13th ICCRTS
"C2 for Complex Endeavors"

"Understanding Information Uncertainty within the Context of a Net-Centric Data Model: A Mine Warfare Example"

Topic(s): 1, 7, 9

Name of Author(s)/Point of Contact: Megan Cramer,
APM for the Mine Warfare Environmental Decision Aids Library (MEDAL)
Name of Organization: PMS495, Mine Warfare Program Office

Complete Address: 490 M St. SW #403W Washington DC 20024
Telephone: 301-509-1450
E-mail Address: megan.cramer@navy.mil

DISTRIBUTION STATEMENT A:
Approved for public release; distribution is unlimited.

Abstract

This paper examines the challenge of assessing operational measures of effectiveness given incomplete and often imperfect information. With the migration of software applications towards a service-oriented architecture and net-centric capability, the ability to capture, quantify, and aggregate uncertainty of information within a semantic framework will be integral to conveying the true operational picture. A potential way to represent the uncertainty of available data is through the incorporation of probabilistic information within a C2-focused semantic data structure. This paper establishes a notional framework for associating probabilities within a content-rich data structure and demonstrates this framework for the Mine Warfare operational measures of effectiveness. The management of multiple variable inputs and the improved bounding of uncertainty over time are developed within a Bayesian context. Finally, the implications of introducing a new method for handling uncertainty within an information-centric data model are explored.

Introduction

With the advent of net-centric technologies and improved data gathering systems, tactical decision aids are achieving greater access to operationally-relevant information. The availability of tactical data provides greater opportunity to convey the level of uncertainty associated with a given mission. In particular cases, uncertainty surrounding an operational Measure of Effectiveness (MOE) may have a significant impact on both the variability of the metric calculation and the accurate communication of progress achieved towards the given metric. This paper utilizes the Naval Mine Warfare (MIW) example to explore the notion of uncertainty within an operational Command and Control (C2) context. Methodologies are considered for presenting uncertainty in information to both a human decision maker as well as to an automated expert system, which may be providing recommendations for potential Courses of Action (COAs). The hypothesis explored is that including probabilistic information within a semantic data model can be a useful tool for considering multiple COAs within an uncertain operational context.

The approach to this analysis is to briefly describe the mine warfare challenge and present a notional semantic framework to support C2 within a future data model. Semantic technologies have been identified as integral components to facilitate migration to a net-centric C2 architecture that can enable future expert systems. The semantic data framework, containing agreed upon definitions and hierarchical relationships to enable information exchange within a net-centric architecture, is based on the tactical contacts that may be found within the area of interest. These tactical contacts include all objects within the area of interest, both mines and non-mines. The number of total contacts in the area is a key assumption in the calculation of the primary MIW MOEs of the estimated risk to a transitor and the expected time required to clear all of the mines. The calculations required to arrive at these two metrics are then explained in detail. To convey the associated uncertainty, upper and lower bounds are drawn around the key metrics to convey the associated uncertainty. A methodology for determining an information score is derived for each MOE by considering both the inherent uncertainty in the probability as well as the underlying assumptions. This probabilistic information and information scoring technique are then tied back within the semantic framework previously described. Finally, using this data structure, several example COAs are generated to show the ability to make trade-offs between the two MOEs, given an uncertain operational context.

The MIW Challenge

The objective of the MIW mission is to reduce the risk of another ship hitting a mine while transiting through identified waterspace. Risk reduction is achieved through conducting mine countermeasures (MCM) effort. The time to perform MCM effort is often limited and therefore becomes an important constraint when considering various courses of action (COAs) to employ MCM effort. Figure 1 illustrates the MIW problem of reducing risk within a specific area to support the reduction of risk to a transiting ship through the area.

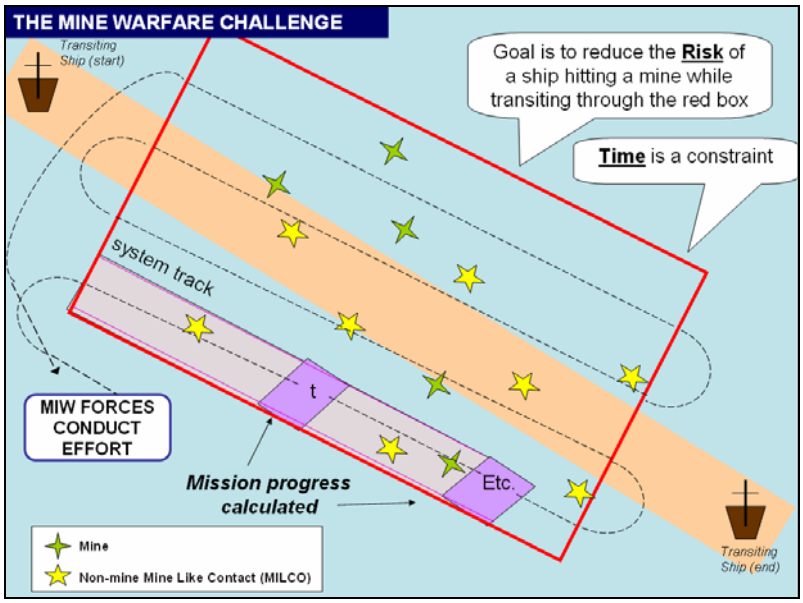

Figure 1: Mine Warfare Problem

In conducting MCM effort, mines and non-mines are discovered and prosecuted. "Risk" is defined as the probability of damage to the transiting ship and can be reduced through MCM effort. The expected time required to perform the MCM mission can be calculated by the number of all mine-like contacts (both mine and non-mine) in the area of interest. The fraction of mines removed, more commonly known as Percent Clearance, is an important underlying factor in determining the likelihood for the number of mines in the area and is a measure of the estimated results of MCM effort conducted in the area of interest. Percent Clearance can be calculated before any MCM effort has been conducted and with no knowledge of the number of mines in the area. This metric of the estimates results of the MCM effort can be updated as the mission progresses.

MIW Measures of Effectiveness (MOEs)

To calculate the Risk to a transiting ship and determine the expected time to conduct the MCM mission, it is useful to consider the underlying tactical contact information that is essential to determination of these operational objectives. The approach will be to create a semantic data model for both the underlying tactical contacts and the overarching MOEs for the MIW mission. This semantic data model will focus on the incorporation of probabilistic information as a method for incorporating uncertainty information within a net-centric architecture. This paper applies and extends concepts described by George Mason doctoral student Paulo Cesar G. da Costa in his doctoral dissertation *Bayesian Semantics for the Semantic Web* (Costa, 2005). In this earlier body of work, an ontology standard is developed for Bayesian probabilistic semantics.

To illustrate the making of a MIW data model, Figure 2 shows an abstraction of the contacts, both mine and non-mine, within the area. Basically, a contact is shown in the area, whether or not it is a mine.

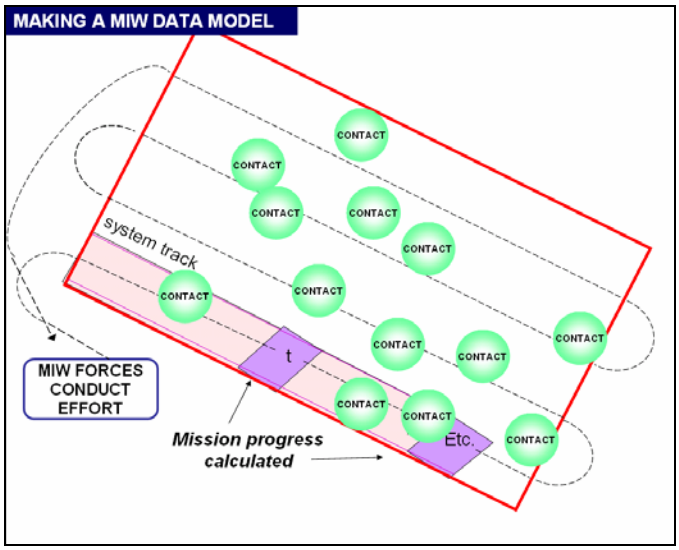

Figure 2: Making a MIW Data Model

Figure 3 focuses on the tactical contact and provides some examples of the types of metadata that could be associated with a tactical contact. These categories are notional only and meant to be representative of the types of metadata that might be included in a semantic data model.

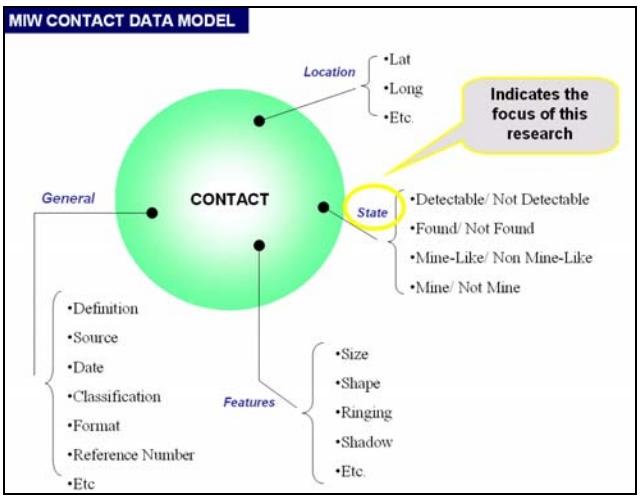

Figure 3: Making a MIW Contact Data Model

For the purposes of incorporating probabilistic information within a semantic data framework for each MIW MOEs, it is necessary to provide "state" information as part of the data model for tactical contacts. A state is defined as the outcome of an event and can therefore be described by random variables. The states that are important to determining the MIW MOEs are:

- *Whether or not a contact is detectable*
- *Whether or not a contact has been found*
- *Whether or not a contact is mine-like (as determined by feature characteristics)*
- *Whether or not a contact is a mine*

Figure 4 depicts the combination of these various states with respect to the entire set of tactical contacts that exist in the area. (Note that contacts that are non-mines and also not detectable are not included as they do not directly impact either MIW MOE.) All other combinations of the states are shown to be mutually exclusive and when considered together, collectively exhaustive of the sample space. The sample space as shown in the diagram, represented by Ω, is the total number of contacts in "ground truth" in the operational area of interest.

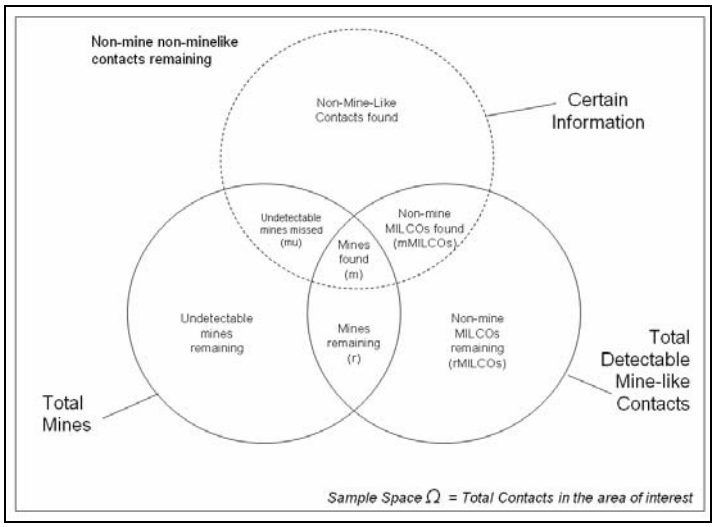

Figure 4: MIW Tactical Contact States

The solid line circle on the left illustrates the total number of mines in the area, which are composed of detectable mines already found, detectable mines remaining, undetectable mines missed, and undetectable mines remaining. The solid line circle on the right describes the set of total mine-like Contacts (MILCOs) that are detectable in the area, which is made up of both mines and mine-like non-mines in the area of interest. Note that mine-like non-mines are described as MILCOs throughout this document. The total set of mine-like contacts in the area is composed of mines found, mines remaining, non-mine mine-like contacts, and non-mine mine-like contacts not yet found. The dotted lined circle at the top of the graphic represents information that is known with certainty. The information that is known by operational forces includes mines found, non-mine mine-like contacts found, a fraction of undetectable mines that have been missed (usually estimated), and non-mine non-mine-like contacts that have been found. As might be expected, information that is known with certainty is influential in determining estimates for uncertain information external to this circle within the sample space.

Calculation of the MIW mission objectives can be calculated within the context of this diagram. Percent Clearance is typically a driver in the MCM effort and is calculated as the estimated fraction of mines removed. It is important to note that this probability can be calculated based on a guess the number of mines that will be cleared through MCM effort. The primary MOE of Risk, or Probability of Damage to a Ship Transitor, is calculated by using information in the highlighted circle on the left, to include both assumed prior information and new information gained throughout the mission. In Figure 5, the determination of Percent Clearance (shaded in gray), based on knowledge of MCM

effort applied in the area, is the common measure utilized by MCM forces to address the MCM problem.

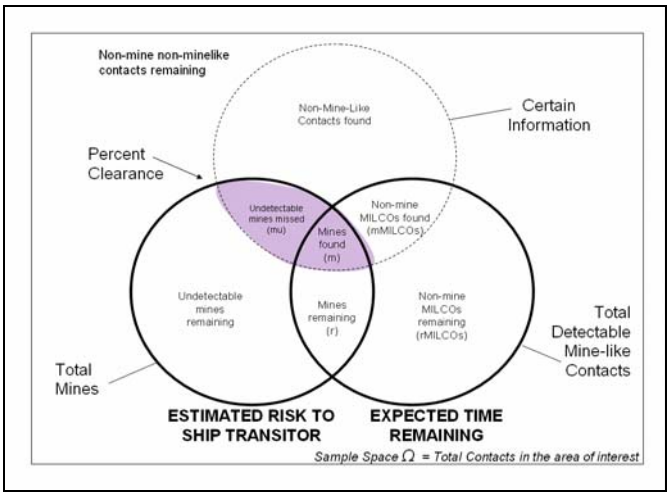

Figure 5: Relationship to MIW MOEs

Similar to the calculation of Risk, the set of information represented by the circle on the right in the above diagram can be utilized to calculate an expectation of the time remaining to complete the operation to the desired level of clearance. For example, the number of false alarms will drive the MCM timeline, if not directly impact the Risk MOE. In the following section, a methodology is described for calculating an expectation for the time remaining in the MCM operation. Upper and lower bounds are provided to qualify a range of uncertainty around this estimate.

As the mission progresses, the set of information that is certain will become proportionally greater compared to the overall sample space representing ground truth. As this circle of certainty expands throughout the mission, the amount of uncertainty surrounding progress towards the mission objectives is correspondingly reduced. This research will present a methodology for calculating an information score associated with both MOEs described above. This information score methodology could serve as a useful tool for conducting the ongoing trade-off analysis between MOEs, using uncertainty as the negotiating factor.

Estimated Risk (Probability of Damage)

A significant amount of work in MIW research as been focused on the determination of Estimated Risk, defined as the probability of damage to a transiting ship caused by a naval mine. The current approach to calculating risk will be discussed, followed by extensions to this work to elaborate upon the presentation of uncertainty information.

It can be observed that the concept of risk, and its calculation as a probabilistic measure, carries with it an inherent association with the notion of uncertainty. The attempt of this research is to communicate to the operational commander an understanding of risk and its associated uncertainty, which is more rich than simply providing a single risk value. This

approach attempts to account for complexities found in a real operational situation, such as the uncertainty associated with known data, accumulation of additional information, and the sensitivity of the metric to assumptions.

From an operational perspective, there are two important points associated with utilizing the current approach. The first point is that there must be some information known about the presence of mines within the area in order to conduct the calculation. This information may be either information known with certainty (preferable) or else a guess may constitute that necessary information. The second point is the sensitivity of this metric to the total number of mines assumed in the area. The number of mines assumed in the area becomes a driver of the metric, although it should be noted that this sensitivity is less as the number of total mines is increased.

To address the second point, the discussion of the presentation of this risk metric to the operation user has been raised, to enable the accurate communication of the metric and its sensitive to the assumptions by the operator. For this reason, it is useful to calculate Uncertainty Bounds around the expectation of Risk.

Literature Review: Calculating the Estimated Risk

The research in this arena of calculating the Estimated Risk to a transiting ship follows a Bayesian approach of determining the *a priori* distribution of mines and determining the likelihood function according to known information as to the number of mines found and clearance operations conducted in the area of interest. This Bayesian approach to calculating risk is described in detail in both the Decision Aid for Risk Evaluation (DARE) algorithm description document (Bryan, 2006) and a recent article published in Military Operations Research by Wagner Associates (Baker and Monach, 2006). Normalization is achieved by dividing by the sample space of all possibilities of the total number of mines in the area. The posterior distribution of mines remaining in the area is therefore determined from the number of total mines estimated.

The variables that are required to determine the probability of the number of total mines in the area, given known information about the number of mines found and the fraction of mines removed are provided below:

- n= total mines in the area in ground truth
- m= mines found
- p=percent clearance
- r= mines remaining.

The inputs into the calculation include $\Pr(n)$ for the prior distribution, $\Pr(m|n,p)$ as the likelihood function, and $\Pr(n|m,p)$ as the posterior distribution. [1]

[1] A research focus for the mine warfare research community has been to determine the appropriate *a prior* distribution to use for the probability mass function of mines assumed in the area. Karna Bryan and Wagner Associates have improved upon solutions for calculating the prior, including implementation of a Dirchelet approach to the prior that is considered superior to a straight multinomial prior. The discussion surrounding the appropriate prior has revolved around the importance that the 'learned' information of the likelihood function should play within the Bayesian update. Because this area is not the focus of this research and for

$$\Pr(n \mid m, p) = \frac{\Pr(m \mid n, p) * \Pr(n)}{\sum_{i=1}^{i=n} \Pr(m \mid n, p) * \Pr(n)} \tag{0.1}$$

The likelihood function is given by $\Pr(m \mid n, p) = \binom{n}{m} P^m (1-P)^{n-m}$. (One adjustment that is often made in practice for the number of mines found *m* is to adjust upwards to *m+1* to error on the conservative side.) The normalization function in the denominator is determined by determining the total probability (sum) for all possible values of n. Pr($r|m,p$) can be inferred directly from the posterior given $r = n-m$.

Once Pr($r|m,p$) is derived, the expected number of mines remaining in the area is the expectation for r given *m* mines are found and *p* percent clearance achieved.

$$\mathrm{E}(R) = \sum_{r=0}^{\infty} r * \Pr(r \mid m, p) \tag{0.2}$$

Risk can be calculated for each transitor given the expected number of mines. The required information for this calculation is the probability of mission abort for each transitor. Probability of Damage can therefore be calculated for multiple mine types and area segments, in addition to multiple transitors.

$$\Pr(D) = \mathrm{E}[\Pr(D \mid r)] \tag{0.3}$$

Probability of Damage, or Estimated Risk, is therefore the expected value of the probability of damage given a certain number of mines remain in the area.

<u>Expounding on Uncertainty</u>

Uncertainty bounds can be determined on the posterior using a standard Bayesian credibility interval approach. Error can be given as an input ε to calculate the range values for the integral for the posterior probability determined above for Pr($r|m,p$), represented here as the function f($r|\ m,p$) where *r* represents a realization of the random variable *R* in the sample space Ω. *R* can be interpreted as the set of all possible outcomes for the number of mines remaining in the area. Note that a Bayesian Uncertainty Bound is analogous to a Confidence Interval in traditional statistics

$$\Pr(a(m,p) < R < b(m,p) \mid m, p) = \int_{a(m,p)}^{b(m,p)} f_{R|m,p}(r \mid m, p)\, dn = 1 - \varepsilon \tag{0.4}$$

---

the purposes of simplicity, a uniform prior distribution is utilized throughout this research. Therefore, Pr(*n*) assumes mines are distributed randomly across the entire operational area.

By drawing a range around the Probability of Damage (Risk) MOE, the uncertainty associated with this metric can be communicated to the operational user. Additionally, the objective then becomes the reduction of uncertainty around the MOE. As uncertainty is reduced, the bounds can be narrowed around the Risk MOE metric, thereby communicating to the user a level of "confidence" in that information.

Expected Time Remaining

Expected Time Remaining to accomplish the mission is an important parameter for MIW operations. Inputs into this expectation must include the number of non-mine MILCOs in addition to the number of actual mines in the area of interest. Expected time is defined as the long-term average time required to identify every remaining MILCO in the area as either a mine or a non-mine.

A similar methodology that has been used in determining the number of mines remaining in an area can be applied to determine the number of detectable mine-like contacts in a given area. As described before, the number of detectable MILCOs can be an important consideration in an operation, even if it does not directly impact the calculation of Risk MOE or the fraction of mines removed (Percent Clearance). The reason for its importance is because the number of MILCOs in an operational area is a tremendous driver in both the timeline to accomplish the mission objectives and the systems that should be utilized to counter the mine threat.

A short discussion on Percent Clearance is warranted here to provide some context to the above:

<u>Literature Review: Percent Clearance</u>
A measure of success in removing the mines in an area, Percent Clearance p is the average cumulative probability that a mine located at any given point within the area has been removed. (Removal implies that the mine must be first detected as a contact, then classified as a MILCO, identified as a mine, reacaquired for purposes of neutralization, and finally, neutralized.) Before the first mine is found, Percent Clearance is estimated according to a level of confidence using a straight-forward negative binomial approach. Once the first mine has been discovered, the determination of Percent Clearance changes to account for effort applied towards reducing the number of mines in the area. Cumulative effort of Percent Clearance p_{cum} towards removing the mines includes the probability of success in using the two kinds of MCM techniques, mine-hunting p_{hunt} and mine-sweeping p_{sweep}. This can be determined by $p_{cum} = 1 - (1 - p_{hunt})(1 - p_{sweep})$.

The calculation of the probability of success in utilizing mine-sweeping techniques assume all mines found are neutralized and also accounts for the reliability of the sweeping system, indicated by Pr*(s)* (according to a conditional survivor function). Note that *mu* is the fraction of undetectables mines. Percent Clearance for sweeping systems p_{sweep} is therefore shown by $p_{sweep} = (1 - mu) * \Pr(s)$.

The calculation of the probability of success in utilizing mine-hunting techniques also includes the fraction of undetectables mines *mu*, the probability of correctly classifying a mine as a mine-like object Pr*(c)*, the reliability of the hunting system Pr*(h)*, and the probability of identification and removal techniques succeeding once the mine has been identified as a mine-like contact *Bn*. For purposes of this discussion, *Bn* described here will be decomposed into Probability of Identification Pr*(id)*, Probability of Reacquisition of the MILCO Pr*(reacq),* and Probability of Neutralization Pr*(neut).* Therefore, calculation of Percent Clearance for mine hunting systems p_{hunt} is
$p_{hunt} = (1 - mu) * \Pr(h) * \Pr(c) * \Pr(id) * \Pr(reaq) * \Pr(neut)$.

There has been much discussion as to the information that should be updated to calculate Percent Clearance in the event of replanning and updating *p* to incorporate new information obtained throughout the operation. The update of *p* and the potential inclusion of conditional probabilities within the stages of MCM effort is an area of future research.

Calculating Expected Time Remaining

To the previous question on calculation of the Expected Time to conduct MCM effort such that every contact in the area is identified as a mine or a non-mine, it is first necessary to determine the Probability of the number of total detectable MILCOs in the operational area. This can be determined according to a Bayesian approach similar to that used to determine the number of total mines in the area. The difference here is that the likelihood is determined according to a multinomial distribution.

The information that is required to determine the (posterior) probability of the number of total detectable MILCOs in the area, given *nMIL* number of MILCOs found, is as follows:

- *fa*= number of false alarms (or MILCOs confirmed not to be mines)
- *m=number of* mines found and confirmed as mines
- *rMIL*= number of detectable non-mine MILCOs remaining
- *r*=number of mines remaining to be found
- *q*=Percent Confirmed

Depicted in Figure 6, Percent Confirmed is a new term and will be described here. Similar to Percent Clearance describing the fraction of mines removed, Percent Confirmed is the fraction of detectable MILCOs that have been confirmed either as mines or non-mines (false alarms). It is useful in determining the number of detectable non-mine MILCOs that will most likely be found in the area, thereby affecting the overall time expected to confirm every detectable MILCO in the area as either a mine or a false alarm.

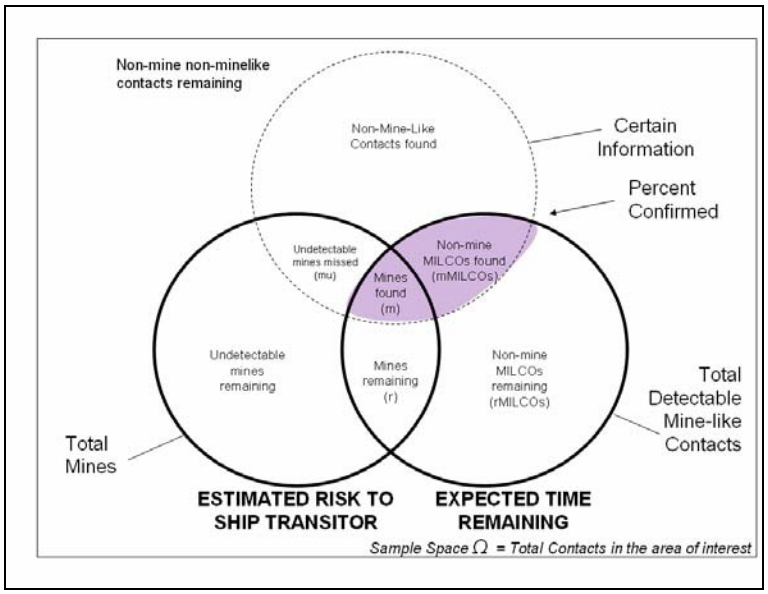

Figure 6: MIW Tactical Contact States

Pr(*fa,m,r*|*nMIL,q*) is the likelihood function and describes the probability of false alarms given that there are *nMIL* total detectable MILCOs in the area and there can be estimated a probability of encountering a false alarm. The likelihood is more complex this time as there is now a multivariate distribution (analogue to the binomial distribution used before). Probabilities for *fa*, *m, rMIL,* and *r* can be determined by dividing each by nMIL, or the estimated total detectable MILCOs in the area. The probabilities will be referred to as P_{fa}, P_m, P_{rMIL}, and P_r. Percent Confirmed is the joint probability of P_{fa} and P_m and therefore, further defined as q= P_{fa} * P_m since the probabilities are independent.

The multinomial distribution for the likelihood is provided below.

$$\Pr(fa,m,r \mid nMIL,q) = \frac{nMIL!}{fa!m!r!(1-fa-m-r)!} P_{fa}{}^{fa} P_m{}^{m} P_r{}^{r} (1-P_{fa}-P_m-P_r)^{(1-fa-m-r)} \quad (0.5)$$

The prior Pr(*nMIL*) describes the probability of the number of detectable total MILCOs in the area. Again, a uniform prior is chosen for simplicity.

The posterior Pr(*nMIL*|*fa,q*) provides the probability of the number of total detectable MILCOs given some information about the probability of false alarms and the number of false alarms already found .

$$\Pr(nMIL \mid fa,m,r,q) = \frac{\Pr(fa,m,r \mid nMIL,q) * \Pr(nMIL)}{\sum_{k=1}^{k=nMIL} \sum_{z=1}^{z=fa} \sum_{y=1}^{y=r} \sum_{x=1}^{x=m} \Pr(fa,m,r \mid nMIL,q) * \Pr(nMIL)} \quad (0.6)$$

The normalization is accomplished by summing over all possible combinations for the number of detectable nMILCOs in the area, the number of false alarms found, the number of mines found, and the number of mines remaining. Pr(*rMIL*|*fa,m,r,q*) can be inferred

directly from the posterior given *rMIL= nMIL-m-fa-r* where the number of mines remaining *r* can be estimated. Additionally, uncertainty bounds can be then calculated around the Pr(*rMIL|fa,m,r,q*) according to the same process described for the Pr(*r|n,P*).

Similar to the analysis for the MOE of Risk, Uncertainty Bounds can be calculated around the probability of the number of detectable non-mine MILCOs remaining in the area. Error can be given as an input ε to calculate the range values for the integral for the posterior probability Pr(*rMILCOs|fa,m,r,q*), shown here as the function f(*rMILCOs*| *fa,m,r,q*) where *rMILCOs* represents a realization of the random variable *RMILCOs* in the sample space Ω. *RMILCOs* can be interpreted as the set of all possible outcomes for the number of detectable non-mine minelike contacts remaining in the area.

$$\Pr(a(fa,m,r,q) < RMILCOs < b(fa,m,r,q) \mid fa,m,r,q) = \int_{a(fa,m,r,q)}^{b(fa,m,r,q)} f_{RMILCOs \mid fa,m,r,q}(rMILCOs \mid fa,m,r,q)dn = 1-\varepsilon \tag{0.7}$$

The sum of the expected time to address all remaining mines and detectable MILCOs can also be used to determine a value for the Time expected to complete the mission to a certain level of Risk. The calculation of this expectation is explained below.

The expectation of the time remaining to conduct MCM effort, or to identify every MILCO as either a mine or a non-mine, is based on both the expected number of mines remaining and the expected number of detectable non-mine MILCOs remaining in the area. The overall expectation is determined by multiplying the expected time to accomplish each task in the MCM sequence by the number of times that each task must be completed for each mine or detectable non-mine MILCO. For every MILCO that is found (either a remaining mine *r* or a detectable non-mine MILCO *rMILCO*), the MCM tasks of Detection, Classification, and Identification must be accomplished. Once a MILCO is positively identified as a mine, the MCM tasks of reacquisition and neutralization most be completed. The times for each MCM task are specifically Average Time for detection *T(det),* Average Time for Classification *T(class),* Average Time for Identification *T(id),* Average Time for Reacquisition *T(reaq),* and Average Time for Neutralization *T(neut).*

The expected mines remaining E(*r*) can be found by $\mathrm{E}(R) = \sum_{r=0}^{\infty} r * \Pr(r \mid m, p)$. The expected detectable non-mine MILCOs E(rMILCOs) can be found similarly by the following equation.

$$\mathrm{E}(rMILCOs) = \sum_{r=0}^{\infty} rMILCOs * \Pr(rMILCOs \mid fa,m,r,q) \tag{0.8}$$

The calculation to determine the expectation for the time remaining for both the expected mines remaining E[*Time_Remaining_r*] and the expected number of detectable non-mine MILCOS remaining E[*Time_Remaining_rMILCOs*] can be found where $E[Time_Remaining_rMILCOs] = \mathrm{E}[rMILCOs] * [T(det) + T(class) + T(id)]$ and

$$E[Time_Remaining_r] = \mathrm{E}[r]*[T(d)+T(class)+T(id)+(T(reaq)+T(neut)].$$

The calculation to determine the total time remaining to conduct the MCM mission *Total_Time_Remaining* can therefore be found from the expectation, where Pr(*rMILCOs*) Pr(*r*) can be found from the fraction of total tactical contacts anticipated. The total expected time remaining to complete the MCM mission is given below.

$$\begin{aligned} E[Total_Time_Remaining] = {} & \mathrm{E}[Time_Remaining_rMILCOs]*\Pr(rMILCOs) \\ & +\mathrm{E}[Time_Remaining_r]*\Pr(r) \end{aligned} \quad . \quad (0.9)$$

The advantage of this determination of the expected time remaining by the average time to complete each MCM task for all remaining MILCOs in the area is that the timelines include the additional time to consider false alarms within an area, in addition to the actual mines. The role of the environment, and particularly a high-clutter environment with many false alarms, is shown to directly influence the MCM MOE of Time.

Relationship between Estimated Risk and Expected Time Remaining

A MIW Commander (MIWC) must consider the relationship between the probability of damage (Estimated Risk) and the Expected Time Remaining to achieve a certain level of risk. The timeframe for an operation may not allow for the identification of every detectable MILCO in the area as either a mine or a false alarm. The MIWC must therefore look to employ the optimum number of assets to achieve a level of risk, often within a given limit of time.

According to the above analysis for each MOE, this trade-off between Time and Risk may also include the uncertainty surrounding both MOEs. This uncertainty can be determined by the bounds calculated for each MOE. For the purposes of this analysis, uncertainty will be shown for risk as that is the primary MOE with which the MIWC is concerned. Risk will be calculated as a function of time, in order to support operational use of this information. Note that the following analysis could also be used to show the uncertainty surrounding the Expected Time remaining as a function of Risk.

In order to show uncertainty around risk as function of time, a Poisson process is set up in MATLAB to simulate an MCM operation. At some constant rate, detectable MILCOs are found and categorized as either a mine or a false alarm. As each MILCO is discovered and appropriately identified, the Probability of Damage and the uncertainty bounds around that probability are calculated. Percent Clearance is assumed to constantly increasing until Percent Clearance of 1 is achieved. The Expected Time Remaining is determined from the expected mines and detectable non-mine MILCOs remaining in the area. In order to show this information most intuitively where time is increasing on the horizontal axis, the time remaining at each point is subtracted from the maximum time remaining that is found. The results of this Simulation #1 are provided in Figure 7.

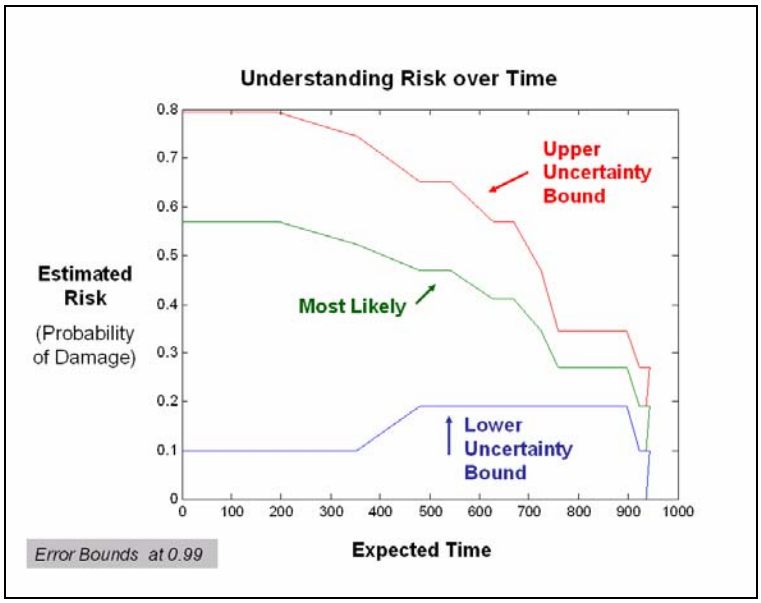

Figure 7: Simulation 1

Assumptions and the data generated as output from Simulation 1 are provided in Appendix A. It is important to note when looking at this process that the assumption of the *a priori* distributions for the number of mines remaining and the number of detectable MILCOs remaining are not updated throughout the simulation. The distributions remain constant in order to show graphically the uncertainty bounds in relationship to the expected time remaining.

Throughout a true operation, however, it would be more realistic to update the assumed prior distributions and consequently updated the expected time remaining in the operation. This is possible to do using the developed Poisson process, but the output does not lend itself to an easy graphical representation due to the changing values for the expected time remaining.[2]

As would be expected, Risk, or Probability of Damage, decreases over time as a larger proportion of the detectable MILCOs are discovered and identified as either mines or false alarms. The Uncertainty Bounds move closer towards the Probability of Damage estimate, thereby decreasing the uncertainty around the Risk MOE as effort is conducted and information is gathered on the MILCOs encountered in the area.

The error ε around the uncertainty bounds is an important input to generate the output in this simulation. Figure 8 shows the output from a second run of the simulation using the same assumptions as inputs, except for where error ε is .05.

[2] This update on the uncertainty bounds and the expected time remaining would be most useful in a replanning, or "running estimate" situation, where the *a priori* distributions are updated and held constant over the expected time remaining, for any given point in time.

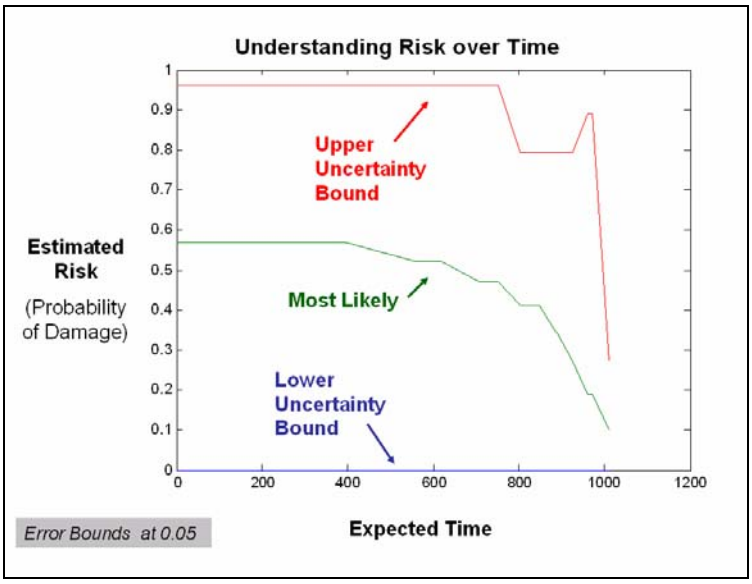

Figure 8: Simulation 2

Due to the random generation of mines or false alarms in the simulation engine, the Risk results are not exactly the same in this second simulation. The effect of the changed error input, however, is very discernable as the uncertainty bounds are now much farther away from the determined Probability of Damage output. The generated data for Simulation 2 is provided in Appendix B.

Information Scoring

Once the framework as been established for conducting the trade-off between Time and Risk MOEs, the question can be posed as to how to most efficiently reduce uncertainty around Risk as a function of Time. The method that is proposed is to determine an overall information score that incorporates both Risk and Time as a mechanism to determine those data inputs that are most important to effect an improvement in the overall information score. The information score is a mechanism for capturing the uncertainty inherent within the joint probability distribution of these two MOEs and in the uncertainty bounds around that probability. This scoring technique would be a useful tool by which to compare the relative information contribution of multiple variable inputs. The results of a sensitivity analysis of multiple data inputs on this overall information score will not be conducted within this paper but is intended as an area of follow-on research. The intent is to propose a mechanism that can be directly applied to convey both the data requirements to most directly reduce uncertainty and the importance of assumptions on the final answer. By utilizing this methodology discussed at a foundational level within this paper, it is proposed that uncertainty can be most efficiently reduced through the gathering of information throughout the operation.

Finding the Joint Probability

In order to determine uncertainty at the mission level, it is necessary to determine a probabilistic statement that encompasses both MOEs and anticipates the remaining MCM effort over which there exists the uncertainty. This can be accomplished by finding the joint probability for all tactical contacts remaining in the operational area. An illustration of this joint probability is shown in Figure 9.

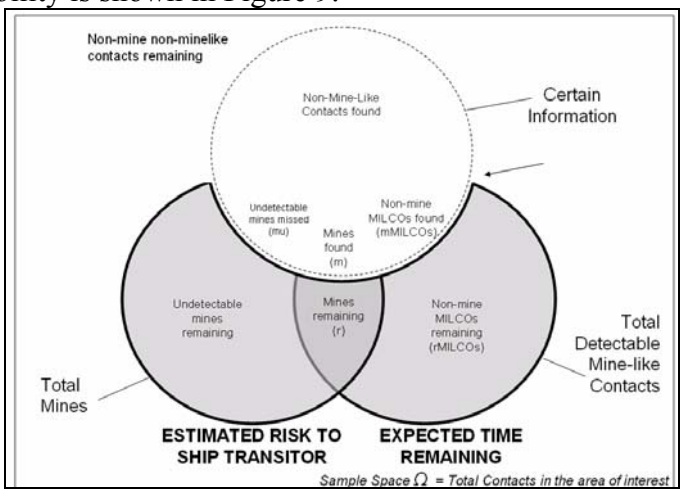

Figure 9: Joint Probability of MCM Effort

This uncertainty of the shaded area can be determined by multiplying Pr(*rMIL*|*fa,m,r,q*) and Pr(*r*|*m,p*) to find the joint probability of the number of mines remaining and the number of detectable MILCOs remaining and subtracting the covariance to account for the fact that these two probabilities are not independent. *p* is again the Percent Clearance and *q* is the Percent Confirmed.

The joint probability will be referred to as the Probability of Effort to conduct MCM Pr(MCM *Effort*) and is given below:

$$\Pr(MCM_Effort) = [\Pr(rMIL \mid fa, m, r, q) * \Pr(r \mid m, q)] - \mathrm{Cov}(rMIL, r) \qquad (4.1)$$

The covariance can be determined by considering the dependency between Pr(*rMIL*|*fa,m,r,q*) and Pr(*r*|*m,p*). The covariance can be determined for *rMIL*) and *r* by

$$\mathrm{Cov}(rMIL, r) = E[(rMIL - v)(r - w)] \qquad (4.2)$$

where *v* and *w* are the expected values for rMIL and r, respectively.

Information Scoring Approach

The Relative Information Scoring methodology described here is derived from a Classical Expert Judgment Model for determining the informativeness of multiple experts providing input on a subject. This Relative Information Scoring approach is described in detail within *Probabilistic Risk Analysis* (Bedford and Cooke, 2001). This situation is analogous to multiple information sources providing input to the overall mission, to

include the Risk and Time MOEs. Note that the information scoring technique is related to the method for calculating entropy, or the amount of uncertainty associated with a random variable.

The scoring approach is to compare the results from multiple data inputs against an empirical background measure. The realizations are found from the previously defined lower-bound, the actual point estimate, and the upper bound for the Probability of Damage, which together specify a 4-bin multinomial distribution. The probabilities for these bins can be determined by applying the previously defined error ε that was used to calculate the range values for the uncertainty bounds around the Probability of Damage. For example, if the previously defined error was 30%, then the 15%, 50%, and 85% percentiles would be specified and the multinomial bins would be distributed as p_i= (p_1, p_2, p_3, p_4) =(0.15, 0.35, 0.35, 0.15). Because there are 4 multinomial bins, then the number of random variables *n* is 4. The variable outcome *v* is the result (realization) of the multinomial experiment with probability distribution p_i, Let $q_i(e)$ denote input *e*'s i percentile where *v* is

v_1 [Interval 1] is [$q_l(e)$, $q_{15}(e)$] with probability p_1
v_2 [Interval 2] is [$q_{15}(e)$, $q_{50}(e)$] with probability p_2
v_3 [Interval 3] is [$q_{50}(e)$, $q_{85}(e)$] with probability p_3
v_4 [Interval 4] is [$q_{85}(e)$, $q_h(e)$] with probability p_4

The lower bound l for Interval 1 and the upper bound for Interval 4 are found where l= min{$q_{15}(1)\ldots q_{15}(m)$,$v$} and h=max{$q_{85}(1)\ldots q_{85}(m)$, v} where m are the number of inputs considered. Therefore, $q_l(e)$ = 1-k(h-l) and $q_h(e)$ = h+k(h-l), where k is a specified overshoot percentage. (k is 10% for this example.) Note that for cases where $q_l(e)$ is found to be less than zero, the value for $q_l(e)$ is constrained at zero.

Relative Information *I* is therefore

$$I(s,p) = \sum_{i=1}^{n=4} s_i Ln(s_i / p_i) \tag{4.3}$$

Assuming independence, (p_1, p_2, p_3, p_4) is the probability for each multinomial bin and (s_1, s_2, s_3, s_4) is the empirical distribution, and (v_1,v_2,v_3,v_4) is the realization of the average joint probability of the two MOEs in the corresponding intervals. s_i is the number of variables in interval *i* divided by the empirical estimate for p_i.

It is worth to note that in using this methodology to calculate the informativeness for every variable input into the joint probability of the MOEs, all information scores are determined against the same uniform empirical estimate and are therefore calculated relative to the other scores determined with respect to a common background measure.

Using this approach, both Percent Clearance and Percent Confirmed are influential in effecting the overall Information score through employing MCM effort. If this same information scoring technique was conducted using only Risk as a driver, with Percent Clearance as the sole motivation for MCM effort, the expected number of detectable non-mine MILCOs in the area would have no impact on the overall information score. This

result would therefore be counter-intuitive as one would expect the information score to improve as information, even contextual information, is discovered.

The usefulness of this Information Scoring methodology is to provide a way of quantitatively evaluating operational courses of action, based on their respective ability to collect additional information, and of then presenting these options as recommendation(s) to the decision-maker. A quantitative approach allows an automated tactical decision aid to interpret the informational value of potential courses of action and provide recommendations as to how to improve situational awareness, even if this is not in direct support to the primary MOEs. An interesting consideration is where the collection of information itself can be a course of action and should be considered as a viable option. A simple example is that a mine warfare commander will often survey an area to gather information, before proceeding to further tactical operations. An experienced commander knows intuitively that gathering information is an important first step in the operation. This Information Scoring approach provides a computer-understandable methodology to arrive at a comparable conclusion and to present this possible course of action to a commander within an automated tactical decision aid. Potential options of operational "next steps" within an automated system would be based on the anticipated value of gathering new information, in addition to options to directly impact the operational MOEs.

Building a Probabilistic Data Model

Once the MOEs, their respective uncertainty bounds, and the overall information score has been determined, this metadata can be incorporated into the data model for the mission. As with the MIW Contact Data Model previously discussed, an abstract data model for the overall MIW mission can now be developed. Utilizing the state information captured into the lower level of the data structure (tactical contact level), probabilistic information can now be derived and aggregated at the higher level (mission level). Figure 9 shows a representative MIW Mission Data Model where "State" is included as a type of metadata, in addition to more traditional metadata types. State is now defined as a random variable for the number of contacts in the area for the previously identified states at the tactical contact level. The states at this more aggregated level are now:

- *Number of contacts that are mines that are detectable/not detectable*
- *Number of contacts that have been found/have not been found*
- *Number of contacts that are mine-like/not mine-like*
- *Number of contacts that are mines/not mines*

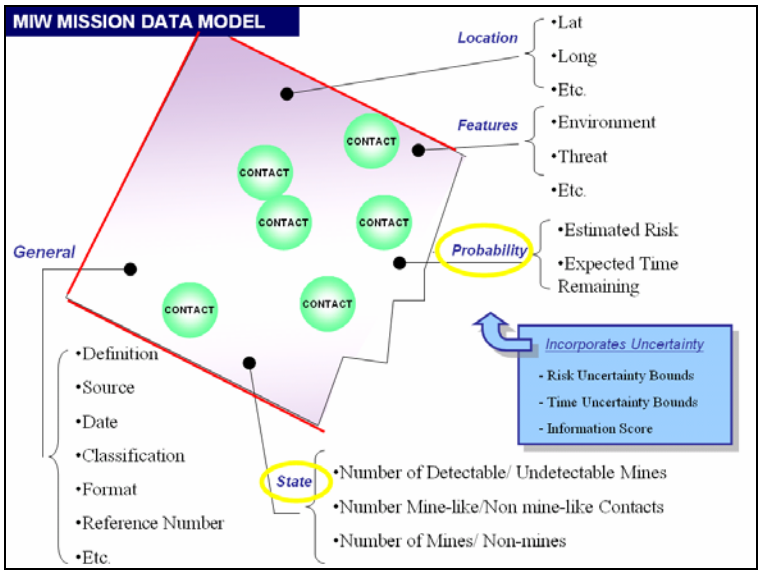

Figure 10: MIW Mission Data Model

Probability has now been added as a metadata type for this mission level data model. Uncertainty information is provided in several ways through the addition of this probabilistic metadata. The first way in which uncertainty information is conveyed is through the probability itself, which inherently conveys a level of uncertainty. The sensitivity of this metric to the underlying assumption of the number of contacts in the area is also communicated through the uncertainty bounds for both MOEs of Estimated Risk and the Expected Time to complete the MCM effort. Finally, an information score is provided to show the level of informativeness known with respect to the primary metric of the Estimated Risk to a transiting ship.

The utility of this probabilistic data model is illustrated in Figure 10. As the mission is conducted and additional information is collected, the data model can enable the recalculation of probabilities to show progress towards the mission objectives and the corresponding reduction in uncertainty over time.

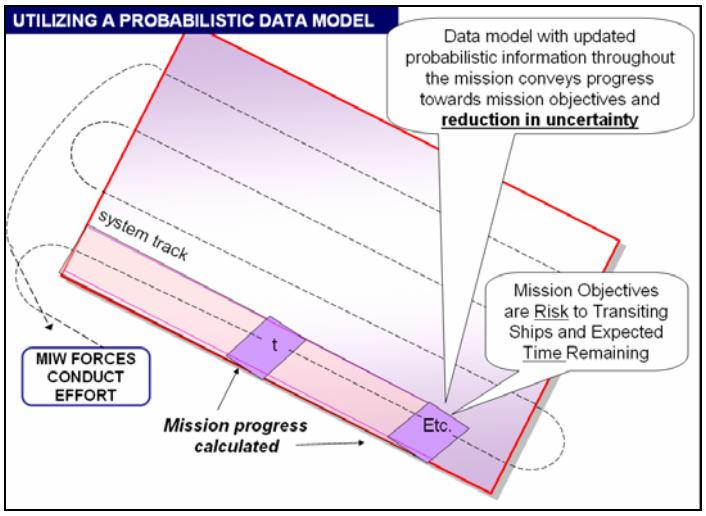

Figure 10: Utilizing a Probabilistic Data Model

This probabilistic information can also be incorporated using other methods beyond direct connection to a data model construct. With the move towards information-centric architectures, however, the incorporation of probabilistic information into semantic data models offers a flexible, robust, and scalable option for managing uncertainty within a net-centric and service-oriented operational environment

Conclusion

This research uses the mine warfare example to examine the importance of uncertainty in assessing Command and Control measures of effectiveness. A method is shown for determining uncertainty bounds for risk, a primary metric for this mission area. To show the trade-off between time and risk, a method is developed for determining the expected time remaining to conduct MCM effort. An information scoring technique is developed to assess the overall uncertainty associated with the mission. This overall information score is potentially useful in generating courses of action to increase understanding of risk within the constraints presented by a limited timeframe to conduct MCM operations. Of note, Percent Confirmed is a driver of contributing information to the MIW mission, in addition to the traditional MCM metric of Percent Clearance.

To support the determination of uncertainty within a net-centric C2 architecture, a strategy is presented to manage this additional information by expanding a semantic data model construct to include probabilistic information. This data-focused construct offers a simple and scalable approach to providing the context of uncertainty within an information-rich environment wherein multiple applications and services might be drawing upon common information.

References

Baker, Joni and W. Reynolds Monach, 2006. Estimating Risk to Transiting Ships Due to Multiple Threat Mine Types. *Military Operations Research* 2, no. 3:35-47.

Beasty, Michael, Art Degroat, Randy Horne, Glynndon Jones, Diane Kosky, Tim Logan, Curtis McVey, Kevin Oakes, Christina Smith, and Fred Sutter. 2007. *Mine Countermeasures (MCM) Theory Primer* (Unclassified). Technical Report NSWC PC/MP/HA12-07-003, Naval Surface Warfare Center- Panama City Division (NSWC-PCD).

Bedford, Tim, and Robert Cooke. 2001. *Probabilistic Risk Analysis: Foundations and Methods.* United Kingdom: Cambridge, University Press.

Bryan, Karna. 2006. *Algorithms for Decision Aid for Risk Evaluation (DARE) version 2.1.* Technical Report NURC-FR-2006-002, NATO Undersea Research Centre (NURC).

Costa, Paulo Cesar G., 2005. Bayesian Semantics for the Semantic Web. Ph. D.diss., George Mason University.

Mazzuchi, Thomas A., "Class Notes on Techniques of Risk Analysis" Washington D.C.: George Washington University. Spring 2007.

Rausand, Marvin and Arnljot Hoyland. 2004. *System Reliability Theory: Models, Statistical Methods, and Applications.* New Jersey: John Wiley & Sons.

Ross, Sheldon. 2006. *A First Course in Probability*. New Jersey: Pearson Prentice Hall.

Washburn, Alan. "Notes on Mine Warfare." Monterey: Naval Post Graduate School. <http://diana.cs.nps.navy.mil/%7Earwashbu/> October 2007.

U.S. Department of Defense. Program Executive Office Littoral and Mine Warfare. 2007. *Mine and Undersea Warfare Measures of Effectiveness.* Instruction 3370.1 (draft).

Appendix A: Simulation 1 Assumptions and Data

I. Simulation 1 Assumptions

Assumption Description	Assumption Vale	Comment
Average Detection Time	*AvgDetT= 30*	Average Time for all MCM systems
Average Classification Time	*AvgClasT= 5*	Average Time for all MCM systems
Average Time to conduct Identification	*AvgIDT=1*	Average Time for all MCM systems
Average Time to conudct Reacquisition	*AvgReacT=1*	Average Time for all MCM systems
Average Time to conduct Neutralization	*AvgNeuT= 2*	Average Time for all MCM systems
Uncertainty Bound Error	*ε= .99*	
Percent Clearance	*p=.95*	Percent Clearance is assumed fully achieved
Number of Mines in the area in Ground Truth	*n=9*	Number of the assumed number of mines in ground truth is not updated throughout the simulation
Number of Detectable MILCOs in Ground Truth in the area (including mines)	*nMILCOs=40*	Number of total detectable MILCOs is not updated throughout the simulation
Number of Mines Initially Found	*m=1*	Process begins after the first mine is found
Ship Damage Distance	*SD = 60*	Ship Damage Distance remains constant throughout the simulation
Channel Width Distance	*CW = 600*	Channel Width Distance remains constant throughout the simulation
Poisson Process for finding mine and non-mine tactical contacts	*constant rate (λ=1/25)*	rate used for simulation has no bearing on the results

II. Simulation 1 Data

Mines Remaining (r)	Detectable non-mine minelike Contacts (rMIL)	Number of mines +/- Risk to determine Probability Bounds (int)	Lower Bound	Risk (Probability of Damage)	Upper Bound	Estimated Time Remaining
8	30	7	0.1	0.5695	0.7941	1051.1
8	29	7	0.1	0.5695	0.7941	982.1
8	28	7	0.1	0.5695	0.7941	916.6
8	27	7	0.1	0.5695	0.7941	855.7
7	27	6	0.1	0.5217	0.7458	696.3
6	27	4	0.19	0.4686	0.6513	570.6
6	26	4	0.19	0.4686	0.6513	508
5	26	3	0.19	0.4095	0.5695	423.8
5	25	3	0.19	0.4095	0.5695	380.8
4	25	2	0.19	0.3439	0.4686	326
3	25	1	0.19	0.271	0.3439	292.8
3	24	1	0.19	0.271	0.3439	261.3
3	23	1	0.19	0.271	0.3439	236.9
3	22	1	0.19	0.271	0.3439	217.4
3	21	1	0.19	0.271	0.3439	201.5
3	20	1	0.19	0.271	0.3439	188.4
3	19	1	0.19	0.271	0.3439	177.4
3	18	1	0.19	0.271	0.3439	168.1
3	17	1	0.19	0.271	0.3439	160.2
3	16	1	0.19	0.271	0.3439	153.4
2	16	1	0.1	0.19	0.271	127.3
2	15	1	0.1	0.19	0.271	119.7
2	14	1	0.1	0.19	0.271	113
2	13	1	0.1	0.19	0.271	107.2
1	13	1	0	0.1	0.19	114.8

Appendix B: Simulation 2 Assumptions and Data

I. Simulation 2 Assumptions

Assumptions remain the same as in Simulation 1, except for the error used to determine the Uncertainty Bounds, which is $\varepsilon = 0.05$.

II. Simulation 2 Data

Mines Remaining (r)	Detectable non-mine minelike Contacts (rMIL)	Number of mines +/- Risk to determine Probability Bounds (int)	Lower Bound	Risk (Probability of Damage)	Upper Bound	Estimated Time Remaining
8	30	23	0	0.5695	0.9618	1051.1
8	29	23	0	0.5695	0.9618	982.1
8	28	23	0	0.5695	0.9618	916.6
8	27	23	0	0.5695	0.9618	855.7
8	26	23	0	0.5695	0.9618	799.5
8	25	23	0	0.5695	0.9618	747.8
8	24	23	0	0.5695	0.9618	700.2
8	23	23	0	0.5695	0.9618	656.4
7	23	24	0	0.5217	0.9618	493.4
7	22	24	0	0.5217	0.9618	460.9
7	21	24	0	0.5217	0.9618	432.9
6	21	25	0	0.4686	0.9618	345
6	20	25	0	0.4686	0.9618	327.4
6	19	25	0	0.4686	0.9618	312.5
6	18	25	0	0.4686	0.9618	299.8
5	18	10	0	0.4095	0.7941	247.7
5	17	10	0	0.4095	0.7941	239.1
5	16	10	0	0.4095	0.7941	231.6
5	15	10	0	0.4095	0.7941	225.2
5	14	10	0	0.4095	0.7941	219.7
5	13	10	0	0.4095	0.7941	214.8
5	12	10	0	0.4095	0.7941	210.7
5	11	10	0	0.4095	0.7941	207.1
5	10	10	0	0.4095	0.7941	203.9
4	10	11	0	0.3439	0.7941	165.2
4	9	11	0	0.3439	0.7941	162.8
3	9	12	0	0.271	0.7941	125.2
2	9	19	0	0.19	0.8906	90.4
2	8	19	0	0.19	0.8906	87.4
2	7	19	0	0.19	0.8906	84.9
2	6	19	0	0.19	0.8906	82.8
2	5	19	0	0.19	0.8906	81
2	4	19	0	0.19	0.8906	79.6
2	3	19	0	0.19	0.8906	78.6
2	2	19	0	0.19	0.8906	77.9
1	2	2	0	0.1	0.271	40.2

Appendix C: MATLAB models for Simulations

Can be provided upon request