \alpha_{j+1}$, so that $\forall i \in [1, L]$ there is $q_i \equiv \alpha_i$.

Step 2 Substituting each $R = q_i \in \mathbf{Q}$ in the Medical Knowledge Base by the corresponding numerical value α_i. Therefore, the representation of a given diagnosis $\mathbf{D}_m$ will be as follows:

$$\mathbf{D}_m = \{(\mathbf{P}_g, v_j^g, \alpha(v_j^g \mid \mathbf{D}_m));\ g = 1 \ldots, G;\ j = 1, \ldots, K_g\}. \tag{19}$$

In fact, the distribution of numerical values at the level of given feature $\mathbf{P}_g$, cannot called a probability distri-

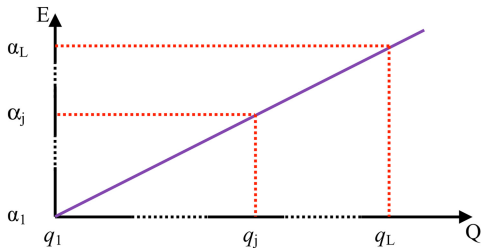


Fig. 3. Projection d'une échelle qualitative en une échelle numérique.

TABLE III
Linguistic Term Substitution by Numerical Ones

	P		
	$v_1^{\mathbf{P}}$	$v_2^{\mathbf{P}}$	$v_3^{\mathbf{P}}$
$\mathbf{D}_1$	always	never	never
$\mathbf{D}_2$	rare	usual	usual
$\mathbf{D}_3$	exceptional	usual	usual
$\mathbf{D}_4$	usual	exceptional	never

bution because the normality condition is not satisfied (i.e., $\sum_{j=1}^{K_g} \alpha(v_j^g \mid \mathbf{D}_m) \neq 1$). For this reason, a normalization operation at the level of feature is necessary.

Step 3 Normalizing the numerical values α at the level of feature, in order to have a conditional probability distribution:

$$\mathbf{D}_m = \{(\mathbf{P}_g, v_j^g, \mathrm{pr}(v_j^g \mid \mathbf{D}_m));\ g = 1\ldots,G;\ j = 1,\ldots,K_g\} \tag{20}$$

so that:

$$\mathrm{pr}(v_j^g \mid \mathbf{D}_m) = \frac{\alpha(v_j^g \mid \mathbf{D}_m)}{\alpha(v_1^g \mid \mathbf{D}_m) + \cdots + \alpha(v_j^g \mid \mathbf{D}_m) + \cdots + \alpha(v_{K_g}^g \mid \mathbf{D}_m)}. \tag{21}$$

For $j = 1,\ldots,K_g$.

Step 4 Applying the Dubois-Prade transformation on the probability distributions in order to construct the conditional possibility distributions. Once the transformation is performed, we obtain the model presented in (18) of $\mathbf{D}_m$.

4.2. Illustrative Example

In order to illustrate the construction of the Possibilistic Knowledge Base, let us consider the following example: Assume that we have a set of four diagnoses $\mathbf{D} = \{\mathbf{D}_1, \mathbf{D}_2, \mathbf{D}_3, \mathbf{D}_4\}$ described using one feature $\mathbf{P}$ of three potential modalities grouped in a set $\mathbf{V}_P = \{v_1^{\mathbf{P}}, v_2^{\mathbf{P}}, v_3^{\mathbf{P}}\}$ (Table III), and the occurrence of these modalities is represented by means of the qualitative scale $\mathbf{Q} = \{q_1 = \textit{never}, q_2 = \textit{exceptional}, q_3 = \textit{rare}, q_4 = \textit{usual}, q_5 = \textit{always}\}$.

TABLE IV
Substituting Linguistic Terms by Numerical Ones

	P		
	$v_1^{\mathbf{P}}$	$v_2^{\mathbf{P}}$	$v_3^{\mathbf{P}}$
$\mathbf{D}_1$	1	0	0
$\mathbf{D}_2$	0.25	0.75	0.75
$\mathbf{D}_3$	0.1	0.75	0.75
$\mathbf{D}_4$	0.75	0.1	0

TABLE V
Resulting Probability Distribution

	P		
	$v_1^{\mathbf{P}}$	$v_2^{\mathbf{P}}$	$v_3^{\mathbf{P}}$
$\mathbf{D}_1$	1	0	0
$\mathbf{D}_2$	0.14	0.43	0.43
$\mathbf{D}_3$	0.06	0.47	0.47
$\mathbf{D}_4$	0.88	0.12	0

In order to construct the possibilistic model of these four diagnoses following the proposed approach, steps from 1 to 4 must be applied as follows:

Step 1 The projection of the qualitative scale $\mathbf{Q}$ (having five linguistic values), onto a numerical scale $\boldsymbol{\alpha}$ (also having five empirical numerical values), will produce $\boldsymbol{\alpha} = \alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5\}$ where $\alpha_i \in [0,1]$, $\alpha_1 = 0,\ldots,\alpha_5 = 1$, and $j \in [1,4] : \alpha_j < \alpha_{j+1}$, so that: $\forall i \in [1,5]$ we obtain $q_i \equiv \alpha_i$ as follows:

—$q_1 = \textit{never} \rightarrow \alpha_i = 0$,
—$q_2 = \textit{exceptional} \rightarrow \alpha_2 = 0.1$,
—$q_3 = \textit{rare} \rightarrow \alpha_3 = 0.25$,
—$q_4 = \textit{usual} \rightarrow \alpha_4 = 0.75$,
—$q_5 = \textit{always} \rightarrow \alpha_5 = 1$.

Step 2 Substituting each linguistic term q_i in Table III by the corresponding numerical value α_i, leads to Table IV.

We note that the sum of numerical values at the level of the feature $\mathbf{P}$ doesn't equal to 1 (for example, $\sum_{j=1}^{3} \alpha(v_j^{\mathbf{P}} \mid \mathbf{D}_2) = 0.25 + 0.75 + 0.75 = 1.75 \neq 1$). For this reason, a normalization operation at the level of feature is necessary.

Step 3 The conditional probability value for each modality according to a given diagnosis is computed according to (21) and shown in Table V. For example, the conditional probability value of modality $v_2^{\mathbf{P}}$ for a given diagnosis $\mathbf{D}_2$ is calculated as follows:

$$\mathrm{pr}(v_2^{\mathbf{P}} \mid \mathbf{D}_2) = \frac{\alpha(v_2^{\mathbf{P}} \mid D_2)}{\alpha(v_1^{\mathbf{P}} \mid \mathbf{D}_2) + \alpha(v_2^{\mathbf{P}} \mid \mathbf{D}_2) + \alpha(v_3^{\mathbf{P}} \mid \mathbf{D}_2)} = \frac{0.75}{0.25 + 0.75 + 0.75} = 0.43.$$

TABLE VI
Conditional Possibility Distribution (Possibilistic Knowledge Base)

	$\mathbf{P}$		
	$v_1^{\mathbf{P}}$	$v_2^{\mathbf{P}}$	$v_3^{\mathbf{P}}$
$\mathbf{D}_1$	1	0	0
$\mathbf{D}_2$	0.14	1	1
$\mathbf{D}_3$	0.06	1	1
$\mathbf{D}_4$	1	0.12	0

The transformation of the probability distribution into a possibilistic one, will be realized by applying the Dubois-Prade's transformation (Step 4), which will finally produce the possibilistic model of this base containing four diagnoses (Table VI). For example, the probability distribution associated with the diagnosis $\mathbf{D}_4\{\text{pr}(v_1^{\mathbf{P}} \mid \mathbf{D}_4) = 0.88;\ \text{pr}(v_2^{\mathbf{P}} \mid \mathbf{D}_4) = 0.12,\ \text{pr}(v_3^{\mathbf{P}} \mid \mathbf{D}_4) = 0\}$ will be transformed according to Dubois-Prade as follows:

—Ranking the probability distribution as follows: $\text{pr}_1 = \text{pr}(v_1^{\mathbf{P}} \mid \mathbf{D}_4) = 0.88 > \text{pr}_2 = \text{pr}(v_2^{\mathbf{P}} \mid \mathbf{D}_4) = 0.12 > \text{pr}_3 = \text{pr}(v_3^{\mathbf{P}} \mid \mathbf{D}_4) = 0$;

—According to (17), we obtain:

$$\pi_1 = \pi(v_1^{\mathbf{P}} \mid \mathbf{D}_4) = 1,$$

$$\pi_2 = \pi(v_2^{\mathbf{P}} \mid \mathbf{D}_4) = \sum_{i=2}^{3} \text{pr}_i = \text{pr}_2 + \text{pr}_3$$

$$= \text{pr}(v_2^{\mathbf{P}} \mid \mathbf{D}_4) + \text{pr}(v_3^{\mathbf{P}} \mid \mathbf{D}_4) = 0.12 + 0 = 0.12$$

$$\pi_3 = \pi(v_3^{\mathbf{P}} \mid \mathbf{D}_4) = \sum_{i=3}^{3} \text{pr}_i = \text{pr}_3 = \text{pr}(v_3^{\mathbf{P}} \mid \mathbf{D}_4) = 0.$$

This table is defined for each feature. If G is the number of features, then we have G tables.

5. POSSIBILISTIC REASONING

Once Possibilistic Knowledge Base is constructed, as detailed in the previous section, the reliability of the possibilistic modeling should be evaluated. This evaluation has to be performed in terms of the quality of different tasks conducted by medical decision support systems. In this paper, we will study the exploitation of our possibilistic model in medical decision support systems adopting the Reasoning by Classification. This reasoning type is based on the comparison of the available information acquired from a patient with the medical a prior knowledge formulated by physicians (i.e., Expert Medical Vision) with the aim to assign potential diagnoses facing this particular patient-case.

Given a new case with an unknown diagnosis B which its description is as follows:

$$B = \{(\mathbf{P}_g, x^g);\ g = 1,\dots G;\ x^g \in \mathbf{V}_g \cup \{0\}\} \quad (22)$$

where

- $\mathbf{P}_g$ represents the feature 'g';
- x^g represents the observed modality of the feature 'g'. If the feature is observed, then x^g take one and only one value from the set of possible modalities $\mathbf{V}_g = \{v_1^g, v_2^g \dots, v_{K_g}^g\}$, and it takes the value 'zero' otherwise (i.e., the feature $\mathbf{P}_g$ is not observed or it is missing data).

In order to classify this case B (i.e., finding its corresponding diagnosis), we have to compare it with all diagnoses included in the knowledge base, through calculating the similarity between this case and each diagnosis, and then ranking the set of obtained potential diagnoses according to the maximum similarity.

The similarity between B and $\mathbf{D}_m$, $m = 1,\dots,M$, is represented in our approach by the possibilistic couple $[N(\mathbf{D}_m \mid B), \Pi(\mathbf{D}_m \mid B)]$ of similarity which can be estimated by performing the following steps:

Step 1 Estimation of the local conditional possibility (i.e., at the level of feature), $\pi(\mathbf{D}_m \mid \mathbf{P}_g)$, $m = 1,2,\dots,M$ and $g = 1,2,\dots,G$. Here, we distinguish two cases:

- The feature $\mathbf{P}_g$ is observed and produced in the case B as the modality x^g: in this case, the local conditional possibility $\pi(\mathbf{D}_m \mid \mathbf{P}_g = x^g)$ will be estimated from the possibilistic knowledge $\pi(x^g \mid \mathbf{D}_m)$ which is available in the possibilistic knowledge base (as we will see later).
- The feature $\mathbf{P}_g$ is not observed or it is a missing data: in this case, the local conditional possibility is considered equal to the unity, $\pi(\mathbf{D}_m \mid \mathbf{P}_g = 0) = 1$. This means that the diagnosis $\mathbf{D}_m$ is considered as possible solution for a given feature $\mathbf{P}_g$.

Step 2 Estimation of the global conditional possibility (i.e., for the set of all features), $\pi(\mathbf{D}_m \mid B)$, $m = 1,\dots,M$, by performing a conjunctive fusion of local conditional possibilities. Indeed, the choice of the conjunctive as a fusion type is justified by the fact that if the diagnosis D_m is impossible to produce as a potential solution, at least for one feature (i.e., $\pi(\mathbf{D}_m \mid \mathbf{P}_g) = 0$), then the diagnosis has to be rejected as an impossible solution to the target case B (i.e., $\pi(\mathbf{D}_m \mid B) = 0$). For example, using the conjunctive operator min, we obtain:

$$\pi(\mathbf{D}_m \mid B) = \min_{g=1}^{G} \pi(\mathbf{D}_m \mid \mathbf{P}_g). \quad (23)$$

After this step, we obtain the conditional possibility distribution defined on the set of diagnoses: $\{\pi(\mathbf{D}_1 \mid B), \dots, \pi(\mathbf{D}_M \mid B)\}$.

Step 3 Using the previous possibility distribution to calculate the conditional possibilistic couple $[N(\mathbf{D}_m \mid B), \Pi(\mathbf{D}_m \mid B)]$, $m = 1,\dots,M$, according to the following formulas:

$$\Pi(\mathbf{D}_m \mid B) = \max_{n=m}(\pi(\mathbf{D}_n \mid B)) = \pi(\mathbf{D}_m \mid B) \quad (24)$$

$$N(\mathbf{D}_m \mid B) = 1 - \Pi(\overline{\mathbf{D}_m} \mid B) = 1 - \max_{\substack{n=1 \\ n \neq m}}^{M}(\pi(\mathbf{D}_n \mid B)). \tag{25}$$

It is clear that the possibilistic couple estimating is essentially based on the availability of the local possibility value, $\pi(\mathbf{D}_m \mid x^g)$ (i.e., more precisely, the possibility value $\pi(\mathbf{D}_m \mid v_j^g)$, $j = 1,2,\ldots,K_g$). However, the real challenge lies in the fact that this value is not available in the possibilistic knowledge base. Indeed, the available information is the local possibility $\pi(v_j^g \mid \mathbf{D}_m)$ (i.e., the possibility of observing a given modality of a certain feature, since the diagnosis $\mathbf{D}_m$). For this reason, the essential question that arises is:

"*How can calculate the conditional possibility* $\pi(\mathbf{D}_m \mid v_j^g)$ *where the information available in the possibilistic knowledge base is the conditional possibility* $\pi(v_j^g \mid \mathbf{D}_m)$, $m = 1,2,\ldots,M$, $g = 1,2,\ldots,G$, $j = 1,2,\ldots,K_g$?"

In order to answer to this question, we use the formula (16) that defines the conditional possibility distribution. From this formula, we can write:

$$\pi(\mathbf{D}_m, v_j^g) = \pi(v_j^g \mid \mathbf{D}_m) * \pi(\mathbf{D}_m) = \pi(\mathbf{D}_m \mid v_j^g) * \pi(v_j^g). \tag{26}$$

From this formula, we notice that:

—The estimating of the conditional possibility $\pi(\mathbf{D}_m \mid v_j^g)$ is based on, beside to the conditional possibility $\pi(v_j^g \mid \mathbf{D}_m)$ which is known, the availability of marginal possibilities $\pi(v_j^g)\pi(\mathbf{D}_m)$ which are unknown.

—Also, this relation does not provide a unique opportunity to build the conditional possibility $\pi(\mathbf{D}_m \mid v_j^g)$.

For these reasons, various rules are proposed in the literature to interpret the relation between the conditional and joint possibility distributions, as well as to define the conditional possibility (i.e., *Zadeh's rule*, *Hisdal's equation*, *Ramer's rule*, etc.) [31].

After having analyzed these rules, two of them can be exploited, *Zadeh's rule* and *Nguyen's rule*, thanks to their good properties and their relevance to the process of medical diagnostic reasoning, because of its capability to estimate the conditional possibility $\pi(\mathbf{D}_m \mid v_j^g)$ using only the conditional possibility $\pi(v_j^g \mid \mathbf{D}_m)$ without any other information as the marginal possibility. In this study, we adopt *Zadeh's rule* defining the conditional possibility as equal to the joint one as follows:

$$\pi_{ZA}(y_n \mid x_m) = \pi_{ZA}(y_n, x_m) = \pi_{ZA}(x_m \mid y_n), \quad \forall x_m \in X \quad \text{and} \quad \forall y_n \in Y. \tag{27}$$

6. MEDICAL APPLICATION AND RESULTS

6.1. Endoscopic Application

The Medical Knowledge Base used in this study is an Endoscopic Knowledge Base [8, 19]. This Base consists of a set of 89 endoscopic findings (diagnoses). Each diagnosis is described using a set of 33 features corresponding with 206 global modalities. The qualitative scale used to express the relationship (Modality)–(Diagnosis) by the physicians consists of the following linguistic values {*never*, *exceptional 2*, *exceptional 1*, *rare 2*, *rare 1*, *usual 2*, *usual 1*, *always*}. Furthermore, the linguistic value *doubtful* that is intermediate between *never* and *exceptional*, is added when the expert has an ignorance about the reality of the modality observation. It is important to notice that there are two importance levels for the three variables *exceptional*, *rare*, and *usual*.

The case base used in this study has been developed in the framework of an endoscopic image analyzes system [19]. It is a decision support system of the diagnosis of endoscopic findings. These findings are described by the physicians, from the endoscopic images, through symbolic terms, which are defined by the Minimal Standard Terminology of the SEGE (European Company of Gastro-enterology). A case (or an object) in the base represents a description of the image (using a set of 33 features, 24 features to describe an object and 9 features to describe a potential sub-object) of an endoscopic lesion.

6.2. Experiments and Results

Before analyzing the results of the proposed approach on the global case base, and in order to have a simple and clear representation of the obtained results, we propose to analyze, as an illustrative example, the classification of a small subset of three cases (i.e., endoscopic lesions), $\mathbf{CB} = \{B_1, B_2, B_3\}$, where the "known" diagnoses of these cases are respectively: *Normal Esophagus*, *Dilated Lumen*, and *Ring*.

The compatibility between each case B_f, $f = 1,2,3$, and each diagnosis $\mathbf{D}_m$, $m = 1,2,\ldots,M$, predefined in the knowledge base, will be estimated according to our possibilistic approach (presented in Section 5). In this approach, the diagnosis $\mathbf{D}_m$ is considered as a potential solution for the case B_f, if the conditional possibilistic couple $[N(\mathbf{D}_m \mid B_f), \Pi(\mathbf{D}_m \mid B_f)]$ is not $[0,0]$.

The results obtained by our approach will be compared with that obtained by the fuzzy approach. In the fuzzy approach, $\mathbf{D}_m$ is considered as a potential solution for the case B_f, if the conditional membership degree $\mu(\mathbf{D}_m \mid B_f)$ is not zero.

Fuzzy Theory uses one measure for uncertainty whereas Possibility Theory uses two measures (i.e., the possibility and necessity measures). So, in order to realize the comparison of our possibilistic approach with other one, we must build one measure Ψ which combine the possibilistic couple $[N, \Pi]$ as follows [24]:

$$\Psi(\mathbf{D}_m \mid B) = \frac{N(\mathbf{D}_m \mid B) + \Pi(\mathbf{D}_m \mid B)}{2}. \tag{28}$$

So, according to this measure, the diagnosis $\mathbf{D}_m$ is considered as a potential solution for the case B_f, if the conditional possibilistic measure $\Psi(\mathbf{D}_m \mid B_f)$ is not zero.

The ranking of the potential solutions will be performed according to: the maximal conditional membership degree in the fuzzy approach, and the maximal measure Ψ in possibilistic approach.

Two types of comparison between these two approaches will be realized: comparison in terms of potential diagnoses ranking, and comparison in terms of decision quality. As an evaluation index of the taken decision quality, we propose the use of the distance between the first two potential solutions, if this distance is great, then the decision is of quality, because the discrimination between the potential solutions is easier.

The obtained results are presented in Table VII. This table shows the first two potential diagnoses proposed for each case B_f, as well as the measure "Dist." which represents the evaluation index of taken decision quality according to the considered decision criteria.

To facilitate the comparison and analysis of results presented in this table, we made a graphic representation in Fig. 4. This figures show a representation of the first two potential diagnoses according to the three approaches as well as the distance Dist., for respectively the cases B_1, B_2, B_3. In these figures, the two potential diagnoses obtained by each approach, are presented in the same color (i.e., the colors green, and blue represent respectively the potential diagnoses obtained by the Fuzzy, and Possibilistic Approach).

By analyzing the table and the figure, we note that:

In terms of potential diagnoses ranking:

- For the case B_1, the two approaches gave the true diagnosis (i.e., diagnosis of the studied case) as the first potential solution.
- For the case B_2, the proposed approach gave the true diagnosis as the unique potential solution, whereas the fuzzy approach gave an additional solution as a second potential solution.
- For the case B_3, the proposed approach gave the true diagnosis as the first potential solution, whereas the fuzzy approach gave two diagnoses as two first potential solutions having the same compatibility degree.

In terms of decision quality:

- For the three cases, the distance between the first two potential solutions obtained by possibilistic approach is greater than that obtained by fuzzy approach.
- For the case B_3, the fuzzy approach could not distinguish between the two potential solutions.

After presenting an illustrative example, we will realize a comparison between two approaches on the global case base containing 4450 cases (lesions). As presented in the previous example, the comparison will be realized in terms of the potential diagnoses ranking, and in terms of the taken decision quality.

In order to realize the comparison in terms of the potential diagnoses ranking, we can distinguish four groups:

—**Found:** represents the number of cases for which the right diagnosis is classified as a potential solution.

TABLE VII
Potential diagnoses of the set **CB** according to two approaches (Fuzzy, Possibilistic)

	Fuzzy Approach	Dist.	Possibilistic Approach	Dist.
	$\mu(\mathbf{D}_m \mid \mathbf{B}_f)$		$\Psi(\mathbf{D}_m \mid \mathbf{B}_f)$	
$\mathbf{B}_1$	$\mathbf{D}_1$ = Normal Esophagus: 0.5 $\mathbf{D}_2$ = Spot: 0.37	0.13	$\mathbf{D}_1$ = Normal Esophagus: 0.94 $\mathbf{D}_2$ = Spot: 0.06	0.88
$\mathbf{B}_2$	$\mathbf{D}_1$ = Dilated Lumen: 0.45	0.45	$\mathbf{D}_1$ = Dilated Lumen: 1	1
$\mathbf{B}_3$	$\mathbf{D}_1$ = Ring: 0.49 $\mathbf{D}_2$ = Web: 0.32	0.17	$\mathbf{D}_1$ = Ring: 0.84 $\mathbf{D}_2$ = Web: 0.15	0.69 2

—**Sole:** represents the number of cases for which the right diagnosis is classified as a unique potential solution.

—**First:** represents the number of cases for which the right diagnosis is classified as the first potential solution.

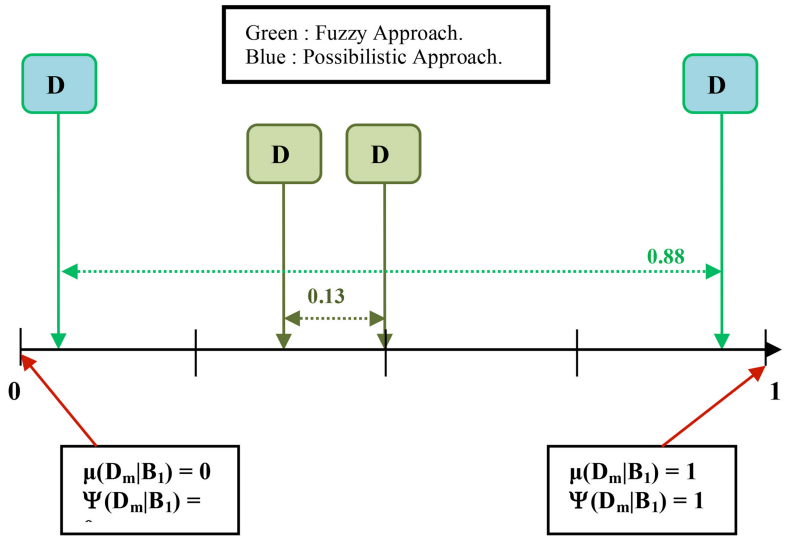
—**Other:** represents the number of cases for which the right diagnosis is classified as a potential solution, but not the first.

We note that the recognition rate associated with diagnostic group called "Found" is always 100% for both fuzzy and possibilistic approaches. This shows that the correct diagnosis still occurs as a potential solution to the target case considered.

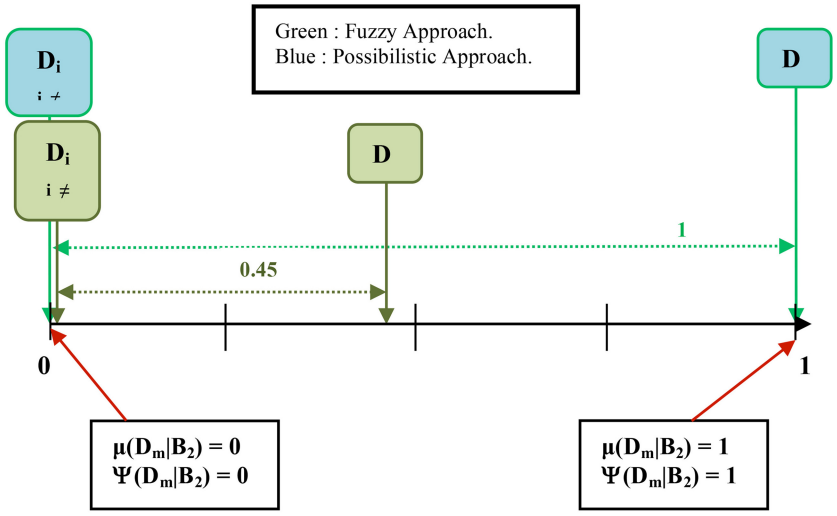
For other groups, we note that the results obtained by the proposed approach are better than those obtained by the fuzzy approach, because the greater recognition rate is devoted to the group First, while this rate is divided in the possibilistic approach for the two groups "Unique" 61.24% and "First" 30.76%.

In order to realize the comparison in terms of the taken decision quality, we apply the following algorithm:

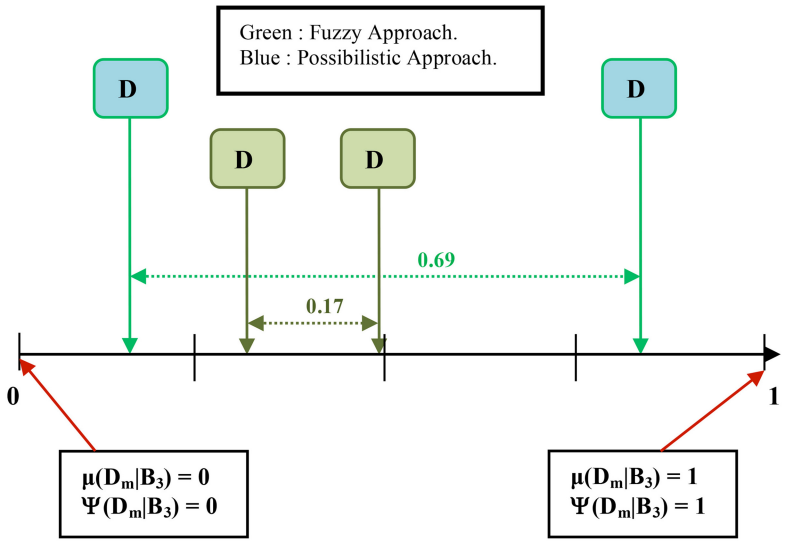
For each approach: *Possibilistic* and *Fuzzy* ***Do***
  From $n = 1$ ***To*** $n = 4450$ ***Do*** (n: means the considered target case)
  1. Calculate the possibilistic couple $[N(\mathbf{D}_m \mid B_n), \Pi(\mathbf{D}_m \mid B_n)]$ for all the diagnoses, $\mathbf{D}_m$, $m = 1, \ldots, M$;
  2. Ranking the set of cases according to the maximum similarity measure;
  3. Identify all cases where the correct diagnosis (i.e., the true diagnosis of considered target case) is the first potential diagnostic obtained by each approach;
  4. Calculate, for each case obtained by the previous step, the distance between the two first potential diagnoses (the true diagnosis and the next diagnosis).
  End
End



(a)



(b)



(c)

Fig. 4. Distance Representation. (a) Case B_1. (b) Case B_2. (c) Case B_3.

TABLE VIII
Comparaison Between the Two Approaches

	Fuzzy Approach	Possibilistic Approach
	$\mu(\mathbf{D}_m \mid \mathbf{B}_f)$	$\Psi(\mathbf{D}_m \mid \mathbf{B}_f)$
Found	100%	100%
Sole	0.2%	61.24%
First	91%	30.76%
Other	8.8%	8%

TABLE IX
Result Obtained by the Possibilistic Approach

	Distance
Superior (Possibilistic > Fuzzy)	3678/4031 = 91.24%
Equal (Possibilistic = Fuzzy)	3/4031 = 0.08%
Lower (Possibilistic < Fuzzy)	129/3207 = 8.68%

After applying the above algorithm, the distances obtained by the possibilistic approach are compared with those obtained by one of the fuzzy approach. Three groups can be distinguished:

—**Superior:** represents the number of cases for which the distance calculated by the possibilistic approach is greater than that calculated by the fuzzy approach.

—**Lower:** represents the number of cases for which the distance calculated by the possibilistic approach is lower than that calculated by the fuzzy approach.

—**Equal:** represents the number of cases for which the distance calculated by the possibilistic approach is equal to that calculated by the fuzzy approach.

We note that the highest rate is dedicated to the group "Superior." This means that the distance characterizing the quality of the solutions obtained by the proposed approach is higher than that obtained by the fuzzy approach.

7. CONCLUSION AND PERSPECTIVES

In this paper, the use of the possibility theory as a global framework is proposed to construct the medical knowledge representation model. This possibilistic model is applied, as a knowledge representation approach, to represent the relationship (Modality)–(Diagnosis), as well as in the construction of the medical knowledge base. Possibilistic reasoning mechanisms are also developed in order to support the case classification by the physician.

This possibilistic representation transforms the expert linguistic knowledge into a model useable by a decision support system. To tackle the case classification issue, the compatibility (based on necessity and possibility measures) has been defined between the target case and different potential diagnoses.

The proposed approach has been applied in the context of Digestive Endoscopic Image Analysis where the medical expert knowledge was successfully modeled with results in full coherence with the expert's expectation.

In this study, we have considered the complete case description (i.e., all features that should have been described by the expert are considered as fulfilled and present). Nevertheless, an important decision making difficulty has not been tackled; this concerns the partial description context where some features, considered by the user (not the expert) as less important, are not filled. This situation makes some the application of decision support systems very difficult, even to the extent of blocking. In the proposed framework, and due to the use of the possibilistic distance, a decision proposition can always be suggested to the user associated with a pertinence value. In a further research work, this pertinence value will be upper and lower bounded allowing thus to improve user confidence in the employed system.

REFERENCES

[1] K. P. Adlassnig
Fuzzy set theory in medical diagnosis.
IEEE Transactions on Systems, Man and Cybernetics, **16**, 2 (1986), 260–265.

[2] K. P. Adlassnig and G. Kolarz
CADIAG-2: Computer-assisted medical diagnosis using fuzzy subsets. M. M. Gupta and E. Sanchez (Eds.), *Approximate Reasoning in Decision Analysis*, North Holland, Amsterdam, 1982, 495–505.

[3] K. Aftarczuk, A. Kozierkiewicz, and N. T. Nguyen
Using representation choice methods for a medical diagnosis problem.
International Conference on Knowledge-Based Intelligent Information and Engineering Systems, Bournemouth, UK, 10 (2006), 805–812.

[4] E. S. Berner
Clinical Decision Support Systems Theory and Practice.
Health Informatics Serie, Springer Science, 2007.

[5] K. Boegl, K. P. Adlassnig, Y. Hayashi, T. Rothenfluh, and H. Leitich
Knowledge acquisition in the fuzzy knowledge representation framework of a medical consultation system.
Artificial Intelligence in Medicine, **30**, 1 (2004), 1–26.

[6] B. Bouchon-Meunier
Théorie des Possibilités et Variables Linguistiques.
Paris: Addison Wesley, 1995, ch. 2.

[7] J. H. Carter
Clinical Decision Support Systems (E. S. Berner, Ed.).
Springer-Verlag, New York, 1999, 169–198.

[8] J. M. Cauvin, C. Le Guillou, et al.
Computer-assisted diagnosis system in digestive endoscopy.
IEEE Transactions on Information Technology in Biomedicine, **7**, 4 (2003), 256–262.

[9] A. Dahabiah, J. Puentes, and B. Solaiman
Digestive database evidential clustering based on possibility theory.
WSEAS Transactions on Systems, **5**, 9 (Sept. 2008), 239–248.

[10] A. Dahabiah, J. Puentes, and B. Solaiman
Possibilistic pattern recognition in a digestive database for mining imperfect data.
WSEAS Transactions on Systems, **8**, 2 (Feb. 2009), 229–240.

[11] A. Dahabiah, J. Puentes, and B. Solaiman
Gastroenterology dataset clustering using possibilistic Kohonen maps.
WSEAS Transactions on Info Science and Applications, **7**, 4 (Apr. 2010), 508–521.

[12] D. Dubois, H. Fargier, et al.
Possibility theory in constraint satisfaction problems: Handling priority, preference and uncertainty.
Applied Intelligence, **6**, 4 (1996), 287–309.

[13] D. Dubois, L. Foulloy, G. Mauris, and H. Prade
Probability-possibility transformations, triangular fuzzy sets, and probabilistic inequalities.
Reliable Computing 10, Kluwer Academic Publishers, printed in the Netherlands, 273–294.

[14] D. Dubois and H. Prade
Possibility theory: an approach to computerized processing of uncertainty.
Gen. Sys. International Journal, *15*, 2 (June 1989), 168–170.

[15] D. Dubois, H. Prade, and S. Sandri
On Possibility/Probability Transformations.
Proc. 4th Int. Fuz. Sys. Assoc. Conf., 1991, 103–112.

[16] L. Hubert
Iterative projection strategies for the least-squares fitting of tree structures to proximity data.
Math. Statist. Psych. Journal, **48**, 2 (1995), 281–317.

[17] T. Kiseliova
A theoretical comparison of disco and CADIAG-II like systems for medical diagnoses.
Kybernetika J., **42**, 6 (2006), 723–748.

[18] G. Kong, D. L. Xu, and J. B. Yang
Clinical decision support systems: A review on knowledge representation and inference under uncertainties.
International Journal of Computational Intelligence Systems, **1**, 2 (2008), 159–167.

[19] C. Le Guillou and J. M. Cauvin
From Endoscopic Imaging and Knowledge to Semantic Formal Images.
Springer, Computer Science, 4370, (2007), 189–201.

[20] R. S. Ledley and L. B. Lusted
Reasoning foundations of medical diagnosis.
Science, **130**, 3366 (1959), 9–21.

[21] H. Leitich, K. P. Adlassnig, and G. Kolarz
Evaluation of two different models of semi-automatic knowledge acquisition for the medical consultant system CADIAGII/RHEUMA.
Artificial Intelligence in Medicine, **25**, 3 (2002), 215–225.

[22] N. H. Phuong and V. Kreinovich
Fuzzy logic and its applications in medicine.
International Journal of Medical Informatics, **62**, 2 (2001), 165–173.

[23] F. Portet and A. Gatt
Towards a Possibility-Theoretic Approach to Uncertainty in Medical Data Interpretation for Text Generation.
D. Riano, A. ten Teije, S. Miksch, and M. Peleg (Eds.), Knowledge Representation for Health-Care: Data, Processes and Guidelines, Berlin and Heidelberg: Springer (LNAI 5943), 2010.

[24] E. Raufaste, R. Da Silvas Neves, and C. Mariné
Testing the descriptive validity of possibility theory in human judgments of uncertainty.
Artificial Intelligence Fuzzy Set and Possibility Theory-Based Methods, *Artificial Intelligence*, **148**, 1–2 (Aug. 2003), 197–218.

[25] R. Salamon, M. Bernadet, and F. Gremy
Ternary Algebra: Theorical Aspects and Possible Applications.
F. T. Dombal and F. Germy (Eds.), Decision Making and Medical Care: Can Information Science Help?, North-Holland, Amesterdam 1976, 505–513.

[26] E. Sanchez
Medical Diagnosis and Composite Fuzzy Relations.
M. M. Gupta, R. K. Ragade, and R. R. Yager (Eds.), Advances in Fuzzy Set Theory and Applications, North-Holland, NY, 1979, 437–444.

[27] R. Seising, C. Schuh, and K. P. Adlassnig
Medical knowledge, fuzzy sets and expert systems.
Workshop on Intelligent and Adaptive Systems in Medicine, Prague 2003. Homepage: http://cyber.felk.cvut.cz/EUNITE 03-BIO/pdf/Seising.pdf.

[28] F. Steimann and K. P. Adlassnig
Fuzzy medical diagnosis.
E. H. Ruspini, P. P. Bonissone, W. Pedrycz (Eds.), Handbook of Fuzzy Computation, IOP Publishing, Bristol, G13,1 (1998), 1–14.

[29] P. Szolovits
Uncertainty and decision in medical informatics.
Methods of Information in Medicine, **34**, 1/2 (1995), 111–121.

[30] M. A. Vila and M. Degado
On medical diagnosis using possibility measures.
Fuzzy Sets and Systems, **10**, 1–3 (1983), 211–222.

[31] P. Walley and G. Decooman
Coherence of rules for defining conditional possibility.
Approx. Reason. International Journal, **21**, 1 (May 1999), 63–107.

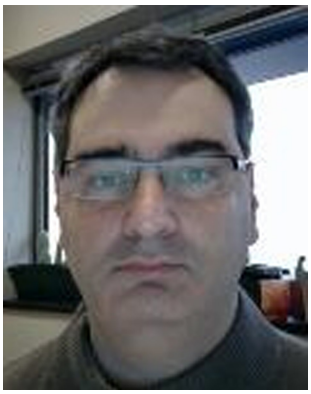
[32] T. Wetter
Medical Decision Support Systems.
Lecture Notes in Computer Science, Frankfurt, Springer-LNCS, 1933, 2000.

[33] L. Zadeh
Fuzzy sets.
Information Control, **8**, 3 (1965), 338–353.

[34] L. Zadeh
The concept of a linguistic variable and its application to approximate reasoning—I.
Information Sciences, **8**, 3 (1975), 199–249.

[35] L. Zadeh
Fuzzy sets as a basis for a theory of possibility.
Fuzzy Sets and Systems, **1**, 1 (1978), 3–28.

Mohammad Homam Alsun (S'07—M'95) received the M.Sc. signal and circuit in and the Ph.D. degree in information and communication sciences and technologies from UBO (University, of Western Brittany), France, in 2007 and 2012, respectively.



Laurent Lecornu (S'90—M'95) received the Diplôme d'Etudes Approfondies and the Ph.D. degree in signal processing from Rennes 1 University, Rennes, France, in 1990 and 1995, respectively.

He is currently an assistant professor at the Image and Information Processing Department, Institut TELECOM, TELECOM Bretagne, Brest, France. His current research interests include image processing and information fusion with medical and environmental applications.

Basel Solaiman (S'84—M'90) received the Telecommun. Eng. degree from the TELECOM Bretagne, Brest, France, and the Diplôme d'Etudes Approfondies and the Ph.D. degree in signal processing from Rennes 1 University, Rennes, France, in 1984 and 1989, respectively.

He is currently a Professor and the Head of the Image and Information Processing Department, Institut TELECOM, TELECOM Bretagne, Brest, France. His current research interests include image processing and information fusion with medical and environmental applications.

Distributed Tracking with a PHD Filter using Efficient Measurement Encoding

BIRUK K. HABTEMARIAM
AMPIKAITHASAN ARAVINTHAN
RATNASINGHAM THARMARASA
KUMARADEVAN PUNITHAKUMAR
THOMAS LANG
THIA KIRUBARAJAN

Probability Hypothesis Density (PHD) filter is a framework for multitarget tracking, which provides estimates for the number of targets as well as the individual target states. Sequential Monte Carlo (SMC) implementation of a PHD filter can be used for nonlinear non-Gaussian problems. However, the application of PHD-based state estimators for a distributed sensor network, where each tracking node runs its own PHD-based state estimator, is more challenging compared with single sensor tracking due to communication limitations. A distributed state estimator should use the available communication resources efficiently in order to avoid the degradation of filter performance. In this paper, a method that efficiently communicates encoded measurements between nodes while maintaining the filter accuracy is proposed. This coding is complicated in the presence of high clutter and instantaneous target births. This problem is mitigated using adaptive quantization and encoding techniques. The performance of the algorithm is quantified using a Posterior Cramér-Rao Lower Bound (PCRLB) that incorporates quantization errors. Simulation studies are performed to demonstrate the effectiveness of the proposed algorithm.

Manuscript received June 21, 2010; released for publication March 21, 2011.

Refereeing of this contribution was handled by Chee-Yee Chong.

Authors' addresses: B. K. Habtemariam, A. Aravinthan, R. Tharmarasa, and T. Kirubarajan, McMaster University, Hamilton, ON, Canada, E-mails: (habtembk@mcmaster.ca, aravinthan@ieee.org, tharman@mail.ece.mcmaster.ca, kiruba@mcmaster.ca); K. Punithakumar, GE Healthcare, London, ON, Canada, E-mail: (Kumaradevan.Punithakumar@ge.com); T. Lang, General Dynamics Canada, Ottawa, ON, Canada, E-mail: (tom.lang@gdcanada.com).

1. INTRODUCTION

The use of a large number of networked sensors, which can be deployed all over the surveillance region, has become feasible in tracking applications because of the availability of cheap sensors. The multisensor data need to be fused in order to fully utilize the information obtained in the network. A common practice in sensor network applications has been to process the collected data in a central processor. This architecture is known as the centralized sensor network [10, 26]. Centralized architectures are generally simpler to execute since the processing of data at one location can reduce the computational requirements of an algorithm. It is theoretically optimal if the network has enough communication bandwidth to send all the sensor data to the fusion node at every sampling time [3].

However, there are several drawbacks associated with the centralized architecture. First, the network relying on one processor to perform the task of every node in the network may result in a single-point failure. Second, in real-time applications, the central node may reside many hops away and sending data from one node to a central node may take too long. This may introduce latency, synchronization problems and imbalanced workload in the network. Further, the centralized architecture may utilize significant resources in communicating the data across the network. Distributed processing over the sensor network can be used to alleviate the problems inherent to the centralized architecture. Further, the distributed architecture requires lighter computational power at each fusion node due to the distribution of processing over multiple nodes.

Distributed algorithms based on particle filters have gained much attention. In [5], methods based on likelihood factorization of particles and adaptive data-encoding scheme are proposed for nonlinear/non-Gaussian systems with distributed architecture. An improvement to the approach proposed in [5] has been presented in [13] using a better encoding scheme and measurement vectorization. More particle-based implementations are given in [18], [22]. The adaptive data-encoding scheme uses the histogram of expected measurements to encode the target-generated measurements effectively. However, the false measurements might end up transmitting a larger number of bits than transmitting measurements without encoding. Hence, the effectiveness of the encoding scheme might degrade dramatically if no method is in place to identify and remove false measurements before transmitting over the network. Also, target birth must be taken care of while removing the false alarms in order to handle the time-varying number of targets.

The primary focus of this paper is on creating distributed algorithms that minimize network communication relating to sensor data fusion when multiple time-varying number of targets are present in the monitored

area. In this paper, a decentralized version of the Probability Hypothesis Density (PHD) filter is used to track multiple targets. The PHD filter eliminates the hard measurement-to-track association problem, unlike the Multiple Hypothesis Tracker (MHT) [4]. Furthermore, the PHD filter has been shown an effective way of tracking time-varying multiple number of targets that avoids model-data association problems [15]. Gaussian mixture implementation of PHD filter (GM-PHD) is presented in [25]. Sequential Monte Carlo (SMC) implementation of the PHD filter is used to handle the nonlinear measurements [24]. There are two options available to perform distributed tracking with a SMC-PHD filter in a sensor network. The first option is to send all the particles that represent the posterior density of targets. However, this option requires high bandwidth communications, which can not be handled by practical wireless sensor networks. The second option is to send the most relevant measurements after eliminating the false alarms to update the global estimates of the targets.

In this paper, measurements are communicated among nodes to update the filters. In this case, data transmission requires higher bandwidth channels unless the quantization of those data are done intelligently [16, 19]. To be effective, non-uniform quantization schemes can be made to match the distribution of the quantity to be discretized. Companding is a widely used method for implementing non-uniform quantizers [17]. It has been observed in non-uniform quantization that the communication can be considerably reduced with the right selection of the compander [16]. Quantized measurements need to be encoded before transmission. It is assumed that an optimal noiseless source code will be employed to minimize transmission needs between nodes. In this paper, Huffman coding is used to encode the quantized measurements. Handling multiple target-originated measurements at the quantization stage and producing identical symbols for encoding and decoding at each node are challenging. This paper proposes "cascaded companders" to nonlinearly quantize multiple target measurements. Predicted probability density is used in generating identical set of symbols and to place the companders at right positions. The measurement quantization and encoding techniques proposed this paper can be applied to distributed tracking with GM-PHD and other PHD filter realization algorithms as well.

Among the various methods to quantify the performance, verifying the closeness of the estimates mean square error matrix to the lower bound is a commonly known method in target tracking applications. The Posterior Cramer-Rao Lower Bound (PCRLB) is defined to be inverse of the Fisher Information Matrix (FIM) for random vector and provides lower bound on the performance of unbiased estimators of the unknown target state [23]. The PCRLB for state estimation with quantized measurements is complicated due to nonlinearity of the quantizer. Previously, in [28] the PCRLB for dynamic target tracking with measurement origin uncertainty and in [8] the PCRLB for state estimation with quantized measurement were developed. In this paper, the PCRLB calculation with quantized measurement is extended to incorporate measurement origin uncertainty for bearing only tracking.

This paper is structured as follows. Section 2 explains the proposed distributed implementation of SMC-PHD filter. Quantization and encoding methods are explaiined in Section 3. Section 4 provides the derivation of the PCRLB with quantized measurements and measurement origin uncertainty. Simulations that demonstrate the effectiveness of the proposed quantization strategy are presented in Section 5. Conclusions are given in Section 6.

2. DISTRIBUTED TRACKING USING SMC-PHD FILTER

2.1. State and Measurement Models

In this paper, the problem of tracking a time-varying number of multiple targets is considered. The general parameterized target dynamics is given by

$$x_{k+1} = F_k x_k + \nu_k \tag{1}$$

where x_k denotes the target state, F_k is a known matrix and ν_k is the process noise at time k.

The measurements originate from either targets or clutter. The target-originated measurement is given by

$$z_k = h_k(x_k) + \omega_k \tag{2}$$

where h_k is a nonlinear function and ω_k is the measurement noise at time k. For simplicity it is assumed that ν_k and ω_k are Gaussian with zero means and covariances Γ_k and Σ_k, respectively.

It is assumed that the number of false alarms is Poisson-distributed with the average rate of λ_k and that the probability density of the spatial distribution of false alarms is $c_k(\mathbf{z}_k)$.

2.2. PHD Filter

In tracking multiple targets, if the number of targets is unknown and varying with time, it is not possible to compare states with different dimensions using ordinary Bayesian statistics of fixed dimensional spaces. However, the problem can be addressed by using Finite Set Statistics (FISST) [15] to incorporate comparisons of state spaces of different dimensions. FISST facilitates the construction of multitarget densities from multiple-target transition functions by computing set derivatives of belief-mass functions [15], which makes it possible to combine states of different dimensions. The main practical difficulty with this approach is that the dimension of the full state space becomes large when many targets are present, which increases the computational load exponentially in the number of targets. Since the PHD is defined over the state space of one target in contrast

to the full posterior distribution, which is defined over the state space of all the targets, the computational cost of propagating the PHD over time is much lower than propagating the full posterior density.

In general, a PHD-based multitarget tracker will experience more difficulty in resolving closely-spaced targets than a tracker based on the full target posterior. However, if the probability density functions of individual targets is highly concentrated around their means compared to the target separation, such that the individual target pdfs do not overlap significantly, it will become possible to resolve the targets using the PHD filter as well. A theoretical explanation on the capability of the PHD filter to resolve closely-spaced targets in Gaussian context is given in [15]. By definition, the PHD $D_{k|k}(x_k \mid Z_{1:k})$, with single target state vector x_k, and given all the measurements up to and time step k, is the density whose integral on any region S of the state space is the expected number of targets $N_{k|k}$ contains in S. That is,

$$N_{k|k} = \int_X D_{k|}(x_k \mid Z_{1:k}) dx_k. \tag{3}$$

This property uniquely characterizes the PHD and the first-order statistical moment of the full target posterior distribution possesses this property. Hence, the first-order statistical moment of the full target posterior, or the PHD, given all the measurement $Z_{1:k}$ up to time step k, is given by the set integral [14]

$$D_{k|k}(x_k \mid Z_{1:k}) = \int f_{k|k}(\{x_k\} \cup Y \mid Z_{1:k}) \delta(Y). \tag{4}$$

More detailed mathematical explanations and derivation of the PHD filter can be found in [14]. The approximate expected target states are given by the local maxima of the PHD. The prediction and update steps of one cycle of PHD filter are given in the following section.

2.2.1. Prediction

In a general scenario of interest, there are target disappearances, target spawning and entry of new targets. The probability that a target with state x_{k-1} at time step $(k-1)$ will survive at time step k is denoted by $e_{k|k-1}(x_{k-1})$, the PHD of spawned targets at time step k from a target with state x_{k-1} by $b_{k|k-1}(x_k \mid x_{k-1})$, and the PHD of newborn spontaneous targets at time step k by $\gamma_k(x_k)$. Then, the predicted PHD, $D_{k|k-1}(x_k \mid Z_{1:k-1})$, at time k given all measurements up to time $k-1$ is given by

$$\begin{aligned} &D_{k|k-1}(x_k \mid Z_{1:k-1}) \\ &\quad = \gamma_k(x_k) + \int [e_{k|k-1}(x_{k-1}) f_{k|k-1}(x_k \mid x_{k-1}) + b_{k|k-1}(x_k \mid x_{k-1})] \\ &\qquad \times D_{k-1|k-1}(x_{k-1} \mid Z_{1:k-1}) dx_{k-1} \end{aligned} \tag{5}$$

where $f_{k|k-1}(x_k \mid x_{k-1})$ denotes the single-target Markov transition density. The prediction equation (5) is lossless since there are no approximations.

2.2.2. Update

The predicted PHD can be corrected with the availability of measurements Z_k at time step k to get the updated PHD. It is assumed that the number of false alarms is Poisson-distributed with the average rate of λ_k and that the probability density of the spatial distribution of false alarms is $c_k(z_k)$. Let the detection probability of a target with state x_k at time step k be $p_D(x_k)$. Then, the updated PHD at time step k is given by

$$\begin{aligned} &D_{k|k}(x_k \mid Z_{1:k}) \\ &\quad \cong \left[\sum_{z_k \in Z_k} \frac{p_D(x_k) f_{k|k}(z_k \mid x_k)}{\lambda_k c_k(z_k) + \psi_k(z_k \mid Z_{1:k-1})} + (1 - p_D(x_k)) \right] \\ &\qquad \times D_{k|k-1}(x_k \mid Z_{1:k-1}) \end{aligned} \tag{6}$$

where the likelihood function $\psi(\cdot)$ is given by

$$\begin{aligned} &\psi_k(z_k \mid Z_{1:k-1}) \\ &\quad = \int p_D(x_k) f_{k|k}(z_k \mid x_k) D_{k|k-1}(x_k \mid Z_{1:k-1}) dx_k \end{aligned} \tag{7}$$

and $f_{k|k}(z_k \mid x_k)$ denotes the single-sensor/single-target likelihood. The update equation (6) is not lossless since approximations are made on predicted multitarget posterior to obtain a closed-form solution. The reader is referred to [14] for further explanations.

2.3. Sequential Monte Carlo PHD Filter

This section describes the SMC approach to the PHD filter [24]. This approach provides a mechanism to represent the posterior probability hypothesis density by a set of random samples or particles, which consist of state information with associated weights, to approximate the PHD. The advantage of this method is that the number of particles can be adaptively allocated such that a constant ratio between the number of particles and the expected number of targets is maintained. This has a significant effect on the computational complexity of the algorithm. The complexity does not increase exponentially, but only linearly with the increasing number of targets. The SMC implementation considered here is structurally similar to the Sampling Importance Resampling (SIR) type of particle filter [2]. Let the posterior PHD $D_{k-1|k-1}(x_{k-1} \mid Z_{1:k-1})$ be represented by a set of particles $\{w_{k-1}^{(p)}, x_{k-1}^{(p)}\}_{p=1}^{L_{k-1}}$. That is,

$$D_{k-1|k-1}(\mathbf{x}_{k-1} \mid Z_{1:k-1}) = \sum_{p=1}^{L_{k-1}} w_{k-1}^{(p)} \delta(\mathbf{x}_{k-1} - \mathbf{x}_{k-1}^{(p)}) \tag{8}$$

where $\delta(\cdot)$ is the Dirac Delta function. In contrast to particle filters, the total weight of the particles $\sum_{p=1}^{L_{k-1}} w_{k-1}^{(p)}$ is not equal to one; instead, total weight gives the expected

number of targets n_{k-1}^X at time step $(k-1)$, which follows from the property that the integral of the PHD over the state space gives the expected number of targets.

2.3.1. Prediction

Importance sampling is applied to generate state samples that approximate the predicted PHD $D_{k|k-1}(x_k \mid Z_{1:k-1})$. State samples $\{x_{k|k-1}^{(s)}\}_{p=1}^{L_{k-1}}$ are generated from the proposal density $q_k(\cdot \mid x_{k-1}, Z_k)$ and i.i.d. state samples $\{x_{k|k-1}^{(p)}\}_{p=L_{k-1}+1}^{L_{k-1}+J_k}$ corresponding to new spontaneously born targets from another proposal density $p_k(\cdot \mid Z_k)$. That is,

$$x_{k|k-1}^{(s)} \sim \begin{cases} q_k(\cdot \mid x_{k-1}, Z_k) & p = 1,\ldots,L_{k-1} \\ p_k(\cdot \mid Z_k) & p = L_{k-1}+1,\ldots,L_{k-1}+J_k \end{cases}. \tag{9}$$

Then, the weighted approximation of the predicted PHD is given by

$$D_{k|k-1}(x_k \mid Z_{1:k-1}) = \sum_{p=1}^{L_{k-1}+J_k} w_{k|k-1}^{(p)} \delta(x_k - x_{k|k-1}^{(p)}) \tag{10}$$

where

$$w_{k|k-1}^{(p)} = \begin{cases} \dfrac{e_{k|k-1}(\mathbf{x}_{k|k-1}^{(p)}) f_{k|k-1}(\mathbf{x}_{k|k-1}^{(p)} \mid \mathbf{x}_{k-1}^{(p)}) + b_{k|k-1}(\mathbf{x}_{k|k-1}^{(p)} \mid \mathbf{x}_{k-1}^{(p)})}{q_k(\mathbf{x}_{k|k-1}^{(p)} \mid \mathbf{x}_{k-1}^{(p)}, Z_k)} w_{k-1}^{(s)} & p = 1,\ldots,L_{k-1} \\ \dfrac{\gamma_k(\mathbf{x}_{k|k-1}^{(p)})}{p_k(\mathbf{x}_{k|k-1}^{(p)} \mid Z_k)} \dfrac{1}{J_k} & p = L_{k-1}+1,\ldots,L_{k-1}+J_k \end{cases}. \tag{11}$$

The functions that characterize the Markov target transition density $f_{k|k-1}(.)$, target spawning $b_{k|k-1}$ and entry of new targets $\gamma_k(\cdot)$ in (11) are conditioned on the target motion model.

2.3.2. Update

With the available set of measurements Z_k at time step k, the updated particle weights can be calculated by

$$w_k^{*(p)} = \left[(1 - p_D(\mathbf{x}_{k|k-1}^{(p)})) + \sum_{i=1}^{N_k^Z} \frac{p_D(\mathbf{x}_{k|k-1}^{(p)}) f_{k|k}(\mathbf{z}_k^i \mid \mathbf{x}_{k|k-1}^{(p)})}{\lambda_k c_k(\mathbf{z}_k^i) + \Psi_k(\mathbf{z}_k^i)} \right] w_{k|k-1}^{(p)} \tag{12}$$

where

$$\Psi_k(z_k^i) = \sum_{p=1}^{L_{k-1}+J_k} p_D(\mathbf{x}_{k|k-1}^{(p)}) f_{k|k}(\mathbf{z}_k^i \mid \mathbf{x}_{k|k-1}^{(s)}), w_{k|k-1}^{(p)} \tag{13}$$

and $f_{k|k}(\cdot)$ is the single-target/single-sensor measurement likelihood function.

2.3.3. Resample

To perform resampling, since the weights are not normalized to unity in PHD filters, the expected number of targets is calculated by summing up the total weights, i.e.,

$$\hat{n}_k^X = \sum_{p=1}^{L_{k-1}+J_k} w_k^{*(p)}. \tag{14}$$

Then the updated particle set $\{w_k^{*(p)}/n_k^X, \mathbf{x}_{k|k-1}^{(p)}\}_{p=1}^{L_{k-1}+J_k}$ is resampled to get $\{w_k^{(p)}/n_k^X, \mathbf{x}_k^{(p)}\}_{p=1}^{L_k}$ such that the total weight after resampling remains n_k^X. Now, the discrete approximation of the updated posterior PHD at time step k is given by

$$D_{k|k}(\mathbf{x}_k \mid Z_{1:k}) = \sum_{p=1}^{L_k} w_k^{(p)} \delta(\mathbf{x}_k - \mathbf{x}_k^{(p)}). \tag{15}$$

2.4. Distributed Architecture

Distributed processing over the sensor network can be used to alleviate the problem inherent to centralized architectures. A sample distributed architecture is shown in Fig. 1, where S indicates the sensor. The underlying sensor network architecture consists of two different types of devices: sensors and nodes. Sensors collect measurements from the targets and report them to computational nodes. Nodes are responsible for running filters to track targets. Information gathered at one node are shared among various nodes. The efficient utilization of communication resources without compromising accuracy is essential.

2.5. Distributed Tracking Algorithm

The objective in this paper is to develop a distributed algorithm based on the SMC-PHD filter while minimizing the communication requirements of the distributed network in the presence of multiple time-varying number of targets and false alarms. It is assumed that the optimization of sensor resources to collect data and communication issues such as network protocols are already efficient enough. The proposed algorithm maintains SMC-PHD filters at all the computational nodes.

There are a number of different options to perform distributed tracking with an SMC-PHD filter in a sensor network. One option is to send all the particles that represent the posterior density of target states. Another is to send Gaussian mixture representation of the posterior density. These two options require high bandwidth communications, which cannot be handled by practical wireless sensor networks. The third option is to send only most relevant measurements after eliminating the

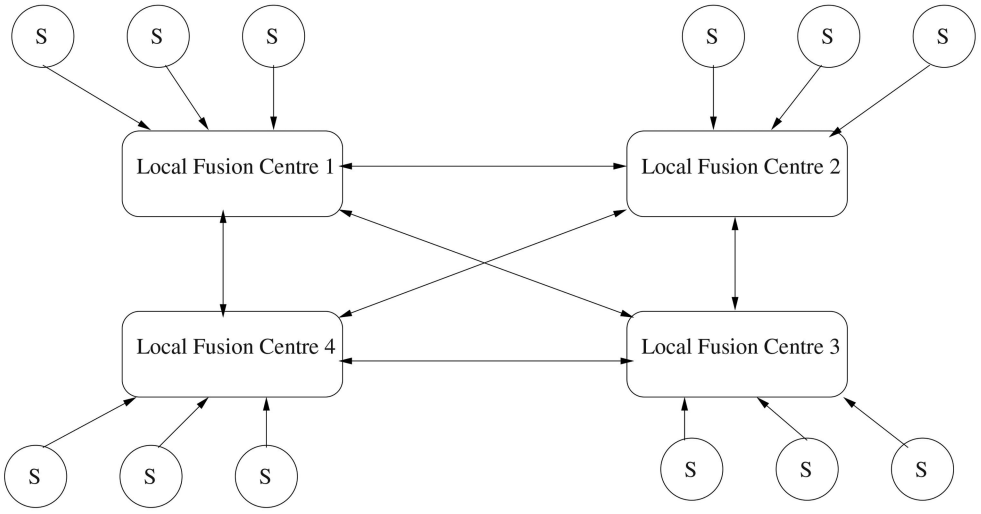


Fig. 1. A sample distributed architecture.

false alarms to update the global estimates of the targets. In this paper, the option of communicating the relevant measurements among nodes to update the filters is used. In a sensor network, it is possible that each node has enough active sensors to track an object by itself with reasonable tracking accuracy. Therefore, a PHD filter can be used to obtain the estimates based on the measurements collected from sensors local to that node. Since these nodes maintain PHD filters based on local measurements, they can also be used in the encoding strategy. The proposed framework will be performed in two layers. The first layer collects measurement data that are local to each node and maintains a local PHD filter using its associated sensors. In the second layer, all measurements are exchanged to all other nodes in the network and the global PHD filters are maintained.

In the proposed algorithm, identical copies of the SMC-PHD filter are maintained at each node. Initially, this is achieved by initializing filters using the same random seed. In order to encode the measurement data, an intelligent quantization and encoding strategy is used. From time step $k-1$ to k, particles are propagated while taking into account the measurement prediction covariance. The range of expected measurements is divided into bins depending on the required accuracy level. The contribution of each propagated particle's distribution is integrated over the bins to form the probability density. The measurements are quantized with a non-uniform quantizer where companders are used to perform non-uniform quantization. The probability density in the measurement space is then transformed to the companded measurement space. Then, the quantized measurements are encoded using Huffman encoding algorithm with the transformed bin probabilities. The encoded measurements are transmitted to all other nodes where each node decodes and decompands the data to obtain the quantized measurements. The details of quantization and encoding strategy used in this algorithm is presented in Section 3.

Each node performs filtering using quantized measurements to obtain the target state estimates. All nodes use the same set of measurement data to update the filter, thereby maintaining the identical copy of filter.

The steps of the distributed SMC-PHD filter are given below.

1) Initialization at $k = 0$:
   - Initialize SMC PHD filter on each node $n = 1,\ldots,N$ using the same random seed to generate identical particle distribution on all the nodes.
   - For each node $n = 1,\ldots,N$
     — Generate samples $\{\mathbf{x}_0^{(p)}\}_{p=0}^{L_0}$
2) Quantization and encoding (For implementation details of this step the reader is referred to Section 3):
   - Local Estimation
     — Perform filtering using the SMC PHD filter acting only on the measurements local to the node.
   - Quantization
     — For each node $n = 1,\ldots,N$
       * For $s = 1,\ldots,L_{k-1}$, predict $\mathbf{x}_{k|k-1}^{(p)}$
       * Calculate the bin probabilities, $p(z_k \mid b_j, \mathbf{z}_{1:k-1}^{(p)})$, in the measurement space using predicted measurements and construct the probability density where b_j is the jth measurement bin.
       * Identify the regions where the companders need to be placed and the number of companders needed. One compander per target is used and the width of the companding region is limited to $3\sigma_p^c$, where the σ_p^c is the standard deviation of the cth cluster. The compander is placed on the mean value, μ_p^c, of the cluster. In other regions linear quantizer is used.
       * Quantize the measurements, $\tilde{\mathbf{z}}_k$

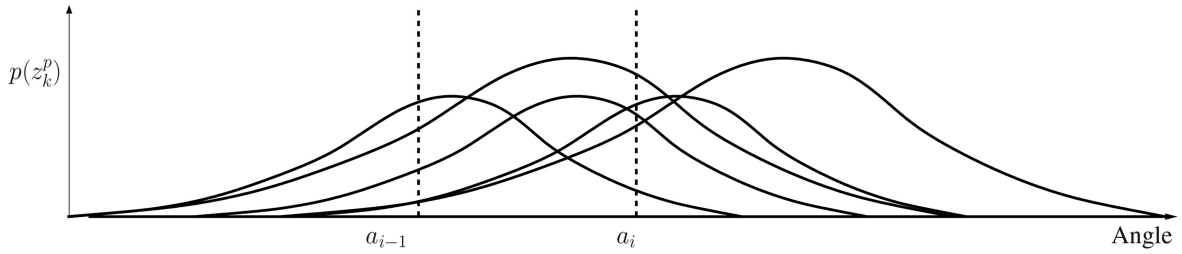


Fig. 2. Calculation of bin probabilities.

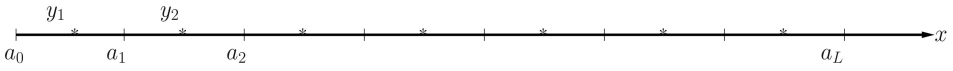


Fig. 3. Quantization.

- Encoding
  - — For each node $n = 1,\ldots,N$
    - * Calculate the bin probabilities, $\tilde{p}(z_k \mid b_j, \mathbf{z}_{1:k-1}^{(p)})$, in the transformed measurement space.
    - * Use the bin probabilities to form Huffman tree H_f^{k-1} and encode quantized measurements.

3) Reducing the false measurements transmitted over the network:
   - Remove the measurements from the queue if the number of bits in each encoded measurement exceeds a predefined threshold, l. This process is done using the local estimates of the target.
4) Global estimate:
   - For each node $n = 1,\ldots,N$, create the Huffman tree H_f^{k-1} and the quantizer to reconstruct the quantized data, $\tilde{\mathbf{z}}_k'$.
   - Using the obtained set of measurements, perform filtering to obtain the global state estimates.

3. QUANTIZATION AND ENCODING

Measurements reported by sensors in a sensor network need to be transmitted in order to perform tracking at high computational nodes called fusion centers. Quantization and encoding play a crucial role whereby measurements are quantized and encoded before being transmitted. Intelligent quantization and encoding schemes are necessary to effectively use the communication resources. This section explains how quantization and encoding can be effectively implemented to perform distributed target tracking with SMC-PHD filters.

The proposed algorithm needs an efficient nonlinear non-uniform quantization for measurements. Therefore, the concept of "cascaded companders," which can quantize measurements from multiple targets, is proposed. This section briefly explains the process of developing the compander. The first step is to construct a probability density of expected measurements to identify the regions where the target originated measurements would lie. The details of this process are given in Section 3.01. Measurements that fall in this region are quantized with minimum quantization error via Gaussian companders. Section 3.13 explains the cascaded companders. Details of encoding and decoding process using Huffman coding are given in Section 3.23. Sections 3.3 and 3.14 provide details on the false alarm elimination process and the incorporation of quantization errors into tracking, respectively.

3.0.1. Construction of a Probability Density

The necessity to have identical and accurate probability densities of targets at each node, where global SMC-PHD filter is running, is clear from the fact that the measurements are quantized, encoded and communicated across these nodes based on the probability density. The construction of probability density begins with propagating the densities of particles from time step $k-1$ to k, taking into account the measurement prediction covariance. The range of expected measurements is divided into bins depending on the required accuracy level. The contribution of each propagated particle's distribution is integrated over the bins to form the probability density. Figure 2 shows the distribution of three sample particles and the quantizer decision boundaries a_{i-1} and a_i. The probability density of predicted particles $p(z_k^s)$ in the measurement space is given by

$$p(z_k^p) = \mathcal{N}(z_k^s; h_k(x_{k|k-1}^p), S_k) \tag{16}$$

where $h_k(.)$ is a nonlinear function and S_k is the measurement prediction covariance. Then the bin probability is given by

$$p(z_k \mid b_j, \mathbf{z}_{1:k-1}^{(p)}) = \sum_{s=1}^{L_{k-1}} \int_{a_{i-1}}^{a_i} p(z_k^s) dz. \tag{17}$$

3.1. Quantization

One dimensional quantizer Q with L levels may be defined by a set of $L+1$ decision levels $a_0, a_1, \ldots, a_L$ and a set of L output levels $y_1, y_2, \ldots, y_L$, as shown in Fig. 3. When a sample x, the quantity to be quantized, lies in the ith quantizer interval $s_i = a_{i-1} < x \leq a_i$ the quantizer produces the output value $Q(x) = y_i$ [9]. The value of y_i is usually chosen to lie within the interval

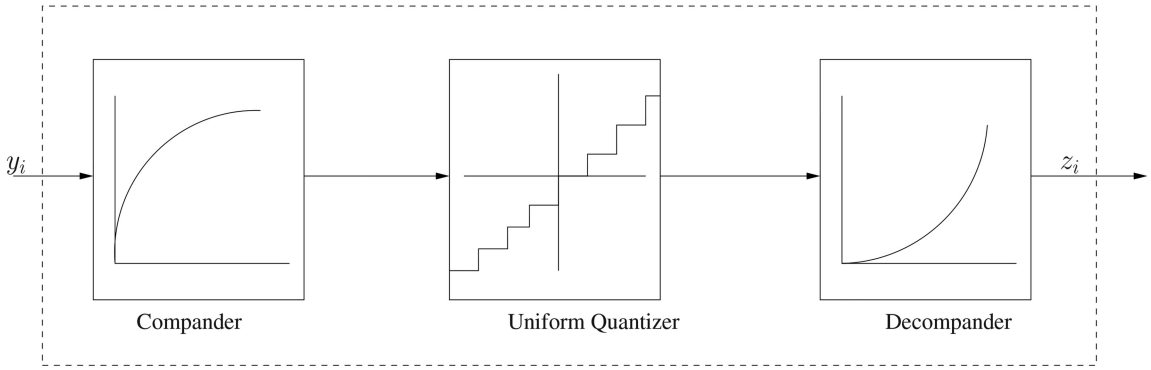


Fig. 4. Nonuniform quantization.

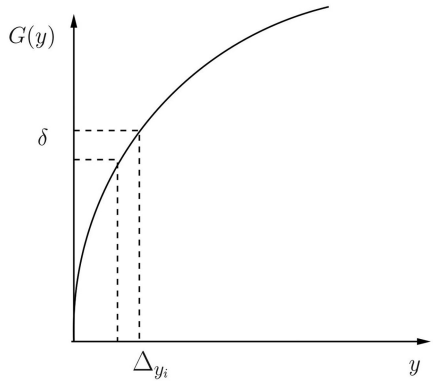


Fig. 5. A typical compander.

s_i. The end levels a_0 and a_L are generally chosen to be the smallest and largest values the input samples may obtain. The L output levels generally have a finite value and if $L = 2^n$, a unique n-bit binary word can identify a particular output level. The input-output characteristics of a one-dimensional quantizer resemble a staircase. The quantizer intervals, or steps, may vary in size.

Uniform and non-uniform quantizer strategies are investigated in this paper.

3.1.1. Uniform Quantization

Uniform quantizer is where the measurement space is divided into equal bins based on the number of bits used to encode. The output points are located at the midpoint of these intervals. If the step size is denoted by $\triangle$, then the maximum absolute error is given by $\triangle/2$. In general, uniform quantization is not the most effective way to obtain good quantizer performance [9].

3.1.2. Non-uniform Quantization

The non-uniform quantization essentially has a non-uniform spacing of decision levels based upon the input probability density [16]. The general model used to represent the non-uniform quantizer is shown in Fig. 4. The combined function of compression, quantization and expansion is termed companding [17]. The quantized samples are transmitted over the network while at the receiver end of the network the quantized samples are decompanded to its original values plus the quantization noise. The variance of the quantization noise associated with the received samples is related to the shape of the companding function $G(.)$ and the number of the bits, n, used for quantization. A typical companding function is shown in Fig. 5. With reference to the figure,

$$G(y + \Delta_y) - G(y) = \delta \tag{18}$$

in which the right hand side is the resolution of the uniform quantizer. Using standard companding techniques, Δ_y can be given as

$$\Delta_y \approx \frac{\delta}{\dot{G}(y)} \tag{19}$$

where $\dot{G}$ denotes differentiation of G.

3.1.3. Measurement Quantization with Cascaded Companders

The non-uniform quantization is performed based on probability density of the targets. Figures 6 and 7 show quantizers at two different time steps, when one and two targets are present in the environment, respectively. The companders are placed in the measurement space such that the target-originated measurements have less quantization errors than other measurements. In this paper, a Gaussian compander law, which is centered on the expected target position and whose curvature is dictated by the standard deviation of the expected position [16], is used. The compander and expander functions are as follows:

- Compander: $\text{erf}(\xi/\sigma\sqrt{6})$
- Expander: $\sigma\sqrt{6}\text{erf}(\xi)$

where $\text{erf}(\xi) = 2/\sqrt{pi}\int_0^{\xi}\exp(-t^2)dt$. One compander per target is used and the width of the companding region is limited to $3\sigma_p^c$, where σ_p^c is the standard deviation of the cth cluster. The compander is placed on the mean value, μ_p^c, of the cluster. A maximum quantization error is set in other regions of the measurement space, where the compander is not placed, by a liner quantizer. The companders are cascaded when multiple targets measurements are to be quantized.

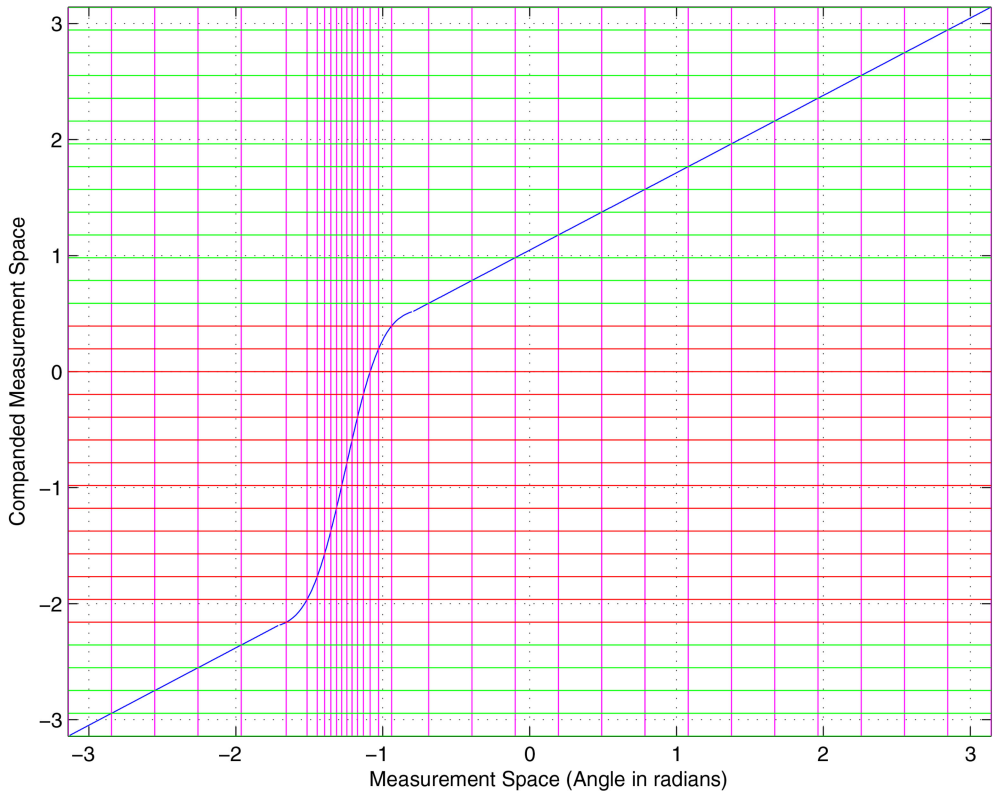


Fig. 6. A 32-bin compander with one target.

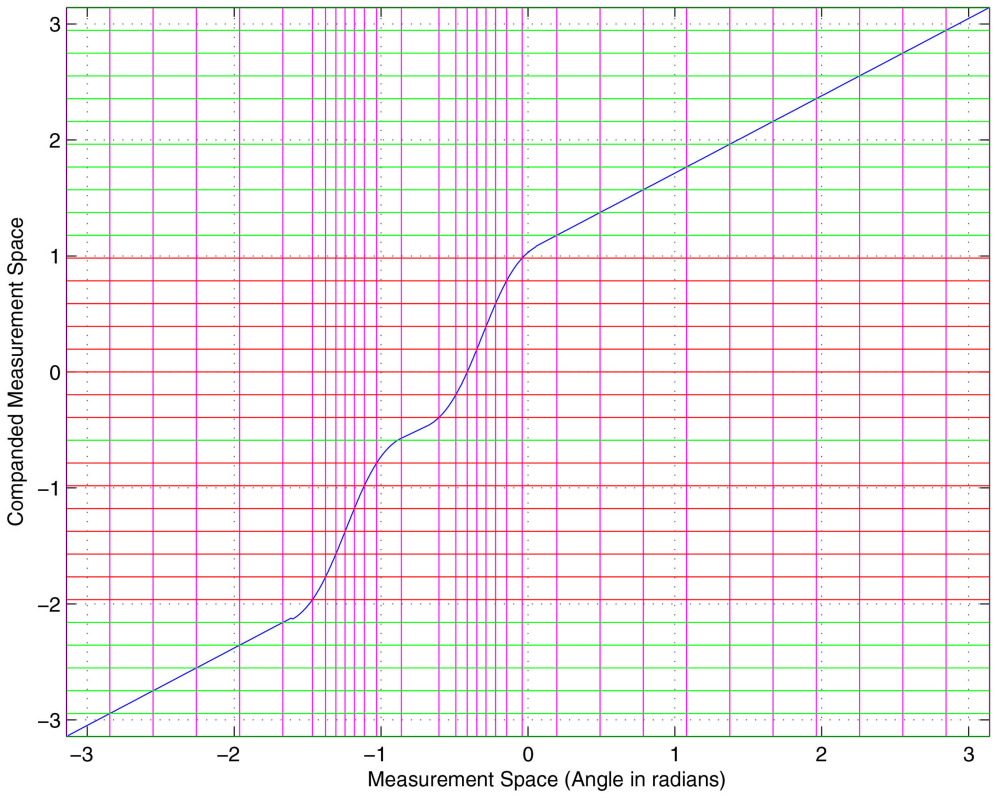


Fig. 7. A 32-bin compander with two targets.

3.1.4. Incorporating Quantization Errors

The insertion of quantized measurement to the SMC-PHD filter is done by updating the current particles by the quantized measurements while taking into account the extra error introduced by the quantization. The error arising from quantization has a uniform distribution.

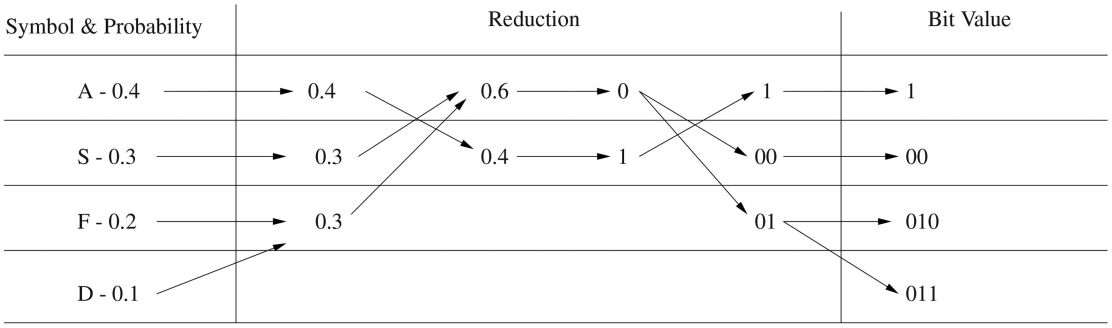


Fig. 8. Construction of Huffman Encoding Table.

The variances of errors introduced due to quantization is given by

- Uniform quantization

$$\mathrm{Var}(z_k^i \mid x_k) = \sigma_w^2 + \frac{\delta^2}{12}. \tag{20}$$

- Non-uniform quantization

$$\mathrm{Var}(z_k^i \mid x_k) = \sigma_w^2 + \frac{\delta^2}{12\dot{G}(y_k^i)^2}. \tag{21}$$

3.2. Encoding

In information theory, an entropy coding is a lossless data compression scheme that is independent of the specific characteristics of the medium. A common method of entropy coding defines a codebook by assigning a code to each symbol. By assigning smaller codes to the more frequent symbols, the average size of each coded symbol can be minimized. This leads to compression over sufficiently large number of encoded symbols. This technique is known as variable length coding. Generally, variable length coding shows a better performance than fixed-length codes where same size is assigned to all symbols [20].

Two widely used entropy coding techniques are Huffman coding [12, 6] and arithmetic coding [27]. Huffman coding is simple to implement and is efficient when the probabilities of symbols to be sent can be calculated in advance. Hence it is best suited for application in this paper.

Encoding will help reduce the communication load only for uniform quantization. In non-uniform quantization, the probability of getting measurement at each measurement bin is almost equal. As a result, it is not possible to achieve communication reduction by encoding for non-uniform quantization.

3.2.1. Huffman Coding

Huffman coding assigns a variable length code to each input symbol where the code and its size are based on the probability of occurrence of the associated symbol. It is necessary to calculate probability of symbols before the assignment and construction of a dictionary. By sorting and analyzing the probability of symbols, a conversion table is constructed so that the symbols with higher probability have the fewer number of bits and no symbol is a prefix to another symbol [20]. Greater compression can be achieved with the accurate estimation of probability distribution.

3.2.2. Building Huffman Codes

The construction of Huffman encoding table is a lengthy process. The probabilities must be sorted so that the two lowest probabilities can be found. These probabilities are added together to create a new probability table. This table is sorted, and the process is repeated until only two probabilities are left. These probabilities are assigned a value of zero and one. The process is now reversed. At each stage the two expanded probabilities are given a one or zero as they are expanded. The process continues until the table is expanded to its original state. For example, assume that the message "ASAFAFDAS" is being encoded. The first step is to find the probability for each symbol. "A" has a probability of 0.4, while S has 0.3, D has 0.1 and F has 0.2. These probabilities are sorted and added to create the table as in Fig. 8. Once the table is constructed, the data can be compressed. The compression process is accomplished by a direct conversion of symbols. The entire message is encoded as "10010101010011100," which requires 17 bits. The unencoded message would normally require 18 bits.

3.2.3. Measurement Encoding and Decoding

The original probability density constructed based on expected measurements is transformed to companded measurement space in order to create a global Huffman dictionary for encoding. The term global refers to the process or information that is related to global SMC-PHD filter running on every node. Companded measurements are encoded and transmitted over the network. In the receiver, measurements are decoded before expanding. The same steps are followed to construct a decoding dictionary.

3.3. False Alarm Elimination

Reducing the number of false measurements communicated over the network is important as they consume most of the communication resources. The number of bits in each encoded measurement, based on the

local Huffman dictionary, can effectively be used to reduce the number of false measurements transmitted over the sensor network. In this approach, it is assumed that since the local PHD filters have the most up-to-date information including the birth of a new target and the target generated measurements are most likely to be in a region in which the value of the probability is high. Thus the target-generated measurements are most likely to have a lesser number of bits in their encoded form compared to false measurements when encoded with local Huffman dictionary. It is reasonable to assume that the measurements that have a higher number of bits are not target generated and, by having a threshold value on the number of bits, they can be removed from the set of measurements that are transmitted over the network. Once the measurements are selected to be transmitted, those measurements are encoded with the global Huffman dictionary in order to transmit over sensor network. However, when measurements corresponding to new targets are encoded with global Huffman dictionary may produce higher number of bits. It could be noted that the new targets can be identified by the global PHD filter quickly. An indicator function, $\mathcal{I}_h^{(k,i)}$ is used to identify whether the measurement has been communicated or not.

$$\mathcal{I}_h^{(k,i)} = \begin{cases} 1 & H_f^{k-1}(\tilde{\mathbf{z}}_k^i) \le l \\ 0 & H_f^{k-1}(\tilde{\mathbf{z}}_k^i) > l \end{cases} \tag{22}$$

$H_f^{k-1}(\tilde{\mathbf{z}}_k^i)$ is a function that generates Huffman codes for each measurement. l is the cutoff number of bits per measurement. If a measurement in its encoded form is less than the cutoff number of bits, then the measurement is communicated and not otherwise.

4. POSTERIOR CRAMER-RAO LOWER BOUND

In this section, the recursive Riccati-like formula for the PCRLB is derived for state estimation using measurements with quantization and origin uncertainty. The Section 4.1 provides a brief review on PCRLB. Incorporating the measurement origin uncertainty in PCRLB is discussed in Section 4.2. In Section 4.3 the PCRLB with quantized measurements is derived.

4.1. Background

Consider the estimation of the state of a dynamical system given by (1) and (2). The quantized measurements at time k are denoted by $\tilde{\mathbf{z}}_k$. Let $\hat{x}_{k|k}$ denote the updated state estimate at time instant k, using measurement $\tilde{\mathbf{z}}_{1:k}$. The estimation error covariance matrix, $P_{k|k}$, for unbiased estimator is bounded as follows:

$$P_{k|k} = E[(x_k - \hat{x}_{k|k})(x_k - \hat{x}_{k|k})^T] \ge J_k^{-1} \tag{23}$$

where J_k is the Fisher information matrix, which is the inverse of PCRLB.

For linear Gaussian systems, Riccati-like recursion is given by [11]

$$J_{k+1} = (Q_k + F_k J_k^{-1} F_k^T)^{-1} + \underbrace{E[-\Delta_{x_{k+1}}^{x_{k+1}} \log p(\tilde{\mathbf{z}}_{k+1} \mid x_{k+1})]}_{J_{zk+1}} \tag{24}$$

with $J_0^{-1} = P_0$.

4.2. Effect of Measurement Origin Uncertainty

Consider n_s (≥ 1) sensors, and let $\tilde{\mathbf{z}}_k^s$ be the quantized measurement vector from sensor s. It is assumed that the measurement noises of sensors are independent. Also, due to false alarms, the total number of measurements can vary among sensors at each time step. Let m_k^s be the total number of measurements from sensor s at time k. Let the observation set at time k from sensor s be

$$\tilde{\mathbf{z}}_k^s = \{\tilde{\mathbf{z}}_k^s(i)\}_{i=1}^{m_k^s} \tag{25}$$

where m_k^s in general is random quantity.

Under the assumption that false alarms are uniformly distributed in the measurement space, and the number of false alarms is Poisson distributed, probability of getting m_k^s is given by [11]

$$p(m_k^s) = (1 - P_D^s)\frac{(\lambda V)^{m_k} \exp(-\lambda V)}{m_k!} + P_D^s \frac{(\lambda V)^{m_k - 1} \exp(-\lambda V)}{(m_k - 1)!} \tag{26}$$

where P_D^s is the probability of detecting the target by sensor s, V is the gated volume of the measurement space.

If false alarms are removed by setting a cut-off length for the number of bits to be sent after encoding, then $\bar{P}_D^s$ must be calculated by considering the possibility of removing a target originated measurement. In the PCRLB calculation, P_D^s must be replaced by $\bar{P}_D^s$. V is must also be calculated using the predicted target distribution and the false alarm removal cut-off limit. Even though the cut-off is set on the number of bits, it can be converted to the probability and can be used to decide the gate size.

Using measurement independent assumption, the measurement information, $J_k(\tilde{\mathbf{z}})$, is given by [11]

$$J_{zk}(\tilde{\mathbf{z}}) = \sum_{s=1}^{n_s} \sum_{m_k^s=0}^{\infty} p(m_k^s) J_{zk}^s(m_k^s) \tag{27}$$

where

$$J_{zk}^s(m_k^s) = E[-\Delta_{x_k}^{x_k} \log p(\tilde{\mathbf{z}}_k^s \mid x_k, m_k^s))] \tag{28}$$

$p(\tilde{\mathbf{z}}_k^s \mid x_k, m_k)$ is given by

$$p(\tilde{\mathbf{z}}_k^s(i)_{i=1}^{m_k^s} \mid x_k) = \left[\frac{(1 - \epsilon(m_k^s))}{V^{m_k^s}} + \frac{\epsilon(m_k^s)}{m_k^s V^{m_k^s - 1}} \sum_{i=1}^{m_k^s} p_1(\tilde{\mathbf{z}}_k^s(i)) \right] \tag{29}$$

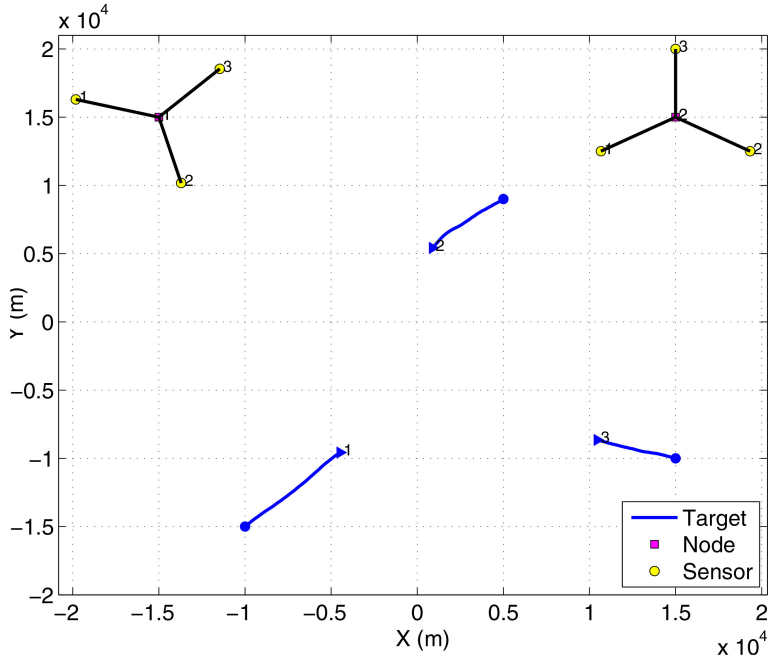


Fig. 9. The simulation environment.

where

$$\epsilon(m_k^s) = \frac{P_D^s}{p(m_k^s)} \frac{(\lambda V)^{m_k-1} \exp(-\lambda V)}{(m_k-1)!} \tag{30}$$

and $p_1(\tilde{\mathbf{z}}_k^s(i))$ is the pdf of the true observation, which is function of x_k. The details of obtaining $p_1(\tilde{\mathbf{z}}_k^s(i))$ with the quantized measurement are given in the following section.

4.3. Effect of Measurement Quantization

Due to the essence of quantization, it is known that $\tilde{\mathbf{z}}_k^s(i)$ has a discrete distribution and the only fact that can be inferred from $\tilde{\mathbf{z}}_k^s(i) = Q(z_k^s(i))$ is that $a_{(i,k)}^s \le \tilde{\mathbf{z}}_k^s(i) < a_{(i+1,k)}^s$ [8]. Under the assumption that the measurement error is a zero-mean Gaussian variable with standard deviation σ_ω^s, $p_1(\tilde{\mathbf{z}}_k^s(i))$ can be written as

$$\begin{aligned} p_1(\tilde{\mathbf{z}}_k^s(i)) &= P\{\tilde{\mathbf{z}}_k^s(i)) = Q(y_k^s(i)) \mid x_k\} \\ &= P\{a_{(i,k)}^s \le h_k(x_k) + v_k^s < a_{(i+1,k)}^s \mid x_k\} \\ &= \int_{a_{(i,k)}^s - h(x_k)}^{a_{(i+1,k)}^s - h(x_k)} \frac{1}{\sigma_w^s \sqrt{2\pi}} \exp\left\{-\frac{t^2}{2(\sigma_w^s)^2}\right\} dt. \end{aligned} \tag{31}$$

It can be shown that

$$\frac{\partial p_1(\tilde{\mathbf{z}}_k^s(i))}{\partial x_k^a} = -\frac{1}{\sigma_w^s \sqrt{2\pi}} \frac{\partial h(x_k)}{\partial x_k^a} \left(\exp\left[-\frac{(a_{(i+1,k)}^s - h(x_k))^2}{2(\sigma_w^s)^2}\right] - \exp\left[-\frac{(a_{(i,k)}^s - h(x_k))^2}{2(\sigma_w^s)^2}\right] \right). \tag{32}$$

From (28) and (32), it can be shown that

$$\frac{\partial \log p(\tilde{\mathbf{z}}_k^s \mid x_k, m_k^s)}{\partial x_k^a} = \frac{\epsilon(m_k^s)}{p(\tilde{\mathbf{z}}_k^s \mid x_k, m_k^s) m_k^s V^{m_k^s-1}} \sum_{i=1}^{m_k^s} \frac{\partial p_1(\tilde{\mathbf{z}}_k^s(i))}{\partial x_k^a}. \tag{33}$$

$J_{zk}^s(m_k^s)$ can be calculated using (33) and

$$\frac{-\partial^2 \log(p(\cdot))}{\partial x^a \partial x^b} = \frac{\partial \log(p(\cdot))}{\partial x^a} \frac{\partial \log(p(\cdot))}{\partial x^b}. \tag{34}$$

5. SIMULATION

In this section, results of the simulation studies for the proposed distributed algorithm with quantization and encoding strategies are presented.

5.1. Simulation Setup

In the simulations studies, a two dimensional tracking example is considered to show the effectiveness of the proposed algorithms. As shown in Fig. 9, it consists of two computational nodes placed at $(-15 \times 10^3, 15 \times 10^3)$ and $(15 \times 10^3, 15 \times 10^3)$. Each node has three sensors reporting bearing-only observations at a time interval of $T = 30$ s. The target motion model, which is nearly constant velocity, has the following linear-Gaussian target dynamics,

$$\mathbf{x}_{k+1} = \mathbf{F}\mathbf{x}_k + \mathbf{v}_k \tag{35}$$

where the target transition matrix F is given by

$$\mathbf{F} = \begin{bmatrix} 1 & T & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & T \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{36}$$

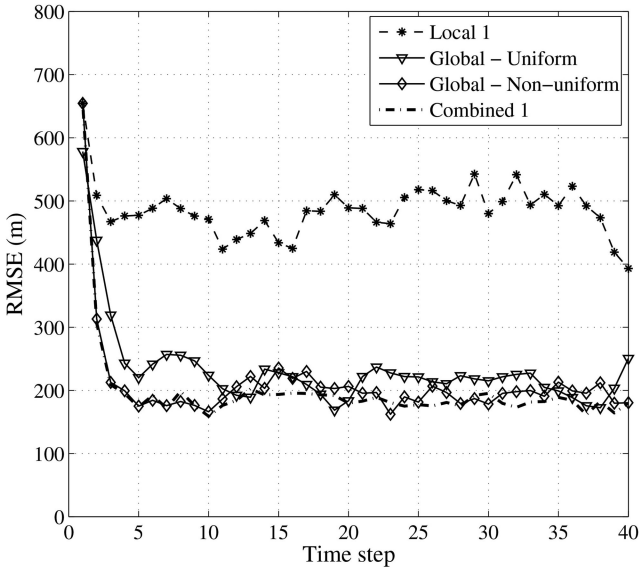


Fig. 10. Position RMSE comparison with 128-bit quantization for target 1.

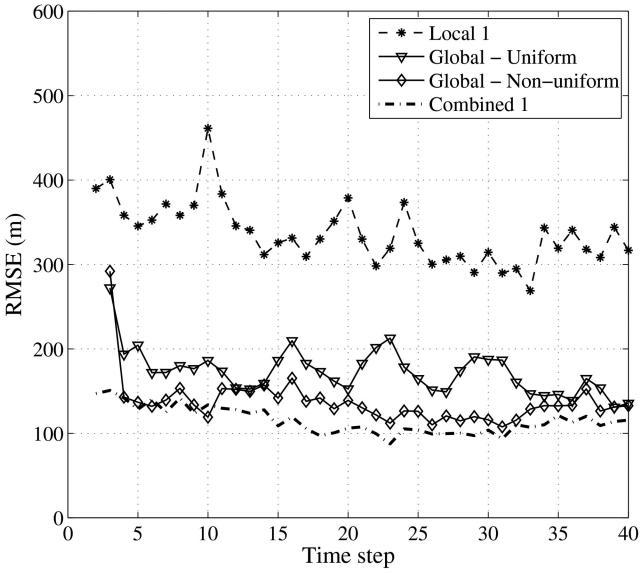


Fig. 11. Position RMSE comparison with 128-bit quantization for target 2.

and $\mathbf{v}_k$ is zero-mean white Gaussian noise with covariance $\mathbf{Q}$ given by

$$\mathbf{Q} = \begin{bmatrix} \frac{1}{3}T^3 & \frac{1}{2}T^2 & 0 & 0 \\ \frac{1}{2}T^2 & T & 0 & 0 \\ 0 & 0 & \frac{1}{3}T^3 & \frac{1}{2}T^2 \\ 0 & 0 & \frac{1}{2}T^2 & T \end{bmatrix} q \tag{37}$$

where $q = 0.001$ is the level of process noise in target motion.

Targets have different stating times and starting positions within the surveillance region. Target 1 and target 2 are present at $k = 0$, and their initial target positions are $(-10 \times 10^3, -15 \times 10^3)$ and $(-5 \times 10^3, 9 \times 10^3)$ m. Target 3 enters later at time $k = 10$ from the position

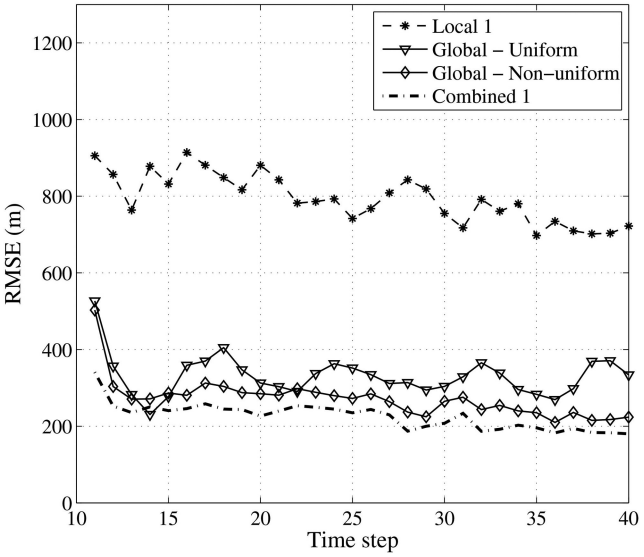


Fig. 12. Position RMSE comparison with 128-bit quantization for target 3.

$(15 \times 10^3, -10 \times 10^3)$ m. The targets' initial velocities are $(5,5)$, $(-4,3)$, $(-5,2)$ ms^{-1}. The target trajectories and sensor network arrangement are shown in Fig. 9.

The target generated measurements corresponding to target j on sensor i

$$z_k^{i,j} = \tan^{-1}\left(\frac{y_k^j - y_S^i}{x_k^j - x_S^i}\right) + v_k^i \tag{38}$$

where v_k^i is an i.i.d. sequence of zero-mean Gaussian variables with standard deviation 0.01 rad. The jth target location is denoted by (x_k^j, y_k^j) and that of ith sensor are denoted by (x_S^i, y_S^i). Additional parameters used in the simulations are: the probability of target survival = 0.99; the probability of target birth = 0.05; the probability of target spawning = 0; number of particles representing one target = 1000; the false alarm density $\lambda = 4 \times 10^{-3}$ rad^{-1}. The simulation results are based on 100 Monte Carlo runs.

5.2. Simulation Results

Figures 10, 11 and 12 show position Root Mean Square Errors (RMSEs) comparison for target 1, 2 and 3, respectively. RMSE values are computed from 100 Monte-Carlo runs. In those figures, 'Local 1' indicates the tracker at the fusion center 1 using only the measurement from local sensors; 'Global-Uniform' and 'Global-Non-uniform' indicate the trackers that use the uniformly and non-uniformly quantized measurement from all the fusion centers, respectively; 'Combined 1' indicates the tracker running at fusion center 1 that uses the quantized measurements from neighboring fusion center and the non-quantized local measurements. As expected, 'Combined 1' gives better performance than all the other trackers. Non-uniform quantization gives better performance than uniform quantization as well. Since the measurements from fusion center 2 are not

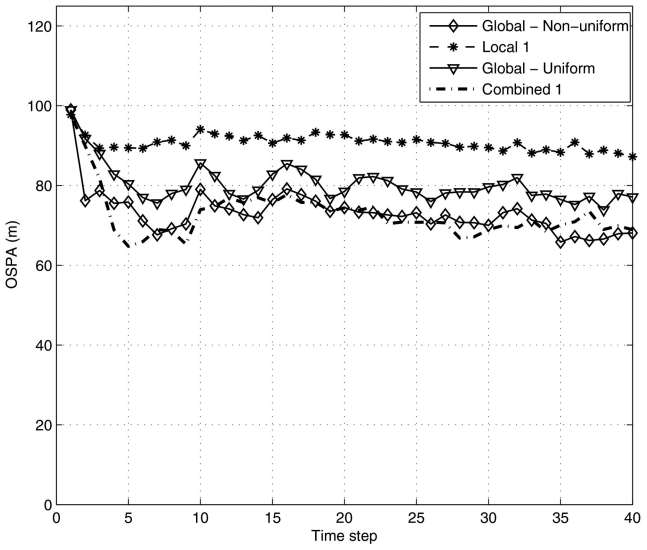


Fig. 13. OSPA comparison with 128-bin quantization.

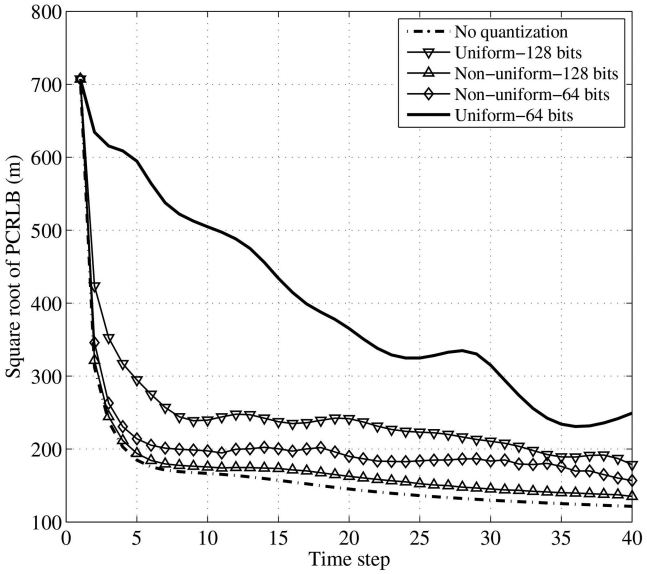


Fig. 14. Position PCRLB comparison for target 1.

used, the performance of the 'Local 1' tracker is the worst. Even though 'Combined 1' gives the best performance, its estimates cannot be used for quantization and encoding, since 'Combined 1's results, which are needed for decoding and decompanding, are not available at fusion center 2. However, 'Combined 1' can be used to eliminate the false alarms. Figure 13 shows Optimal Subpattern Assignment (OSPA) [21] distance comparison of the aforementioned approaches with OSPA parameters, $c = 25$ m, $p = 1$.

The PCRLB comparison for target 1 with various approaches is given in Fig. 14. From this figure, it can be noticed that non-uniform quantization performs close to the optimal performance, i.e., without quantization. Also, non-uniform quantization with 64 bits performs better than uniform quantization with 128 bits. Hence, non-uniform quantization can also be used to reduce the communication load in addition to improving the tracking performance.

The numbers of bits transmitted with and without Huffman coding are shown in Figs. 15, 16 and 17, where the overhead bits are not included. The effect of false alarms on communication load is shown in Fig. 15. In general, most of the false alarms are away from the target originated measurements. Hence, the number of bits allocated for the false alarms using Huffman coding, which used the probability density function of the target originated measurement, is very high. As a result Huffman coding will result in poor performance unless the false alarms are not eliminated.

After false alarms are eliminated as explained in Section 3.3, the number of bits transmitted is significantly reduced when Huffman coding is used with uniform quantization (see Fig. 16). When a new target enters and is detected by the local fusion center, the number of bit allocated for the new target originated measurement is high as the global estimate does not have information about the new target. Once the target is initialized the Huffman dictionary takes into account the new target so the encoded measurements have fewer bits. This can be observed at time step 11. Also, it is not possible to eliminate all the false alarms at all the times. Especially, it is hard to eliminate a false alarm if it falls close to any of the existing targets. This could be the reason for the slight increase in the number of bits at time step 24.

The number of bits transmitted using Huffman coding with non-uniform quantization is shown in Fig. 17. During the non-uniform encoding, the probability distribution is uniform over the measurement bins. As a result, there is no reduction in the number of bits transmitted. Hence, it is better to use no encoding with non-uniform quantization.

6. CONCLUSIONS

In this paper, a distributed implementation of SMC-PHD filter and an efficient quantization and encoding for communicating measurements were considered. Communication resources need to be handled efficiently in sensor networks while maximizing the tracking performance. False alarms take significant communication resources unless their communication is handled properly. A non-uniform quantization via companding was implemented to take advantages of the filter properties. It ensures that the target-originated measurements are quantized with less errors than others. An effective way of eliminating false alarms was also implemented. Posterior covariance was derived to access the algorithm using a recursive formula for the Fisher Information Matrix. Simulation studies confirm that the proposed quantization, encoding and false alarm elimination techniques are shown to be more efficient in terms of communication resource utilization and tracking performance than unencoded techniques. The pro-

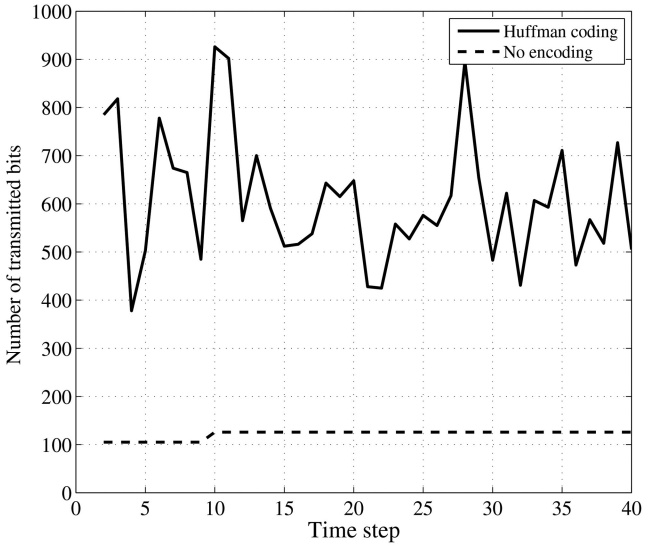


Fig. 15. Number of bits transmitted without false alarms elimination (with 3 false alarms at each time step).

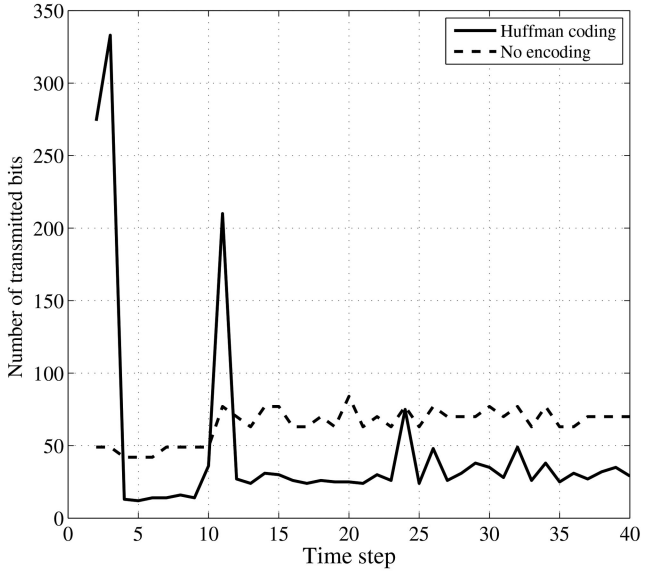


Fig. 16. Number of bits transmitted with uniform quantization and false alarm elimination.

posed distributed algorithm for SMC-PHD filter is also shown effective when the results were compared to its performance bound.

REFERENCES

[1] A. Aravinthan
Distributed tracking with probability hypothesis density filters using efficient measurement encoding.
Open Access Dissertations and Theses, Paper 4322, 2009.

[2] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp
A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking.
IEEE Transactions on Singal Processing, **50**, 2 (Feb. 2002), 173–188.

[3] Y. Bar-Shalom and X. R. Li
Multitarget-Multisensor Tracking: Principles and Techniques.
Storrs, CT: YBS Publishing, 1995.

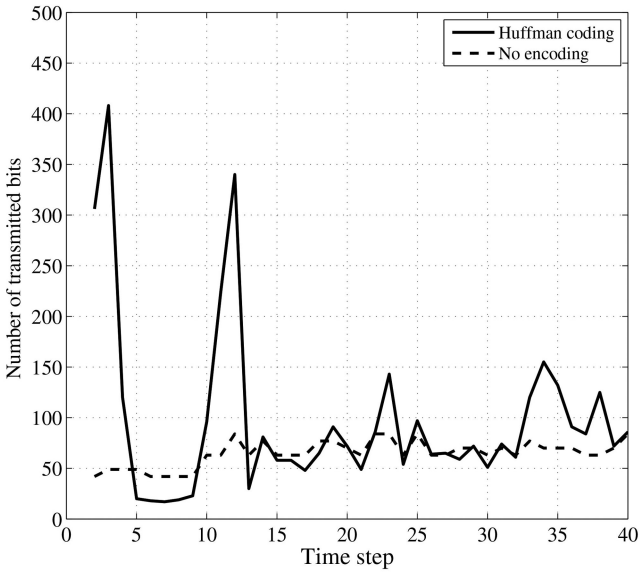


Fig. 17. Number of bits transmitted with non-uniform quantization and false alarm elimination.

[4] S. S. Blackman
Multiple hypothesis tracking for multiple target tracking.
IEEE Aerospace and Electronic Systems Megazine, **19**, 1 (Jan. 2004), 5–18.

[5] M. J. Coates
Distributed particle filters for sensor networks.
Proceedings of Third International Symposium on Information Processing in Sensor Networks, Apr. 2004, pp. 99–107.

[6] T. M. Cover and J. A. Thomas
Elements of Information Theory.
New York: Wiley, 1991.

[7] A. Doucet, N. De Freitas, and N. Gordon
An Introduction to Sequential Monte Carlo Methods.
New York: Springer-Verlag, 2001.

[8] Z. Duan, V. P. Jilkov, and X. R. Li
Posterior Cramer-Rao bounds for state estimation with quantized measurement.
Proceedings of 40th Southestern Symposium on System Theory, New Orleans, LA, Apr. 2008, pp. 376–380.

[9] R. M. Gray and D. L. Neuhoff
Quantization.
IEEE Transactions on Information Theory, **44**, 6 (Oct. 1998).

[10] D. L. Hall
Mathematical Techniques in Multisensor Data Fusion.
Norwood, MA: Artech House, 1992.

[11] M. L. Hernandez, T. Kirubarajan, and Y. Bar-Shalom
Multisensor resource deployment using posterior Cramer-Rao bounds.
IEEE Transaction on Aerospace and Electronic Systems, **40**, 2 (Apr. 2004), 399–416.

[12] D. A. Huffman
A method for the construction of minimum-redundancy codes.
Proceedings of the I.R.E., **40** (Sept. 1952), 1098–1101.

[13] G. Ing and M. J. Coates
Parallel particle filters for tracking in wireless sensor networks.
Proceedings of IEEE 6th Workshop on Signal Processing Advances in Wireless Communications, June 2005, pp. 935–939.

[14] R. P. S. Mahler
Multitarget Bayes filtering via first-order multitarget moments.
IEEE Transactions on Aerospace and Electronics Systems, **39**, 4 (Oct. 2003), 1152–1178.

[15] R. P. S. Mahler
Multitarget moments and their application to mulitarget tracking.
Proceedings of the Workshop on Estimation, Tracking and Fusion: A Tribute to Yaakov Bar-Shalom, Monterey, CA, 2001, pp. 134–166.

[16] F. Palmieri, S. Marano, and P. Willett
Measurement fusion for target tracking under bandwidth constraints.
Proceedings of the 2001 IEEE Aerospace Conference on Information Fusion, Big Sky, MT, Mar. 2001.

[17] P. Peebles
Digital Communication Systems.
Upper Saddle River, NJ: Prentice-Hall, 1987.

[18] M. Rosencrant, G. Gordon, and S. Thrun
Decentralized sensor fusion with distributed particle filters.
Proceedings of the Conference on Uncertainty in Artificial Intelligence, Acapulco, Mexico, Aug. 2003.

[19] Y. Ruan, P. Willett, A. Marrs, F. Palmieri, and S. Marano
Practical fusion of quantized masurements via particle filtering.
IEEE Transactions on Aerospace and Electronic Systems, **44**, 1 (Jan. 2008), 15–29.

[20] K. Sayood
Introduction to Data Compression.
Morgan Kaufman Publishers, 2nd ed., 2000.

[21] D. Schuhmacher, B-T. Vo, and B-N. Vo
A consistent metric for performance evaluation in multi-object filtering.
IEEE Transactions on Signal Processing, **56**, 8 (Aug. 2008), 3447–3457.

[22] X. Sheng, Y. H. Hu, and P. Ramanathan
Distributed particle filter with GMM approximation for multiple targets localization and tracking in wireless sensor network.
Proceedings of Fourth International Symposium on Information Processing in Sensor Networks, Apr. 2005, pp. 181–188.

[23] H. Van Trees
Detection, Estimation and Modulation Theory, vol. I.
New York: Wiley, 1968.

[24] B-N. Vo, S. Singh, and A. Doucet
Sequential Monte Carlo methods for multitarget filtering with random finite sets.
IEEE Transactions on Aerospace and Electronic Systems, **41**, 4 (Oct. 2005), 1224–1245.

[25] B-N. Vo and W-K. Ma
The Gaussian mixture probability hypothesis density filter.
IEEE Transactions on Signal Processing, **54**, 11 (Nov. 2006), 4091–4104.

[26] E. Waltz and J. Llinas
Multisensor Data Fusion.
Norwood, MA: Artech House, 1990.

[27] I. H. Witten, R. M. Neal, and J. G. Cleary
Arithmetic coding for data compression.
Communications of the ACM, **30**, 6 (June 1987), 520–540.

[28] X. Zhang and P. Willett
Cramer-Rao bounds for discrete-time linear filtering with measurement origin uncertainties.
Proceedings of the Workshop on Estimation, Tracking and Fusion: A Tribute to Yaakov Bar-Shalom, Monterey, CA, May 2001, pp. 546–560.

Biruk K. Habtemariam received the B.Sc. degree in electrical engineering from Mekelle University, Ethiopia, in 2007, and M.A.Sc. degree in electrical and computer engineering from McMaster University, Canada, in 2010.

Currently he is a research assistant/Ph.D. student in the Electrical and Computer Engineering Department at McMaster University. From 2007 to 2008 he was a graduate assistant in the Electrical Engineering Department at Mekelle University. His research interests include information fusion, detection/estimation theory, and target tracking. He is a recipent of International Excellence Award in 2011.

Ampikathasan Aravinthan received the B.Sc.Eng. and M.Sc.Eng. degrees in electronic and telecommunication engineering from University of Moratuwa, Sri Lanka, in 2002 and 2005, respecteively. He received his M.A.Sc. degree in electrical and computer engineering from McMaster University, Canada in 2009.

His research interests include estimation, target tracking and information theory.

Ratnasingham Tharmarasa received the B.Sc.Eng. degree in electronic and telecommunication engineering from University of Moratuwa, Sri Lanka in 2001, and the M.A.Sc. and Ph.D. degrees in electrical engineering from McMaster University, Canada in 2003 and 2007, respectively.

From 2001 to 2002 he was an instructor in electronic and telecommunication engineering at the University of Moratuwa, Sri Lanka. During 2002–2007 he was a graduate student/research assistant in ECE Department at McMaster University, Canada. Currently he is working as a research associate in the Electrical and Computer Engineering Department at McMaster University, Canada. His research interests include target tracking, information fusion and sensor resource management.

Kumaradevan Punithakumar received the B.Sc.Eng. (with First class Hons.) degree in electronic and telecommunication engineering from the University of Moratuwa, Moratuwa, Sri Lanka, in 2001, and the M.A.Sc. and Ph.D. degrees in electrical and computer engineering from McMaster University, Hamilton, ON, Canada, in 2003 and 2007, respectively.

From 2002 to 2007, he was a teaching assistant in the Electrical and Computer Engineering Department, McMaster University, where he became a postdoctoral research fellow in 2008. He is currently an imaging research scientist at GE Healthcare, London, ON, Canada. His research interests include medical image analysis, target tracking, sensor management, and computer vision.

Dr. Punithakumar was the recipient of the Industrial R&D Fellowship by the National Sciences and Engineering Research Council of Canada in 2008.

Tom Lang received Bachelor (1983) and Master (1985) of Engineering degrees in engineering physics from the Faculty of Engineering at McMaster University in Hamilton, Ontario, Canada.

Since 1985, he has been employed by General Dynamics Canada in Ottawa, Ontario, where he pursues research and development in sonar signal and data processing. In 2007, he was appointed an adjunct professor in the Department of Electrical and Computer Engineering at McMaster University. His primary research interests include sonar signal processing, multitarget tracking, and multisensor data fusion. He currently holds memberships in the IEEE Aerospace and Electronic Systems Society and the Acoustical Society of America.

Thiagalingam Kirubarajan received the B.A. and M.A. degrees in electrical and information engineering from Cambridge University, England, in 1991 and 1993, and the M.S. and Ph.D. degrees in electrical engineering from the University of Connecticut, Storrs, in 1995 and 1998, respectively.

Currently, he is a professor in the Electrical and Computer Engineering Department at McMaster University, Hamilton, Ontario. He is also serving as an Adjunct Assistant Professor and Associate Director of the Estimation and Signal Processing Research Laboratory at the University of Connecticut. His research interests are in estimation, target tracking, multisource information fusion, sensor resource management, signal detection and fault diagnosis. His research activities at McMaster University and at the University of Connecticut are supported by U.S. Missile Defense Agency, U.S. Office of Naval Research, NASA, Qualtech Systems, Inc., Raytheon Canada Ltd. and Defense Research Development Canada, Ottawa. In September 2001, Dr. Kirubarajan served in a DARPA expert panel on unattended surveillance, homeland defense and counterterrorism. He has also served as a consultant in these areas to a number of companies, including Motorola Corporation, Northrop-Grumman Corporation, Pacific-Sierra Research Corporation, Lockhead Martin Corporation, Qualtech Systems, Inc., Orincon Corporation and BAE systems. He has worked on the development of a number of engineering software programs, including BEARDAT for target localization from bearing and frequency measurements in clutter, FUSEDAT for fusion of multisensor data for tracking. He has also worked with Qualtech Systems, Inc., to develop an advanced fault diagnosis engine.

Dr. Kirubarajan has published about 100 articles in areas of his research interests, in addition to one book on estimation, tracking and navigation and two edited volumes. He is a recipient of Ontario Premier's Research Excellence Award (2002).

Association Performance Enhancement Through Classification

QUIRIN HAMP
LEONHARD REINDL

Association of spatial information about targets is conventionally based on measures such as the Euclidean or the Mahalanobis distance. These approaches produce satisfactory results when targets are more distant than the resolution of the employed sensing principle, but is limited if they lie closer. This paper describes an association method combined with classification enhancing performance. The method not only considers spatial distance, but also information about class membership during a post-processing step. Association of measurements that cannot be uniquely associated to only one estimate, but to multiple estimates, is achieved under the constraint of conflict minimization of the combination of mutual class memberships.

With Monte Carlo simulations the performance of this new method is compared with a Kalman filter. This evaluation is performed in a multi-target environment with unknown correspondence between measurements and targets. The evaluation can not be only based on the root mean square error (RMSE) of the position estimate, but requires a performance assessment of the underlying target number estimation and the association. Therefore, two new measures are introduced.

The new method outperforms the Kalman filter approach with respect to association performance and RMSE.

Manuscript received December 22, 2010; revised June 20, 2011; January 12, 2012; released for publication April 4, 2012.

Refereeing of this contribution was handled by Huimin Chen.

We gratefully acknowledge financial support from the *German Federal Ministry of Education and Research* (support code: 13N9759).

Authors' address: University of Freiburg–IMTEK, Department of Microsystems Engineering, Laboratory for Electrical Instrumentation, Georges-Köhler-Allee 106, D-79110 Freiburg, Tel: +49 761/203-7158, Fax: +49 761/203-7222, Email: (quirin.hamp@imtek.uni-freiburg.de).

1. INTRODUCTION

This paper describes the improvement of localization of targets through information fusion. In particular, information about static targets is fused, that are closely lying below surface, for so named ***G**round **P**enetrating **L**ocalization* (GPL).

GPL is used in various fields such as in geology, mine sweeping, and ***U**rban **S**earch **a**nd **R**escue* (USAR). This paper focuses on USAR. The employed methods for GPL can be classified into three categories illustrated in Fig. 1: detection, localization, and verification methods.

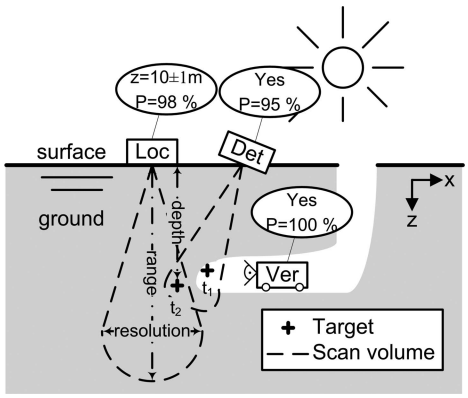


Fig. 1. Ground-penetrating localization (GPL) of targets t_i by different types of methods in xz-plane: **Det**ection, **Loc**alisation, and **Ver**ification. In the scan volume of the detection method, two targets can not be discerned because they lie within the resolution of the sensing principle.

Detection methods produce a binary result with a given detection probability of whether there is a target within their scan volume. The detection probability density is uniform within the scan volume as detection methods are unable to localize targets. A detection is determined by the signal to noise ratio and the detection threshold. An example detection method is a search dog sniffing out a victim.

In contrast, a localization method not only detects (with a given detection probability) a target within its scan volume, but can also localize its position. Localization can consist of a one-dimensional range measurement, a direction, or three-dimensional coordinates. Examples for localization methods are: ***G**round **P**enetrating **R**adar* (GPR) systems [29], cellular phone localization [13].

Whereas both of the previous methods can only produce uncertain results with respect to the existence of a target, verification methods can provide evidencethrough visual or physical inspection. The problem with verification methods consists of knowing where evidence about a target has been collected. This is often not trivial during ground-penetrating exploration that is often operated in an unknown environment without a map. Endoscopes and rescue robots can for instance be classified under verification methods since they can penetrate into a rubble pile and provide evidence about the whereabouts of victims [15].

GPL with detection and localization methods corresponds to remote sensing because the perceptive organs or devices remain at the surface while the targets are in the ground (see Fig. 1). In such remote sensing situation, a single sensor often only provides an incomplete, imprecise, and uncertain (for definitions refer to [5]) picture about the target location. Therefore, it is common practice to employ multiple sensors of different types—i.e., heterogeneous sensors—and to fuse the readings. However, if multiple targets lie close to each other, the resolution capabilities are limited to that of each single sensor.

The idea behind the proposed performance enhancement solution is that the sensing by heterogeneous sensors is based on target attributes that are not necessarily the same among all targets and can be sensed by search methods. These attributes are for instance the health condition, respiration frequency, and size of a buried victim. For instance, the GPR is capable of measuring the respiration frequency which might be different for every individual. Association is compromised if a measurement can be associated to multiple closely lying position estimates of a target. In this case, if the association method not only considers spatial aspects but previously detected unique attributes of the targets as well, measurements can be unambiguously attributed to the appropriate targets. The hypothesis is that implementing this idea enhances localization performance.

Unknown material between the surface and the target impedes not only the target position estimation, but also the fundamental detection accuracy. The scan volume is commonly unknown as well. The sensing range corresponds to the maximal distance between target and surface for which a true-positive detection occurs with sufficient probability. Resolution expresses the minimal mutual distance at which two targets can be detected individually. For instance, the distribution of steel reinforced walls may be unknown, but will influence the range and resolution of GPR. If the localization method, despite some expectable position measurement error, is capable of identifying the targets within its scan volume, the distinction of targets is easy, but this is often not possible. In order to circumvent these limitations, sensing can be carried out multiple times at different locations and orientations, as presented in Fig. 1.

When errors are random, redundant fusion may be used to increase the reliability, accuracy of and confidence in the information [21]. Systematic errors can be minimized using heterogeneous GPL methods. For instance, a systematic error may only affect one sensor type, but not others. Applying multiple sensing principles may allow decreasing, or even detecting and eliminating systematic errors.

In terms of the ***J**oint **D**irectors of **L**aboratories* (JDL) Data Fusion Working Group, this paper is limited to Level 1 Object Refinement [12]. In multi-target tracking applications using multiple sensors, several fusion architectures exist [11, 21]. The particularity of the ***A**ssociation with **C**lassification* (*AC*) method presented in this paper, is that it combines centralized fusion of locational and feature information. However, we restrict the focus to features that do not allow for the identification of a target. *AC* only demonstrates its full potential, if the sensing by heterogeneous sensors is based on multiple attributes. It is applicable for GPL of multiple, unique targets, such as USAR of trapped victims or for geologic exploration.

In Section 2, the state of the art is presented. In Section 3, the problem of uncertain and imprecise information association is formulated. The *AC* method based on an initial probabilistic association and post-processing using a possibilistic approach, is presented in Section 4. The simulation environment and the performance measures used for benchmarking *AC* with a Kalman Filter with an association gate are described in Section 5. The results of a Monte Carlo simulation are presented in Section 6 and discussed in Section 7.

2. STATE OF THE ART

Combining data originating from different sources with the goal to improve the quality of information, is called data fusion [28]. In pattern recognition, the data originates mostly from a single source and is "complete" when processed [16]. This paper focuses on scenarios where information can be "incomplete" because it is collected on demand.

Data fusion techniques have been developed for tracking mobile targets. Hence, they are particularly suited for dynamic worlds. In early work, the filters were restricted to situations where a single target is being monitored. At every time step, only one measurement was selected to update the state. Filters such as ***N**earest **N**eighbor* or ***S**trongest **N**eighbor* satisfied these requirements. However, multiple observations may be available to improve the accuracy of the estimate for the single target. The ***P**robabilistic **D**ata **A**ssociation* (PDA) filter is an "all-neighbor" approach that uses all neighboring observations within some gated region and improves the state estimation ([2], p. 299). However, the tracking performance of multiple targets using the PDA filter is poor, since "…the computation of the association probabilities separately for each target is not effective in the presence of a neighboring target." ([3], p. 325) The stability of target tracks while crossing is compromised by the persistent interference of the measurements of neighboring targets. This applies also to static worlds.

The ***J**oint **PDA*** (JPDA) filter suggested by Bar-Shalom jointly calculates the probability of measurements belonging to targets, to overcome the mentioned limitations of PDA filters. However, the underlying assumption of a JPDA filter is that a measurement only originates from one target at every time step of a scan.

In static worlds, multiple measurements may support a single target. The bijective constraint between measurement and target has to be withdrawn to be able

to associate multiple measurements with a single target. Probabilistic association techniques for multi-target environments in data fusion systems rely on validation gates.[1] The gates around position estimates allow the association of measurements based on a probabilistic approach. An alternative is ***M**ultiple **H**ypothesis **T**racker* (MHT) that represents a measurement centric approach. Because it is difficult to consider all association vectors, this approach requires pruning [27]. Furthermore, it is designed to make a single inference on one target object [17]. Hence, it would need to be extended to handle multiple targets.

To overcome the resolution limitation in dense multi-target environments, the **JPDA** with **M***erged measurements filter* (JPDAM) has been suggested. It accounts for situations where a measurement may have originated from the detection of multiple targets that are indistinguishable. The membership of a merged measurement to a track is calculated depending on its respective signal strength ([3], p. 366), i.e., an *unresolved or merged measurement* carries less information (relative to its signal strength) than a *resolved measurement*. In GPL this approach is often not possible since a measurement's relative signal strength can not be expressed. Hence, hard association is favored.

The possibilistic association of Ayoun, et al. [1] based on the ***T**ransferable **B**elief **M**odel* (TBM) of Smets [24] is limited on the one hand by a discretized resolution grid and on the other by the computational effort [14], but has the interesting distinction of allowing to search for the target location based on conflict minimization within a finite set of measurements. Possibilistic approaches of association based on class membership for target type estimation are presented in Chapter 13 of [23], but are limited to a single target and are designed for dynamic worlds.

Clustering methods such as *k-means* are inappropriate because they require the number of expected targets which is unknown. There are methods to iteratively evaluate the measurements with varying estimated target number and choose the one which minimizes or maximizes a given measure. However, this represents a bigger computational burden which decreases efficiency. Furthermore, the results of *k-means* clustering are not reproducible since they are dependent upon the initial conditions. Methods of statistical pattern recognition such as *Expectation Maximization* or *density estimation* have the advantage of considering class membership, but need a considerable amount of information to work properly, their computational effort is greedy [16]. Furthermore, their design is particularly adapted for static fusion and not for dynamic fusion.

3. FORMULATION OF THE ASSOCIATION PROBLEM

The association aims to link ***m**easurements **to** a position **e**stimate* (M2E). A measurement $\vec{r}$ and a position estimate $\hat{\vec{r}}$ are both described by a position (x,y,z) and a class membership mass function m. In the following, the vector sign is omitted.

The challenge in M2E association lies in initiating and revising a position estimate.[2] If the sparsity of a scenario (i.e., the minimal distance between targets) is smaller than the resolution, accurate position estimates are challenging[3] [6].

The description of the association problem will be twofold. First, we explain the processing of measurements without considering their class membership, represented by the orientation of the shapes of the measurements in Fig. 2. Second, the same case is revisited considering the class membership.

Association without classification: The following example is based on Bar-Shalom's terminology [3] and illustrates the complexity of the association problem without classification consideration. The binary validation matrix Ω_{ij} (see Eq. 2) expresses all feasible association events between measurements r and estimates $\hat{r}$ and is illustrated in Fig. 2. The rows represent the measurements ($r_1,\ldots,r_i,\ldots,r_8$ from top to bottom), the middle column the estimate $\hat{r}_1$, and the right column $\hat{r}_2$. The left column of the matrix indicates that every measurement may originate from a spurious source. A conjunctive area ($A \cap B$) is determined by intersecting validation gates. Measurements that are within this area, such as r_5, r_6, r_7 in Fig. 2, require particular attention because there is more than one possible way of associating them.

A measurement that could be simultaneously associated to multiple estimates is called ***un**resolved measurement* and is member of set C^{ur}. A measurement which can be associated to only one estimate $\hat{r}_j$ is called ***re**solved measurement* and is member of set C_j^{re}. Consequently, if there is no associated estimate the measurement is called ***una**ssociated measurement* and is member of set C^{ua} (see r_4 in Fig. 2). Whether r_4 will initiate a new estimate is determined by its detection probability.

The cases in following expression reflect these three conditions quantitatively. The distinction is based on a sum for a row of the validation matrix over the columns from the first estimated target to the total estimated number of targets $\hat{N}_{\mathrm{t}}$.

[1] The validation gate is depicted by an ellipse centered about the nominal position estimate that represents the contour of constant probability (in two dimensions) for a multivariate Gaussian distribution [26]. To filter measurements that are not associated with an estimate (i.e., the information source) is the purpose of a validation gate that corresponds to the volume around the position estimate ([20], p. 157). If the Mahalanobis distance between an estimated target and a measurement is smaller than a predefined threshold, the latter is associated with the former.

[2] Tracking of mobile targets i.e., a dynamic state consists of revising the state up to date, since it changes. However, the belief about the static state is **revised** (not updated), because the state does not change.

[3] See measurements r_5, r_6, r_7 in Fig. 2 which could be associated either to estimate $\hat{r}_1$, to $\hat{r}_2$, to both, or to none.

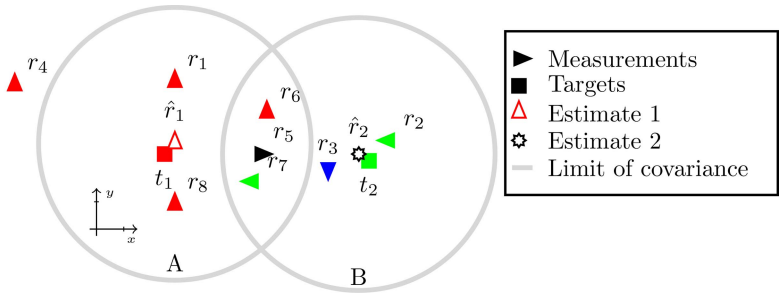


Fig. 2. Measurements (triangles r_i) originating from two unknown targets (squares: t_1 from **top** class and t_2 from **left** class). Intersecting validation regions (solid circles) around position estimates $\hat{r}_1$ (not filled triangle) categorized as **top** class member and uncategorized $\hat{r}_2$ (not filled star) based on resolved measurements (xy-plane). The validation regions A, B have the same size because the distribution of measurements is unknown and presumed to be equal for any estimate.

$$r_i \in \begin{cases} C_j^{\text{re}} & \text{if} \quad \sum_{j=2}^{\hat{N}_t} \Omega_{ij} = 1; \\ C^{\text{ur}} & \text{if} \quad \sum_{j=2}^{\hat{N}_t} \Omega_{ij} > 1; \\ C^{\text{ua}} & \text{if} \quad \sum_{j=2}^{\hat{N}_t} \Omega_{ij} = 0. \end{cases} \tag{1}$$

Association with classification: This processing reduces the feasible association events as presented in Eq. 2:

$$\Omega_{ij} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix} \overset{\text{processing}}{\longrightarrow} \Omega_{ij}^* = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & \mathbf{0} \\ 1 & \mathbf{0} & 1 \\ 1 & 1 & 0 \end{bmatrix}. \tag{2}$$

In the following, three mutually exclusive classes will be examined corresponding to the orientation of the triangles in Fig. 2: **top**, **bottom**, **left**.

The relative complement areas of A\B and of B\A are predominant for classification of estimates $\hat{r}_1$ and $\hat{r}_2$, respectively. The measurements r_1 and r_8 are located in the predominant area of estimate $\hat{r}_1$ (A\B). Based on its membership to the **top** class, it can be inferred that the estimate $\hat{r}_1$ most likely belongs to class **top** as well. In contrast, no stringent inference on the class membership of estimate $\hat{r}_2$ can be achieved, because in its predominant region (B\A), conflicting classes are present: **left** and **bottom** of r_2 and r_3, respectively. However, it can be assumed that with an increasing number of measurements present in the predominance area, the classification errors become negligible.

Unresolved measurements for which association is not unique can actually be assessed with respect to the estimates' class that was determined by the *resolved measurements*. **Top**, *unresolved measurement* r_6 can be associated uniquely to $\hat{r}_1$ and becomes *resolved*. For measurement r_7 the situation is more complex since $\hat{r}_2$ has conflicting class membership. $\hat{r}_2$ could be **left** or **bottom**. However, the **left** class corresponds more to $\hat{r}_2$ than to $\hat{r}_1$. Hence, *unresolved measurement* r_7 can be associated uniquely to $\hat{r}_2$, and thus is *resolved*.

The class membership of measurement r_5 is not known. Hence, it can not be associated uniquely to any of the two estimates and remains *unresolved*.

Ω becomes Ω^* (see Eq. 2) when considering class membership. It contains less feasible association events than Ω.

These examples convey the complexity of association considering class membership. *Resolved measurements* must allow the determination of the estimate's class. The estimates must be of different classes in order to be able to associate *unresolved measurements* uniquely. If these constrains are fulfilled, unresolved measurements in the conjunctive area can be *resolved*.

4. DESCRIPTION OF AC METHOD

Before presenting the two main steps of the AC method, which consist of an initial probabilistic association followed by a post-processing of *unresolved measurements*, the underlying assumptions and simplifications will be explained.

4.1. Assumptions and Simplifications

The presented method can only be applied under the following assumptions:

The localization performed by the sensors may be based on different attributes of the targets. Since these attributes may not be common to all targets, classification becomes feasible depending on the detected attributes of the target. A sensor may be able to detect multiple attributes at once, or heterogeneous sensor technologies can be employed, which offer the opportunity for complementary fusion,[4] and can thus recognize different classes. In this preliminary paper, the classification capability of sensors is assumed to be perfect even if this can not be expected in reality.

[4]Definition provided in [8].

Furthermore, not only do the sensors need to recognize different classes, but targets must be of different classes. It is also presumed that a measurement only originates from one information source, unlike with a JPDAM filter where (to some extent) the model tries to associate a measurement to two targets.

The distribution of measurements around a target is unknown, but is constant in static worlds, and its estimation depends upon the number of measurements available. The measurement noise is usually modeled by a Gaussian distribution ([27], p. 154). An erroneous estimation of the number and position of targets can have two sources. Either there are not enough measurements concerning all targets, or the processing is incorrect. In order to focus on latter, we consider only situations where sufficient measurements for each target are available to determine their number and position.

In situations where not many measurements are to be expected such as during USAR, the initiation can not be based on multiple measurements. Hence, a position estimate is initiated as soon as a new measurement is generated that can not be associated.

4.2. Initial Step of Association Method

Association of measurements is based on the Mahalanobis distance, and consists of finding the correlation between measurements and an information source. This statistical squared distance indicates the probability that measurement r_i belongs to estimate $\hat{r}_j$. In the following, matrices are represented by uppercase letters while vectors are lowercase.

Assuming that a positional error of a measurement r in two dimensions (x, y) can be described with a Gaussian distribution, the following covariance matrix can be used, where ρ is the correlation coefficient for x and y [26]:

$$C = \begin{pmatrix} \sigma_x^2 & \rho\sigma_x\sigma_y \\ \rho\sigma_x\sigma_y & \sigma_y^2 \end{pmatrix}. \tag{3}$$

Since the distribution of measurements around the target is unknown but constant, it is assumed to be Gaussian. The validation gate is based on an estimate of this distribution. The distance between measurement and estimate is based on the sum of the covariance matrices of the gate G and of the measurement R_i. Unlike Kalman filters, G remains constant. Otherwise it would converge, because fusion leads to a reduction of uncertainty ([3], p. 444), i.e., the fused uncertainty ellipse is strictly contained in the intersection of the two ellipses prior to fusion (see Eq. 6). Hence, for an increasing number of measurements, the probability of association would decrease considerably and hinder the association.

The Mahalanobis distance d_{ij}^2 is defined in Eq. 4. The gate threshold γ gives the maximal distance for the association of measurements to an estimate.

$$d_{ij}^2 = (r_i - \hat{r}_j)^T (R_i + G)^{-1} \underbrace{(r_i - \hat{r}_j)}_{\text{innovation}} \leq \gamma. \tag{4}$$

Following Smith and Cheeseman [26] the revision of an estimate's position $\hat{r}'_j$ is calculated by:

$$\hat{r}'_j = \hat{r}_j + R_i(R_i + \Sigma_j)^{-1}(r_i - \hat{r}_j) \tag{5}$$

and its corresponding covariance matrix expressing imprecision is:

$$\Sigma'_j = R_i - R_i(R_i + \Sigma_j)^{-1} R_i. \tag{6}$$

The standard application of Kalman filters is to process random signals that are represented "…as the output of a linear dynamic system excited by independent or uncorrelated random signals ("white noise")" [18]. This application differs from the standard application, because it is not a dynamic system.

The processing order of measurements fusion, particularly when targets are spaced close together, influences the association and leads to erroneous associations. This is especially true with *unresolved measurements*. In order to correct these erroneous associations, a post-processing step is performed.

4.3. Post-Processing of Initial Association

The post-processing is based on fusion of class membership information, where the chosen framework is TBM. The advantage of this approach is that inference from contradictory class information is possible. Therefore, the finite set of propositions Θ is extended to the set of all subsets 2^Θ by conjunction and to the empty set. The set of all subsets is also called frame of discernment. With this extension, the exhaustive and the exclusive assumptions on the proposition framework can be disregarded. This disregard is equivalent to the open world assumption. This assumption accepts that none of the propositions may be true [24]. No matter how large the frame of discernment is, it contains three different proposition classes: the empty set $\emptyset$, the atomic propositions corresponding to Θ, and conjunctive propositions (number: $2^\Theta - \Theta - 1$). Belief in a proposiiont is quantized by a function called basic belief assignment (bba) m(A), $A \in \Theta$. Let $m : 2^\Theta \rightarrow [0, 1]$, where following constraints hold:

$$\sum_{A \subseteq \Theta} m(A) = 1, \qquad m(\emptyset) = 0. \tag{7}$$

(8)

It is worth noting that the conjunction of all atomic propositions corresponds to the vacuous function, also referred to as total ignorance. The main difference of the seminal proposition of Dempster and Shafer [22] is that the conjunctive combination of beliefs

$$m_{12}(A) = \sum_{B \cap C = A} m_1(B) m_2(C) \tag{9}$$

in the TBM is not normalized, since:

$$\sum_{B\cap C=\emptyset} m_1(B)m_2(C) > 0. \tag{10}$$

For this reason, TBM is able to express conflict among beliefs. The conflict is given by the mass allocated to the empty set. It is advantageous that the combination is commutative and associative. Conflict, reliability and uncertainty can be expressed quantitatively.

The method is also attractive because of the ease of implementation by matrices, as presented in [25].

4.3.1. Class Membership of Predominant Areas

Class membership is computed by combining all class membership mass vectors m_i of the *resolved measurements* ($r_i \in C_j^{\text{re}}$) (i.e., measurements in the predominant area of an estimate). The result is again a class membership mass vector for $\hat{r}_j$: m_j^{p}. If estimates are very close, there are most likely no *resolved measurements* available. This will prove to be a limitation, and will be evaluated in Monte Carlo simulations.

4.3.2. Resolving *Unresolved Measurement*

To determine the most likely association, the mass vector m of an *unresolved measurement* ($r \in C^{\text{ur}}$) is combined with the mass vector m_j^{p} of each $\hat{r}_j$, and the minimal conflict is determined. The bijective association is given by the combination with the minimum conflict.

$$m_{ij}^{\text{c}} = m_i \cdot m_j^{\text{p}} \qquad \forall r_i \in C^{\text{ur}}. \tag{11}$$

$$\Omega_{ij} = \begin{cases} 1 & \text{if} \quad j = \min m_{ij}^{\text{c}}(\emptyset); \\ 0 & \text{if} \quad j \neq \min m_{ij}^{\text{c}}(\emptyset). \end{cases} \tag{12}$$

ALGORITHM 1 *Association with classification (AC)*

Input: *association threshold* γ, r_i
Output: *position estimates* $\hat{r}_j$

Evaluate number $\hat{N}_{\text{t}}$ *and* $\hat{r}_j \qquad \forall\{r_i \in C_j^{\text{re}}\}$
$\Omega_{ij} \leftarrow 1$ *if* $d_{ij}^2 \leq \gamma$
if $\exists r_i \in C^{\text{ur}}$ **then**
  $m_j^{\text{p}} \leftarrow m_i \cdot m_j^{\text{p}} \qquad \forall\{r_i \in C_j^{\text{re}}\}$
  for $\forall r_i \in C^{\text{ur}}$ **do**
    for $j = 1$ *to* $\hat{N}_{\text{t}}$ **do**
      $m_{ij}^{\text{c}} \leftarrow m_i \cdot m_j^{\text{p}}$
      {*Save the association of minimal conflict*}
      if $\min m_{ij}^{\text{c}}(\emptyset)$ **then**
        $k \leftarrow j$
      end if
    end for
    {*Resolve measurement*}
    $\Omega_{ij}^* \leftarrow 0 \qquad \forall j \neq k \wedge j > 1$
  end for
end if
return $\hat{r}_j \leftarrow \bar{r}_i$
$\forall\left\{r_i \mid \Omega_{ij}^* = 1 \wedge \sum_{j=2}^{\hat{N}_{\text{t}}} \Omega_{ij}^* = 1\right\}$

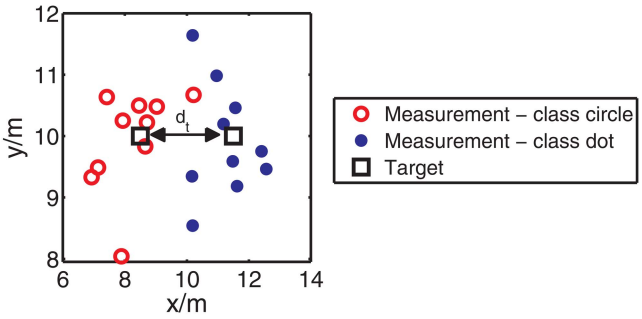


Fig. 3. Measurements ($N_{\text{m}} = 20$ with normal spatial error $\sigma_{\text{e}} = 1$ m) around two targets (squares, distance $d_{\text{t}} = 4$ m). The difference in the class of the measurements (circle or dot) indicates the original target (xy-plane).

The estimates' positions of the post-processed measurements are calculated by taking the average instead of using Eq. 5, as for Kalman filters.

5. SIMULATION

The evaluation of the performance of the proposed fusion method *AC* will be carried out in a Monte Carlo simulation. It is compared with a Kalman filter approach which is called Kalman Gating (*KG*). It corresponds to the *AC* method, with the difference that the post-processing step is omitted. The position of the estimates is calculated using the sequential Kalman update Eq. 5.

Two targets will be considered in this simulation, but the methods can cope with multiple targets. The distance d_{t} between the two targets is the main evaluation vector, because it simulates various densities which may occur in real scenarios. The scale of the scenario is assumed to be in the order of meters.

Parameters determining *AC* and *KG* are the standard deviation $\sigma = 1$ m of the constant gate covariance matrix G and the gate threshold $\gamma = 4$.

5.1. Generation of Simulated Measurements about the World State

Measurements around targets are presumed to be subject to a normal measurement error with standard deviation σ_{e}. Each target is supported by the same number of measurements. Each measurement's spatial uncertainty is assumed radially symmetric, and is expressed by a covariance matrix R_i.

There are no outliers and no classification errors. Only two classes are considered. Figure 3 represents such a scenario.

The scenario and the simulation parameters are given in Tables I and II, respectively.

5.2. Performance Measures

The quantitative evaluation of fusion methods requires performance measures which are general enough to be applied to all types of scenarios, including those with more than two targets. The determination of the estimates' positions relies on a consecutive order of pro-

TABLE I
Parameters of the GPL Scenario

N_t	N_m	σ of R_i/m	σ_e/m
2	20	0.4	0.5

TABLE II
Parameters of the Monte Carlo Simulation

d_t/m	Step Size/m	Repititions
0–5.5	0.1	1500

cessing steps. First, the number of targets has to be estimated. Then the measurements have to be associated to these estimates and finally, the positions of the estimates are calculated based on the associated measurements. The accuracy of the position estimate is expressed with the RMSE, which depends on the correct estimation of the number of targets and the correct association of measurements to the estimates.

It is worth noting that the performance can only be evaluated if all knowledge about the world is available. The so called *ground truth* gives information about the target position and the origin of a measurement.

Even if the RMSE is conditioned by the underlying association performance, the performance of estimating the number of targets and of associating measurements to these position estimates will be evaluated individually.

5.2.1. Target Estimation

The fusion methods may generate a number of position estimate $\hat{N}_t$ that do not correspond to the real number of targets N_t.

A method's performance in estimating the correct number of targets is given by T: the quotient of the number of estimated $\hat{N}_t$ over the number of real targets N_t as given in Eq. 13. For values $T < 1$ the method underestimates and for $T > 1$ it overestimates the number of targets.

$$T = \frac{\hat{N}_t}{N_t}. \tag{13}$$

5.2.2. Association

The association performance A is the quotient of the correctly associated measurements $(N_m - N_{err})$ over the total measurements N_m originating from a target (Eq. 14). A is strictly bound in the interval $(1/N_t, 1]$.

$$A = \frac{N_m - N_{err}}{N_m}. \tag{14}$$

The evaluation of N_{err} requires particular attention. When the measurements are created, the index of their original target is saved in an array. The aim of any association method is to reconstruct this array (or "code word"), but there are two complicating factors. First, there might be N_{err} errors that must be detected in the reconstructed code and second, the sequence could possibly be generated with a different character concordance.

For instance, following two code words[5] $c_1 = 123123$ and $c_2 = 321321$ contain the same information. However, characters "1" and "3" in c_1 correspond to character "3" and "1" in c_2, respectively. The code words use different permutations of the character set Ψ:

$$\Phi_1 = [1,2,3] \neq \Phi_2 = [3,2,1].$$

Since the concordance is unknown, the generated code word is transcribed with any possible concordances that are all possible permutations of Ψ. The number of all possible permutation is: $q!$ (in the example: $3! = 6$).

For each of these possible permutations, the Damerau-Levenshtein distance d_{DL} is computed [4, 9]. The Damerau-Levenshtein distance is particularly suited for the evaluation of q-ary codes. The number of association errors corresponds to the smallest d_{DL} distance between these transcribed code words $c_{trans}^{\Phi_i}$ and the original code word c_{orig} (see Eq. 15).

$$N_{err} = \min_i d_{DL}(c_{orig}, c_{trans}^{\Phi_i}) \qquad \forall \Phi_i \in \Psi. \tag{15}$$

5.2.3. Root Mean Square Position Error

The position estimate error is given by the mean of the root mean square distance between estimated positions $\hat{r}_j$ and target positions t_k given in Eq. 16. Since in GPL practice, the correspondence between estimate and target remains unknown, only the most likely association can be considered. At most, the presumption can be made that the correct correspondence is the one with the minimal Euclidean distance among their positions. The significance of the RMSE is limited, even if the number of estimated targets corresponds to the number of targets $(\hat{N}_t = N_t)$ because two estimates can be associated with the same target. It is even more detrimental if the estimation of the number is erroneous $(\hat{N}_t \neq N_t)$, which is why previously presented performance indicators about target estimation and association are crucial.

$$\mathrm{RMSE} = \sqrt{\frac{1}{\hat{N}_t - 1} \sum_{j=1}^{\hat{N}_t} \min_k (\|\hat{r}_j - t_k\|)^2}. \tag{16}$$

6. RESULTS

The results of the Monte Carlo simulation are presented in Figs. 4, 5, and 6. The analysis of the results represented in Figs. 4 and 5 is performed by comparison with the cumulative intersection distribution function P_{int}, i.e., the separation of two Gaussian distributions. Eq. 17 expresses the cumulative intersection probability of two uni-variate normal distributions with equal

[5]Length $n = 6$, character set $\Psi = \{1,2,3\}$, character set cardinality $q = |\Psi| = 3$.

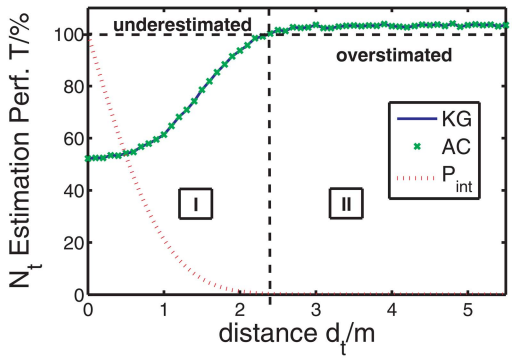


Fig. 4. The target estimation performance T evaluated through a Monte Carlo simulation is represented over the distance d_t between two targets for the method *KG* and *AC* (lines interpolated). The intersection probability P_{int} represents the overlap of the measurements' distribution. The methods' performances are identical.

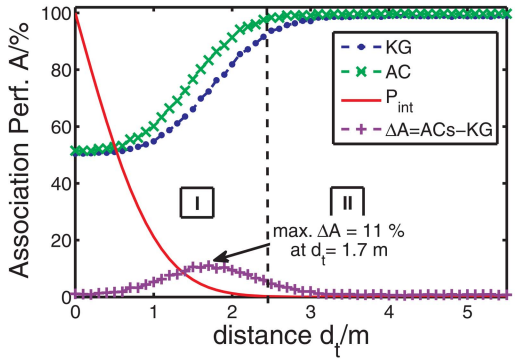


Fig. 5. The association performance A is represented over the distance d_t between two targets for the *KG* and the *AC* method (lines interpolated). *AC* outperforms the *KG* with respect to A when their validation intersect, but do not match.

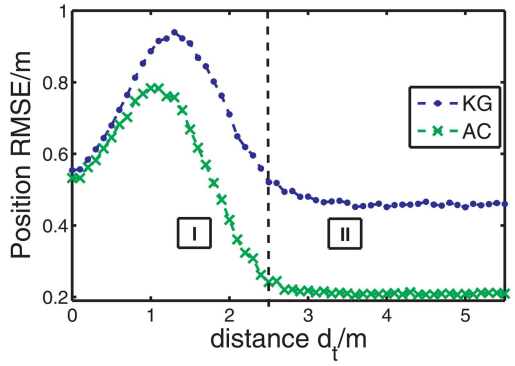


Fig. 6. The RMSE is represented over the distance d_t between two targets for *KG* and *AC* (lines interpolated). The characteristic peak in the RMSE is dependent on the methods' capability to detect the two targets.

standard deviation σ as a function of d_t:

$$P_{int} = 1 - \mathrm{erf}\left(\frac{d_t}{2\sqrt{2\sigma^2}}\right). \tag{17}$$

As *AC* is based on the *KG* for target number estimation, their performances with respect to T are identical as visible in Fig. 4. The graph is classified into two regions depending on whether the target number N_t is underestimated (region I) or overestimated (region II).

The *AC* method enhances the performance of correct association to estimates whenever validation gates intersect, but do not match. An increase in the distance d_t in region I (see Fig. 5) improves the association performance A. The maximal difference in association performance (ΔA) is reached at an intersection probability of $P_{int} = 3\%$.

The positional RMSE for both methods increases in region I with an increasing distance d_t. However, the maximal error is reached for both methods before the limit of region I. *AC* reaches its maximum RMSE around $d_t = 1.0$ m and *KG* around $d_t = 1.3$ m. After the maximum the RMSE monotonically decreases. The RMSE of *AC* is always smaller that of *KG*.

7. DISCUSSION AND OUTLOOK

The performance of any parametric fusion method depends upon an accurate estimation of parameters. To avoid, for instance, an overestimation of the number of targets ($T > 1$), the standard deviation of the measurements σ_e around the targets has to correspond to the standard deviation of validation gate G. If σ of G is larger than σ_e, the association performance A deteriorates. In contrast, if σ of G is smaller than σ_e, the number of estimates N_t is overestimated at a smaller distance of d_t, enhancing A.

Comparing Fig. 5 with Fig. 6 allows to reason that the smaller RMSE of *AC* in region II—in average 0.25 m better—is not due to better association performance, since in region II, A converges to the optimum for both. Hence, the smaller position RMSE must be due to the difference on how the estimates' position is evaluated. In region I however, there is an improvement of association performance with *AC* which explains why the maximal RMSE is reached at smaller d_t and why it constantly decreases in comparison to *KG*.

The presented method is based on Kalman filtering which intrinsically is based on a motion model for the targets. Despite the focus of this paper on static states, the authors hope to extend the presented framework to dynamic states where situations of limited resolution may also arise. The Kalman filter in the static case behaves like a median or low-pass filter.

The evaluation of the position error (Eq. 16) could be based on an optimal assignment method such as the Hungarian method [19]. However, the Hungarian method may as well produce unsatisfactory results due to its global optimization nature, especially in cases where the estimated target number does not correspond to the true one.

The revisable reasoning characteristic of the TBM method is questionable because the mass of the empty set corresponding to the conflict is a strictly monotonic increasing function. More combinations of conflicting

class information decrease the specificity of information. Robust combination rules such as suggested by Florea or Dezert might be beneficial [7, 10]. What is missing in the TBM combination rule is an expression of the reliability dependent on the total amount of fused information.

8. CONCLUSION

This paper demonstrates the potential of basing association not only on spatial aspects, but on classification attributes as well. The post-processing of the *AC* method enhances the association performance for closely lying targets while diminishing the RMSE. The belief combination framework of the TBM is a valuable tool to quantitatively express conflict, and to infer from class membership even for contradictory class information, which is to be expected for closely lying targets.

REFERENCES

[1] A. Ayoun and P. Smets
Data association in multi-target detection using the transferable belief model.
International Journal of Intelligent Systems, **16**, 10 (Oct. 2001), 1167–1182.

[2] S. S. Backman
Mulitple-Target Tracking with Radar Applications.
Artech House, 1986.

[3] Y. Bar-Shalom and X-R. Li
Multitarget Multisensor Tracking: Principles and Techniques.
Danvers, MA: YBS, 1995.

[4] F. J. Damerau
A technique for computer detection and correction of spelling errors.
Communication ACM, **7**, 3 (1964), 171–176.

[5] S. Das
High-Level Data Fusion.
Norwood, MA: Artech House, Inc., 2008.

[6] S. Deb, M. Yeddanapudi, K. Pattipati, and Y. Bar-Shalom
A generalized S-D assignment algorithm for multisensor-multitarget state estimation.
IEEE Transactions on Aerospcae and Electronic Systems, **33**, 2, Part 1, (Apr. 1997), 523–538.

[7] J. Dezert, F. Smarandache, A. Tchamova, and P. Konstantinova
Target Type Tracking with PCR5 and Dempster's rules: A Comparative Analysis.
Technical report, ArXiv, 2006.

[8] H. F. Durrant-Whyte
Sensor Models and Multisensor Integration.
The International Journal of Robotics Research, **7**, 6 (1988), 97–113.

[9] G. Fazekas and V. I. Levenshtein
On upper bounds for code distance and covering radius of designs in polynomial metric spaces.
Journal of Combinatorial Theory, Series A, **70**, 2 (1995), 267–288.

[10] M. C. Florea, A-L. Joussel, E. Bossé, and D. Grenier
Robust combination rules for evidence theory.
Information Fusion, **10** (Apr. 2009), 183–197.

[11] D. L. Hall and S. A. H. McMullen
Mathematical Techniques in Multisensor Data Fusion.
Artech House, Inc., 2004.

[12] D. L. Hall and J. Llinas
An introduction to multisensor data fusion.
In *Proceedings of the IEEE*, **85** (Jan. 1997), 6–23.

[13] Q. Hamp, R. Zhang, L. Chen, O. Gorgis, T. Ostertag, Y. Yankelevich, J. Pavlina, M. Loschonsky, F. Schilling, and L. M. Reindl
Results from German research project I-LOV.
In *International Conference on Wireless Technologies for Humanitarian Relief (ACWR2011)*, 2011.

[14] Q. Hamp, M. Eitelberg, B-S. Lee, T. Becker, D. Wiebeck, and L. M. Reindl
Information Fusion based on graph analysis during Urban Search and Rescue.
In *Information Fusion*, 2010.

[15] Q. Hamp, A. Kleiner, and L. M. Reindl
Lessons Learned from German Research for USAR.
In *International Symposium on Safety, Security, and Rescue Robotics*, 2011.

[16] A. K. Jain, R. P. W. Duin, and J. Mao
Statistical Pattern Recognition: A Review.
IEEE Transactions on Pattern Analysis and Machine Intelligence, **22**, 1 (Jan. 2000), 4–37.

[17] P. Jensfelt and S. Kristensen
Active global localization for a mobile robot using multiple hypothesis tracking.
IEEE Transactions on Robotics and Automation, **17**, 5 (Oct. 2001), 748–760.

[18] R. E. Kalman
A new approach to linear filtering and prediction problems.
Journal of Basic Engineering, 82 (Series D) (1960), 35–45.

[19] H. W. Kuhn
The Hungarian method for the assignment problem.
Naval Research Logistics Quarterly (NRL), 2 (1955), 83–97.

[20] H. Mitchell
Multi-Sensor Data Fusion.
Berlin: Springer, 2007.

[21] E. F. Nakamura, A. A. F. Loureiro, and A. C. Frery
Information fusion for wireless sensor networks: Methods, models, and classifications.
ACM Computing Surveys, **39**, 3 (2007).

[22] G. Shafer
The combination of evidence.
International Journal of Intelligent Systems, 1 (1986), 155–179.

[23] F. Smarandache and J. Dezert (Eds.)
Advances and Applications on DSmT for Information Fusion (Collected Works), Vol. 2.
American Research Press, 2006.

[24] P. Smets and R. Kennes
The Transferable Belief Model.
Artificial Intelligence, **66**, 2 (Apr. 1994), 191–234.

[25] P. Smets
The application of the matrix calculus to belief functions.
International Journal of Approximate Reasoning, **31**, 1–2 (2002), 1–30.

[26] R. C. Smith and P. Cheeseman
On the representaiton and estimation of spatial uncertainty.
International Journal of Robotics Research, **5**, 4 (1986), 56–68.

[27] S. Thrun, W. Burgard, and D. Fox
Probabilistic Robotics.
Boston, MA: MIT Press, 2005.

[28] L. Wald
Some terms of reference in data fusion.
IEEE Transactions on Geoscience and Remote Sensing, **37**, 3, Part 1 (May 1999), 1190–1193.

[29] E. Zaikov
UWB radar for detection and localization of trapped people.
In *11th International Radar Symposium (IRS)*, 2010, 1–4.

Quirin Hamp graduated 2005 from the Technical University Munich, Germany and the Ecole Centrale Paris, France. He holds a Dipl. Ing. degree in mechanical engineering and a M.Sc. degree in industrial management.

During research stays at the Massachusetts Institute of Technology (2005) and Columbia University (2006), he worked on soft biomimetic robots for autonomous exploring of obstructed environments. Since 2008, he is with University of Freiburg, Germany, and works on information fusion during disaster response. His research interest in disaster response is motivated by his experiences as paramedic he gained at the German Red Cross.

Leonhard M. Reindl received the Dipl.Phys. degree from the Technical University of Munich, Germany in 1985 and the Dr. sc. techn. degree from the University of Technology Vienna, Austria in 1997.

Since 2003, he chairs as full professor the Laboratory for Electrical Instrumentation at the Department of Microsystems Engineering (IMTEK), University of Freiburg, Germany. He is member of the IEEE and of the European Security Research and Innovation Forum and holds more than 30 patents and has authored or coauthored more than 150 papers. His research interests include wireless sensor and identification systems, Surface Acoustic Wave devices and materials as well as microwave communication systems based on SAW devices.

A Fusion Analysis and Evaluation Tool for Multi-Sensor Classification Systems

ROMMEL NOVAES CARVALHO
KUOCHU CHANG

Multi-Sensor Fusion is founded on the principle that combining information from different sensors will enable a better understanding of the surroundings. However, it would be desirable to evaluate how much one gains by combining different sensors in a fusion system, even before implementing it. This paper presents a methodology and tool that allows a user to evaluate the classification performance of a multi-sensor fusion system modeled by a Bayesian network. Specifically, we first define a generic global confusion matrix (GCM) to represent classification performance in a multi-sensor environment, we then develop a methodology with analytical convergence bounds to estimate the performance. The resulting system is designed to answer questions such as: (i) What is the probability of correct classification of a given target using a specific sensor individually? (ii) What if a specific set of sensors combined together are used instead? (iii) What is the performance gain by adding another sensor to this set? and (iv) Which sensors provide a better cost/benefit ratio? These questions are answered based on the probability of correct classification that can be analytically estimated using Bayesian inference with the given sensor models defined by confusion matrices. The principle that combining information enhances the understanding of the surroundings is also supported by the analysis made in example models for air target tracking and classification using the developed tool.

Manuscript received November 2, 2009; revised January 18, 2011, March 28, 2011, April 22, 2012; released for publication April 23, 2012.

Refereeing of this contribution was handled by Mujdat Cetin.

Authors' address: Department of Systems Engineering & Operations Research, Volgenau School of Information Technology and Engineering, George Mason University, Fairfax, VA 22030, E-mail: (rommel.carvalho@gmail.com, kchang@gmu.edu).

1557-6418/12/$17.00

1. INTRODUCTION

Fusion of information from multiple sources to achieve performances exceeding those of individual sources has been recognized in diverse areas [17] such as reliability, forecasting, pattern recognition, neural networks, decision fusion, and statistical estimation. In engineering systems, the fusion methods have proven to be particularly important since they can provide system capabilities with multiple sensors significantly beyond those of single sensor systems. Multi-sensor data fusion allows the combination of information from sensors with different physical characteristics to enhance the understanding of the surroundings and provide the basis for planning, decision-making, and control of autonomous and intelligent machines. It seeks to combine information from multiple sensors and sources to achieve inferences that are not feasible from a single sensor or source.

To fully exploit the capabilities of a fusion system, modeling and performance evaluation methodologies are critical in order to optimally design and effectively evaluate fusion performance of multiple heterogeneous sensor data. In particular, a systematic approach to evaluate the overall performance of the system is indispensable. To allow developers and users to assess their fusion system performance under various conditions before a data fusion system is deployed, a tool based on the Fusion Performance Model (FPM) [8] was developed with a focus on one of the most important performance measures, spatial and classification performance modeling and prediction. Note that the purpose of the FPM is to predict performance given sensor suite and operating conditions.

For a sensor fusion system, typical questions that could be asked would be "what is the best achievable performance, and is it good enough?" The FPM will be able to answer the first question and if the answer is "not good enough," a sequence of "what if" scenarios can be added for FPM to conduct new assessments. Those scenarios may include changing operating conditions, such as signal-to-noise ratio (SNR), geometry, and revisit rate, to name a few of the existing sensors or adding new sensors. The assessment results can then be used to better manage sensors and allocate system resources.

While the FPM model described in [8] developed a kinematic performance prediction methodology and defined the classification performance model and [7] described an analytical method to predict classification performance and an efficient approximate formula to estimate the average probability of correct classification given sensor characteristics, there is still a lack of effective tools to evaluate a fusion system performance as described in [8] and [7] in an easy and accessible way in order to make the assessment results promptly available to better control sensors and allocate system resources.

In particular, for the case of discrete reporting elements—the sensor exploitation system's estimate of target type or activity, it is more complicated to predict fusion performance for target/activity identification or situation assessment. For example, observations of the target's attributes, such as feature-modulation frequency, radar emissions characteristics, and visual image, may be used to establish target identity. They are based on a transformation between observed target attributes and a labeled identity. Methods for identity estimation involve [15] pattern recognition techniques based on clustering algorithms, neural networks, or decision-based methods such as Bayesian inference, Dempster-Shafer's method, or weighted decision techniques.

The focus of this paper is on developing a methodology and software tool to model and evaluate performance of a multi-sensor classification system. Specifically, we define a generic classification performance metric for multi-sensor fusion, called global confusion matrix (GCM), from the local sensor confusion matrices described by Bayesian network models. We then develop a stochastic simulation methodology with analytical convergence bounds to estimate the performance. Based on the methodology, a software tool is developed to help a decision maker answer the following questions: (i) What is the probability of correct classification of a given target using a specific sensor individually? (ii) What if a specific set of sensors combined together is used instead? (iii) What is the performance gain by adding another sensor to this set? and (iv) Which sensors provide a better cost/benefit ratio? We apply Bayesian network (BN) to model the relationship between target variable and various levels of observables and compute the defined performance accordingly. We assume that the BN model is given, should that be created from expert knowledge, learning from past data or any other method, and our goal is to assess its performance. In other words, the methodology is generic and independent of the source of the model. However, if the model itself is inaccurate due to limited training data or insufficient domain expertise, then we have to take into account the model uncertainty when assessing the performance. We demonstrate the types of evaluation and conclusions that can be achieved with this tool using an example model from a model-based identification (MBID) component described in [5].

The evaluation process described was implemented as an extension of a free, Java based, and open-source probabilistic network framework, UnBBayes [2–4, 10, 13]. This framework proved to be an interesting alternative since it already had Bayesian networks representation, simulation, and inference algorithms built-in, making the development of the evaluation module much easier and faster.

This paper is organized as follows. Section 2 reviews the main concepts concerning the Fusion Performance Model described in [7, 8]. Section 3 describes the technical approach used to implement the method to predict classification performance based on the FPM. Section 4 derives an analytical convergence property of the evaluation methodology and predicts number of simulation trials needed in order to achieve a desirable error bound. Section 5 shows an overview of the probabilistic network framework used, UnBBayes, to implement the evaluation module. Section 6 presents the evaluation tool and its use in several example models. Finally, Section 7 relates the main contributions of this paper and some future work.

2. MODELING CLASSIFICATION FUSION PERFORMANCE

Currently, data fusion systems are used extensively for target tracking, automated identification of targets, situation assessments, and some automated reasoning applications [15]. This paper uses a Bayesian network model that is a part of a model-based identification (MBID) component of an effort to design a decision-theoretic sensor management system. This model, described in [5], is used for incorporating target identification (ID) into a multiple-hypothesis tracking (MHT) system in a multi-sensor environment.

Bayesian networks [9] are directed acyclic graphs (DAGs), where the nodes are random variables, and the arcs specify the independence assumptions that must hold between the random variables (the arc points from the parent to the child node). These independence assumptions determine what probability information is required to specify the probability distribution among the random variables in the network.

To specify the probability distribution of a Bayesian network, one must give the prior probabilities of all root nodes (nodes with no parents) and the conditional probabilities of all other nodes given all possible combinations of their parents. Bayesian networks allow one to calculate the conditional probabilities of the nodes in the network given that the values of some of the nodes have been observed.

In addition to the convenient and flexible representation, a major benefit of using BNs is the existence of many powerful probabilistic inference algorithms, such as the distributed algorithm [16], the influence diagram algorithm [18], the evidence potential algorithm [14], simulation algorithms [11, 20], and the symbolic probabilistic inference (SPI) algorithm [19].

For MBID system mentioned earlier [5], the BN is used to relate the target states to the detected measurements at the sensors. Each evidence node represents the detected observation from a source at a given sensor. The conditional probabilities depend on the propagation from the target to the sensor, array gain, detection thresholds, etc. Other information such as relative geometry between target and sensor, the strength of the target, and the transmitted energy can also be summarized in the conditional probability of the received measurement given the target state.

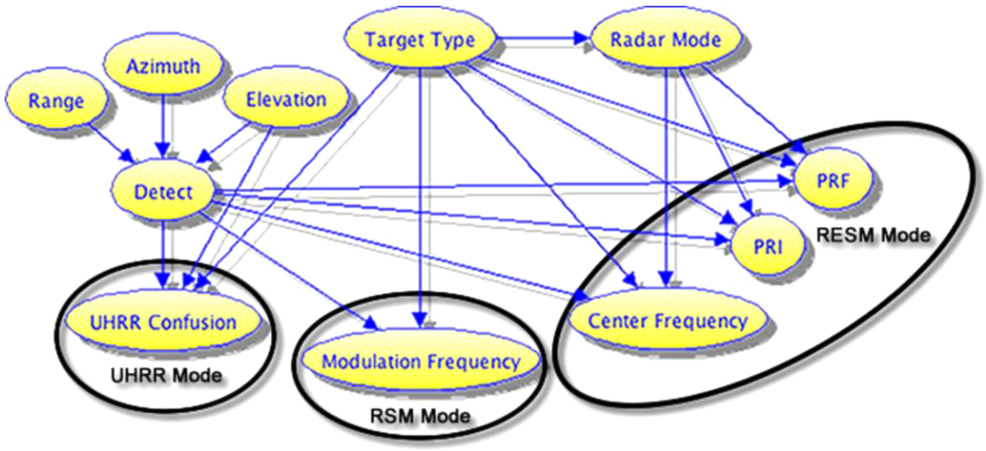


Fig. 1. A Bayesian network model example for a model-based identification (MBID) system.

It is assumed that in the system, there are two types of sensors: electronically scanning radar (ESA) and Infrared search and track (IRST). In addition to the regular search and update capabilities, the ESA radar is modeled to have three identification modes: ultra-high-resolution radar (UHRR), radar signal modulation (RSM), and radar electronic support mode (RESM). Since the radar detection and observation processes are fairly complicated and cannot be easily expressed in a simple form, a Bayesian network is used to model the processes and compute the association likelihoods as well as manipulate the target state distribution. In this system, a centralized fusion architecture is assumed, i.e., data collected from multiple sensors are pooled together in a central site where they are combined.

UHRR is an active technique and is basically an imaging technique that will be able to identify features of an airplane and therefore infer the target type. RSM is an active technique that can detect a target feature-modulation frequency. RESM is a passive technique that can observe the characteristics of the target's radar emissions. Based the observed features of the radar signal, the MBID system will infer a radar mode that will, in turn, be an evidence for a target type.

Figure 1 shows the BN that contains all three identification modules. Note that at any given moment, only one module can be active. In other words, only one type of evidence can be attached to the network. In the figure, the UHRR module is represented by the node "UHRR Confusion," RSM is represented by the node "Modulation Frequency" (MF), and RESM is represented by the remaining three nodes, "Center Frequency" (CF), "PRI," (pulse repetition interval) and "PRF" (pulse repetition frequency). Note that in all three radar modes, the observation is dependent on the probability of detection represented by the node "Detected." Probability of detection is a function of target range and aspect angle. The measurements from the three radar modes can either be discrete or continuous values. For example, the observation of UHRR is the actual target type, which can only be one of the given values. On the other hand, the observations for other radar modes have continuous values and can assume any value within the defined ranges.

Despite these qualitative notions and quantitative calculations of improved system operation by using multiple sensors and fusion processes, actual implementation of effective data fusion systems is far from simple. In practice, the combination of sensor data may actually produce worse results than could be obtained by tasking the most appropriate sensor in a sensor suite. This is caused by the attempt to combine accurate (i.e., good data) with inaccurate or biased data, especially if the uncertainties or variances of the data are unknown [15]. Before a data fusion system is deployed, developers and users need to be able to assess their fusion system performance under various conditions. This paper develops a tool based on the Fusion Performance Model (FPM) described in [8]. The focus is on predicting the classification performance. To do so, first we define the following terminologies [7].

Local Confusion Matrix (LCM): Local confusion matrices are the ones based on single sensor classification system observations. There are two types of LCMs—feature level LCM defined as the conditional probability/likelihood tables of the observable/evidence nodes given their parent nodes, and label (target ID) level LCM, $Pr(\text{Obs} \mid T = j)$, defined as the conditional probability tables of the observable/evidence given true target ID. It is the latter that will be used to obtain the global confusion matrix (GCM).

Global Confusion Matrix (GCM): Global confusion matrix is obtained based on multiple sensor classification observations with a given LCM over time. Each element in the GCM is defined as the probability of inferred[1] target class given true target class, $\text{GCM}(i, j) = \Pr(I = i \mid T = j)$. Note that the GCM is applicable to both single sensor and multiple sensor systems.

[1]The posterior probability of the target class being true given the sensor observations.

TABLE I
"Target Type" Node's Global Confusion Matrix Example

	B1	B2	R1	R2	N	Unk
B1	0.80	0.02	0.10	0.02	0.02	0.04
B2	0.02	0.70	0.02	0.14	0.08	0.04
R1	0.10	0.02	0.80	0.02	0.02	0.04
R2	0.02	0.14	0.02	0.70	0.08	0.04
N	0.02	0.15	0.04	0.15	0.60	0.04
Unk	0.04	0.04	0.04	0.04	0.04	0.80

Average Probability of Correct Classification (Pcc): Average probability of the class corresponding to the true class of the associated target. It is the average value of the diagonal elements of GCM, computed as,

$$Pcc = \frac{1}{k}\sum_i \text{GCM}(i,i) = \frac{1}{k}\text{Trace(GCM)} \qquad (1)$$

where k is the total number of target classes.

Table I illustrates a global confusion matrix example for the "Target Type" node in Fig. 1, where the columns represent the predicted class and the rows the true class.[2] Therefore, the diagonal elements represent the probabilities of correct classification. In this example, the probability of correct classification is 0.733.

3. COMPUTING CLASSIFICATION FUSION PERFORMANCE

To compute the GCM is to infer the target ID/type based on a series of sensor reports. This is similar to state-estimation, where the state of a system is estimated based on observed measurements. Similar to the Kalman filter, which allows for off-line estimation of the expected tracking performance (covariance matrix), there is a need for a systematic approach to evaluate the classification performance of a sensor, or multiple sensors.

In order to predict classification performance with a Bayesian network model, we need $\Pr(\text{Obs} \mid T = j)$, the sensor target ID level LCM. This can be done by "predicting" the observation distribution using forward inference given a target ID with either a simulation method or an exact (e.g., Junction tree [14]) algorithm depending on the network configuration. Given the LCM, the GCM can then be computed as,

$$\begin{aligned}\text{GCM}(i,j) &= \Pr(I = i \mid T = j) \\ &= \sum_{\text{Obs}} \Pr(I = i \mid \text{Obs}, T = j)\Pr(\text{Obs} \mid T = j) \\ &\approx \sum_{\text{Obs}} \Pr(I = i \mid \text{Obs})\Pr(\text{Obs} \mid T = j) \qquad (2)\end{aligned}$$

where I is the inferred target ID, T is the true target ID, and Obs is the sensor observation.[3]

[2]"B1" and "B2" are for Blue classes, "R1" and "R2" are for red classes, "N" is for neutral class, and "Unk" is for unknown class that do not belong to any other class.

[3]Note that to derive (2), a Markov chain property of T_{Obs} I is assumed.

Note that this will give us a square matrix where each row indicates that if the target $T = j$ is true, what is the probability of the sensor/classifier inferring it as $I = i$ given a single report Obs from a sensor. The performance measure can then be defined as the average correct classification probability as described earlier. When a total of n observations are reported by the sensors, the expression in (2) will need to be summed over all possible realizations of Obs, namely, an exponential enumeration of all $|\text{Obs}|^n$ possible realizations.

In general, the calculations for the elements of the GCM are computationally extensive. There are two ways to do so. One is to use the Monte Carlo approach where we randomly simulate the sensor observations based on a given BN model and recursively update the target state probability. Another way is to use analytical performance model.

Our goal here is to develop a mixed approach where analytical calculations will be performed whenever feasible. Otherwise, a stochastic simulation will be used. We have developed a very efficient polynomial-time analytical approach to approximate the GCM based on the assumption that the sensor observations are conditionally independent given the target ID [21]. Due to potentially high model complexity, we also developed an approximation method based on stochastic simulation. The algorithm is briefly summarized in Figure 2.

1. Select the target node, the corresponding evidence nodes, the condition on which we want to evaluate the BN, and the number of simulation trials to be generated from the model. Note that, in this implementation, only one target node is allowed.

2. Simulate the data based on the number of trials desired from the BN model using the stochastic sampling algorithm (see [6, 11–12, 20–21] for details). Note that in general the error of the estimated probabilities is inversely proportional to the sample size. More detail is given in Section 4.

3. Compute the approximate conditional probabilities, based on its frequency of occurrence, of the evidence nodes given target node, $\Pr(\text{Obs} \mid T = j)$.

4. Compute probabilities for predicted target ID given evidences using the probabilities computed in step 3.

$$\begin{aligned}\Pr(I = i \mid \text{Obs}) &= \frac{\Pr(\text{Obs} \mid I = i)\Pr(I = i)}{\Pr(\text{Obs})} \\ &= \frac{\Pr(\text{Obs} \mid I = i)\Pr(I = i)}{\sum_j \Pr(\text{Obs} \mid I = j)\Pr(I = j)} \qquad (3)\end{aligned}$$

where the prior probability $\Pr(I = i)$ is retrieved from the marginal distribution of the target node.

5. Finally, compute the global confusion matrix, as described in equation (2), by using the values computed in steps 3 and 4.

Fig. 2. Estimating GCM with stochastic simulation.

Some of the nodes in the network are neither target nor evidence nodes, however these "hidden" nodes are essential to describe the sensor models. For sensitivity analysis, they can be defined by "conditioning" on a set of specific values. Although the estimation error is inversely proportional to the sample size, its true value also depends on the number of nodes (evidences and target) considered and their size (number of sates). Therefore, the estimation error obtained here might be much smaller than its actual value.

4. ALGORITHM CONVERGENCE ANALYSIS

As mentioned above, due to high complexity and possible hybrid nature (mixed variables) of the model, we may have to apply approximate method to estimate $\Pr(\text{Obs} \mid T = j)$ based on stochastic simulation and subsequently compute the posterior probability according to the Bayes rule, namely, (3). To do so, one critical question to be answered is how fast the simulation algorithm converges when estimating the GCM and does it converge to the correct probability. This section derives an analytical convergence rate of the FPM algorithm and estimates the number of simulation trials needed in order to achieve a desirable accuracy level (error bounds).

With (2), let Y_{Oj} be the indicator function for estimating $\Pr(\text{Obs} \mid T = j)$, i.e., $Y_{Oj} = 1$ when "Obs" is realized given $T = j$ in a particular simulation trial and $Y_{Oj} = 0$ otherwise. Then it can be easily shown that when the sample size n is large, the probability distribution of $X_{Oj} \equiv \sum Y_{Oj}/n$ can be approximated by the normal distribution,

$$p(X_{Oj}) \sim N[X_{Oj}; \bar{X}_{Oj}, \sigma^2_{Oj}] \tag{4}$$

where $\bar{X}_{Oj} = P_{Oj} \equiv \Pr(\text{Obs} \mid T = j)$ is the mean and $\sigma^2_{Oj} = \text{var}(X_{oj}) \approx P_{Oj}(1 - P_{Oj})/n$ is the variance of the random variable X_{oj} respectively.

To estimate the probability of correct classification, as shown in (1), we only need to focus on the diagonal elements of the GCM. From (1)–(3), it can be easily shown that,

$$\begin{aligned}\text{GCM}(i,i) &= \Pr(I = i) \sum_{\text{Obs}} \frac{\Pr(\text{Obs} \mid I = i)^2}{\Pr(\text{Obs})} \\ &= \Pr(I = i) \sum_{\text{Obs}} \frac{\sigma^2_{Oi}}{\Pr(\text{Obs})} \left[\frac{\Pr(\text{Obs} \mid I = i)}{\sigma_{Oi}}\right]^2.\end{aligned} \tag{5}$$

Assuming the size of the state space of Obs is relatively large, i.e., $|\text{Obs}| \equiv m \gg 1$, and assuming that $P_{Oi} \ll 1$ and $P_O \equiv \Pr(\text{Obs}) \approx P_{Oi}$,[4] then

$$\text{GCM}(i,i) = \frac{\Pr(I = i)}{n} \sum_{\text{Obs}} \left[\frac{\Pr(\text{Obs} \mid I = i)}{\sigma_{Oi}}\right]^2. \tag{6}$$

[4]This approximation could be poor.

Let $G_{i,i}$ be the indicator function of $\text{GCM}(i,i)$, then the sample mean of $G_{i,i}$ can be approximated by,[5]

$$\bar{G}_{i,i} \approx \frac{\Pr(I = i)}{n} \sum_{\text{Obs}} \left[\frac{X_{Oi}}{\sigma_{Oi}}\right]^2 \tag{7}$$

where $Z \equiv \sum_{\text{Obs}} [X_{Oi}/\sigma_{Oi}]^2$ follows a non-central Chi-square distribution with $\bar{Z} = m + \lambda$ and $\sigma_z^2 = 2(m + 2\lambda)$, where

$$\begin{aligned}\lambda &= \sum_{\text{Obs}} \left[\frac{\bar{X}_{Oi}}{\sigma_{Oi}}\right]^2 = \sum_{\text{Obs}} \left[\frac{\bar{X}^2_{Oi}}{\bar{X}_{Oi}(1 - \bar{X}_{Oi})/n}\right] \\ &= n \sum_{\text{Obs}} \left[\frac{\bar{X}_{Oi}}{(1 - \bar{X}_{Oi})}\right]\end{aligned} \tag{8}$$

In the case when P_{Oj} is approximated by a uniform distribution, $\bar{X}_{Oi} \approx 1/m$, then $\lambda = mn/(m-1) \le mn$, and $\sigma_z^2 = 2(m + 2\lambda) \le 2m(1 + 2n)$. Therefore, with $P_i \equiv \Pr(I = i)$, the variance of $\bar{G}_{i,i}$ is,

$$\sigma^2_{\bar{G}_{i,i}} = (P_i/n)^2 \sigma_z^2 < P_i^2 \frac{2m(1 + 2n)}{n^2}. \tag{9}$$

Finally, assume $\Pr(I = i)$ is approximately uniform, since from (1), $P_{\text{CC}} = (1/k) \sum_{i=1}^{k} \text{GCM}(i,i)$, the variance of the estimate $\hat{P}_{\text{CC}}$ is $\sigma^2_{\text{CC}} = (1/k^2) \sum_{i=1,\ldots,k} \sigma^2_{\bar{G}_{i,i}}$, where k is the state space size of the ID node, then

$$\sigma^2_{\text{CC}} = \frac{1}{k^2} \sum_i (P_i/n)^2 \sigma_z^2 < \frac{2m(1 + 2n)}{k^3 n^2}. \tag{10}$$

With Chebyshev's inequality, for any $\varepsilon > 0$,

$$P_r(|P_{\text{CC}} - \hat{P}_{\text{CC}}| < \varepsilon) \ge 1 - \frac{\sigma^2_{\text{CC}}}{\varepsilon^2}. \tag{11}$$

Equations (10)–(11) provide a performance bound for the absolute error of the average correct classification probability given the observation state space size, the target state space size, and the number of simulation trials. When P_{CC} is normally distributed, then a tighter bound can be obtained as follows,

$$P_r(|P_{\text{CC}} - \hat{P}_{\text{CC}}| < \varepsilon) = \alpha \tag{12}$$

where $\varepsilon = \Phi^{-1}((1 + \alpha)/2)\sigma_{\text{CC}}$, α is the confidence level, and Φ is the CDF of the standard normal distribution.

5. PERFORMANCE EVALUATION

Recall that the main objective of the tool developed in this paper is to evaluate the fusion performance and quantify how much one gains by combining different sensors in a fusion system. Specifically, the system is designed to help a decision maker answer the following questions: (i) What is the probability of correct classification of a given target using a specific sensor individually? (ii) What if a specific set of sensors combined together is used instead? (iii) What is the performance gain by adding another sensor to this set? and (iv) Which sensors provide a better cost/benefit ratio?

[5]Likewise, this could be a rough approximation.

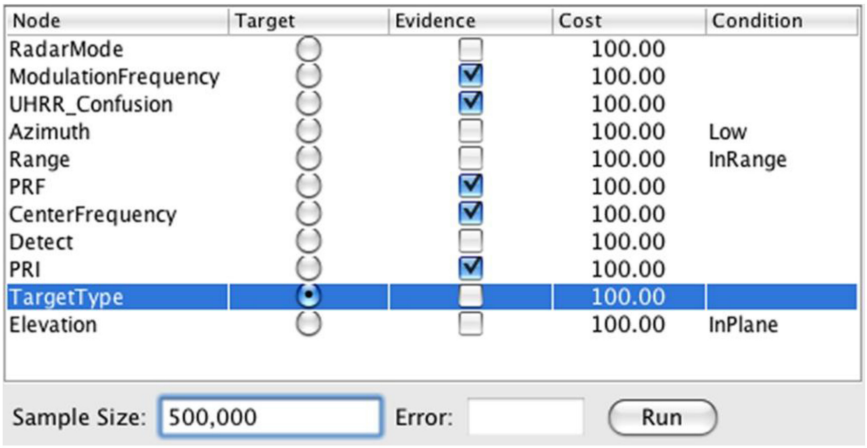


Fig. 3. UnBBayes' evaluation module input panel.

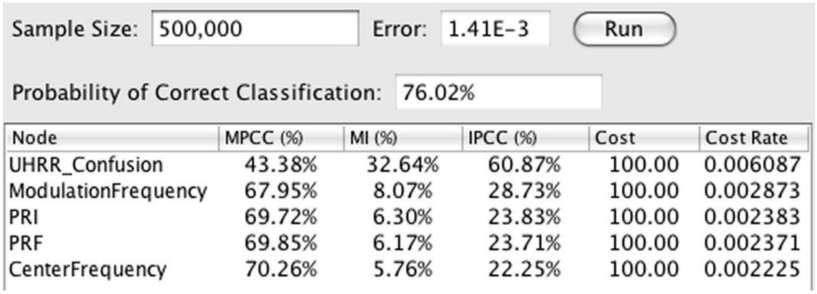


Fig. 4. UnBBayes' evaluation module output panel.

These questions can be answered by using the technical approach described in Sections 3–4. Figure 3 shows the necessary inputs that need to be specified for the FPM evaluation module. The inputs are:

Target node: Select the target node of interest.

Evidence nodes: Choose the evidence nodes, they are the sensor observables in a Multi-Sensor Classification System.

Cost: The cost associated with the evidence node. It is assigned heuristically or based on *a priori* knowledge of the cost of allocating the corresponding sensor resource.

Condition: Choose the state for the conditioning nodes, which represent an optional artificial context for sensitivity analysis. These nodes cannot be target or evidence.

Sample size: The number of trials to be generated from the model. The larger the number the more accurate the result, however the longer it will take to compute.

Error bound: In case the error bound is given, the sample size will be automatically computed based on the analysis given in Section 4.

Figure 4 presents the outputs computed in UnBBayes' evaluation module that can answer such questions. The outputs include:

GCM: The global confusion matrix computed for the selected target node and all the chosen evidence nodes.

Error: As explained in Section 4, the error can be approximately computed by equations (11) and (12).

Probability of Correct Classification (Pcc): The probability of correct classification computed from the GCM considering all evidence nodes.

Marginal PCC (MPCC): The probability of correct classification computed from the GCM given all evidence nodes other than the one presented in the row (see "Node" column).

Marginal Improvement (MI): The probability of correct classification gained by adding the node presented in the row to the rest of other nodes,

$$\text{MI} = \text{PCC} - \text{MPCC}. \tag{13}$$

Individual PCC (IPCC): The probability of correct classification computed from the LCM considering only the evidence presented in the row.

Cost Rate: The individual probability of correct classification over the cost,

$$\text{Cost Ratio} = \frac{\text{IPCC}}{\text{Cost}}. \tag{14}$$

Using the tool and its output just presented we were able to analyze the example model described in Section 2 for air target tracking and classification. Recall that in this example we have three identification modules that represent the evidence nodes. UHRR is an active technique and is basically an imaging technique that will be able to identify features of an airplane and

therefore infer the target type. RESM (represented by nodes "Center Frequency"—CF, "PRF," and "PRI") is an passive technique that can detect a target feature-modulation frequency. RSM (represented by the node "Modulation Frequency"—MF) is an active technique that can observe the characteristics of the target's radar emissions. Based the observed features of the radar signal, the MBID system will infer a radar mode that will in turn be an evidence for a target type.

In the BN model, the target type node has 6 different possible labels: [*B*1, *B*2, *R*1, *R*2, *Neutral*, *Unknown*] with a given prior probability distribution. The feature of UHRR is an identification of target type with conditional probability of a UHRR ID given the true target type. This confusion matrix is indexed by relative target elevation, which we assume is known at the time of the UHRR action. The detailed description of each node and their conditional probability tables were given in [5].

As shown in the BN model, for all identification modules, the observation is dependent on the probability of detection represented by the node "Detected." The detection probability of each target is a function of the relative geometry between the target and the sensor. In reality, the values of the kinematic states need to be assigned dynamically for each target. In the test scenario, we selected the values of three kinematic state nodes, "Range," "Azimuth," and "Elevation," such that the detection probability is approximately 0.95.

Table II shows different sets of evidence nodes used to detect the node "Target Type" using exact and an approximate (with an error strictly lower than 2%) computation. To be concise we did not include all the information computed in UnBBayes in this table, but most of them can be derived from the table. For instance, the MI for the node UHRR in the evidence set UHRR+RSM is 65.36% minus the IPCC of the RSM, which is also, in this case, the MPCC of the node UHRR in this set. So MI for UHRR = 65.36% − 29.27% = 35.69%, while the MI for RSM = 65.36% − 61.90% = 3.46%.

With the information obtained in Table II, the decision maker is able to understand how the system works and which set of sensors work better together by comparing individual performance as well as marginal improvements when more than one sensor is used. For example, with RESM (CF+PRI+PRF) alone, the Pcc is about 33%; with RSM alone, the Pcc is about 29%; and with UHRR alone, the Pcc is about 62%. With the first two together, the Pcc increases to only 43%, while with all three of them, the Pcc increases to over 71%.

Another benefit of using the tool is that the decision maker could analyze the cost/benefit ratio of each sensor resource to determine the best allocation strategy. The cost ratio can also be integrated into an automatic sensor resource management (SRM) algorithm for changing the sensor mode dynamically on a real time basis.

The Pcc is used to evaluate the performance of the model given that the model is available and assumed correct. However, it is not used to judge the accuracy of the model. To test the robustness of the FPM evaluation methodology, we evaluated the same structure model but with different parameter values, meaning we kept the nodes and arcs the same but changed the conditional probability tables (CPT) by adding some noise to them. The goal is to verify that even if the model is somewhat imprecise, we can still apply the evaluation process to obtain a reasonable result. The results in Table III show that the estimated Pcc performance is relatively insensitive to the model uncertainty. Note that the noise level in Table III represents the uncertainty magnitude in the model quantified by the random variations in percentage of the conditional probabilities.

To compare the analytical performance bounds derived in Section 4 and the simulation results, Figs. 5–7 show the relationship between sample size and the absolute estimation error given different observation state space sizes with 99% confidence bounds ($\alpha = 0.99$) predicted by (12). The target state space size is assumed to be fixed ($k = 6$). As can be seen from the figures, the theoretical analysis provides good performance bounds

TABLE II
Classification Performance Prediction with Different Evidence Sets

# Ev. Nodes	Evidence Set	Pcc Exact	Pcc 2% Error
1	RESM (CF)	22.28%	22.88%
	RESM (PRF)	23.73%	24.36%
	RESM (PRI)	23.82%	25.01%
	RSM	28.67%	29.27%
	UHRR	61.02%	61.90%
2	RESM (CF+PRI)	27.85%	29.75%
	RESM (CF+PRF)	27.81%	29.08%
	RESM (PRI+PRF)	28.91%	30.18%
	UHRR+RSM	65.48%	65.36%
3	RESM (CF+PRI+PRF)	31.72%	32.93%
	RSM+RESM (CF+PRI)	38.76%	39.85%
	RSM+RESM (CF+PRF)	38.72%	39.72%
	RSM+RESM (PRI+PRF)	39.67%	40.70%
	UHRR+RESM (CF+PRI)	65.61%	66.15%
	UHRR+RESM (CF+PRF)	65.58%	66.37%
	UHRR+RESM (PRI+PRF)	66.16%	66.39%
4	RSM+RESM (CF+PRI +PRF)	42.14%	43.30%
	UHRR+RESM (CF+PRI+PRF)	67.55%	68.54%
	UHRR+RSM+RESM (CF+PRI)	69.32%	70.14%
	UHRR+RSM+RESM (CF+PRF)	69.31%	69.99%
	UHRR+RSM+RESM (PRI+PRF)	69.78%	70.61%
5	UHRR+RSM+RESM (CF+PRI+PRF)	70.95%	71.72%

TABLE III
Pcc for Models with Different Evidence Nodes

Noise	Pcc	
	Exact	Approximate
0%	70.95%	72.73%
5%	67.82%	69.79%
10%	66.36%	68.37%

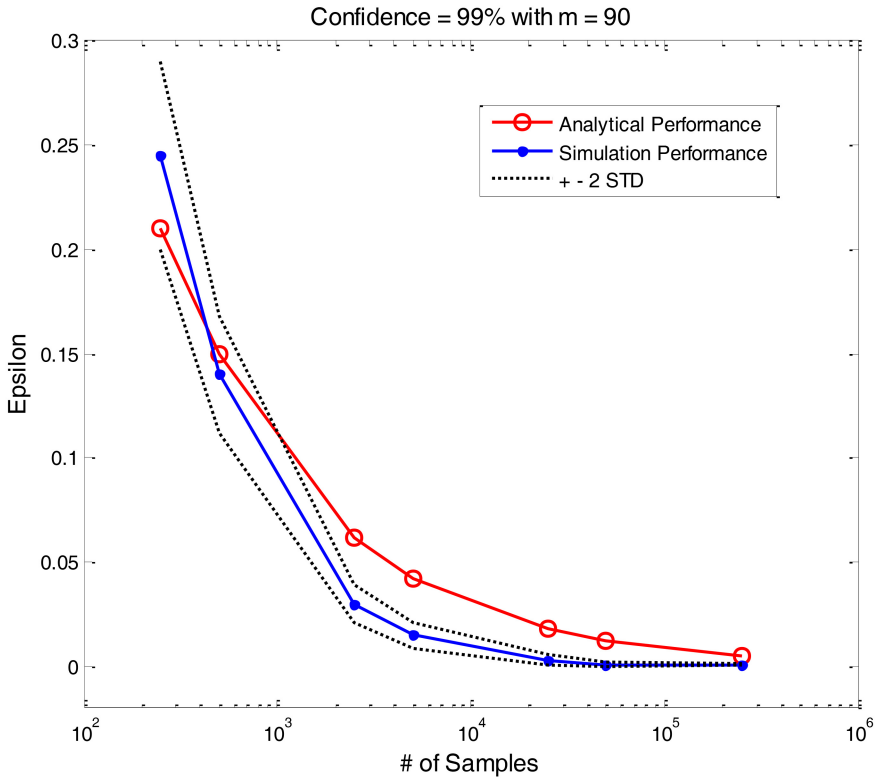


Fig. 5. Absolute error as a function of sample size with $m = 90$.

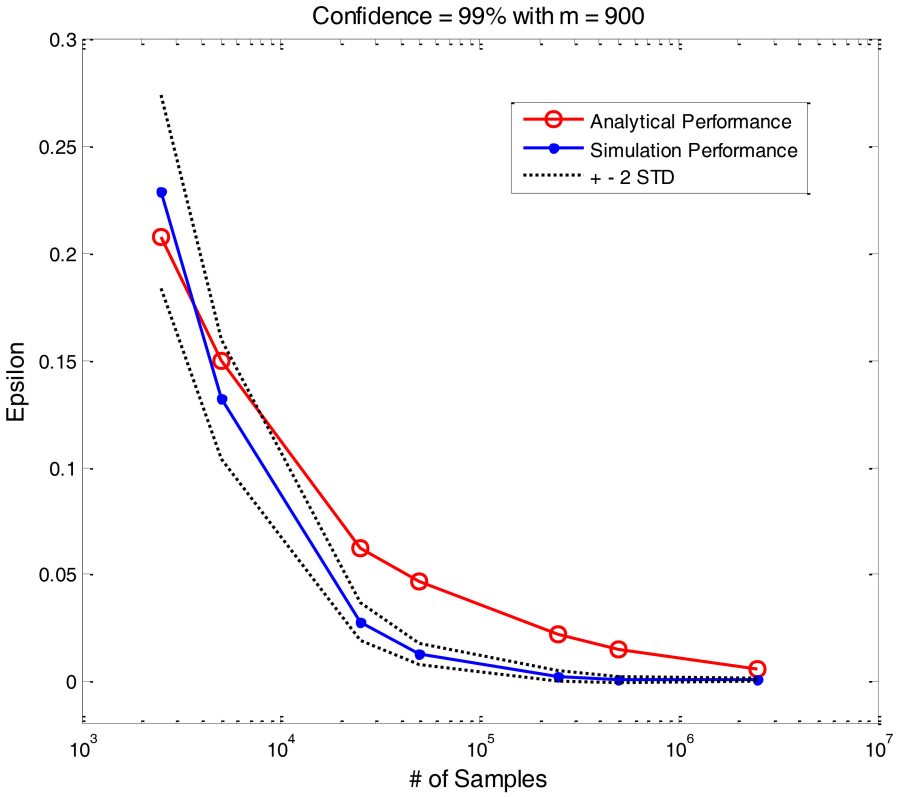


Fig. 6. Absolute error as a function of sample size with $m = 900$.

under different conditions (i.e., various observation state space sizes). The bounds work well particularly when the sample sizes are relatively small or relatively large. However, the analytical bounds are somewhat conservative otherwise due to various approximations employed in the analysis.

Finally, we evaluated the fusion performance of a classification system based on a Bayesian network

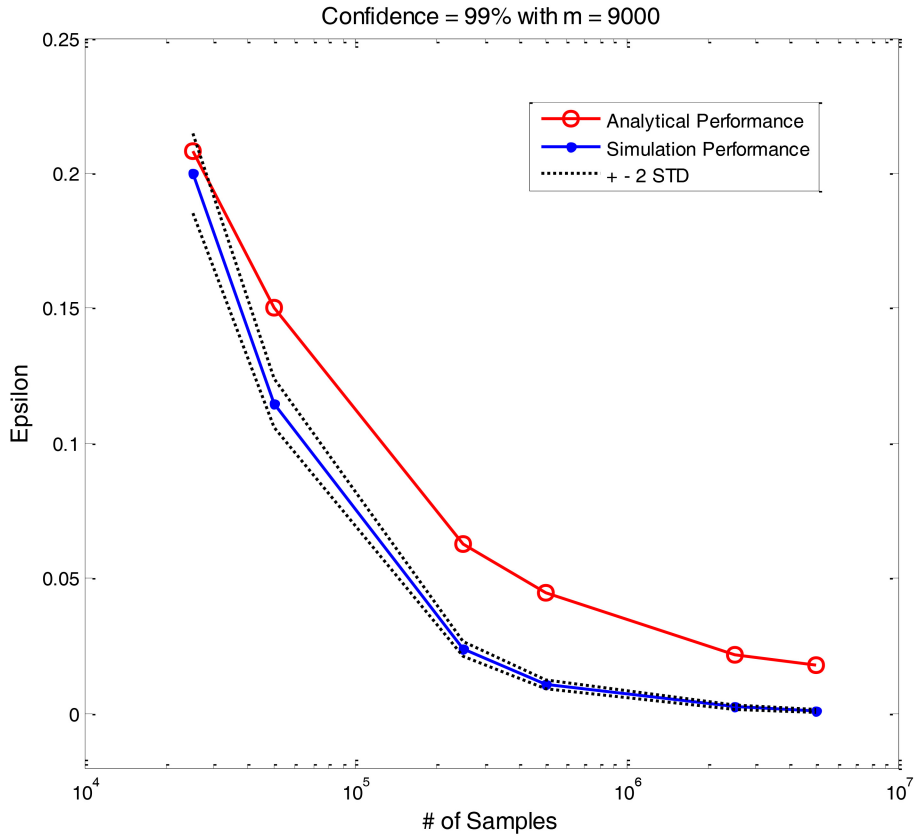


Fig. 7. Absolute error as a function of sample size with $m = 9000$.

model used for Combat ID and threat assessment described in [1] (see Fig. 8). According to [1] the model can be used in a number of ways to infer information about the Hostility, ID Platform and Threat variables.

Table IV shows the Pcc for different target nodes given all evidence nodes or a subset of them, which has the nodes IFF, ATR, ESM, Speed, OnDataLink, and Intelligence. The Pcc values were all computed with a sample size of 2.5M. If the classifications were selected at random the Pcc for Hostility, ID Platform, and Threat would be 16.66%, 12.50%, and 50.00%, respectively, since Hostility has 6 possible states, ID Platform has 8, and Threat has 2 states. Therefore, it can be seen that this model has almost the same performance for the classification of Hostility and Threat as if the classification were being selected at random. Although the Pcc for the ID Platform is slightly better, the decision maker might still consider it low. Note that, however, the correct classification performance is estimated with a single sample observation from each sensor. For multiple observations from different sensors or from a single sensor over multiple sampling times, additional analysis is required. For details, see [21].

With these two models analyzed, we can see the full benefit of using our tool for performance evaluation. We were not only able to detect when a model is useful for classification but also able to detect when it is not as efficient. Furthermore, the same model might not have a good performance for classifying a specific target node, but it might be good for classifying a different one.

6. CONCLUSIONS

In this paper, we present a tool that allows a user to evaluate the classification performance of a multi-sensor fusion system modeled by a Bayesian network. With the Fusion Performance Model (FPM) described in [7] and [8], we developed a new module and integrated it with the free, open-source, and platform independent probabilistic network framework UnBBayes.

We demonstrate the functionalities of the tool with a model-based ID example for air target tracking and classification. We were able to answer questions related to probability of correct classification of a given target using a specific individual sensor resource or a set of resources. We were also able to evaluate the marginal performance gain and cost/benefit ratio of each sensor resource. This tool is very valuable for a decision maker to analyze trade-off between performance and costs and to select proper sensor suites according to requirements and constraints. As far as we could tell, there is no other tool available for evaluating a fusion system performance as described in this paper.

We developed an analytical convergence analysis where we derived theoretical formulae to estimate the convergence rate and predict the number of simulation trials needed in order to achieve a desirable accuracy level (error bounds). We also compared the simulation

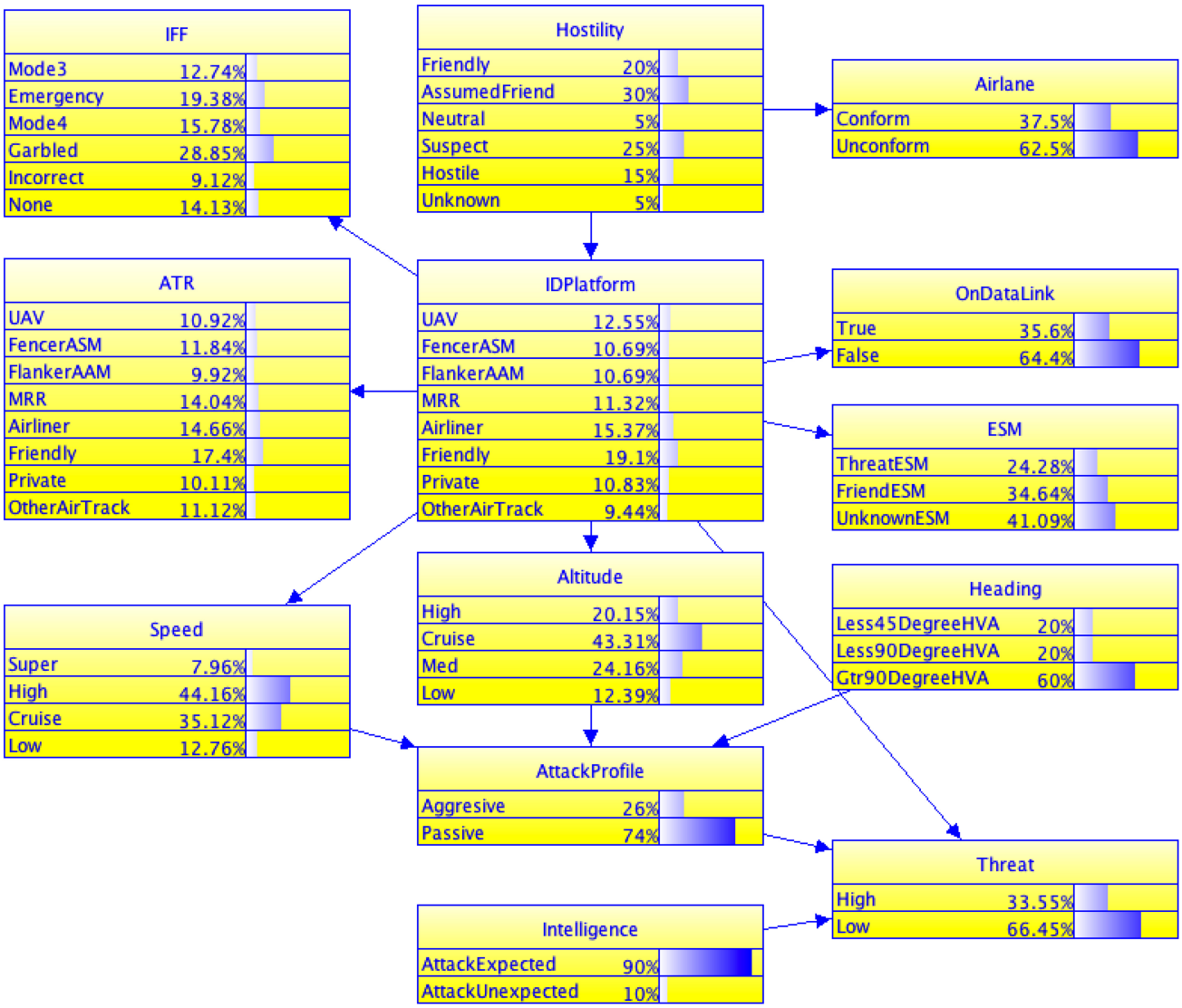


Fig. 8. A simplified model of a Bayesian network used for Combat ID and threat assessment.

results with the analytical ones and we showed that the bounds work well under different conditions.

We also showed that the tool developed is useful to identify the quality of the classification models. Moreover, the performance of the model depends on the node to be classified, i.e., the same model might not have a good performance for classifying a specific target node, but it might be good for classifying a different one.

This research was conducted as part of a larger effort to design an integrated multi-sensor tracking and ID performance evaluation system. A major advantage of using the current approach is the flexibility of modifying the Bayesian models to account for various potential environmental or sensor changes. One important future research direction is to integrate the kinematic tracking module into a combined track/ID performance evaluation system and to extend the system to accommodate for high level fusion. Additionally, we intend to incorporate other efficient analytical or simulation algorithms to improve the computational efficiency of the tool. Finally, since the model itself might be inaccurate due to limited training data or insufficient domain expertise, it is important to take into account the model uncertainty while assessing its performance.

TABLE IV
Pcc for Different Target Nodes Given all Evidence Nodes or a Subset of Them (IFF, ATR, ESM, Speed, OnDataLink, and Intelligence)

Target Node	Pcc Given Evidence Nodes	
	All	Subset
Hostility	38.99%	19.18%
ID Platform	61.35%	48.16%
Threat	65.64%	52.84%

REFERENCES

[1] P. Bladon, P. S. Day, T. Hughes, and P. Stanley
High-level fusion using Bayesian networks: Applications in command and control.
In *Information Fusion for Command Support*, Proceedings RTO-MP-IST-055, Neuilly-sur-Seine, France, 2006.

[2] R. N. Carvalho, M. Ladeira, L. L. Santos, and P. C. Costa
A GUI tool for plausible reasoning in the semantic web using MEBN.
In *Proceedings of the Seventh international Conference on Intelligent Systems Design and Applications*, pp. 381–386; ISDA, IEEE Computer Society, Rio de Janeiro, Brazil, 2007.

[3] R. N. Carvalho, M. Ladeira, L. L. Santos, S. Matsumoto, and P. C. Costa
UnBBayes-MEBN: Comments on implementing a probabilistic ontology tool.
In *Proceedings of the IADIS International Conference on Applied Computing*, Portugal, 2008, pp. 211–218.

[4] R. N. Carvalho, M. Ladeira, L. L. Santos, S. Matsumoto, and P. C. Costa
A GUI tool for plausible reasoning in the semantic web using MEBN.
Book Innovative Applications in Data Mining, DOI: 10.1007/978-3-540-88045-5_2, Springer Berlin/Heidelberg, 2009, pp. 17–45.

[5] KC. Chang and R. Fung
Target identification with Bayesian networks in a multiple hypothesis tracking system.
Optical Engineering, **36** (Mar. 1997), 684–691.

[6] KC. Chang and D. He
Inference with importance sampling for dynamic Bayesian networks.
In *Proceedings of Fusion 2005*, Philadelphia, C2-1, July 2005.

[7] KC. Chang, E. Sivaraman, and M. Liggins
Performance modeling for multisensor tracking and classification.
In *Proceedings of SPIE Defense and Security Symposium*, Vol. #5429, Orlando, FL, Apr. 2004.

[8] KC. Chang, Y. Song, and M. Liggins
Performance modeling for multisensor data fusion.
In *Proceedings of SPIE AeroSense*, Orlando, FL, Apr. 2003.

[9] E. Charniak
Bayesian network without tears.
AI Magazine, 1991, pp. 50–63.

[10] P. C. Costa, M. Ladeira, R. N. Carvalho, L. L. Santos, S. Matsumoto, and K. B. Laskey
A first-order Bayesian tool for probabilistic ontologies.
In *Proceedings of the Twenty-First International Florida Artificial Intelligence Research Society Conference*, Menlo Park, CA: The AAAI Press, 2008, pp. 631–636.

[11] R. Fung and KC. Chang
Weighting and integrating evidence for stochastic simulation in Bayesian networks.
In *Proceedings of the 5th Workshop on Uncertainty in AI*, Windsor University, Aug. 1989.

[12] M. Henrion
Propagating uncertainty in Bayesian networks by probabilistic logic sampling.
In *Proceedings of Uncertainty in Artificial Intelligence*, J. Lemmer and L. Kanal (Eds.), North-Holland, Amsterdam, 1986, pp. 317–324.

[13] M. Ladeira, D. C. Silva, M. H. P. Vieira, R. N. Carvalho, M. S. Onishi, and W. T. Silva
Platform independent and open tool for probabilistic networks (in Portuguese, "Ferramenta Aberta e Independente de Plataforma para Redes Probabil´sticas").
In *Proceedings of XXIII Congresso da Sociedade Brasileira de Computação*, Campinas, Brazil, 2003.

[14] S. L. Lauritzen and D. J. Spiegelhalter
Local computations with probabilities on graphical structures and their application in expert systems.
Journal Royal Statistical Society B, 50 (1988).

[15] M. Liggins, D. Hall and J. Llinas (Eds.)
Handbook of Multisensor Data Fusion: Theory and Practice (2nd ed.).
Sept. 2008.

[16] J. Pearl
Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference.
Morgan Kaufmann Publishers, 1988.

[17] N. Rao
Multisensor fusion under unknown distributions: Finite sample performance guarantees.
Multisensor Fusion, Kluwer Academic Publishers, 2002.

[18] R. D. Shachter
Intelligent probabilistic inference.
In *Proceedings of the Uncertainty in Artificial Intelligence*, L. N. Kanal and J. F. Lemmer (Eds.), North-Holland, Amsterdam, 1986.

[19] R. Shachter, A. D. Favero, and B. D'Ambrosio
Symbolic probabilistic inference: a probabilistic perspective.
In *Proceedings of the Association for the Advancement of Artificial Intelligence*, 1990.

[20] R. Shachter and M. Peot
Simulation approaches to general probabilistic inference on belief networks.
In *Proceedings of the 5th Workshop on Uncertainty in Artificial Intelligence at the Association for the Advancement of Artificial Intelligence*, Windsor University, Aug. 1989.

[21] E. Sivaraman and KC. Chang
Performance evaluation of multi-sensor classification systems.
IEEE Transactions on Aerospace and Electronic Systems, **43** (Oct. 2007), 1265–1281.

Rommel N. Carvalho is now a postdoctoral research associate at the Center of Excellence in Command, Control, Communications, Computing and Intelligence (C4I) at George Mason University (GMU). During the 3 years of his Ph.D., he was a graduate research assistant in the Department of Systems Engineering and Operations Research at GMU, VA. He received his Master in Computer Science and his Bachelor of Computer Science from University of Bras´lia, DF, Brazil, in 2008 and 2003, respectively. He is an Artificial Intelligence (AI) researcher with focus on uncertainty in the Semantic Web using Bayesian Inference, Data Mining, Software Engineering and Java Programming. Awarded programmer with experience in implementation of Bayesian Network systems (UnBBayes), Multi-Entity Bayesian Network and Probabilistic Web Ontology Language (PR-OWL), and various web-based applications. Rommel N. Carvalho has been working for the Brazilian Government at the Office of the Comptroller General (CGU) as an Information Technology (IT) expert since 2005. He has also done extensive research on fraud detection and prevention for the Brazilian Government and situation awareness for the U.S. Navy. In most of the systems he helped develop he was the project manager, which gave him the experience necessary to get the Project Management Professional (PMP) certificate. During his Ph.D., he has published over 15 papers, among conference and workshop papers, book chapters, journal papers, and workshop proceedings.

Kuo-Chu Chang received the M.S. and Ph.D. degrees in electrical engineering from the University of Connecticut in 1983 and 1986 respectively.

From 1983 to 1992, he was a senior research scientist in Advanced Decision Systems (ADS) division, Booz-Allen & Hamilton, Mountain View, CA. In 1992, he joined the Systems Engineering and Operations Research Department, George Mason University where he is currently a professor. His research interests include estimation theory, optimization, signal processing, and multisensor data fusion. He is particularly interested in applying unconventional techniques in the conventional decision and control systems. He has more than 25 years of industrial and academic experience and published more than one hundred and fifty papers in the areas of multitarget tracking, distributed sensor fusion, and Bayesian Networks technologies. He was an associate editor on Tracking/Navigation Systems from 1993 to 1996 and on Large Scale Systems from 1996 to 2006 for *IEEE Transactions on Aerospace and Electronic Systems*. He was also an associate editor of *IEEE Transactions on Systems, Man, and Cybernetics, Part A*, from 2002 to 2007.

Dr. Chang is a member of Etta Kappa Nu and Tau Beta Pi.

Modified Scoring in Multiple-Hypothesis Tracking

STEFANO CORALUPPI
CRAIG CARTHEL

Track-oriented multiple-hypothesis tracking is a powerful and widely-accepted methodology in multi-target tracking. We show that the target-death problem inherent in the probability hypothesis density filter does not arise in the MHT. However, the MHT suffers from a problem of its own: excessive competition for measurements from tentative tracks. We introduce a mechanism to mitigate this effect by favoring confirmed tracks in the association process. A heuristic justification for the technique is that it mitigates the sub-optimality associated with hypothesis pruning and sequential track extraction. Perhaps more convincingly, the modification to the MHT equations is provably optimal in the limiting case of cardinality tracking with unity detection probability. We show that modified-scoring MHT improves upon standard MHT in several benchmark studies.

Manuscript received April 1, 2011; revised August 12, 2011, January 12, 2012, and May 15, 2012; released for publication May 18, 2012.

Refereeing of this contribution was handled by William Blair.

Author's address: Compunetix Inc., 2420 Mosside Boulevard, Monroeville, PA 15146, E-mail: (stefano.coraluppi@compunetix.com, craig.carthel@compuneitx.com).

1. INTRODUCTION

Track-oriented multiple hypothesis tracking (MHT) is well-established as a paradigm for multi-sensor multi-target tracking. The fundamental approach includes many variants. Hypothesis-oriented MHT was first proposed by Reid [10]. The initial integer-programming formulation of the problem is due to Morefield [8]. The hybrid-state decomposition that allows for computationally-efficient track-oriented MHT is due to Kurien [7]. An efficient solution to the optimization problem required for n_{scan} hypothesis pruning via Lagrangian relaxation is due to Poore and Rijavec [9]. A linear-programming based relaxation approach to the same optimization problem was proposed independently by Coraluppi et al [3] and by Storms and Spieksma [12].

In practice, MHT implementations must limit the number of local (or track) hypotheses. This can be achieved by measurement gating, by limiting hypothesis generation, and by pruning or merging existing hypotheses. Additionally, sequential track extraction schemes are adopted in lieu of optimal (batch) track extraction [1]. These techniques, while necessary for computationally-realizable and real-time MHT processing, lead to suboptimal data association decisions and track extraction. In this paper, we show that the suboptimality can be mitigated by favoring nearly-confirmed and confirmed tracks over tentative ones in the data-association process with suitable modification to the MHT track scoring equations.

This paper is organized as follows. In Section 2, we summarize the hybrid-state derivation of the track-oriented MHT, with some modifications with respect to the original derivation [7]. In Section 3, we address briefly the *target-death problem* that arises in *probability hypothesis density* (PHD) filtering as discussed in Erdinc et al [6], and show that it does not arise in track-oriented MHT. In Section 4, we introduce the modified-scoring MHT equations and considering a limiting case of the general tracking problem that we call *cardinality tracking*. Section 5 provides simulation results that demonstrate the improved performance of modified-scoring MHT over standard MHT. Concluding remarks are in Section 6.

2. MULTIPLE-HYPOTHESIS TRACKING

A key challenge in multi-sensor multi-target tracking is *measurement origin uncertainty*. That is, unlike a classical nonlinear filtering problem, we do not know how may objects are in the surveillance region, and which measurements are to be associated. New objects may be born in any given scan, and existing objects may die.

We assume that for each sensor scan, contact-level (or detection-level) data is available, in the sense that signal processing techniques are applied to raw sensor data yielding contacts for which the detection and localization statistics are known. We are interested in a scan-based (or real-time) approach that, perhaps with

some delay, yields an estimate of the number of objects and corresponding object state estimates at any time.

Several approaches to contact-level scan-based tracking exist. In this section, we employ a hybrid-state formalism to describe the track-oriented multiple-hypothesis tracking approach. Our approach follows closely the one introduced in [7]. We assume Poisson distributed births at each scan with mean λ_b, Poisson distributed false returns with mean λ_{fa}, object detection probability p_d, object death or termination probability p_χ at each scan. (We neglect the *time-dependent* nature of birth and death probabilities as would ensue from an underlying continuous-time formulation, and we neglect as well inter-scan birth and death events.)

We have a sequence of sets of contacts $Z^k = (Z_1, \ldots, Z_k)$, and we wish to estimate the state history X^k for all objects present in the surveillance region. X^k is compact notation that represents the state trajectories of targets that exist over the time sequence $(t_1, \ldots, t_k)$. Note that each target may exist for a subset of these times, with a single birth and a single death occurrence i.e. targets do not reappear. We introduce the auxiliary discrete state history q^k that represents a full interpretation of all contact data: which contacts are false, how the object-originated ones are to be associated, and when objects are born and die. There are two fundamental assumptions of note. The first is that there are no target births in the absence of a corresponding detection, i.e. we do not reason over new, undetected objects. The second is that there is *at most* one contact per object per scan.

We are interested in the probability distribution $p(X^k \mid Z^k)$ for object state histories given data. This quantity can be obtained by conditioning over all possible auxiliary states histories q^k.

$$\begin{aligned} p(X^k \mid Z^k) &= \sum_{q^k} p(X^k, q^k \mid Z^k) \\ &= \sum_{q^k} p(X^k \mid Z^k, q^k) p(q^k \mid Z^k). \end{aligned} \tag{1}$$

A pure MMSE approach would yield the following:

$$\hat{X}_{\text{MMSE}}(Z^k) = E[X^k \mid Z^k] = \sum_{q^k} E[X^k \mid Z^k, q^k] p(q^k \mid Z^k). \tag{2}$$

The track-oriented MHT approach is a mixed MMSE/MAP one, whereby we identify the MAP estimate for the auxiliary state history q^k, and identify the corresponding MMSE estimate for the object state history X^k conditioned on the estimate for q^k.

$$\hat{X}(Z^k) = E[X^k \mid Z^k, \hat{q}_k] \tag{3}$$

$$\hat{q}^k = \hat{q}_{\text{MAP}}(Z^k) = \arg\max p(q^k \mid Z^k). \tag{4}$$

The MHT recursion. Each feasible q^k corresponds to a global hypothesis. (The set of global hypotheses is generally constrained via measurement gating and hypothesis generation logic.) We are interested in a *recursive* and *computationally efficient* expression for $p(q^k \mid Z^k)$ that lends itself to maximization without the need for explicit enumeration of global hypotheses. We do so through repeated use of Bayes' rule. Note that $f(\cdot)$ denotes the probability density function and $p(\cdot)$ denotes the probability mass functions. The normalizing constant c_k does not impact MAP estimation.

$$\begin{aligned} p(q^k \mid Z^k) &= \frac{f(Z_k \mid Z^{k-1}, q^k) p(q^k \mid Z^{k-1})}{c_k} \\ &= \frac{f(Z_k \mid Z^{k-1}, q^k) p(q_k \mid Z^{k-1}, q^{k-1}) p(q^{k-1} \mid Z^{k-1})}{c_k} \end{aligned} \tag{5}$$

$$c_k = f(Z_k \mid Z^{k-1}) = \sum_{q^k} f(Z_k \mid Z^{k-1}, q^k) p(q^k \mid Z^{k-1}). \tag{6}$$

Recall that we assume that in each scan the number of target births is Poisson distributed with mean λ_b, the number of false returns is Poisson distributed with mean λ_{fa}, targets die with probability p_χ, and targets are detected with probability p_d. The recursive expression (5) involves two factors that we consider in turn.

Computation of $p(q_k \mid Z^{k-1}, q^{k-1})$. It will be useful to introduce the aggregate variable ψ_k (consistent with the approach in [7]) that accounts for the number of detections d for the τ existing tracks, the number of track deaths χ, the number of new tracks b, and the number of false returns $r - d - b$, where r is the number of contacts in the current scan.

$$p(q_k \mid Z^{k-1}, q^{k-1}) = p(\psi_k \mid Z^{k-1}, q^{k-1}) p(q_k \mid Z^{k-1}, q^{k-1}, \psi_k) \tag{7}$$

$$\begin{aligned} p(\psi_k \mid Z^{k-1}, q^{k-1}) = &\left\{ \binom{\tau}{\chi} p_\chi^\chi (1 - p_\chi)^{\tau - \chi} \right\} \\ &\cdot \left\{ \binom{\tau - \chi}{d} p_d^d (1 - p_d)^{\tau - \chi - d} \right\} \\ &\cdot \left\{ \frac{\exp(-p_d \lambda_b) p_d^b \lambda_b^b}{b!} \right\} \\ &\cdot \left\{ \frac{\exp(-\lambda_{fa}) \lambda_{fa}^{r-d-b}}{(r-d-b)!} \right\} \end{aligned} \tag{8}$$

$$p(q_k \mid Z^{k-1}, q^{k-1}, \psi_k) = \frac{1}{\binom{\tau}{\chi} \binom{\tau - \chi}{d} \left(\frac{r!}{(r-d)!} \right) \binom{r-d}{b}}. \tag{9}$$

Substituting (8–9) into (7) and simplifying yields the following.

$$p(q_k \mid Z^{k-1}, q^{k-1}) = \left\{ \frac{\exp(-p_d\lambda_b - \lambda_{fa})\lambda_{fa}^r}{r!} \right\} \cdot p_\chi^\chi((1-p_\chi)(1-p_d))^{\tau-\chi-d} \cdot \left(\frac{(1-p_\chi)p_d}{\lambda_{fa}} \right)^d \left(\frac{p_d\lambda_b}{\lambda_{fa}} \right)^b. \tag{10}$$

Computation of $f(Z_k \mid Z^{k-1}, q^k)$. This quantity is given by (11), where $Z_k = \{z_j, 1 \le j \le r\}$, J_d is the set of measurements associated with detections of existing tracks, J_b is the set of measurements associated with target births, J_{fa} is the set of measurements hypothesized as false, $|J_d| + |J_b| + |J_{fa}| = r$, and the factors on the R.H.S. are derived from filter innovations, filter initiations, and the false contact distribution (generally uniform over measurement space).

$$f(Z_k \mid Z^{k-1}, q^k) = \prod_{j \in J_d} f_d(z_j \mid Z^{k-1}, q^k) \cdot \prod_{j \in J_b} f_b(z_j \mid Z^{k-1}, q^k) \prod_{j \in J_{fa}} f_{fa}(z_j \mid Z^{k-1}, q^k). \tag{11}$$

For example, in the linear Gaussian case, $f_d(z_j \mid Z^{k-1}, q^k)$ is a Gaussian residual, i.e. it is the probability of observing z_j given a sequence of preceding measurements. If there is no prior information on the target, $f_b(z_j \mid Z^{k-1}, q^k)$ is generally the value of the uniform density function over measurement space. Similarly, $f_{fa}(z_j \mid Z^{k-1}, q^k)$ is as well usually taken to be the value of the uniform density function over measurement space, under the assumption of uniformly distributed false returns. Note that the expressions given here are general and allow for quite general target and sensor models.

Final form of the MHT recursion. Substituting (10–11) into (5) and simplifying results in (12–13). This expression is the key enabler of track-oriented MHT. In particular, it provides a recursive expression for $p(q^k \mid Z^k)$ that consists of a number of factors that relate to its constituent local track hypotheses.

$$p(q^k \mid Z^k) = p_\chi^\chi((1-p_\chi)(1-p_d))^{\tau-\chi-d} \cdot \prod_{j \in J_d} \left[\frac{(1-p_\chi)p_d f_d(z_j \mid Z^{k-1}, q^k)}{\lambda_{fa} f_{fa}(z_j \mid Z^{k-1}, q^k)} \right] \cdot \prod_{j \in J_b} \left[\frac{p_d\lambda_b f_b(z_j \mid Z^{k-1}, q^k)}{\lambda_{fa} f_{fa}(z_j \mid Z^{k-1}, q^k)} \right] \frac{p(q^{k-1} \mid Z^{k-1})}{\bar{c}_k} \tag{12}$$

$$\bar{c}_k = \frac{c_k}{\left\{ \frac{\exp(-p_d\lambda_b - \lambda_{fa})}{r!} \lambda_{fa}^r \right\} \prod_{j \in J_d \cup J_b \cup J_{fa}} f_{fa}(z_j \mid Z^{k-1}, q^k)}. \tag{13}$$

Note that the constant $\bar{c}_k$ normalizes the recursion with respect to the case in which all returns in the current scan are false. That is, for the case $b = 0$ (no births) and $\tau = d = \chi = 0$ (no current tracks, and correspondingly no detections or terminations on current tracks), we have

$$p(q^k \mid Z^k) = \left\{ \frac{\exp(-p_d\lambda_b - \lambda_{fa})}{r!} \lambda_{fa}^r \right\} \cdot \prod_{j \in J_d \cup J_b \cup J_{fa}} f_{fa}(z_j \mid Z^{k-1}, q^k) \frac{p(q^{k-1} \mid Z^{k-1})}{c_k}.$$

That is, the denominator in (13) is precisely the product of the probability of no detected births, i.e. $\exp(-p_d\lambda_b) \cdot p_d^0\lambda_b^0/0!$, the probability of r false alarms, i.e. $(\exp(-\lambda_{fa})/r!)\lambda_{fa}^r$, and the filter residuals associate with all measurements being false.

An implicit reduction in the set of hypotheses in (12–13) is that target births are assumed to occur only in the presence of a detection (i.e. there is no reasoning over un-detected births). Correspondingly, the factor p_d reduces the effective birth rate to $p_d\lambda_b$ (though surprisingly the factor is absent in [7]). Further, in the first scan of data, it would be appropriate to replace $p_d\lambda_b$ by $p_d\lambda_b/p_\chi$ to account properly for the steady-state expected number of targets. (More generally, target birth and death parameters should reflect sensor scan rates, as the underlying target process is defined in continuous time.) Further reduction in the set of hypotheses is generally achieved via *measurement gating* procedures [1]. Finally, for a given track hypothesis, one usually applies rule-based spawning of a missed detection or termination hypothesis, but not both (e.g. only spawn a missed detection hypothesis until a sufficiently-long sequence of missed detections is reached).

One cannot consider too large a set of scans before pruning or merging local (or track) hypotheses in some fashion. A popular mechanism to control these hypotheses is n_{scan} *pruning*. This amounts to solving (4), generally by a relaxation approach to an integer programming problem [3, 8–9, 12], followed by pruning of all local hypotheses that differ from $\hat{q}^k$ at a depth of n_{scan}. That is, all remaining global hypotheses are identical up to scan $k - n_{\text{scan}}$. Note that, if one were to set $n_{\text{scan}} = 0$, this would amount to immediate resolution of association hypotheses up to the current time. The n_{scan} pruning methodology is applied after each new scan of data is received, resulting in a fixed-delay solution to the tracking problem.

Often, n_{scan} pruning is referred to as a *maximum likelihood* (ML) approach to hypothesis management.

ML estimation is closely related to *maximum a posteriori* (MAP) estimation. In particular, we have:

$$\hat{X}_{\text{MAP}}(y) = \arg\max f(y \mid X) f(X) \tag{14}$$

$$\hat{X}_{\text{ML}}(y) = \arg\max f(y \mid X). \tag{15}$$

Note that ML estimation is a *non-Bayesian* approach as it does not rely on a prior distribution on X. ML estimation can be interpreted as MAP estimation with a uniform prior. In the track-oriented MHT setting, n_{scan} pruning relies on a single parent global hypothesis, thus the ML and MAP interpretations are both valid.

Once hypotheses are resolved, in principle one has a state of object histories given by $\hat{X}(Z^k)$. In practice, it is common to apply track confirmation and termination logic to all object histories [1]. A justification for this is that it provides a mechanism to remove spurious tracks induced by the sub-optimality inherent in practical MHT implementations that include limited hypothesis generation and hypothesis pruning or merging. Further, sequential track extraction allows for real-time processing which optimal (batch) track extraction would not.

Given the use of post-association track confirmation and termination logic, a reasonable simplification that is pursued in [3] is to employ equality constraints in the data-association process, which amounts to accounting for all contact data in the resolved tracks. Spurious tracks are subsequently removed in the track-extraction stage.

To summarize, at each stage of processing, track-oriented MHT maintains a set of track trees with depth n_{scan}. When a scan of measurements is received, each measurement is compared with each (local) track hypothesis, and a new level of leaf nodes is created. All track hypotheses continue as well in the absence of a measurement. Additionally, each measurement defines the root of a new track tree. Following hypothesis generation, the MAP global hypothesis is determined via a linear programming relaxation approach [3]. Correspondingly, the set of track trees is pruned so that a single global hypothesis exists at depth $n_{\text{scan}} + 1$. The process then repeats for the next scan. (If one were to set n_{scan}, the procedure reverts to a standard 2D assignment solution.)

Data association is followed by track extraction. Tentative tracks are reported at the tracker output only once a suitable track-confirmation criterion is achieved. Similarly, once a track has degraded sufficiently (or once it is determined that a still-tentative track cannot achieve the confirmation criterion), the track is terminated. This information flows back to the data association module from the track extraction module, invalidating subsequent association hypotheses for the terminated track.

3. THE TARGET-DEATH PROBLEM

A useful re-interpretation of the *probability hypothesis density* (PHD) filter, known as the bin-occupancy filter, is given in [6]. This paper describes as well the target death problem that the authors had earlier identified, and which in turn has led to the *cardinalized* PHD (CPHD) filter.

Consider the single-target case with no false alarms. In the absence of a target measurement, the PHD surface follows (16). Note that the PHD surface $D_{k|k}(x)$ at each time t_k is a function of all data received up to t_k and is computed recursively. $D_{k|k}(x)$ admits the interpretation that it identifies the probability of target presence at a given state.

$$D_{k|k}(x) = (1 - p_d(x)) D_{k|k-1}(x). \tag{16}$$

While (16) may appear reasonable, it can be shown that it is inconsistent with the following simple Bayesian argument. Let Y_{k-1} be the existence state for the target at scan $k-1$, and assume that the death probability at any scan is given by p_χ, as before. The updated probability of existence after a missed detection is given by (17).

$$\begin{aligned} p(Y_k = 1 \mid |Z_k| = 0) &= p(Y_k = 1 \mid Y_{k-1} = 1, |Z_k| = 0) p(Y_{k-1} = 1) \\ &= \frac{p(Y_k = 1, |Z_k| = 0 \mid Y_{k-1} = 1) p(Y_{k-1} = 1)}{p(|Z_k| = 0 \mid Y_{k-1} = 1)} \\ &= \frac{(1 - p_\chi)(1 - p_d)}{1 - (1 - p_\chi) p_d} p(Y_{k-1} = 1). \end{aligned} \tag{17}$$

Comparing (16) and (17), we see that the PHD filter penalizes missed detections too heavily; it is claimed in [6] that the CPHD appears to follow (17).

What happens with the track-oriented MHT approach? We compare the ratio of the probability associated with the track *coast* hypothesis (track is alive in the absence of a measurement) with the probability of track *coast* or *death*. That is, in the numerator we want the case "no detection and target alive," and in the denominator we want the case "no detection (target alive or dead." Let q_i^k and q_j^k denote global hypotheses that include coast and death events, respectively, for the target of interest. From (12), we see that (18) follows immediately. Indeed, all factors in the global hypothesis probability cancel except for those associated with the (undetected) track.

$$\begin{aligned} \frac{p(q_i^k \mid Z^k)}{p(q_i^k \mid Z^k) + p(q_j^k \mid Z^k)} &= \frac{(1 - p_\chi)(1 - p_d)}{(1 - p_\chi)(1 - p_d) + p_\chi} \\ &= \frac{(1 - p_\chi)(1 - p_d)}{1 - (1 - p_\chi) p_d}. \end{aligned} \tag{18}$$

Note that this validation is quite general, and in particular it is directly applicable to the multi-target case, under the assumption that no contacts satisfy the hypothesis gating criterion for the (undetected) object of interest here. We conclude that track-oriented MHT properly handles missed detections, and no target-death problem is observed.

4. MODIFIED-SCORING MHT AND CARDINALITY TRACKING

In scan-based processing, assume that a track hypothesis is *confirmed* when it achieves an M-of-N criterion, with the start of the (tentative) track defined with the first of the relevant M measurements, and that tentative tracks that have no chance to achieve the M-of-N criterion are discarded. Also, a track hypothesis is *terminated* if K missed detections are exceeded. Note that, under multiple-hypothesis processing, a confirmed track may be pruned under a hypothesis-reduction scheme such as n_{scan} pruning. Following hypothesis resolution, a single global hypothesis exists that is composed of a number of *resolved* tracks. Next, the set of resolved tracks undergoes a *track extraction* process based on the same M, N, and K parameters. With these notions of confirmed, resolved and extracted tracks, we now introduce a modification to (12) that will prove useful. In particular, confirmation reward factors $\xi_2 > \xi_1 > 1$ are applied to track updates for *confirmed* and *nearly-confirmed* track hypotheses. The later refers to tentative tracks that reach confirmation in the current scan. The measurement sets for confirmed, nearly-confirmed, and tentative tracks are denoted by J_d, J_c, and J_t, respectively.

$$\begin{aligned} p(q^k \mid Z^k) = {} & p_\chi^\chi((1-p_\chi)(1-p_d))^{\tau-\chi-d} \\ & \cdot \prod_{j\in J_t}\left[\frac{(1-p_\chi)p_d f_d(z_j \mid Z^{k-1},q^k)}{\lambda_{fa} f_{fa}(z_j \mid Z^{k-1},q^k)}\right] \\ & \cdot \prod_{j\in J_c}\left[\frac{(1-p_\chi)\xi_1 p_d f_d(z_j \mid Z^{k-1},q^k)}{\lambda_{fa} f_{fa}(z_j \mid Z^{k-1},q^k)}\right] \\ & \cdot \prod_{j\in J_d}\left[\frac{(1-p_\chi)\xi_2 p_d f_d(z_j \mid Z^{k-1},q^k)}{\lambda_{fa} f_{fa}(z_j \mid Z^{k-1},q^k)}\right] \\ & \cdot \prod_{j\in J_b}\left[\frac{p_d \lambda_b f_b(z_j \mid Z^{k-1},q^k)}{\lambda_{fa} f_{fa}(z_j \mid Z^{k-1},q^k)}\right] \frac{p(q^{k-1} \mid Z^{k-1})}{\bar{c}'_k}. \end{aligned} \tag{19}$$

We refer to *standard MHT* as solution to (4) based on (12–13) with a fixed hypothesis tree depth (n_{scan}). We refer to *modified-scoring MHT* as the solution to (4) based on (13, 19) with a fixed hypothesis tree depth (n_{scan}). Note that the normalization factor $\bar{c}'_k$ in (19) differs slightly from the normalization factor in (12), since the track hypothesis scores have been modified with the confirmation reward factors.

The use of the reward factors $\xi_2 > \xi_1 > 1$ amounts to favoring confirmed and nearly-confirmed tracks in the association process. While this appears reasonable in (sub-optimal) MHT processing, we provide justification for the procedure on two grounds: (1) optimality of modified-scoring MHT in the limiting case of the tracking problem known as *cardinality tracking*; (2) simulation results for the general case. We address (1) next, while (2) is treated in Section 5.

Let us consider now the case where measurements are not informative with respect to target state: we are given only a sequence of cardinality measurements. Assume we are given birth, death, detection and false alarm statistics as well as a sequence that specifies the number of measurements received. An example might be (1, 2, 3, 3, 1…). We must decide how many targets there are as a function of time.

Note that all filter residuals in (12) are identical, leading to (20); correspondingly, (19) leads to (21) with c_1 the number of nearly-confirmed tracks and c_2 is the number of confirmed tracks. In this context, note that measurement gating is not a meaningful concept as all track updates are equivalent. Then, *cardinality tracking* involves identifying the sequence of target cardinalities $|X|^k$ given the sequence of measurement cardinalities $|Z|^k$.

$$p(q^k \mid Z^k) = \frac{p_\chi^\chi (1-p_\chi)^{\tau-\chi}(1-p_d)^{\tau-\chi-d} p_d^{b+d} \lambda_b^b}{\lambda_{fa}^{d+b}} \frac{p(q^{k-1} \mid Z^{k-1})}{\bar{c}_k} \tag{20}$$

$$\begin{aligned} p(q^k \mid Z^k) = {} & \frac{p_\chi^\chi (1-p_\chi)^{\tau-\chi}(1-p_d)^{\tau-\chi-d} \xi_1^{c_1} \xi_2^{c_2} p_d^{b+d} \lambda_b^b}{\lambda_{fa}^{d+b}} \\ & \cdot \frac{p(q^{k-1} \mid Z^{k-1})}{\bar{c}_k}. \end{aligned} \tag{21}$$

For purposes of the ensuing analysis, it is useful to introduce some assumptions regarding the parameters in (20) so that the form of the optimal solution to (4) yields a reasonable structure as explained below. It will be useful to represent a tracking solution X^k in a compact manner, where each track is represented as a sequence of existence states, with 1 denoting measurement update, 0 denoting existence with no measurement update i.e. a track *coast*, and x denoting non-existence. For example, $(x,x,1,1,0)$ represents a track that exists beginning with the third sensor scan, involves two measurements and one track coast and then terminates.

As a reminder, the assumptions below apply only to the cardinality-tracking problem.

Assumption 1 (preference for longer tracks).

$$\frac{p_d}{\lambda_{fa}}(1-p_\chi) > 1.$$

Consider $|Z|^k = \{1,1\}$. Assumption 1 insures that solution $\bar{X}^k = \{(1)\}$ or $\bar{X}^k = \{(x,1)\}$ has lower probability than $\tilde{X}^k = \{(1,1)\}$, i.e. $p(\tilde{X}^k \mid Z^k) > p(\bar{X}^k \mid Z^k)$.

Assumption 2 (singleton tracks discarded).

$$\lambda_b p_\chi \frac{p_d}{\lambda_{fa}} < 1.$$

Consider $|Z|^k = \{1\}$. Assumption 2 insures that solution $\bar{X}^k = \{(1)\}$ has lower probability than $\tilde{X}^k = \emptyset$, i.e. $p(\tilde{X}^k \mid Z^k) > p(\bar{X}^k \mid Z^k)$.

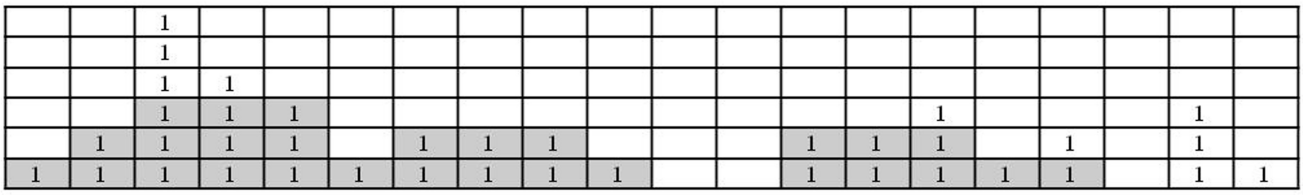


Fig. 1. For $|Z|^k = \{1,2,6,4,3,1,2,2,2,1,0,0,2,3,2,1,2,0,3,1\}$, a Tetris solution with parameter 3 is illustrated above, for which $|X|^k = \{1,2,3,3,3,1,2,2,2,1,0,0,2,3,2,1,1,0,0,0\}$. Measurements that are part of the solution are denoted by grey cells.

…	0	1_A	1_B	0	…
…	1_C	1_A	0	…	…

Fig. 2. A violation of Tetris structure.

Assumptions 1–2 imply the following (preference for track association). $\lambda_b p_\chi < 1 - p_\chi$.

Consider $|Z|^k = \{1,1\}$. The above inequality insures that solution $\bar{X}^k = \{(1),(x,1)\}$ has lower probability than $\tilde{X}^k = \{(1,1)\}$, i.e. $p(\tilde{X}^k \mid Z^k) > p(\bar{X}^k \mid Z^k)$.

Note that Assumptions 1–2 place limits on allowable clutter rates for non-empty optimal tracking solutions; the interested reader is referred to [2] for further discussion of this issue.

We consider now a special case of cardinality tracking ($p_d = 1$) for which a number of results can be established. First, we define a *Tetris solution* to be the solution obtained with sequential track extraction maximizing track length with contiguous sequences of measurements. The Tetris solution is parameterized by a minimum track length parameter, such that tracks shorter than a specified threshold are extracted. The solution is best described by illustration: see Fig. 1.

Result 1 (structure of optimal solution). Let $p_d = 1$. An optimal solution to (4) is given by the Tetris solution with minimum track length parameter

$$k_0 = \min_i \left\{ i, \frac{\lambda_b (1 - p_\chi)^{i-1} p_\chi}{\lambda_{fa}^i} \geq 1 \right\}.$$

Indeed, since tracks with length less than k_0 contribute a score less than unity to the posterior probability, a Tetris solution with parameter less than k_0 is not optimal. Similarly, a Tetris solution with parameter greater than k_0 will not include tracks that contribute a score greater than unity to the posterior probability. Thus, such a Tetris solution is not optimal either. It remains to show that a non-Tetris solution cannot outperform the Tetris solution with parameter k_0. Assume a non-Tetris optimal solution exists, and that it cannot be re-expressed as a Tetris solution by a re-ordering of entire rows (else the solution is equivalent to a Tetris one). In particular, the non-Tetris solution must contain two (possibly non-neighboring) row portions that are as shown in Fig. 2, where each cell denotes a sequence of zeros or ones of arbitrary dimension.

We now show that dropping 1_B to the lower row (i.e. partial row reordering) yields a posterior probability that is equal or higher. Indeed, if 1_A is a track of length k_0 or greater, the two posterior probabilities are the same (Fig. 3-top). If 1_A is a track of length less than k_0, the solution with the top-row 1_A replaced by zeros has larger posterior probability (Fig. 3-bottom). Thus, by a sequence of steps of this kind, we recover a Tetris structure. This shows that a non-Tetris solution cannot outperform the optimal Tetris one.

…	0	1_A	0	0	…
…	1_C	1_A	1_B	…	…

…	0	0	0	0	…
…	1_C	1_A	1_B	…	…

Fig. 3. Equivalent solution (top) or improved solution in the case of a short track 1_A (bottom).

Result 2 (optimality of modified MHT). The modified MHT solution with $M = k_0$, $N = k_0$, $K = 0$ and $n_{\text{scan}} \geq k_0 - 3$ is optimal.

Result 2 is best illustrated by example. First, assume that the target and sensor parameters are such that $k_0 = 3$ in Result 1. According to Result 2, modified MHT with $n_{\text{scan}} \geq 0$ is optimal. Consider the measurement sequence $|Z|^k = \{1,2,1\}$. With standard MHT with arbitrary n_{scan}, one obtains either the set of tentative track $X^k_{\text{tentative}} = \{(1,1,1)\}$ or $X^k_{\text{tentative}} = \{(1,1),(x,1,1)\}$. Indeed, there is no preference in terms of posterior probabilities in associating the measurement in the third scan with the longer or shorter tentative track. Correspondingly, after track extraction, one obtains either $X^k = \{(1,1,1)\}$ or $X^k = \emptyset$. The cardinality-tracking result is thus either $|X|^k = \{1,1,1\}$ or $|X|^k = \{0,0,0\}$. With modified-scoring MHT, one is guaranteed that the measurement in the third scan is associated to the tentative track of length two; indeed, this track is nearly-confirmed, and $\xi_1 > 1$ in (21) insures that the solution to (4) yields $|X|^k = \{1,1,1\}$. Thus, modified-scoring MHT achieves optimality while standard MHT is not guaranteed to do so.

Next, assume once again that $k_0 = 3$ in Result 1 and consider the measurement sequence $|Z|^k = \{1,2,2,1\}$. By similar reasoning, we see that standard MHT results in either $X^k_{\text{tentative}} = \{(1,1,1,1),(x,1,1)\}$ or $X^k_{\text{tentative}} = \{(x,1,1,1),(1,1,1)\}$. After track extraction, one thus obtains either $X^k = \{(1,1,1,1)\}$ or $X^k = \{(1,1,1),(1,1,1)\}$. The posterior probability associated with the latter solution is the same as for the solution $X^k = \{(1,1,1,1),(x,1,1)\}$. This is immediately seen to have posterior

TABLE I
Simulation Parameters for the Cardinality Tracking Problem

Parameter Description	Setting
Target birth rate	1
Target death probability	0.1
Sensor probability of detection	1
Sensor false alarm rate	0.33
Track initiation	3-of-3
Hypothesis tree depth (n_{scan})	1
Confirmation reward factors (modified-scoring MHT)	2 (confirmed tracks), 1.5 (nearly-confirmed tracks)
Track termination (maximum missed detections)	0
Number of scans in each scenario realization	100
Number of realizations	1000

probability that does not exceed that of $X^k = \{(1,1,1,1)\}$, since a track of length two does not contribute to the posterior probability (see definition of k_0 in Remark 1). With modified-scoring MHT, the measurement in the fourth scan is guaranteed to be associated to the confirmed track rather than to the nearly confirmed one, since $\xi_2 > \xi_1$ in (21) insures that the solution to (4) yields $|X|^k = \{1,1,1,1\}$.

Finally, assume that the $k_0 = 4$ in Result 1 and consider the measurement sequence $|Z|^k = \{1,2,1,1\}$. Result 2 tells us that modified-scoring MHT requires $n_{\text{scan}} \geq 1$ to insure optimality. Indeed, with $n_{\text{scan}} = 0$, the measurement in the third scan will be associated with either the shorter or longer track, since neither is nearly confirmed. Thus, either $X^k_{\text{tentative}} = \{(1,1,1,1),(x,1)\}$ or $X^k_{\text{tentative}} = \{(1,1),(x,1,1,1)\}$ results, from which we have $X^k = \{(1,1,1,1)\}$ or $X^k = \emptyset$, respectively. This in turn leads either to solution $|X|^k = \{1,1,1,1\}$ or $|X|^k = \{0,0,0,0\}$. Conversely, with $n_{\text{scan}} = 1$, we do not decide on which track is updated with the measurement in the third scan until the fourth scan is received. Accordingly, $\xi_1 > 1$ in (21) insures that the solution to (4) yields $X^k = \{(1,1,1,1)\}$ and thus $|X|^k = \{1,1,1,1\}$.

The importance of this section is that it demonstrates the superiority of modified MHT over standard MHT in a limiting case. For this case, we are able to show that modified MHT with a sufficiently large hypothesis tree depth achieves optimality in the sense of maximizing the posterior probability over all hypotheses. Modified MHT processing introduces a mechanism whereby preference is given to tracks that have achieved or will achieve track confirmation. This is an interesting result in its own right, and provides motivation for use of modified MHT in a more general setting.

We now illustrate the performance of modified-scoring MHT and standard MHT approaches to the cardinality tracking problem for a specific numerical example. The example provides experimental validation of the claims in Results 1–2. A nice aspect of evaluating cardinality tracking is that it is much easier to provide statistically significant results for which tracking parameters are matched to target and sensor characteristics. Indeed, ground truth is obtained via a Poisson birth-death process and kinematic-space realizations are absent, so that we are not limited to a small set of benchmark scenarios. The simulation parameters are captured in Table I.

The parameters in Table I satisfy Assumption 1–2. Note that, as sensor measurements are not informative with regard to target state and are only relevant to target existence, the tolerable false alarm rates are quite low compared to a general tracking problem. The tracking initiation and termination settings and the choice of n_{scan} are consistent with the requirements for Result 1–2:

$$\frac{\lambda_b(1-p_\chi)^2 p_\chi}{\lambda_{fa}^3} > 1 > \frac{\lambda_b(1-p_\chi)p_\chi}{\lambda_{fa}^2} \Rightarrow k_0 = 3,$$

$$n_{\text{scan}} \geq k_0 - 2 \Rightarrow n_{\text{scan}} \geq 1.$$

An illustration of one realization is given in Fig. 4, along with the corresponding modified-scoring MHT output. Note that we provide a compact representation of ground truth X^k, sensor measurements Z^k, and tracker output $\hat{X}^k$: we illustrate the sequence of cardinalities $|X|^k$, $|Z|^k$, and $|\hat{X}|^k$.

Statistical performance results are based on computation of the posterior probability $p(q^k \mid Z^k)$. We find as expected that the modified-scoring MHT is optimal in the posterior-probability sense. Standard MHT suffers a performance loss resulting in a (normalized) posterior probability of 0.958. (By normalized posterior probability, we mean the ratio of the probabilities associated with the standard and modified MHT solutions, respectively.)

5. SIMULATION RESULTS FOR THE GENERAL TRACKING PROBLEM

We now evaluate modified MHT and standard MHT approaches to the general tracking problem for several scenarios of interest. First, we identify the performance metrics for this analysis. Our approach to tracker performance evaluation is somewhat novel as we do not identify a global mapping of tracks to targets. Indeed, a global mapping can be problematic due to track swap phenomena, true tracks that are seduced by false contacts and become false tracks, etc. Instead, we rely on a scan-based association of tracks to targets consistent with the recently-introduced Optimal Subpattern Assignment (OSPA) metric [11].

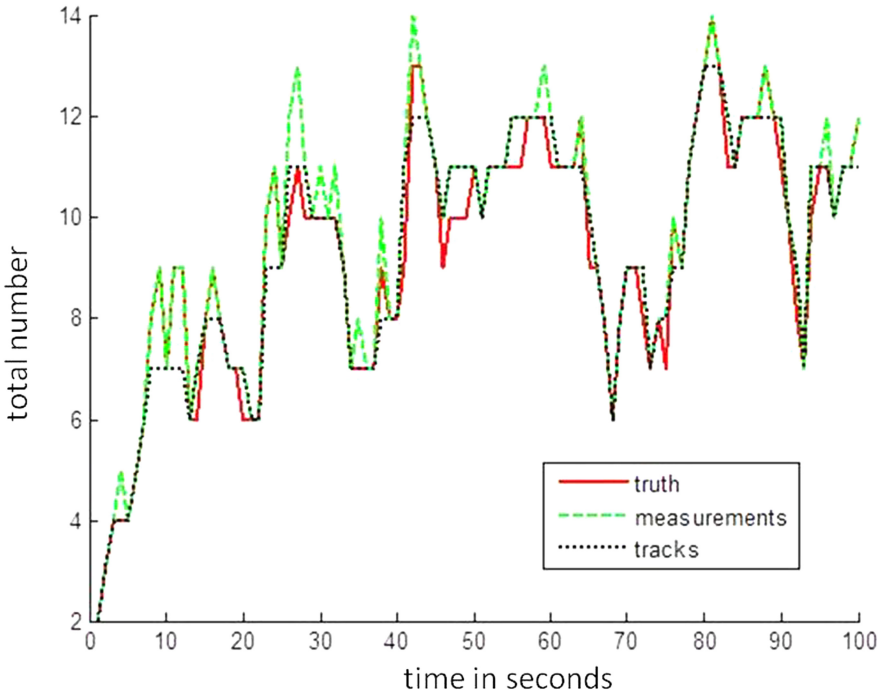


Fig. 4. One realization of truth, measurement and optimal track cardinality sequences.

For each scan time $t_i \in (t_1,\ldots,t_N)$, we have an $N_i \times M_i$ cost matrix A_i where N_i and M_i are the number of targets and tracks in existence, respectively. Given a distance threshold ξ on feasible track-truth assignments, we determine the optimal OSPA assignment between tracks and targets for each scan in a given dataset as described in [11]. For a given scan, those tracks that are assigned to targets are deemed to be *true track instances*. Correspondingly, there is a *detected target instance*. Let g_i denote the number of such (target, track) pairs for the scan at time t_i. Next, we compute the following metrics for each scenario realization:

- *Track PD*—ratio of total number of true track instances (summed over all scan times) to total number of target existence instances (summed over all targets and all scan times): $\sum_{i=1}^{N} g_i / \sum_{i=1}^{N} N_i$;
- *Track quality*—ratio of number of true track instances and total number of track instances: $\sum_{i=1}^{N} g_i / \sum_{i=1}^{N} M_i$;
- *Track purity*—ratio of number of true track instances that are as well from the mode assignment (i.e. from the most frequently associated target) and total number of track instances: $\sum_{i=1}^{N} \bar{g}_i / \sum_{i=1}^{N} M_i$; here, $\bar{g}_i \leq g_i$ is the number of truth-track assignments where truth is the *mode target* for the track, i.e. the target to which the track is associated the most.
- *Track rate*—ratio of total number of tracks to total number of targets;
- *Track localization error*—average displacement between true track instances and corresponding target location that we denote by σ_T.

Since our metrics do not rely on classifying each track as *true* or *false*, the false track statistics are understood as follows. First, the *track rate* metric answers the question: how many tracks does the system generate, relative to the true number of targets? Secondly, the *track quality* metric answers the question: for any given track at any given time, what is the probability that it is target originated? That is, track quality is the total duration of good tracks as a fraction of the overall duration of all tracks. Thus, these metrics provide an assessment of how much false track (both in number and in duration) is generated by the system, without the need for global track assessment that is often problematic when tracks are partially target-originated and partially false.

We report here on our metrics, where for each of three benchmark scenarios the metrics are averaged over multiple Monte Carlo realizations. Complete simulation parameters are identified in Table II. Illustrations of one modified MHT tracker output realization of each of three scenarios are given in Figs. 5–7; Fig. 8 illustrates the realization of the corresponding measurement data for the third scenario. Scenario 1 includes three linear-motion targets that move with identical speeds but displaced in the y dimension and with different birth and death times. Scenario 2 includes a single maneuvering target. Scenario three includes three maneuvering targets that are matched in birth and death times and in velocities and are displaced in the x dimension.

Monte Carlo performance results are given in Tables III–IV.

Encouragingly, for all scenarios we find improved performance with respect to all performance metrics of interest for modified MHT processing over standard

TABLE II
Simulation Parameters for the General Tracking Problem

Parameter Description	Setting (for scenario 1; 2; 3 respectively)
Monte Carlo realizations	100 (for each of the three scenarios)
Scenario duration	150 sec
Number of targets	3; 1; 3
Target birth (x,y) positions	$(-40,5),(-20,0),(-40,-5)$; $(-40,-5)$; $(-40,-5),(-35,-5),(-30,-5)$ m/sec
Target (x,y) velocities	$(0.5,0)$;$(0.5,0.5)$ or $(0.5,-0.5)$; $(0.5,0.5)$ or $(0.5,-0.5)$ in m/sec-all turns after 25 s
Target birth times from start	$(10,50,10)$; 10; $(10,10,10)$ sec
Target death times from start	$(140,140,100)$; 110; $(110,110,110)$ sec
Sensor footprint	2000 m^2
Sensor revisit rate	1 Hz
Sensor probability of detection	0.8
Sensor false alarm rate	5
Sensor measurement error covariance in x,y	1 m^2
Track initiation	6-of-6
Hypothesis tree depth (n_{scan})	1
Target birth rate	0.01
Target death probability	0.01
Confirmation reward factor (modified MHT)	2 (confirmed tracks), 1.5 (nearly-confirmed tracks)
Track termination (maximum missed detections)	2
Prior velocity covariance in x,y	1 m^2/s^2
Filter process noise in x,y	0.001 m^2/s^3
Data association gate	99%
Distance threshold for track-truth association	2

TABLE III
Performance Results for the Benchmark Scenarios
(Numerical results are based on 100 Monte Carlo realizations for each of the three scenarios.)

Scenario (description)	Tracker Modality	Track PD	Track Quality	Track Purity	Track Rate	Track Loc. Error
1 (3 linear)	standard MHT	0.915	0.850	0.780	1.90	0.825
1 (3 linear)	modified MHT	0.925	0.918	0.867	1.58	0.755
2 (1 maneuvering)	standard MHT	0.739	0.847	0.847	3.45	1.111
2 (1 maneuvering)	modified 3.45	1.111				
2 (1 maneuvering)	modified MHT	0.759	0.869	0.869	3.41	1.092
3 (3 maneuvering)	standard MHT	0.780	0.778	0.682	4.20	1.242
3 (3 maneuvering)	modified MHT	0.809	0.820	0.727	3.97	1.235

TABLE IV
Incremental Performance Benefit of Modified MHT, Averaged Across Scenarios, with Respect to all Metrics of Interest: Higher Track PD, Track Quality, and Track Purity; Lower Track Rate and Track Localization Error
(Numerical results accounts for all 300 Monte Carlo realizations.)

Metric	Track PD	Track Quality	Track Purity	Track Rate	Track Loc. Error
Percent change	2.43%	5.34%	8.10%	−6.09%	−3.02%

MHT. (Note that for track rate and track localization error, a reduction indicates improved performance.) Not surprisingly, since the scenarios are of increasing complexity we find consistently lower performance as we move from scenario 1 to scenario 2, and again from scenario 2 to scenario 3, as can be seen in the track quality, track rate, and track localization error. The one exception to the trend is track PD as we go from scenario 2 to scenario 3, though this can be explained: for multi-target scenarios, it is sufficient for a track instance by be close to *any* target to be deemed a detection, thus the presence of multiple nearby targets makes this easier to achieve.

Track purity is the same as track quality in the single-target scenario (scenario 2); in multi-target scenarios, track purity is lower than track quality as we require not only good track instances but from the same target as well. Indeed, track purity reflects the impact of track switching, whereby the target associated with a track may change over time. If no switching occurs, track purity equals track quality. Figure 6 illustrates what occasionally occurs, even in single-target settings: track fragmentation whereby the first track is seduced by false returns, and a second track is initiated. Note that our OSPA-based metrics correctly classify at most one track update as associated with each target at any time.

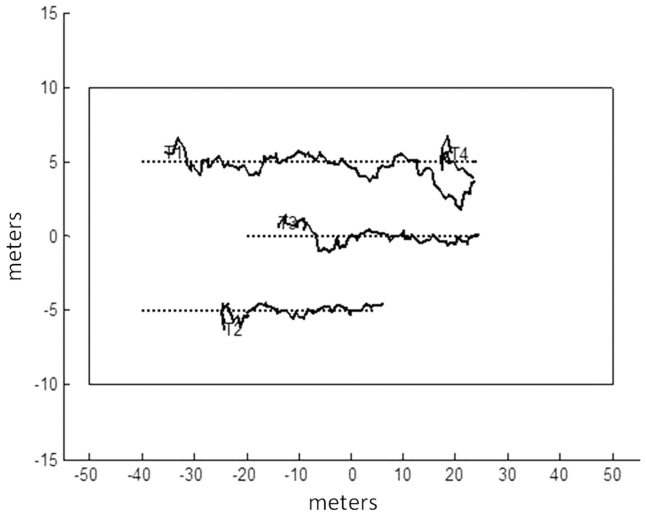


Fig. 5. An example realization of scenario 1 (target trajectories are dotted lines and modified MHT tracks are solid lines).

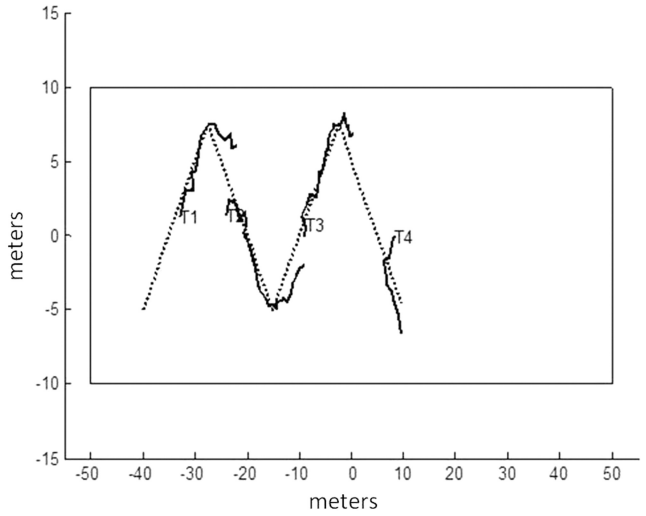


Fig. 6. An example realization of scenario 2 (target trajectories are dotted lines and modified MHT tracks are solid lines).

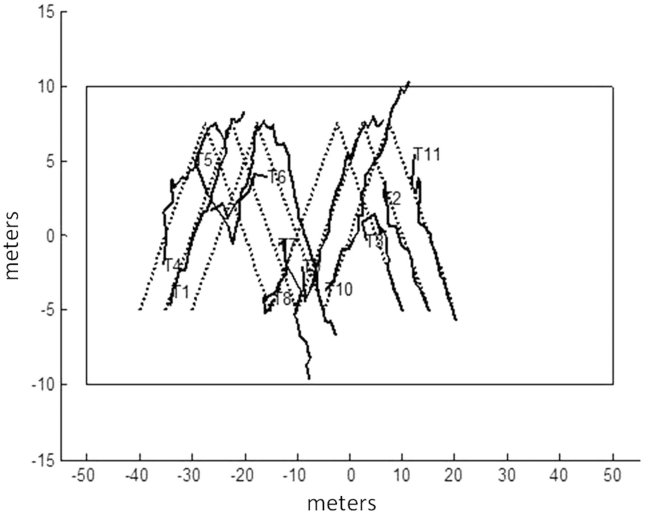


Fig. 7. An example realization of scenario 3 (target trajectories are dotted lines and modified MHT tracks are solid lines).

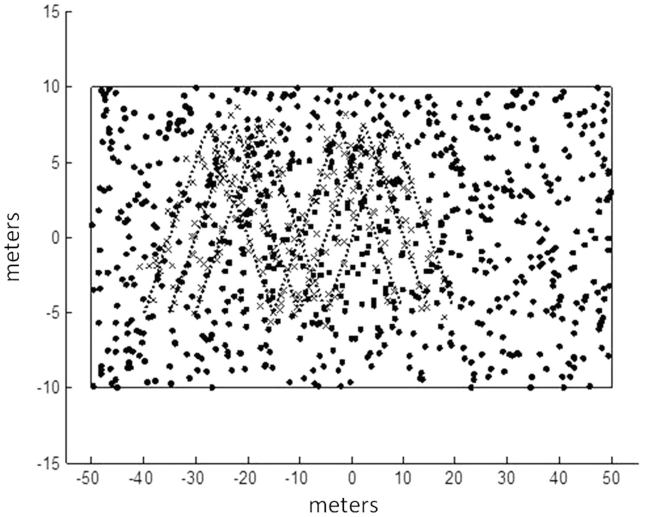


Fig. 8. Measurement data for one run of scenario 3 (crosses are target-originated returns and dots are false alarms).

TABLE V
Fusion Gain Computation (1.18 = 0.959/0.813)

Tracker Modality	IQ-scenario 1	IQ-scenario 2	IQ-scenario 3	IQ-average
Standard MHT	1.249	0.686	0.510	0.813
Modified MHT	1.610	0.729	0.532	0.959

Thus, the fragmentation and track redundancy observed here are reflected in degraded *track rate*, *track quality*, and *track purity* values.

It is helpful to capture tracker performance improvement with a single scalar metric, using the notion of *information quality* [4]; as discussed in [4], this notion can be related to the *information reduction factor* discussed in [15]. *Information quality* (IQ) is the average information content (in a Fisher information sense) of an arbitrarily-selected output track instance. With some probability, the track is associated with a true target: this is given by the *track quality* metric. Correspondingly, the information content is given by the Fisher information, which in turn is the inverse of the track error covariance matrix. Thus, IQ is the product of track quality and Fisher information:

$$\mathrm{IQ} = \frac{\sum_{i=1}^{N} g_i}{\sigma_T^2 \sum_{i=1}^{N} M_i}. \tag{22}$$

Further, we can evaluate *fusion gain* as the IQ ratio of two competing tracking solutions. For the results captured in Table III, we find that modified MHT provides a

fusion gain of 1.18 (or 18%) over standard MHT. This is obtained by computing the ratio of average IQ for modified MHT and average IQ for standard MHT. Further details are in Table V. Note that, not surprisingly, for both tracker solutions the IQ metric degrades with increasing scenario difficulty.

6. CONCLUSIONS

This paper provides a compact and accessible introduction to track-oriented multiple-hypothesis tracking (MHT). It shows that track-oriented MHT does not suffer from the so-called *target-death problem* that has been observed in the probabilistic hypothesis density (PHD) filter. Unfortunately, the MHT exhibits a problem of its own, whereby unconfirmed tracks are often found to take contacts away from confirmed or nearly-confirmed tracks, degrading their quality.

We first study this problem in a simplified context with no measurement state information: this formulation reduces to the cardinality tracking problem. For this problem, and with the further assumption of unity detection probability, we are able to establish structural results for the optimal tracking solution and, remarkably, we find that the modified MHT solution with appropriately-selected track initiation and termination criteria and with sufficient hypothesis tree depth is guaranteed to achieve optimality. Simulation results are consistent with our theoretical findings.

The performance characteristics of modified-scoring MHT in the simplified cardinality-tracking context motivate its use for more general tracking problems. We study several benchmark scenarios and find improved performance for modified-scoring MHT over standard MHT. It is important to note that in all scenarios, all targets die before the scenario end; thus, there appear to be no negative side-effect to modified-scoring MHT processing whereby confirmed tracks are kept alive despite target death.

In a nutshell, modified MHT scoring is needed since we cannot perform batch track extraction from the set of track hypotheses. Indeed, batch extraction would incur computational infeasibility (unbounded n_{scan}) as well as large reporting latency. Accordingly, we must use (suboptimal) sequential track extraction. Favoring good (i.e. confirmed or nearly-confirmed) tracks over tentative ones in the extraction process can be motivated on two grounds: (1) empirically, as the scheme is found to perform better; (2) in a limiting (albeit simplistic) case—cardinality tracking, the scheme matches the performance of optimal batch extraction, provided n_{scan} is large enough (where the lower bound is quantified). While (1–2) do not *prove* that modified MHT is better than standard MHT, they do provide meaningful practical & theoretical motivation.

Our scheme is similar in its effects to the two-stage assignment scheme that has been adopted in an MHT setting [13–14]; indeed there is a need to balance track initiation and maintenance. A merit of our work, we think, is to emphasize an often-ignored aspect of making MHT work well in practice. Interestingly, giving preference to established tracks is a scheme whose applicability is not limited to the MHT approach; a recent example in the context of the *Histogram Probabilistic Multi-Hypothesis Tracker* (H-PMHT) may be found in [5].

REFERENCES

[1] S. Blackman and R. Popoli
Design and Analysis of Modern Tracking Systems.
Artech House, 1999.

[2] S. Coraluppi and C. Carthel
Aggregate surveillance: A cardinality tracking approach.
In *Proceedings of the 14th International Conference on Information Fusion*, Chicago IL, July 2011.

[3] S. Coraluppi, C. Carthel, M. Luettgen, and S. Lynch
All-source track and identity fusion.
In *Proceedings of the National Symposium on Sensor and Data Fusion*, San Antonio TX, June 2000.

[4] S. Coraluppi, M. Guerriero, and C. Carthel
Fusion gain in multi-target tracking.
In *Proceedings of the 13th International Conference on Information Fusion*, Edinburgh, Scotland, July 2010.

[5] S. Davey
Histogram PMHT with particles.
In *Proceedings of the 14th International Conference on Information Fusion*, Chicago IL, July 2011.

[6] O. Erdinc, P. Willett, and Y. Bar-Shalom
The bin-occupancy filter and its connection to the PHD filters.
IEEE Transactions on Signal Processing, **57**, 11 (Nov. 2009).

[7] T. Kurien
Issues in the design of practical multitarget tracking algorithms.
In *Multitarget-Multisensor Tracking: Advanced Applications*, Y. Bar-Shalom (Ed.), Artech House, 1990.

[8] C. Morefield
Application of 0-1 integer programming to multitarget tracking problems.
IEEE Transactions on Automatic Control, **22**, 3 (June 1977).

[9] A. Poore and N. Rijavec
A Lagrangian relaxation algorithm for multidimensional assignment problems arising from multitarget tracking.
SIAM J. Optimization, **3**, 3 (Aug. 1993).

[10] D. Reid
An algorithm for tracking multiple targets.
IEEE Transactions on Automatic Control, **24**, 6 (Dec. 1979).

[11] D. Schuhmacher, B.-T. Vo, and B.-N. Vo
A consistent metric for performance evaluation of multi-object filters.
IEEE Transactions on Signal Processing, **56**, 8 (Aug. 2008).

[12] P. Storms and F. Spieksma
An LP-based algorithm for the data association problem in multitarget tracking.
In *Proceedings of the 3rd International Conference on Information Fusion*, Paris, France, July 2000.

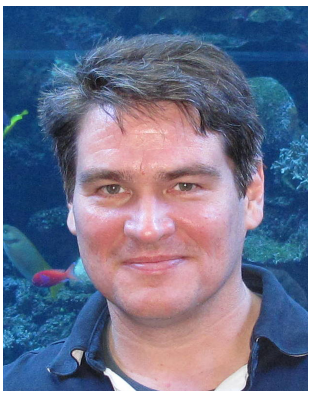
[13] A. Sinha, Z. Ding, T. Kirubarajan, and M. Farooq
Track quality based multitarget tracking algorithm.
To appear in *IEEE Transactions on Aerospace and Electronic Systems*.

[14] A. Sinha, Z. Ding, T. Kirubarajan, and M. Farooq
Track quality based multitarget tracking algorithm.
In *Proceedings of the SPIE Conference on Signal and Data Processing of Small Targets*, Orlando FL, Apr. 2006.

[15] X. Zhang, P. Willett, and Y. Bar-Shalom
Dynamic Cramer-Rao bound for target tracking in clutter.
IEEE Transactions on Aerospace and Electronic Systems, **41**, 4 (Oct. 2005).

Stefano Coraluppi received the B.S. degree in electrical engineering and mathematics from Carnegie Mellon University (1990), and M.S. and Ph.D. degrees in electrical engineering from the University of Maryland (1992, 1997).

Dr. Coraluppi has worked on the research staff at ALPHATECH Inc. (1997–2002), the NATO Undersea Research Centre (2002–2010), and Compunetix Inc. (since 2010). He has contributed to programs in ground, undersea, maritime and air surveillance for security and defense applications. In 2006, he was general cochair (with Professor Peter Willett) for the ISIF/IEEE 9th International Conference on Information Fusion in Florence, Italy. Currently, he serves on the Board of Directors of the International Society of Information Fusion (ISIF), for which he served as president in 2010. He is associate editor for target tracking and multisensor systems for the *IEEE Transactions on Aerospace and Electronic Systems* and area editor for tracking for the *ISIF Journal of Advances in Information Fusion*. His research interests include multi-target tracking, data fusion, detection and estimation theory, and optimal and stochastic control.



Craig Carthel received B.S. degrees in physics and mathematics in 1988, a M.S. in mathematics in 1992, and a Ph.D. in mathematics in 1995, all from the University of Houston.

He is a principal scientist at Compunetix Inc. in Monroeville, PA. He did research in numerical analysis and optimization theory at the University of Houston. From 1995 to 1997, he worked at the Institute for Industrial Mathematics at Johannes Kepler University, in Linz, Austria on parameter identification and inverse problems. From 1998 to 2002, he was a senior mathematician at ALPHATECH Inc. in Burlington, MA, where he worked on image processing, multisensor data fusion, and ground target tracking. From 2002 to 2010, he was a senior scientist in the Applied Research Department at the NATO Undersea Research Centre in La Spezia, Italy, where he worked on military operations research, simulation, optimization, and data fusion problems associated with maritime environments. In 2006, he served as the technical program chair for the 9th International Conference on Information Fusion.

INTERNATIONAL SOCIETY OF INFORMATION FUSION

ISIF Website: http://www.isif.org

2012 BOARD OF DIRECTORS*

2010–2012	2011–2013	2012–2014
Simon Maskell	Sten F. Andler	Darin T. Dunham
Peter Willett	Yvo Boers	Fredrik Gustafsson
Wolfgang Koch	Lyudmila Mihaylova	Lance M. Kaplan

*Board of Directors are elected by the members of ISIF for a three year term.

PAST PRESIDENTS

Joachim Biermann, 2011	Pierre Valin, 2006	Pramod Varshney, 2001
Stefano Coraluppi, 2010	W. Dale Blair, 2005	Yaakov Bar-Shalom, 2000
Elisa Shahbazian, 2009	Chee Chong, 2004	Jim Llinas, 1999
Darko Musicki, 2008	Xiao-Rong Li, 2003	Jim Llinas, 1998
Erik Blasch, 2007	Yaakov Bar-Shalom, 2002	

SOCIETY VISION

The International Society of Information Fusion (ISIF) is the premier professional society and global information resource for multidisciplinary approaches for theoretical and applied information fusion technologies.

SOCIETY MISSION

Advocate

To advance the profession of fusion technologies, propose approaches for solving real-world problems, recognize emerging technologies, and foster the transfer of information.

Serve

To serve its members and engineering, business, and scientific communities by providing high-quality information, educational products, and services.

Communicate

To create international communication forums and hold international conferences in countries that provide for interaction of members of fusion communities with each other, with those in other disciplines, and with those in industry and academia.

Educate

To promote undergraduate and graduate education related to information fusion technologies at universities around the world. Sponsor educational courses and tutorials at conferences.

Integrate

Integrate ideas from various approaches for information fusion, and look for common threads and themes–look for overall principles, rather than a multitude of point solutions. Serve as the central focus for coordinating the activities of world-wide information fusion related societies or organizations. Serve as a professional liaison to industry, academia, and government.

Disseminate

To propagate the ideas for integrated approaches to information fusion so that others can build on them in both industry and academia.

Call for Papers

The Journal of Advances in Information Fusion (JAIF) seeks original contributions in the technical areas of research related to information fusion. Authors of papers in one of the technical areas listed on the inside cover of JAIF are encouraged to submit their papers for peer review at http://jaif.msubmit.net.

Call for Reviewers

The success of JAIF and its value to the research community is strongly dependent on the quality of its peer review process. Researchers in the technical areas related to information fusion are encouraged to register as a reviewer for JAIF at http://jaif.msubmit.net. Potential reviewers should notify via email the appropriate editors of their offer to serve as a reviewer.

Journal of Advances in Information Fusion Volume 7 Number 2 December 2012 ISSN 1557-6418