A Risk-Based Approach to Sensor Resource Management

Dimitri Papageorgiou and Maxim Raykin

The Raytheon Company
Integrated Defense Systems
Woburn, MA 01801
{Dimitri_J_Papageorgiou,Maxim_Raykin}@raytheon.com

Abstract. We investigate the benefits of employing a suitable risk-based metric to determine in real-time the high level actions that an agile sensor should execute during a mission. Faced with a barrage of competing goals, a sensor resource manager must optimize system performance while simultaneously meeting all requirements. Numerous authors advocate the use of information-theoretic measures for driving sensor tasking algorithms, wherein the relative value of different sensing actions is calculated in terms of the expected gain in information. In this chapter, motivated by the sensor resource allocation problem in missile defense, we deviate from the information-based trend and propose an approach for determining sensor tasking decisions based on risk, or expected loss of defended assets. We present results of a missile defense simulation that illustrate the advantages of our risk-based objective function over its information-theoretic and rule-based counterparts.

Keywords: Sensor Resource Management, Risk, Missile Defense[1].

1 Introduction

This chapter addresses what we refer to as the sensor resource allocation problem, the problem of tasking a multi-modal sensor to perform high level sensing actions, e.g., search, track maintenance, and discrimination, over the course of a mission. A multi-modal sensor can collect data on objects and areas of interest using a variety of sensing modalities. This flexibility has led to marked improvements in detection of targets, kinematic estimation, and classification capabilities. At the same time, this large palette of sensing actions has also introduced challenges concerning the timely and efficient use of limited sensing resources. This chapter focuses on a specific sensor resource management problem that appears in the context of ballistic missile defense (BMD).

To date, a majority of sensor systems employ a prioritization scheme to determine which actions should be taken during a data collection interval. In this approach, a slew of Boolean conditions are quickly checked and then actions are

[1] The United States Missile Defense Agency approved this work for public release (07-MDA-2387).

M.J. Hirsch et al. (Eds.): Adv. in Cooper. Ctrl. & Optimization, LNCIS 369, pp. 129–144, 2007.
springerlink.com

chosen based on some pre-specified action plan. While this method has its advantages – requirements can be plainly stated, Boolean conditions are typically easy to verify in real-time, and the decision chain is traceable – it suffers in several notable ways. First, an approach based on a fixed set of rules cannot be completely adaptable to the evolving battlespace environment and can therefore be far from optimal. Second, adjustment of priorities is a delicate, time-consuming procedure. As new requirements are introduced into the system, it is difficult to ensure that the conditions are appropriately agreed upon so that a new requirement's priority is properly set. Third, glitches and irregularities in algorithm behavior may be difficult to diagnose due to tacit assumptions and makeshift implementation choices.

To circumvent these deficiencies, numerous suggestions have been propounded that provide a single metric able to automatically and simultaneously capture the complex tradeoffs involved when choosing between sensor allocations. A metric that has received considerable attention is entropy, which attempts to measure the uncertainty associated with random variables of interest. Within this information-theoretic framework, authors typically focus on Shannon entropy [6,12,15], Kullback-Leibler divergence [8,11,17,18], and Rényi divergence [9,10]. In this approach, sensor tasking decisions are made based on the principle that actions should be chosen to maximize the information expected to be extracted from the scene of interest. Within a Bayesian estimation framework, a good measure of the quality of a sensing action is the reduction in entropy of the posterior distribution that is expected to be induced by a measurement. For instance, when evaluating the benefits of a track update (or propagation without an update), these algorithms use the logarithm of the ratio of the determinants of the a priori and a posteriori covariance matrices as a measure of sensor effectiveness.

Although the use of entropy for judging a sensor's performance can be justified for a conventional battlefield situation, we believe that it is not the most suitable metric for BMD. Indeed, for traditional military applications, e.g., surveillance of enemy troops, ground target tracking, etc., the battlespace has infinite variability and the ultimate objective often cannot be stated precisely. The goal of a sensor or system of sensors in this situation may only be to maximize the amount of information collected for subsequent use in decision making. In contrast, the situation arising in BMD can be stated in precise terms (i.e., we have finite number of objects, each with finite degrees of freedom) and the underlying goal of BMD is clear and always the same – to minimize our losses from an enemy's missile attack. We will use the term *risk* for the expected value of this loss and consider risk reduction as a driver for a sensor's action, by which its performance should be judged. We briefly note that other authors [3,4,19] have considered a similar metric, but they may define it in different terms or apply it in different contexts.

The chapter is organized as follows. In the next section, we introduce discrimination risk and discuss the meaning of cost coefficients. The expressions for risk reduction are derived in Section 3. In Section 4, we describe a myopic approach to scheduling based on risk reduction and a heuristic approach to non-myopic scheduling. In Section 5, we present the results of a missile defense simulation that compares our risk-based approach with other competing methods. Conclusions are presented in Section 6.

2 Discrimination Risk and the Cost Coefficients

Consider the following situation: an object is to be classified into one of two classes, C_1 (threat) or C_2 (nonthreat). Let p_i be the current probability that the object belongs to class C_i, $i = 1, 2$. An object that is classified as a threat will be fired upon and destroyed by an interceptor, while an object that is classified as a nonthreat will be left unscathed. Let c_{12} denote the cost of an interceptor and c_{21} the cost of leakage, i.e., the cost of misclassifying a threat as a nonthreat. Then, the risk of declaring an object as belonging to class i, for $i = 1, 2$, is given by

$$\begin{aligned} R_1 &= c_{12}\,, \\ R_2 &= c_{21} p_1\,. \end{aligned} \tag{2.1}$$

Note that while R_2 depends on the probability p_1, R_1 does not depend on a probability because once an interceptor is launched, its cost is incurred regardless of whether or not the object was a threat.

The decision rule for class selection minimizes the risk R; that is, the object is declared to belong to the class C_i with the smallest R_i:

$$R \,=\, \min_{i=1,2} R_i \,=\, \min\,(c_{12},\, c_{21} p_1). \tag{2.2}$$

We assume that $c_{12} < c_{21}$, i.e., the cost of an interceptor is less than the cost of leakage, otherwise the decision is always made in favor of C_2 and the problem becomes trivial. Observe that while p_1 grows from zero to the "critical value" c_{12}/c_{21}, the decision is made in favor of C_2 (nonthreat) and the risk of this decision grows linearly from zero to c_{12}. Similarly, while p_1 grows from c_{12}/c_{21} to 1, the decision is made in favor of C_1 (threat), and its risk remains constant at c_{12}, representing the loss of an interceptor.

Regarding the origin and value of the cost coefficients c_{12} and c_{21}, it is a common misconception that the cost of an interceptor is just the monetary price of its production and is, therefore, negligible with respect to the potential loss of defended assets (quantified by a cost of leakage). We argue, however, that this line of reasoning is incorrect. Indeed, interceptors are our last defense against a missile attack. Moreover, at any given moment, we have a limited supply of them, which cannot be increased instantaneously. Expending interceptors now depletes their availability for future defense. Consequently, the cost of interceptors should regulate their use and reflect the balance between the demand for them now (or in the near future, before new interceptors can be produced) and their current supply. As such, the cost of interceptors has nothing or very little to do with the price of their production; rather, this cost is just a parameter, which should be selected by a commanding entity in such a way that expending interceptors at their current cost would be optimum with respect to the current military and political situation. Guidelines for how cost coefficients should be set are suggested in [14].

3 Risk Reduction

In this section, we derive expressions for the risk reduction due to two critical sensor actions, discrimination and tracking, and show that the expected value of the discrimination part of risk reduction is always nonnegative. In the first subsection we consider discrimination risk when the target can be classified into one of two classes, e.g., lethal and nonlethal. We then proceed by incorporating the risk due to the uncertainty in a target's kinematic state, which we call *track risk*, and conclude by considering the combined influence of classification and kinematic uncertainties on risk estimation in a general case of n classes.

3.1 Discrimination Risk in the Case of Two Classes

Suppose a sensor is trying to classify an object into one of two possible classes. If the sensor has an opportunity to collect an additional measurement on this object before making a classification decision, the risk associated with this object may be reduced. We now derive an expression for the expected value of this risk reduction. Let x be the feature which we measure and let $p(x|i)$, $i = 1, 2$ be the corresponding class-conditional probability density functions (PDFs). We assume that a sufficient amount of time has passed from the previous measurement of x so that the new measurement can be considered independent from the previous one. Then after the new value of x is measured, the probabilities are updated according to Bayes' rule and the new probabilities become

$$p_i' = \frac{p(x|i)p_i}{p(x)}, \quad i = 1, 2, \tag{3.1}$$

where $p(x) = \sum_{i=1}^{2} p(x|i)p_i$ is the PDF of the feature x. The updated risk of a classification decision, which is based on probabilities p_i', is

$$R' = \min(c_{12},\, c_{21}p_1') = \min\left[c_{12},\, c_{21}\frac{p(x|1)\,p_1}{p(x)}\right], \tag{3.2}$$

and its expected value is

$$\langle R' \rangle = \int R' p(x)\,dx = \int \min\left[\,c_{12}p(x),\, c_{21}p_1 p(x|1)\,\right] dx. \tag{3.3}$$

Using the normalization of $p(x|i)$, we have from Equation (3.3)

$$\langle R' \rangle \le \min\left[\int c_{12}p(x)\,dx,\, \int c_{21}p_1 p(x|1)\,dx\right] = \min(c_{12},\, c_{21}p_1) = R, \tag{3.4}$$

which means that the expected value of the new risk after an additional measurement is never larger than the old risk. This is a desirable mathematical property as we never anticipate, in expectation, to increase risk by collecting more information. Note that risk itself (as opposed to its expected value) can increase after an additional measurement due to an atypical result of the measurement.

3.2 Track Risk

Imperfect knowledge of a target's kinematic state may lead to an additional risk, which we term *track risk*. As before, we assume that the object may be either a threat (C_1) or a nonthreat (C_2). The case of several classes may be considered in a similar fashion (see Section 3.3). Since in the case of a nonthreat decision we will not shoot at the target, the risk of this decision remains the same as in Equation (2.1). The risk of a threat decision, however, will change. Namely, if we make this decision and shoot at the target, there is still some probability p_{miss} that the interceptor will miss, in which case, with probability p_1, we will suffer a loss of c_{21}. Correspondingly, the term $c_{21}p_1p_{miss}$ should be added to the risk of a threat decision, where we assume that only one interceptor is fired at the target. As a result, with track risk taken into account, instead of Equation (2.1), we will have

$$\begin{aligned} R_1 &= c_{12} + c_{21}p_1p_{miss} , \\ R_2 &= c_{21}p_1 . \end{aligned} \tag{3.5}$$

Apparently, $R_1 = c_{12} + p_{miss}R_2$. Therefore, if p_{miss} is sufficiently large, R_1 might become larger than R_2 even when p_1 is large (e.g., even when $p_1 = 1$). In particular, this will always be the case when $p_{miss} = 1$. In this situation, a nonthreat decision should be made regardless of the value of p_1. Thus, as one would expect, we should not shoot (and waste) an interceptor if the interceptor is guaranteed to miss the target in the first place.

The probability p_{miss} depends on, among other factors, a state estimation error covariance matrix Σ. With each successive track measurement, Σ changes as described by Kalman filtering equations, and so p_{miss} and R_1 will change accordingly. This change will measure the risk reduction utility of a track measurement. Namely, the corresponding risk reduction is $\Delta R = R(\Sigma) - R(\Sigma')$, where Σ' is a state error covariance matrix after the measurement and $R(\Sigma) = \min\left[c_{21}p_1 , c_{12} + c_{21}p_1p_{miss}(\Sigma)\right]$.

3.3 Modifications to Discrimination Risk Due to the Presence of Track Risk

Here we consider the situation when an object is to be classified into one of n classes $C_1, \ldots, C_n$. We denote the current probabilities as p_k, $k = 1, \ldots, n$, and introduce the set of nonnegative costs c_{kl} of declaring an object a member of class k when in fact it belongs to class l. Consequently, $R_k = \sum_{l=1}^{n} c_{kl}p_l$ is the risk of declaring an object a member of class C_k. In keeping with the convention that C_1 represents the class of lethal objects, we will set $c_{1l} = c_{int}$ for all $l = 1, \ldots, n$, where c_{int} is the cost of an interceptor. The risk of a threat decision R_1 will then be the same as in Equation (2.1).

Taking track risk into account implies corrections to our expressions for expected risk after an additional discrimination measurement. Indeed, during the time interval between discrimination measurements, the error covariance matrix evolves from its current value Σ to some new value Σ', as is typically described by Kalman filtering equations *without* a track measurement. The current risk of declaring the object as a member of class k is

$$R_k = \sum_{l=1}^{n} c_{kl} p_l + \delta_{1k} c_{leak} p_{miss}(\Sigma) p_1 \,, \tag{3.6}$$

where δ_{1k} is the Kronecker delta, equal to 1 for $k = 1$ and 0 otherwise, and c_{leak} is the cost of leakage. Correspondingly, the current risk is

$$R(\Sigma, p) = \min_k R_k \; = \min_k \left[\sum_{l=1}^{n} c_{kl} p_l + \delta_{1k} c_{leak} p_{miss}(\Sigma) p_1 \right] . \tag{3.7}$$

After a discrimination measurement, class probabilities get updated and Σ gets propagated. The new risk becomes

$$\begin{aligned} R(\Sigma', p') &= \min_k \left[\sum_{l=1}^{n} c_{kl} p'_l + \delta_{1k} c_{leak} p_{miss}(\Sigma') p'_1 \right] \\ &= \min_k \left[\sum_{l=1}^{n} c_{kl} \, \frac{p(x|l)\, p_l}{p(x)} + \delta_{1k} c_{leak} p_{miss}(\Sigma') \, \frac{p(x|1)\, p_1}{p(x)} \right] , \end{aligned} \tag{3.8}$$

and its expected value is

$$\begin{aligned} \langle R(\Sigma', p') \rangle &= \int R(\Sigma', p')\, p(x)\, dx \\ &= \int \min_k \left[\sum_{l=1}^{n} c_{kl} p(x|l) p_l + \delta_{1k} c_{leak} p_{miss}(\Sigma') p(x|1) p_1 \right] dx \,, \end{aligned} \tag{3.9}$$

while the expected risk reduction is $\langle \Delta R \rangle \; = R(\Sigma, p) - \langle R(\Sigma', p') \rangle$. Following the same logic as in the derivation of Equation (3.4), one can show that the expected value of the discrimination part of the decision risk [represented by the first term in Equation (3.9)] never grows as a result of a discrimination measurement.

4 Sensor Resource Management Algorithms

Having derived expressions for risk and risk reduction associated with kinematic estimation and classification, we now incorporate these calculations into various sensor resource management (SRM) algorithms, wherein a resource manager tasks a sensor to perform actions in an effort to minimize expected risk. After describing myopic and far-sighted SRM algorithms, we outline how a far-sighted risk-based approach can be extended to facilitate hierarchical control in a multisensor system.

4.1 Myopic Sensor Resource Management

Using the expressions for risk reduction derived in the previous section, it is straightforward to suggest a myopic resource management algorithm for a single sensor which strives to achieve the fastest possible rate of risk reduction (RRR) over the next data collection interval. Prior to every data collection interval, we assume the sensor has the choice of applying one of several waveforms to any target. If there are n_w available waveforms and n_t targets, then there are a total of $n_t n_w$ action-object pairs from which to choose. For each pair we can calculate the fraction $f_{ij} = \mathrm{ERR}_{ij}/d_i$, where ERR_{ij} is the expected risk reduction due to the application of waveform i to target j, and d_i is the amount of sensor resources or duty required to perform action i. Obviously, f_{ij} represents the rate at which the risk is expected to decrease due to resources spent. Being myopic, we would like to maximize this rate, and so, the algorithm selects the action-object (here, the waveform-object) pair that maximizes f_{ij}.

4.2 Far-Sighted Sensor Resource Management

The myopic algorithm just described minimizes expected risk after the next sensor action is taken. If that were the time when a final decision had to be made, then this algorithm would be optimal. This, however, is rarely the case, as the information collected now is usually used (much) later. An ideal planner would, instead, have a far-sighted planning horizon and be able to enumerate all possible action-object pairs up to some future deadline for all threats. For each threat, it would compute the expected risk resulting from a sequence of actions taken up to that deadline. Finally, based on the risk associated with the various action sequences, it would then task the sensor with the best possible action-object pair for next planning interval, allow the system to evolve, and then repeat the process. A standard approach to tackling such a problem is to formulate it as a finite-horizon Markov Decision Process, also known as Stochastic Dynamic Program, although some authors reserve the latter name to characterize solution methods for this class of problems. Classic references include [1,2,13,20]. Although we have investigated a number of approximate dynamic programming approaches, our formulations and solution methods lie outside the scope of this discussion. Instead, we briefly describe a heuristic approach for far-sighted SRM, akin to the "critical ratio" algorithm given in Feinberg et al. [5], which will also set the stage for our discussion of hierarchical control in multisensor resource management.

Our heuristic approach is based on the observation that the myopic algorithm leads to the appearance of an expected residual loss, or residual risk, i.e., a risk which is impossible or very difficult to eliminate once it has been incurred. For example, residual risk appears if the sensor fails to detect a new target before it leaves a search volume, or fails to collect enough information about a target which is due to be intercepted. Obviously, resource management should be done in such a way as to avoid the appearance of residual risk. The reason it appears in the myopic approach is that the sensor fails to accomplish some goals by their corresponding deadlines. We therefore conclude that for critical tasks similar to those just mentioned, goals and corresponding deadlines should be imposed on the resource manager in addition to the objective of maximizing RRR.

Let N be the total number of tasks the sensor is executing, where by task we mean a particular sensor activity, such as tracking or discrimination, performed on a particular object. For every task i, let d_i be the remaining time until the deadline by which this task should be accomplished, i.e., its goal should be achieved. Goals and deadlines are set by a so-called battle manager. We will assume that given the goal i for task i and the current state of our knowledge, we have some predictive capability to determine a conservative estimate of the expected time t_i the sensor needs to spend on the corresponding task in order to accomplish it. On every iteration, the algorithm orders tasks according to their deadlines, so that $d_1 \leq d_2 \leq \cdots \leq d_N$, and for every $k = 1, \ldots, N$, it checks if there is enough time left to accomplish the first k tasks with some safety margin. In other words, the algorithm verifies if $\alpha \sum_{i=1}^{k} t_i < d_k$, $k = 1, \ldots, N$, where α is a "safety factor" which should be greater than 1. If this inequality holds for all k, then there is no need to worry about deadlines, and the algorithm follows the original RRR logic. If, however, for some k the inequality is violated, then the algorithm finds the "most critical" task index $\hat{k}$ such that

$$\hat{k} = \arg\max_{k} \left(\alpha \sum_{i=1}^{k} t_i - d_k \right) \tag{4.1}$$

and schedules a measurement required by the task $\hat{k}$. If this measurement takes time τ to be executed, then after this measurement both $t_{\hat{k}}$ and $d_{\hat{k}}$ become smaller by τ, and $\alpha \sum_{i=1}^{\hat{k}} t_i - d_{\hat{k}}$ becomes smaller by $(\alpha - 1)\tau$. Since $\alpha > 1$, the task is now less critical than it was before the measurement. In the event it is impossible to satisfy all deadlines, the algorithm first sacrifices the task with the smallest residual risk.

4.3 A Hierarchical Multisensor Control Architecture

Thus far, we have limited our discussion to risk-based resource management algorithms for a single sensor in missile defense. We assumed that for each task, a sensor has a corresponding goal and deadline, which can be incorporated into a risk-based approach for optimizing the set of actions taken in the subsequent data collection interval. In this section, we describe a natural extension of our risk-based approach to a hierarchical decision-making architecture for multisensor resource management. Such hierarchical approaches have gained increasing attention over the past decade in the reinforcement learning domain [16], and are well suited for the missile defense problem in which a distributed architecture is already in place. This hierarchical architecture facilitates solution of the (intractable) global problem of assigning all sensors a set of actions to perform over the course of a mission by decomposing the larger long-term problem into smaller short-term problems. In this way, a hierarchical architecture exploits a "divide-and-conquer" mentality for solving complex, large-scale problems.

In the current missile defense decision-making architecture, a Battle Manager (BM) acts as a commanding entity that tasks participants (e.g., sensors, platforms, and interceptors) throughout a given mission to collect information or execute some plan. A BM maintains an integrated picture of the battlespace, or in dynamic programming terminology, the state of the system, including (1) target-related information like track accuracy, classification, predicted impact point, and estimated time to impact, (2) sensor-related information, including sensor capabilities and current tasking, and (3) weapon system-related information, including the number of available interceptors and interceptor capabilities.

The ability of our approach towards sensor resource management to accommodate goals and deadlines allows us to naturally insert it into a general hierarchical framework of system management for coordinating multiple sensors. Given all available information about existing missile complexes, which consist of one or more targets spawned from the same object, the BM interacts with individual sensors with some periodicity, known as a Battle Manager Planning Interval (BMPI), collecting information obtained during the previous data collection interval, and giving assignments for the next interval. Based on known trajectories of the missile complexes relative to the positions of the sensors and known performance characteristics of the sensors, the BM creates a battle plan. For each BMPI and for each missile complex, the plan specifies which sensor or sensors will observe this missile complex, and with which task (tracking and/or discrimination). Search can be considered on the same grounds as a sensor activity related to a potential additional threat. Included in the plan, therefore, is the expected improvement of our knowledge of this missile complex's tracking and classification characteristics (or of the presence of a threat in a search volume). The plan is designed in such a way as to provide the smallest expected loss of defended assets and gets updated every BMPI according to the evolving situation. The generation of a battle plan is a separate (and complex) problem, which is not considered here. We do, however, consider the interaction of the BM with individual sensors assuming this problem has been solved.

In our hierarchical structure this interaction is organized as follows. At the beginning of each BMPI, after computing its own long-term plan, the BM assigns each sensor a set of targets and search volumes along with the associated cost coefficients and search/track/discrimination goals corresponding to the expected improvements mentioned above. The natural deadline for these goals is the end of this BMPI, although in certain situations the deadline could vary. Since it is possible for a sensor to achieve all of its goals by the corresponding deadlines and still have some remaining resources, the BM also informs each sensor about other targets, not assigned to it, and the value of their cost coefficients. Now each sensor finds itself in a situation described in the previous section: it faces a number of targets with associated cost coefficients, goals, and deadlines. Accordingly, each sensor acts as described above, without regard to the presence of other sensors, by attempting to determine an optimal plan over a shorter time horizon (a BMPI) that simultaneously meets all goals and deadlines while minimizing risk.

The results are reported to the BM at the end of a BMPI and will be used for updating the battle plan and creating an assignment list for the next BMPI.

5 Simulation Results

A low fidelity simulation environment was constructed in MATLAB to test and compare myopic sensor resource management algorithms that assign track and discrimination actions to multiple objects based on a heuristic policy or the maximization of a single objective function. The application of far-sighted approaches, which incorporate search management, is not considered here, but is discussed in [14].

We assume a single sensor has just begun tracking N objects. There is only one lethal target, known as a re-entry vehicle (RV), amongst the targets. At each time step, an action-object pair is selected depending on the algorithm used, and that action is then performed on that object. Each action takes the same amount of time to complete. Each object is assumed to belong to one of three classes (lethal objects belong to class 1), and classification is based on the measurement of three independent features. We assume that class-conditional PDFs are known and are Gaussian for each feature (see Figure 1). The sensor can make an observation on exactly one feature at a time when performing a discrimination action. This assumption could easily be relaxed. If the manager decides to measure a particular feature of an object, then the result of this measurement is generated as a random variable whose distribution corresponds to this feature's distribution for the true class of the observed object. An object's posterior probability of belonging to any class is then computed using Bayes' rule, where we have made the simplifying assumption that observations are independent from one time step to the next.

We assume a simple tracking model of a target moving with a constant velocity without a process noise. We model the probability that an interceptor successfully "kills" a target (*probability of kill* for short) as a function of a track's error covariance matrix. In particular, we used a sigmoidal function of the form

$$p_{kill}(x) = \frac{1 + \exp(-m/s)}{1 + \exp((x - m)/s)}$$

to determine the probability of kill, where x denotes the Euclidean norm of a track's position error, m defines the "midpoint," i.e., the point at which $p_{kill}(x) \approx 1/2$, and s represents the "spread" of the curve. Note that small values of s result in near step functions where p_{kill} is either close to one or zero. Such a function could easily be extended to incorporate additional factors beyond just the position error of track (e.g., velocity errors, classification information, etc.). In fact, Kalandros and Pao [7] give several examples of why more information may be necessary.

As described in Section 2, cost coefficients are needed to compute the risk of making a particular decision. We set the cost of incorrectly declaring a threat a nonthreat to 6, the cost of incorrectly declaring a nonthreat as a threat to 1, and

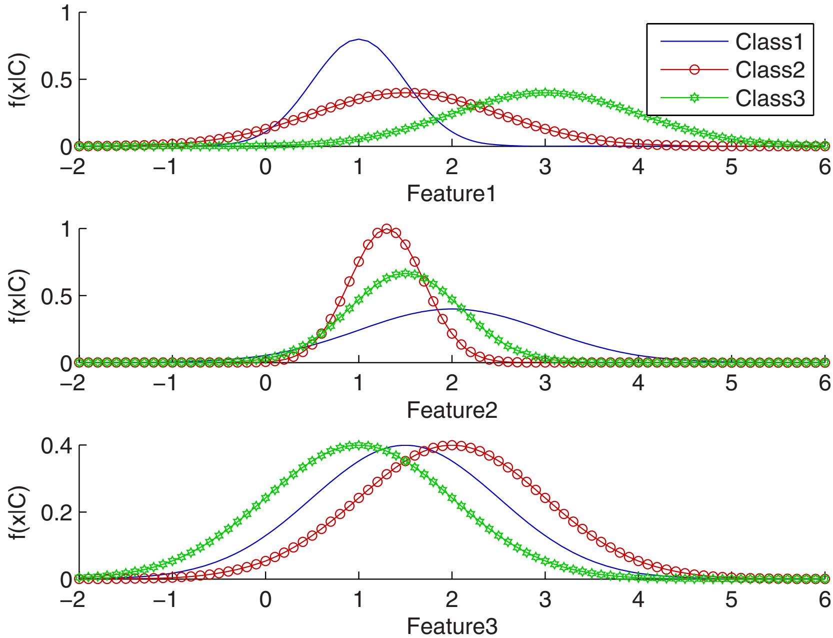

Fig. 1. Class-conditional PDFs used in simulations

the cost of incorrectly classifying a nonthreat as a different type of nonthreat to 0.2. We assume there are three objects, and, in truth, object i belongs to class $i,\ i = 1, 2, 3$. (Initial results with more than three objects demonstrated that our conclusions remain the same.)

For general class-conditional PDFs, the expected risk reduction from an additional measurement cannot be computed analytically. Thus, we turned to numerical integration techniques in our computations.

There are six different planners (or resource management algorithms) that we tested for comparison. At each planning interval (each time step), the planner assigns the sensor to perform a single action on a specific object during the subsequent time interval. The planners (and their symbols used in the figure legends) are:

1. Risk Reduction Planner (maxRRR): Enumerates all action-object pairs and determines which action-object pair will yield the largest expected reduction in risk.
2. Information Gain Planner (maxInfoG): Enumerates all action-object pairs and determines which action-object pair will yield the largest expected information gain.
3. Improved Information Gain Planner (maxIInfoG): Operates exactly like the Information Gain Planner except there is no information gain for performing

a track update on an object whose track error is below a pre-defined threshold.

4. Highest Probability of RV Planner (maxPRV): Identifies the object with the largest current probability of lethality (i.e., of being a re-entry vehicle) and randomly generates an action to be performed on this object. The purpose of this planner is to dispel the oft-held belief that spending the most resources on the most threatening object is an optimal use of resources.
5. Highest Risk Object Planner (maxRiskObject): Identifies the object with the largest current risk and randomly generates an action to be performed on this object.
6. Round Robin Planner (RoundRobin): First performs action 1 on all objects, then action 2, and so on.

The different resource management algorithms were compared with respect to five different metrics: (1) average loss; (2) average probability of correct classification of the lethal object; (3) average probability of correct classification of all nonlethal objects; (4) average track quality of the lethal object; and (5) average track quality of all nonlethal objects. In general, we found that all planners maintain a very high track quality on the object it believes to be lethal and a sufficient track quality on all remaining objects. Results with respect to the first three metrics are described below.

As one would expect, the risk reduction planner, which strives to reduce risk as quickly as possible, outperforms all other planners in the average loss category (see Figure 2). What is interesting is that the planner that attempts to maximize pure information gain (maxInfoG) over the course of the mission dedicates the majority of its resources to performing track maintenance actions. Under the assumptions of this simulation, this corroborates our initial statement that metrics based on pure information gain may not be well suited in the context of missile defense.

To give a more mathematical explanation as to why a purely information-based approach may yield inferior results, consider the following classification problem involving an object that can belong to one of n possible classes $C_1, \ldots,$ C_n, where C_1 is the class of lethal objects and all other classes represent various nonlethal objects. The object's class can be represented as a discrete random variable X, which must take on one of the values $x_1, \ldots, x_n$ with probabilities $p_1, \ldots, p_n$, respectively. It is well known that the (Shannon) entropy of the object's class, $H(X) = -\sum_{i=1}^{n} p_i \log_2 p_i$, is maximized when all of the p_i are equal because the object is equally likely to belong any of the n classes. In a similar way, suppose that an object has been perfectly classified as a nonlethal object, i.e., $p_1 = 0$, but that the exact type of nonlethal object is completely unknown, i.e., $p_2 = \ldots = p_n = 1/(n-1)$. Then, the entropy associated with this object is still relatively large. However, from the standpoint of risk, or expected loss, this object is of little concern. One could then argue that it would be an inappropriate use of scarce resources to determine precisely what class of nonlethal object it is, when its associated risk has already been determined to be zero.

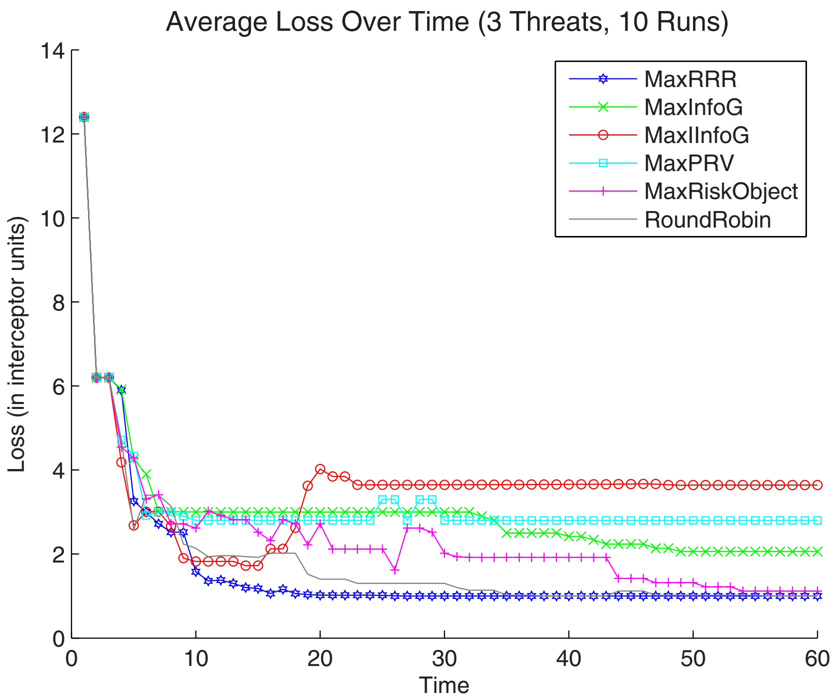

Fig. 2. Performance comparison of different resource managers

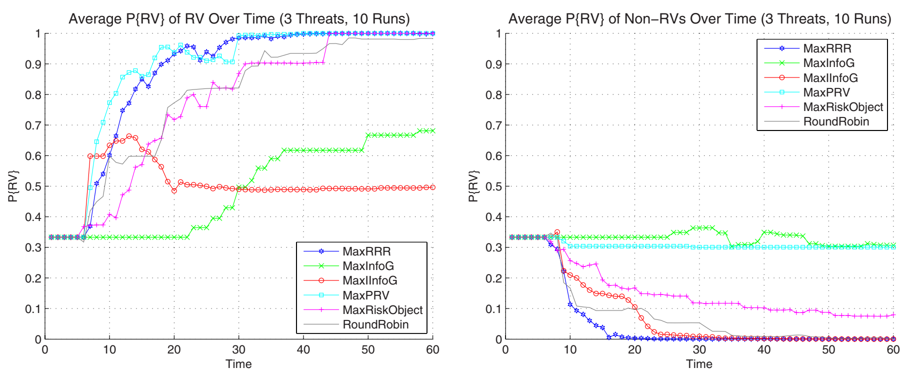

Fig. 3. Discrimination performance of different resource managers. P{RV} is the probability that an object is lethal.

Besides average loss, it is illustrative to compare the various planners based on two important questions related to classification: (1) How well were the targets classified? (2) How long did it take the sensor to classify the targets? Focusing solely on classification, one would expect an ideal sensor to quickly classify threatening objects as threatening and identify nonthreatening objects as nonthreatening. The sensor could then provide higher quality information in less time to an interceptor whose goal is to prosecute all threatening targets and reduce expected loss. It turns out that this goal is achieved as a byproduct of the risk reduction planner, and is illustrated in Figure 3. To understand this, recall that a nonzero probability of lethality directly contributes to the risk of a nonthreat decision. Consequently, if the cost of leakage is relatively high and several objects have a probability of lethality well above zero, then it is beneficial to perform additional classification measurements in order to reduce this probability. Reducing this probability is one way to possibly decrease total risk. Thus, a natural consequence of the risk reduction planner is to reduce the probability of lethality on all nonthreatening objects by performing additional discrimination actions.

6 Conclusions and Future Work

This chapter advocates the use of a risk-based objective function for sensor resource management in the context of missile defense. After presenting a formal description of the equations and update formulas needed to compute risk and risk reduction quantities, we outlined a risk-based approach to single-sensor resource management as well as a hierarchical approach to multisensor control. We performed a comparative analysis of various myopic approaches for tasking a sensor to track and discriminate targets (without the presence of deadlines) and found that maximizing the expected rate of risk reduction produced superior results.

Although not presented in this work, we have conducted an investigation of a modified rate of risk reduction method when presented with Battle Manager goals and deadlines [14]. Future research includes refinement of a mathematical solution methodology as needed to solve a finite-horizon dynamic program. Likewise, we are currently working to incorporate the risk due to possible misassociation of closely-spaced targets, which adds yet another layer of complexity into our formulation. It can be shown that our risk-based approach is also applicable in a situation when some contextual information is available for discrimination. Although the calculations become more involved, the results will be very similar. Finally, we are continuing to develop and test our proposed hierarchical resource management approach and the associated dynamic programming techniques involved.

Acknowledgements

The authors are grateful to Fred Daum, Herb Landau, and Manny Silvia for many valuable discussions, which stimulated a number of new ideas.

References

1. D.P. Bertsekas, *Dynamic Programming and Optimal Control*, Athena Scientific, Belmont, MA, 1995.
2. D.P. Bertsekas and J.N. Tsitsiklis, *Neuro-Dynamic Programming*, Athena Scientific, Belmont, MA, 1996.
3. F. Bolderheij and P. van Genderen, "Mission Driven Sensor Management," Proceedings of the 7th International Conference on Information Fusion, pp. 799-804, Stockholm, Sweden, 2004.
4. F. Bolderheij, F.G.J. Absil, and P. van Genderen, "A Risk-Based Object-Oriented Approach to Sensor Management," Proceedings of the 7th International Conference on Information Fusion, vol. 1, pp. 598-605, July 2005.
5. E.A. Feinberg, M.A. Bender, M.T. Curry, D. Huang, T. Koutsoudis, and J.L. Bernstein, "Sensor Resource Management for an Airborne Early Warning Radar," In O.E. Drummond, editor, Signal and Data Processing of Small Targets, Proceedings of SPIE vol. 4728, pp. 145-156, 2002.
6. K.J. Hintz, "A Measure of the Information Gain Attributable to Cueing," IEEE Transactions on Systems, Man, and Cybernetics, vol. 21, no. 2, pp. 237-244, 1991.
7. M. Kalandros and L.Y. Pao, "Covariance Control for Multisensor Systems," IEEE Transactions on Aerospace and Electronic Systems, vol. 38, No. 4, pp. 1138-1157, 2002.
8. K. Kastella, "Discrimination Gain to Optimize Detection and Classification," IEEE Transactions on Systems, Man, and Cybernetics, Part A: Systems and Humans, 27:1, pp. 112-116, 1997.
9. C. Kreucher, A.O. Hero, and K. Kastella, "A Comparison of Task Driven and Information Driven Sensor Management for Target Tracking," Proceedings of the 44th IEEE Conf. on Decision and Control, Seville, Spain, pp. 4004-4009, 2005.
10. C. Kreucher, "An Information-Based Approach to Sensor Resource Allocation," Ph.D. dissertation, The University of Michigan, 2005.
11. R. Mahler, "Global Optimal Sensor Allocation," in Proceedings of the Ninth National Symposium on Sensor Fusion, vol. I, pp. 167-172, Monterey, CA, 1996.
12. G.A. McIntyre and K.J. Hintz, "An Information-Theoretic Approach to Sensor Scheduling," Proceedings of SPIE, 2755, pp. 304-312, 1996.
13. M.L. Puterman, *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. John Wiley & Sons, Inc., 1994.
14. M. Raykin and D. Papageorgiou, "A Risk-Based Approach to Sensor Resource Management," Raytheon Technical Report HQ0006-06-D-0001, 2007.
15. M. Rudary, D. Khosla, J. Guillochon, P.A. Dow, and B. Blyth, "A Sparse Sampling Planner for Sensor Resource Management," In I. Kadar, editor, Proceedings of SPIE vol. 6235, Signal Processing, Sensor Fusion, and Target Recognition XV, pp. 62350A1-62350A9, 2006.
16. M. Ryan. Hierarchical Decision Making. In J. Si, A.G. Barto, W.B. Powell, and D. Wunsch, editors, "Handbook of Learning and Approximate Dynamic Programming," IEEE Press Series on Computational Intelligence, 2004.
17. W.W. Schmaedeke, "Information-based Sensor Management," Proceedings of SPIE vol. 1955, Signal Processing, Sensor Fusion, and Target Recognition II, pp. 156-164, April 1993.

18. W.W. Schmaedeke and K. Kastella, "Information-based Sensor Management and IMMKF," Proceedings of the SPIE Conference on Signal Processing of Small Targets, vol. 3373, Orlando, FL, 1998.
19. M.K. Schneider and C. Chong, "A Rollout Algorithm to Coordinate Multiple Sensor Resources to Track and Discriminate Targets," In I. Kadar, editor, Proceedings of SPIE vol. 6235, Signal Processing, Sensor Fusion, and Target Recognition XV, pp. 62350E1-62350E10, 2006.
20. R.S. Sutton and A.G. Barto, *Reinforcement Learning: An Introduction*. The MIT Press, Cambridge, MA, 1998.