AD-A228 680

Underwater Tracking in the Presence of Modeling Uncertainty

J. G. Baylog
A. A. Magliaro
S. M. Zile
K. F. Gong
Combat Control Systems Department

Naval Underwater Systems Center
Newport, Rhode Island / New London, Connecticut

Approved for public release; distribution unlimited.

Reprint of a paper published in the IEEE 1987 Conference
Record of the 21st Asilomar Conference on Signals, Systems,
UNDERWATER TRACKING IN THE PRESENCE OF MODELING UNCERTAINTY

by

John G. Baylog Antonio A. Magliaro Suzanne M. Zile Kai F. Gong

Naval Underwater Systems Center Newport Laboratory Newport, Rhode Island 02841

ABSTRACT

This paper addresses the problem of passive underwater contact tracking in the presence of mismodeling. A configuration is proposed which includes contact state estimation and hypothesis testing coupled with an evidential reasoning capability to resolve uncertainty in modeling assumptions. The Dempster-Shafer theory of evidence is applied to direct the generation of evidence required to classify modeling anomalies and provide compensation. Experimental results are presented that demonstrate the reduction of modeling uncertainty with the accumulation of evidence.

I. INTRODUCTION

Contact tracking in the underwater environment deals primarily with processing passively observed measurement data under a variety of operating conditions [1]. The efficient extraction of contact-related parameters from measurement data is contingent upon judicious modeling of the tracking environment. The tracking model should provide an accurate but concise description of the underlying principles and assumptions associated with contact, observer, and sensor dynamics; acoustic sound propagation; and the measurement generation process. Encumbrances arise when abnormalities occur within the estimation process. These abnormalities are exhibited as irreconcilable trends in measurement residuals, indicative of mismodeling that has occurred.

Recent work has addressed the detection of modeling anomalies through analysis of measurement residuals. In particular, hypothesis testing techniques have been employed to detect bias within temporally delineated segments of residuals [2]. These testing techniques have been integrated with contextual information processing in a prototype knowledge-based system for targeting solution integration and assessment (TARSIA). TARSIA [3] focused on the intelligent control of a generalized maximum likelihood estimator and the use of numerically derived solution attributes as the basis for heuristic analysis in assessing solution quality. As such, TARSIA relies on synergistic coupling between numeric and symbolic components as indicated in figure 1.

This paper extends these previous efforts by addressing the problem of uncertainty in the classification of data and modeling anomalies and its impact on solution quality assessment through the application of evidential reasoning methods. Evidence pertaining to the nature of data and modeling anomalies may be deduced from several sources, including intelligence and tracking data, sonar information, and sound channel characterization. It is manifested as patterns within time series (e.g., measurements and measurement residuals) and trends in solutions. Often, more than one anomaly type might be considered based on a given set of evidence. This results in uncertainty as to the nature of the anomaly and a need for additional information to resolve the ambiguity.

The reasoning technique considered for the interpretation and combination of evidence is Dempster-Shafer theory [4,5]. An important feature of the theory is its ability to narrow the set of likely hypotheses with the accumulation of evidence, while retaining the notion of ignorance when insufficient evidence is presented. This is in contrast with Bayesian techniques, which fail to discriminate between equal uncertainty and lack of knowledge. With the utilization of Dempster-Shafer theory, a set of plausible anomaly types can be
formulated that allows for the simultaneous pursuit of confirming evidence to further reduce the size of the set, and the inclusion of candidate anomaly types in the formulation of the parameter uncertainty descriptions.

A brief description of Dempster-Shafer theory is provided in the next section. A conceptual model of a processing configuration that incorporates Dempster-Shafer theory is then presented. This is followed by a mathematical formulation of evidence generated subsequent to data processing. Experimental results based on angle-of-arrival measurements are presented.

II. DEMPSTER-SHAFER THEORY

The Dempster-Shafer theory of evidence has as its basis the formulation of a frame of discernment. This is a mutually exclusive and exhaustive set of allowable hypotheses, denoted by

$$\Theta = \{a_1, a_2, \ldots, a_l\}$$

where \(a_i\), \(i=1, \ldots, l\) denote the individual (singleton) hypotheses. The power set associated with \(\Theta\) is the set of all possible subsets of \(\Theta\). This is illustrated in figure 2 for a four-element frame of discernment.

![Figure 2. The Power Set](image)

As evidence regarding the validity of certain hypotheses becomes available, probability mass (or belief) is distributed to appropriate subsets within the power set with the stipulation that

$$\sum_{A_j} m(A_j) = 1,$$

where \(A_j\) represents an element of the power set, and \(m(A_j)\) represents the non-negative basic probability assignments (BPA). Belief not assigned to particular subsets is assigned to \(\Theta\) and represents uncertainty in the interpretation of evidence. Associated with each \(A_j\) is an evidential interval \([S, P]\) where \(S\) represents the support for the element and \(P\) represents its plausibility [6]. This interval corresponds to a minimum and maximum likelihood for the given subset. When evidence confirms a subset, \(S\) is calculated for each \(A_j\) by summing the BPA of \(A_j\) and all its subsets, i.e.,

$$S(A_j) = \sum_{A_j \subseteq A_i} m(A_j).$$

Plausibility, an indication of the lack of disconfirming evidence, is computed as

$$P(A_j) = 1 - S(A_j^c),$$

where \(A_j^c\) denotes the complement of \(A_j\).

Ignorance regarding the validity of a particular hypothesis is represented by the evidential interval \([0,1]\), while \([1,1]\) and \([0,0]\) represent the Boolean notions of true and false, respectively.

As multiple bodies of evidence become available, the need exists to merge the evidence into a single belief representation. This is accomplished through the use of Dempster's rule of combination, given by

$$m(\Theta) = \frac{1}{K} \sum_{X \cap Y = \emptyset} m(X)m(Y),$$

where \(K\) denotes the total number of distinct frames under consideration.

Dempster's combination rule is both commutative and associative, providing an order-independent method for merging evidence. Normalization by \(K\) disallows any probability mass being assigned to the null set, since this would violate the exhaustive set requirement for the frame of discernment.

A prime concern when applying Dempster-Shafer theory is the potential computational explosion that occurs when the size of the frame of discernment becomes large. One recent approach to this problem is to construct multiple frames of discernment linked together through compatibility relations [7]. That is, in lieu of defining a single frame that encompasses all possible hypotheses, the hypothesis space is delimited into natural partitions, each of which can be addressed by a reduced order frame. A network of compatibility mappings is then constructed such that evidence producing BPAs in one frame, \(\Theta_A\), results in analogous BPAs assigned in another frame, \(\Theta_B\). A compatibility relation, \(\Theta_A \times \Theta_B\), defines which elements in two distinct frames are true simultaneously, with

$$\Theta_A \times \Theta_B \subseteq \Theta_A \times \Theta_B.$$

Compatibility mappings are given by

$$C_{A \times B}(A_k) = \{b_j | (a_i, b_j) \in \Theta_A \times \Theta_B, a_i \in A_k\}.$$

where at least one pair \((a_i, b_j)\) is specified for each \(a_k \in \Theta_A\). In this way, any frame may be linked to any other frame in the network.

III. PROPOSED PROCESSING CONFIGURATION

A configuration for integrating Dempster-Shafer evidential reasoning with statistical parameter estimation for anomaly classification is illustrated in figure 3. When anomalies are detected, the data are reprocessed with suspect data segments omitted to preclude their influence on the generated solution. They are monitored, however, for clues as to the nature of the anomaly. In this way, residual characteristics are based on a pivotal unbiased solution derived from a consistent data set. Preliminary evidence is generated via characterization of suspect data segments in terms of their perceived features. Refinement of these features is achieved iteratively by resegmenting...
suspect data to resolve events and event boundaries. Additional evidence is furnished by examining solution features. These include comparison of parameter values with a priori information and the formulation of distance measures between alternative solutions. A more detailed description of these sources of evidence is provided in the following section.

The evidential reasoning component contains a network of frames of discernment, where a frame is constructed for each measurement type under consideration. Measurement frames contain only those anomaly types to which the measurement in question is sensitive. A hierarchically related solution frame is linked to each measurement frame; and, as evidence is applied to measurement frames, compatibility mappings translate the evidence to the solution frame. In this way, evidence across measurement sets is merged. Solution extracted evidence is applied to the solution frame. Inferences regarding anomaly classification are produced primarily through examination of the solution frame.

The classification of anomalies generally results in modification of modeling assumptions over segments of data. In the event that the anomaly type remains unresolved, a plausibility ranking motivates the search for confirming evidence or, if this is unattainable, directs the inclusion of these anomaly types into the formulation of uncertainty region descriptions.

IV. MATHEMATICAL FORMULATION OF EVIDENCE

The estimator represented in figure 1, an iterative maximum likelihood estimator (MLE) used for mapping data into solutions and integrating solutions, is described by Murphy [8]. Its outputs include the solution in the specified coordinate system and a square root realization of the Fisher information matrix (FIM). Based on these, various pieces of evidence are created. To facilitate subsequent discussion, let \( \mathbf{z} \) be defined as an N-dimensional measurement vector:

\[
\mathbf{z} = \begin{bmatrix}
  z_1 & z_2 & \ldots & z_L
\end{bmatrix}^T,
\]

where \( z_i \), \( i=1,\ldots,L \), denote partitions of arbitrary length. Further, let each individual measurement, \( z_j(k) \), identified by measurement type \( j \) and time index \( k \), be governed by plant and measurement models:

\[
\bar{x}(k+1) = g(x(k), y(k), k, k+1),
\]

\[
z_j(k) = f_j(x(k), y(k)) + \eta_j(k),
\]

where \( \bar{x}(k) \) is an n-dimensional state vector, \( \eta(k) \) is a system parameter vector, and \( \eta_j(k) \) denotes a zero-mean, independent, white noise sample with \( N(0,\sigma_j^2) \). The MLE yields an estimate \( \hat{x} \) that minimizes the performance index (PI):

\[
PI = (1/N)(\mathbf{z} - \mathbf{z})^T D^{-1} (\mathbf{z} - \mathbf{z}),
\]

where \( \mathbf{z} \) denotes the measurement residual vector with elements

\[
z_{ej}(k) = z_j(k) - f_j(x(k), y(k)),
\]

and

\[
D = E(\eta_j^2) = \text{Diag}(\sigma_j^2).
\]

Note that for unbiased estimates the PI becomes \( \chi^2 \)-distributed with an expected value of unity. In a similar fashion, the data partitions of equation (9) yield local values of the PI.

Solution Consistency

The primary attributes used to ascertain the consistency of solutions include both local and global assessments of the PI, and the output of a hypothesis testing technique in which measurement residuals are interrogated for bias. In the former case, thresholds are set to detect the occurrence of improbable values of PI based on a specified probability of false alarm. In the latter case, two hypotheses are formulated. The null hypothesis \( H_0 \) assumes that the residuals within a given data segment are zero mean. Under the alternative hypothesis \( H_1 \), the data residuals exhibit a significant bias. That is,

\[
H_0: z_{ej}(k) = \{ \Phi_j / 3x(k) \} \delta x(k) + \eta_j(k),
\]

\[
H_1: z_{ej}(k) = \{ \Phi_j / 3x(k) \} \delta x(k) + \eta_j(k) + m_j(k),
\]

where

\[
\delta x(k) = \bar{x}(k) - \hat{x}(k),
\]

and \( m_j(k) \) denotes a least-squares fit to the residuals within the data segment. The Neyman-Pearson test procedure is applied [2] to yield the test criteria

\[
\frac{P(\mathbf{Z} | H_1)}{P(\mathbf{Z} | H_0)} > \lambda,
\]

where \( \lambda \) is chosen to satisfy a specified probability of false alarm. A data segment is deemed consistent when the hypothesis test is passed, indicating unbiased residuals in accordance with the theoretical model. In general, bias
exhibited within one or more data segments can cause the solution \( \hat{x} \) to become biased and, in the process of minimization, results in other data segments appearing biased. Thus, when a particular data segment is deemed to be the source of modeling inconsistency, it is removed from the processed data set, while its associated attributes continue to be monitored. In this way, features extracted from monitored segments are based on consistent solutions and provide evidence as to the nature of detected anomalies.

**Perturbation Analysis**

Given an estimate \( \hat{x} \) deemed consistent prior to the detection of an anomaly, the effects of mismodeling on the residuals is now examined. Equations (10) and (11) can be approximated in a Taylor series expansion as

\[
\begin{align*}
\hat{x}(k+1) & = \hat{x}(k) + \delta x(k) + \frac{\partial f}{\partial x}(k) \delta x(k) + \frac{\partial f}{\partial y}(k) \delta y(k), \\
\delta y(k) & = \hat{y}(k) - \hat{y}(k),
\end{align*}
\]

This yields the approximate residual sequence

\[
\begin{align*}
\delta e_j(k+1) & = \delta e_j(k) + \frac{\partial f}{\partial x}(k) \delta x(k) + \frac{\partial f}{\partial y}(k) \delta y(k),
\end{align*}
\]

Consider the case where system parameter mismodeling commences at time \( k_m \). The effect of modeling changes after time \( k_m \) can be computed from

\[
\begin{align*}
\delta e_j(k+1) & = \delta e_j(k) + \frac{\partial f}{\partial x}(k) \delta x(k), \\
\delta e_j(k+1) & = \delta e_j(k) + \frac{\partial f}{\partial y}(k) \delta y(k),
\end{align*}
\]

where

\[
\delta x(k+1) = [\frac{\partial f}{\partial x}(k)] \delta x(k) + [\frac{\partial f}{\partial y}(k)] \delta y(k),
\]

\[
\delta y(k+1) = [\frac{\partial f}{\partial x}(k)] \delta x(k) + [\frac{\partial f}{\partial y}(k)] \delta y(k),
\]

Equations (12) and (13) can be approximated in a Taylor series expansion as

\[
\begin{align*}
\delta x(k+1) & = \delta x(k) + \frac{\partial f}{\partial x}(k) \delta x(k) + \frac{\partial f}{\partial y}(k) \delta y(k), \\
\delta y(k+1) & = \delta y(k) + \frac{\partial f}{\partial x}(k) \delta x(k) + \frac{\partial f}{\partial y}(k) \delta y(k),
\end{align*}
\]

Note that when \( \hat{x} \) is unbiased, the expected value of the residual becomes that of the modeling error mapped into measurement space. Depending on operating conditions, the modeling error can easily dominate the residuals. Hence, for a given class of modeling errors, the effect on measurement residuals is predictable. Note, however, that modeling parameters are subject to observability requirements just as are state parameters.

**Solution Features**

Solutions derived from independent data sets can be integrated by forming and processing pseudo-measurement and pseudo-residual sets of the form

\[
\begin{align*}
\hat{z}_j = \hat{x}_j \hat{x}_j, \\
\hat{z}_j = \hat{z}_j - \hat{f}(\hat{x}_j),
\end{align*}
\]

where \( \hat{x}_j \) and \( \hat{z}_j \) denote local solutions and their square root information matrices, respectively, and \( \hat{x} \) denotes the global (integrated) solution. \( \hat{x}_j \) and \( \hat{x} \) need not be of the same dimension or coordinate system as indicated by the mapping function \( f \).

Further evidence can be generated through a partitioning of the pseudo-measurements into reduced dimension components and forming partitioned distance measures of the form

\[
\begin{align*}
\hat{z}_j = [\hat{z}_{j1}, \hat{z}_{j2}, \ldots, \hat{z}_{jT}].
\end{align*}
\]

Examples: **Angle-of-Arrival Measurements**

The effects of certain anomaly types on angle-of-arrival measurements are now provided. Consider the measurement types bearing and depression/elevation (D/E) angle given by

\[
\begin{align*}
\theta_v &= \tan^{-1}(r_y/r_x), \\
\theta_{D/E} &= \tan^{-1}(r_y/r_x),
\end{align*}
\]

where \( r(k) \) is defined in a Cartesian coordinate system with

\[
\begin{align*}
\hat{x}_T = [r_{x1}, r_{y1}, \ldots, r_{xT}, r_{yT}]^T, \\
\hat{x}_o = [r_{x0}, r_{y0}, \ldots, r_{x0}, r_{y0}]^T.
\end{align*}
\]
\[
X = [r_X, r_Y, v_X, v_Y, r_Z]^T = \bar{X}_T - \bar{X}_0,
\]

\[
\Phi(k, k+1) = \begin{bmatrix}
    1 & 0 & 0 & 0 \\
    0 & 1 & 0 & 0 \\
    0 & 0 & 1 & 0 \\
    0 & 0 & 0 & 1
\end{bmatrix},
\]

(30)

and where \(\bar{X}_T, \bar{X}_0, \) and \(\bar{X}\) denote target, observer, and relative state vectors, respectively. Here, \(r_Z\) is defined as the depth of a target image when reflections of sound occurs on the bottom or surface of the sound channel. Also consider the conical angle measurement type as observed with a linear array:

\[
\theta_C = \tan^{-1}(r_h/\sqrt{r_h^2 + r_s^2}),
\]

(31)

where \(r_h\) and \(r_s\) denote parallel and perpendicular displacements, respectively, of the target from the linear array. The modeling perturbation \(\delta y(k)\) is now defined for certain anomalies. Consider first an unmodeled target maneuver. Often, a constant-velocity assumption is placed on target motion, and a change in target course or speed violates this assumption. This can be related as plant inputs

\[
v_{xt}(k+1) = v_{xt}(k) + \Delta v_x,
\]

(32)

\[
v_{yt}(k+1) = v_{yt}(k) + \Delta v_y,
\]

(33)

where \(\Delta v_x\) and \(\Delta v_y\) are normally zero and, for an instantaneous change in velocity, are non-zero only at time \(k_m\). Similarly, consider an unmodeled change in relative image depth \(\Delta r_z\) occurring at time \(k_m\) due to either varying bottom conditions, a change in propagation path, or a target depth change. These modeling perturbations can be represented as

\[
\delta y(k) = [\Delta v_x, \Delta v_y, \Delta r_z]^T \quad \text{(for } k = k_m\).
\]

(34)

Finally, consider an additive bias placed on the respective measurement types and commencing at time \(k_m\). Here

\[
\delta y(k) = \Delta r_z(k) \quad \text{(for } k > k_m\).
\]

(35)

Then, it is straightforward to show that

\[
\Delta \delta y(k) = \Delta r_z(k) + [\Delta \delta y(k)](k - k_m)^T,
\]

\[
\Delta \delta \theta_{D/E}(k) = \Delta \theta_{D/E}(k)
\]

\[
+ \Delta r_z(k) \cos \theta_{D/E}(k)/r_s(k) - [\Delta \delta \theta_{D/E}(k)](k - k_m)^T,
\]

\[
\Delta \delta \theta_C(k) = \Delta \theta_C(k)
\]

\[
+ \Delta r_z(k) \sin \theta_{D/E}(k)/[r_s(k) \tan \theta_C(k)]
\]

(36)

\[
+ [\Delta \delta \theta_C(k)](k - k_m)^T,
\]

where

\[
\Delta \delta y(k) = [\Delta v_x \sin \theta(k) - \Delta v_y \cos \theta(k)]/r(k),
\]

\[
\Delta \delta \theta_{D/E}(k) = [\Delta \theta_x \sin \theta(k) + \Delta \theta_y \cos \theta(k)]
\]

\[
- \sin \theta_{D/E}(k)/r_s(k),
\]

\[
\Delta \delta \theta_C(k) = [\Delta \theta_x \cos \theta(k) \cos \theta_C(k)
\]

- \Delta \theta_C(k)/r_s(k),
\]

\[
r(k) = [r_x^2(k) + r_y^2(k)]^{1/2},
\]

\[
r_s(k) = [r_x^2(k) + r_y^2(k)]^{1/2},
\]

\[
\varphi(k) = \tan^{-1}[r_y(k)/r_x(k)],
\]

(37)

and where solution error and measurement noise are assumed to be dominated by the modeling error.

Note that the perturbations consist of measurement bias and geometry-dependent components. Furthermore, image depth change is constant after time \(k_m\), yet its effect on measurements is geometry-dependent. The target maneuver, however, while also modulated by the geometry, has a linear variation in time. These features can be postulated as jumps, drifts, and curvature in the measurement residulas.

If independent solutions are produced based on data sets before and after time \(k_m\), it is seen that

\[
\bar{X}_e(k_m) - \bar{X}_e(k_m) = [0, 0, \Delta v_x, \Delta v_y, \Delta r_z]^T
\]

(38)

for the target maneuver and image depth change. The effect of bias is unpredictable and depends on the precise nature of the bias. Bias removal or compensation techniques can be evaluated by examining how well solutions cluster with compensation. Finally, if image depth is not estimated or presumed known, then modeling error will contaminate the other target parameters and the effect is again unpredictable. Once more, solution clustering can be assessed to validate the assumed image depth.

V. EXPERIMENTAL RESULTS

Experimental results are presented to demonstrate the utility of the proposed concept. The network of frames shown in figure 4 was constructed with the frames of discernment defined in tables 1 and 2. Static bias is defined as occurring instantaneously and maintained constant over a segment while dynamic bias is defined as time-varying but with no discontinuities in the waveform. Since bias can occur in any measurement set, both static and dynamic bias for each measurement set is linked to systematic error in the solution frame. Image depth changes are presumed to occur instantaneously. Sensor mismodeling applies only to the linear array and solutions derived from conical angle data. Basic probability assignments were developed based on the anomaly set defined above, with separate BPAs defined for interpreting the features of measurement residuals for each respective measurement type.
The integration of the Dempster-Shafer theory of evidence with statistical parameter estimation techniques has been demonstrated. Doing this requires development of a frame-of-discernment hierarchy, feature extraction, basic probability assignments, and formulation of methods of searching for confirming evidence. In addition, the utility of forming a plausible set of anomaly types from which to build inferences regarding solution quality has been demonstrated. That is, compensation techniques should accommodate all anomaly types deemed plausible, given that anomalies are detected.

It should be noted that while BPA distribution for the experiments presented was based on analysis of evidence, it can still be primarily subjective in nature. Dempster-Shafer theory accommodates the use of known a priori statistics, but also allows for conservative distribution of belief as a function of the integrity of the evidence. That is, by assigning more likelihood to the frame of bearing and D/E angle are processed jointly. The target maneuver presents drift characteristics that become increasingly apparent with time. This is reflected in figure 7 in the convergence of support for this anomaly type.

Figure 8 describes results obtained when static bias in D/E angle and image depth change are considered. In these cases, the anomaly types exhibit nearly identical features at the occurrence of the anomaly and become only marginally discernable, given sufficient time and a suitable tracking geometry. Hence, figure 8 depicts support for the individual anomaly types as well as their union. Figure 8a describes results when static bias is the actual anomaly, and figure 8b describes the image depth change case. These results accentuate the utility of Dempster-Shafer theory for representing modeling uncertainty.

Three anomaly types were simulated in these experiments: target maneuver, static bias in D/E angle, and image depth change. Test procedures to generate confirming evidence by compensating for these anomaly types were developed with subsequent BPA distribution. The geometry considered is shown in figure 5, where the maneuvering target trajectory is indicated by a dashed line. Examples of residual features in angle-of-arrival measurement residuals that are characteristic of these forms of mismodeling are shown in figure 6. Experimental results for the case of the target maneuver are shown in figure 7, where support for that anomaly type is plotted as a function of the size of the data window following anomaly detection that is available for feature extraction. Support is plotted for low, medium, and high noise cases. Figure 7a depicts the belief in target maneuver as the exhibited anomaly type when only bearing is processed, while figure 7b depicts the case where
discernment itself, the onus for hypothesis resolution is placed on the evidence gathering function. Similarly, a priori information obtained in-situ should be allowed to effect the distribution of belief.

The simplified examples presented are indicative of the processing capabilities required to generate and interpret evidence for the general case of parameter estimation in the presence of modeling uncertainty. Feature extraction should be matched to the modeled frame of discernment and can require a high degree of sophistication. Temporal reasoning capabilities may be required to ascertain event occurrence and event boundaries. This is especially true when multiple events can occur either simultaneously or sequentially. Future efforts will address this latter aspect.

REFERENCES


