The Value of Delayed Information in Tracking with Distributed Sensor Networks

A Thesis

Presented in Partial Fulfillment of the Requirements for the Degree Bachelor of Science with Honors Research Distinction in the College of Engineering at The Ohio State University

By

Christopher Dean

Undergraduate Program in Computer Science and Engineering

The Ohio State University

2012

Examination Committee:

Prof. Emre Ertin, Advisor

Prof. James W. Davis

Abstract

Tracking the movement of objects using multiple sensors (such as tracking vehicles or people using a distributed set of video cameras) is a topic of importance in security and automation applications. A prime consideration in these multi-sensor tracking networks is the timeliness of estimates, as tracking is a time-sensitive task. Current methods for handling delayed observations can be computationally intensive while minimally improving the estimate. This work aims to develop methods for estimating the value of delayed information. The value of observations from a variety of sensor locations at time points in the past are evaluated using the proposed strategy. We derive the relationships and trends that relate the 'age' of past sensor data to its marginal contribution toward improving the accuracy of an estimate of the current location of an object. Low-complexity methods for assessing the value of delayed sensor information can impact applications in defense and national security, personal electronics, and financial engineering where timely decisions are paramount.

Contents

Page

Abstract ii

List of Tables v

List of Figures vi

1. Introduction 1

2. Background 2
   - 2.1 Bayesian Filtering 4
   - 2.2 Sensor Selection and Utility 5

3. Approach 7
   - 3.1 Sensor Selection 7
   - 3.2 Expected Information Utility 9
   - 3.3 Conceptual Example 10

4. Application 16
   - 4.1 Problem Formulation 16
     - 4.1.1 Object Dynamics 16
     - 4.1.2 Sensor Model 17
   - 4.2 Filtering Techniques 20
   - 4.3 Sensor Selection 25
   - 4.4 Experimental Results 27

5. Discussion 32

5.1 Conditions for Integrating Delayed Information 32
5.2 The Value of Delayed Information 35

6. Conclusion 37

Bibliography 38

List of Tables

Table **Page**

4.1 Sum of squared errors performance indicators. 28

4.2 Number of tracks retained out of the total. 28

List of Figures

Figure **Page**

3.1 The state transition model for tracking a target down a tree. 11

3.2 Figure 3.2a shows the tracker attempting to integrate a current observation from the green node. Figure 3.2b shows the posterior state distribution after discovering that the (previously) green node did not detect the target. 12

3.3 Posterior state distribution at time step k (middle row). Each state at time k could transition to one of its four children at time $k+1$. These possible $k+1$ states are shown in the bottom row. 13

3.4 Suppose at time $k+1$ we select the green node shown at time $k+1$ (in the bottom row). 14

3.5 The expected utility of querying for a measurement at time t_k and t_{k+1}, and the change with respect to turning probability P. 15

4.1 Plots of detection probability vs. the probability of a false alarm for several distances. 21

4.2 Plot of detection probability vs. distance. (Here, $R = 50$ meters). . . 22

4.3 Target tracking with delayed observations permitted. Sensors arranged as a grid. 29

4.4 Target tracking with delayed observations prohibited. Sensors arranged as a grid. 29

4.5 Target tracking with standard EIF approach. Sensors arranged as a grid. 30

4.6 Target tracking considering missed detections and both old and new observations. 30

4.7 Target tracking considering missed detections and only new observations. 31

Chapter 1: Introduction

A common concern in tracking with sensor networks is the tradeoff between integrating informative observations and conserving power through reducing communications. This thesis considers selecting the observation we integrate using *a priori* knowledge to predict the most informative observation. Specifically, it aims to accelerate the selection process when delayed observations are among those considered for inclusion.

Chapter 2 examines the problem in more detail and summarizes existing related literature. Chapter 3 details the proposed selection strategy and provides a theoretical example an example. Chapter 4 applies this strategy to a realistic tracking scenario with multiple sensors and range-bearing observations. Chapter 5 evaluates the performance of the proposed strategy. Chapter 6 draws conclusions and proposes avenues of possible future research.

Chapter 2: Background

This chapter covers a wide breadth of background knowledge used in our approach to this problem.

Kalman filtering, a technique commonly used in tracking, iteratively refines a statistical model and constructs probabilistic estimates of the target state as sensor information becomes available. This statistical model consists of two parts: a dynamics model, which describes the evolution of the target state over time, and a linear observation model, which linearly relates the observations to the current state. Each model accounts for uncertainties in either the target dynamics or the observations using a Gaussian random variable. The amount of uncertainty is retained in an error covariance matrix for both the target dynamics and the observation model. Observations may be delayed in time before they are received; the delay may be caused, for example, by communication delays if the observation is taken by a distant sensor node and is transmitted to the filter. Bayesian theory provides a mechanism for incorporating delayed information, but this approach is computationally intensive and typically results in negligible improvement in the estimation accuracy.

Distributed sensor networks are commonly used for tracking-oriented applications. These networks are comprised of nodes that are each capable of sensing, processing, storing data, and communicating with other nodes. Power conservation is a ubiquitous concern in these systems, where the largest power draw is communication between nodes [13].

'Sleeping' sensor networks are a particular type of network that attempts to minimize the amount of communication between nodes. Consider a network where all nodes except one (called the leader node) are in a power-saving 'sleep' mode. The leader node is responsible for generating a local estimate of the target. In addition to the leader node's current observation, this estimate can also incorporate past or present observations from other nodes. At each time step, the leader node has to select the next leader node and select a subset of all available observations to incorporate into the local estimate. The subset of observations should be chosen in a principled manner, as the acquisition and processing of each additional observation is very costly in terms of both time and energy.

The value of information, defined by the utility function φ, is central to both of the decisions to be made. As utility can't be explicitly calculated without knowing the observation (which is not freely available to all nodes), we must use an estimate of the utility function to aid in these decisions. The proposed work aims to develop methods for estimating the value of delayed information. More specifically, we seek to develop low-complexity methods to indicate when assimilating a delayed measurement will improve the estimate by some non-negligible amount. These methods for assessing the value of delayed information could impact many fields where such evaluation would save time, power, money, or other valuable resources. For example, certain applications in national defense, security systems, and consumer electronics may benefit from of the low-complexity methods to be developed in this work.

2.1 Bayesian Filtering

We assume that a target is moving throughout some finite state space $\{\mathbf{x}_i\}_{i=1}^N$. At each time step t we denote the true target state by $\mathbf{x}^k$. Each of the M sensors makes a noisy observation of the target state $\mathbf{z}_j^k$; a measurement may or may not be added to a set of measurements maintained by the tracker, $\mathcal{Z}^k = \left\{\mathbf{z}^0, \mathbf{z}^1, \ldots, \mathbf{z}^k\right\}$.

The state transition model, generally, is given as

$$\mathbf{x}^{k+1} = F\left(\mathbf{x}^k, \mathbf{v}^k\right) \rightarrow f\left(\mathbf{x}^{k+1} \mid \mathbf{x}^k\right). \tag{2.1}$$

Similarly, the general form of the observation model is

$$\mathbf{z}_j^k = H_j\left(\mathbf{x}^k, \mathbf{w}^k\right) \rightarrow h_j\left(\mathbf{z}_j^k \mid \mathbf{x}^k\right). \tag{2.2}$$

In Bayesian filtering frameworks, we compute a predicted target state distribution for time step $k+1$:

$$p\left(\mathbf{x}^k \mid \mathcal{Z}^{k-1}\right) = \int_i f\left(\mathbf{x}^k \mid \mathbf{x}_i^{k-1}\right) p\left(\mathbf{x}_i^{k-1} \mid \mathcal{Z}^{k-1}\right)\, d\mathbf{x}^{k-1}. \tag{2.3}$$

This is followed by computing the posterior distribution, or the estimate of the target state at time $t+1$ based on the observation from time $t+1$:

$$p\left(\mathbf{x}^k \mid \mathcal{Z}^k\right) \propto h_j\left(\mathbf{z}^k \mid \mathbf{x}^k\right) p\left(\mathbf{x}^k \mid \mathcal{Z}^{k-1}\right).$$

The Bayesian criterion is based on minimizing the expected uncertainty at each time step.

This work investigates the inclusion of delayed or old measurements into this framework. The intuition here is that we have to ‘roll back’ the tracker to the time step of the new measurement, and then ‘roll forward’, propagating the new estimate and re-incorporating all of the previous data. The recursive nature of integrating observations of more than one time-step behind is clearly computationally intensive. This can be partially mitigated by allowing the sensor to have some memory stack in which it can store n prior state estimates for use in updating. The computational demands of this approach drive the search for a low-complexity metric for the usefulness of old measurements.

2.2 Sensor Selection and Utility

Tracking with sensor networks often requires the system to make decisions about which sensors to query at each time. As specified by the Bayesian tracking criterion, we select the sensor that maximally reduces our uncertainty. The value of information, defined by the utility function φ, is an expression of that reduction in uncertainty provided by querying a specific sensor.

Durrant-Whyte, in [9] and [3], advocates for the use of a distribution's log-likelihood as its utility. Then, its expected utility is its negative entropy or Shannon information. This entropy-based definition is both simple and intuitive, yet impractical as it requires the measurement itself to evaluate its information content.

In [16], Zhao *et al.* examine several utility measures within their Information-Driven Sensor Querying (IDSQ) approach. Instead, [16] proposes to select the next leader node based on the expected state estimate of the next time step. Ertin *et al.* extend this work in [4], demonstrating that selection based on the expected updated state estimate at the next time step is just an indirect way of selection based on the predicted state estimate at the next time step. Instead they propose selecting the sensor $\hat{j}$ such that

$$\hat{j} = \arg\max_{j \in V} E\left\{-H\left(p\left(\mathbf{x}^k \mid \mathcal{Z}^{k-1} \cup \mathbf{z}_j^k\right)\right) \mid \mathcal{Z}^{k-1}\right\}. \tag{2.4}$$

It is shown that it is more computationally tractable to select the next leader node based on the expected updated state covariance, which simplifies to selecting the sensor with maximum mutual information with the predicted state estimate.

Several other papers have improved performance by using these approaches over a range of time steps. In doing so, they try to maximize the long-term information gain of the tracker. Williams *et al.* build on this by using a dynamic programming algorithm approximation to simultaneously solve for the next leader node and the subset of other nodes over a rolling time horizon [13]. Similarly, Liu *et al.* apply the IDSQ approach to a series of values and obtain similar improvement over the single-step, 'greedy' alternative. None of these works provide explicit treatment of the delayed sensor measurements considered here.

Chapter 3: Approach

Our approach expands on the approaches discussed above in two ways. First, we apply principles of utility to sensor models where a sensor is not guaranteed to observe the target. Then we discuss the extension of these principles to delayed data. Finally, our approach is illustrated in an example.

In this work we are interested in evaluating the performance of a tracker that is able to use both current and past information. The problem described here is tailored to facilitate this comparison but the results should apply generally.

3.1 Sensor Selection

We consider a system that is tracking a single target as it progresses through some space. The true target state at time t_k is denoted x_k. A sensor network is deployed over this space in a distributed manner, that is, the state estimate is maintained by a 'leader node.' At each time step k, which is the discretized time t_k, each sensor node j makes a noisy measurement z_k^j of the target's location. Additionally, we assume that every sensor node has some finite memory stack and retains both its current observation z_k^j and its prior observation z_{k-1}^j.

At each time step k, the current leader node must choose one observation to incorporate into the estimate. The leader node, which at time k has the updated state estimate from the previous time step $k-1$, can choose any observation from the current time step k or from the previous time step $k-1$. In either case, the leader node uses some utility measure φ to select an observation to integrate. As in most real-world tracking scenarios, this utility measure must be able to estimate the information content of this observation without explicit knowledge of the observation. This constraint aims to reduce the amount of communication between nodes and assumes that the sensor modalities and locations are known by all nodes.

Regardless of whether a current or prior observation was selected, the leader node transmits the state estimate to the sensor node that made the observation. An indication of whether it has selected the current or prior observation for integration is also sent. The receiving node becomes the new leader node; if the leader node selects its own current observation it remains the leader node.

If a current observation is desired, the new lead node propagates the state estimate from time step $k-1$ to time step k and then updates the state estimate using the observation. Propagating the state estimate reduces its information content; updating the state estimate with the current observation then increases the state estimate's information content. If a prior observation is desired, the lead node updates the state estimate using the observation from time step $k-1$ and then propagates this updated state estimate forward to time step k. Updating the state estimate with the old observation increases the estimate's information content; some portion of this information gain is lost when the estimate is propagated forward to time step k. Once the state estimate has been propagated and updated (or updated and propagated) the system advances to time step $k+1$ and must select one observation to incorporate into the estimate.

3.2 Expected Information Utility

In many sensor networks there is no guarantee that every sensor can detect the target over the entire state space. This is caused by a variety of factors including occlusion and signal attenuation. Many standard tracking algorithms simply refrain from updating upon not receiving an observation. We, however, argue that there is actually information to be obtained in this case. Specifically, a non-detection tells us where the target *is not*.

We consider this information when deriving a utility measure. Assume the system is currently at time step k and has the most recent posterior distribution, $p(\mathbf{x}_{k-1} \mid \mathcal{Z}_{k-1})$. Then the *actual* information utility from integrating an observation $\mathbf{z}_{k-\gamma}^{j}$ from sensor node j and from time $k-\gamma$, where $\gamma \in \{0, 1\}$, is given by

$$\begin{aligned}\varphi\left(\mathbf{x}_k \mid \mathcal{Z}_{k-1} \cup \mathbf{z}_{k-\gamma}^{j}\right) &\triangleq \varphi\left(p\left(\mathbf{x}_k \mid \mathcal{Z}_{k-1} \cup \mathbf{z}_{k-\gamma}^{j}\right)\right) \\ &= \log_2 p\left(\mathbf{x}_k \mid \mathcal{Z}_{k-1} \cup \mathbf{z}_{k-\gamma}^{j}\right).\end{aligned}$$

Ideally we would like to find the sensor j^* for which

$$j^* = \arg\max_{j \in V} \varphi\left(\mathbf{x}_k \mid \mathcal{Z}_{k-1} \cup \mathbf{z}_{k-\gamma}^{j}\right),$$

but this is unreasonable to compute directly. Instead, we consider the *expected information utility* , given by

$$\begin{aligned}\Phi\left(\mathbf{x}_k \mid \mathcal{Z}_{k-1} \cup \mathbf{z}_{k-\gamma}^{j}\right) &\triangleq E\left\{\varphi\left(\mathbf{x}_k \mid \mathcal{Z}_{k-1} \cup \mathbf{z}_{k-\gamma}^{j}\right)\right\} \\ &= \sum_{\text{Detect?}} p\left(\mathbf{z}_{k-\gamma}^{j\,(?)}\right) \log p\left(\mathbf{x}_k \mid \mathcal{Z}_{k-1} \cup \mathbf{z}_{k-\gamma}^{j\,(?)}\right) \\ &= p\left(\mathbf{z}_{k-\gamma}^{j\,(+)}\right) \varphi\left(\mathbf{x}_k \mid \mathcal{Z}_{k-1} \cup \mathbf{z}_{k-\gamma}^{j\,(+)}\right) \\ &\quad + p\left(\mathbf{z}_{k-\gamma}^{j\,(-)}\right) \varphi\left(\mathbf{x}_k \mid \mathcal{Z}_{k-1} \cup \mathbf{z}_{k-\gamma}^{j\,(-)}\right) \qquad (3.1)\end{aligned}$$

where the superscript $j\,(+)$ means that sensor j detected the target and, likewise, $j\,(-)$ means that sensor j did not detect the target. For a shorthand notation we can write

$$\mathbf{\Phi}\left(\mathbf{x}_k \mid \mathcal{Z}_{k-1} \cup \mathbf{z}^j_{k-\gamma}\right) = P_D \varphi\left(\mathbf{x}_k \mid \mathcal{Z}_{k-1} \cup \mathbf{z}^{j\,(+)}_{k-\gamma}\right) + (1 - P_D)\varphi\left(\mathbf{x}_k \mid \mathcal{Z}_{k-1} \cup \mathbf{z}^{j\,(-)}_{k-\gamma}\right) \quad (3.2)$$

where $P_D = p\left(\mathbf{z}^{j\,(+)}_{k-\gamma}\right)$ is the probability that sensor j will detect the target and $(1 - P_D) = p\left(\mathbf{z}^{j\,(-)}_{k-\gamma}\right)$ is the probability that sensor j will not detect the target. Including both detections and non-detections in the expected utility estimate should facilitate more reasonable sensor selection and improved track retention.

It should be noted that the definition of expected utility for detections and non-detections are relatively convenient quantities for certain types of probability distributions. For discrete state spaces, the expected utility for the detection and non-detection posteriors is simply the posterior entropy. For multivariate Gaussian distributions the expected utility is, similarly, the determinant of the inverse covariance matrix.

3.3 Conceptual Example

To motivate our approach we first consider a simple example in a discrete state space. Suppose we have some target that is traversing a tree. Specifically, at each time step, it transitions from its current node to one of that node's children. Additionally, without loss of generality, assume that each node has four children and that the target transitions to its leftmost child with probability $1 - P$. The probability of transitioning to one of the three rightmost branches is equally split between them. This is the state dynamics model state dynamics model $f\left(\mathbf{x}^{(t+1)} \mid \mathbf{x}^{(t)}\right)$ illustrated in Figure 3.1. Assume for simplicity that $0 < P < 0.5$.

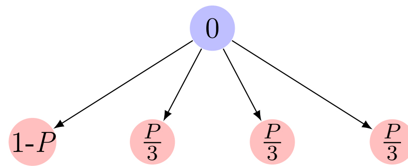


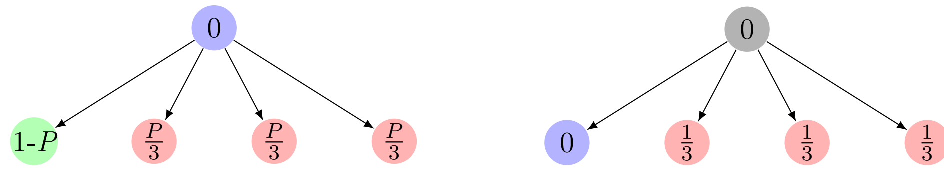
Figure 3.1: The state transition model for tracking a target down a tree.

Assume that we have one sensor in each of the discrete nodes. These sensors either detect the target if the target is in the same node as the sensor, or not if the target is in some other node. The observation model $h_j\left(\mathbf{z}_j^{(t)} \mid \mathbf{x}^{(t)}\right)$ returns a 1 if the target is in the same state as the sensor and a 0 otherwise. At each time t_k, we may query a single sensor to receive an observation from time t_k or t_{k-1}.

Assume that at time t_{k-1} we knew that the target was at the blue node in Figure 3.1. Now, at time t_k, we are trying to select a sensor to query. Since $0 < P < 0.5$, the sensor with the highest probability of detection is the green node in Figure 3.2a. If querying that sensor reveals that the target was not in that node then we can infer that it was in one of the red children at time t_k. The posterior distribution for the target state at time t_k is shown by nodes in the bottom row of the white nodes in Figure 3.2b.

The time advances to t_{k+1} after we query the sensor at t_k. We now need to select a sensor to query from either time t_k or t_{k+1}. The proposed expected utility metric is used to determine which sensor and time provides the most informative observation.

We will first consider the utility of old measurements from time t_k as shown in Figure 3.3. The green and red nodes in the middle row are all possible target locations at t_k. As they are equally likely and equally informative, we arbitrarily select the green node. The definition for expected utility from Equation 3.2, reproduced below, is used



(a) (b)

Figure 3.2: Figure 3.2a shows the tracker attempting to integrate a current observation from the green node. Figure 3.2b shows the posterior state distribution after discovering that the (previously) green node did not detect the target.

to compute the estimated utility of selecting this node.

$$\mathbf{\Phi}\left(\mathbf{x}_k \mid \mathcal{Z}_{k-1} \cup \mathbf{z}^j_{k-\gamma}\right) = P_D\varphi\left(\mathbf{x}_k \mid \mathcal{Z}_{k-1} \cup \mathbf{z}^{j\,(+)}_{k-\gamma}\right) + (1-P_D)\varphi\left(\mathbf{x}_k \mid \mathcal{Z}_{k-1} \cup \mathbf{z}^{j\,(-)}_{k-\gamma}\right)$$

The likelihood of the target being at the green node at time t_k is $\frac{1}{3}$, so $P_D = \frac{1}{3}$. If the target was at the green node at time t_k there are four possible nodes that it could have transitioned to by time step $k+1$: these are the four light-green nodes in the bottom row of Figure 3.3. This is the effect of propagation, or making a state prediction, that introduces additional uncertainty into the estimate. We quantify the potential information content $\varphi\left(\mathbf{x}_k \mid \mathcal{Z}_{k-1} \cup \mathbf{z}^{j\,(+)}_{k-\gamma}\right)$ of this distribution using the negative entropy

$$H(\mathcal{X}) = -\sum_i p(x_i)\log p(x_i),$$

where $\mathcal{X} = \{x_1, x_2, \ldots, x_n\}$ is the set of possible states and the probabilities $p(x_i)$ have been normalized so that $\sum_i p(x_i) = 1$.

The likelihood of the target not being at the green node at time step k is equivalent to the likelihood that it is at either of the two red nodes in Figure 3.3. This gives us $(1 - P_D) = \left(1 - \frac{1}{3}\right) = \frac{2}{3}$. In this case, by time t_{k+1}, the target may be in any of the eight light red nodes on the bottom row. This uncertainty is also quantified using

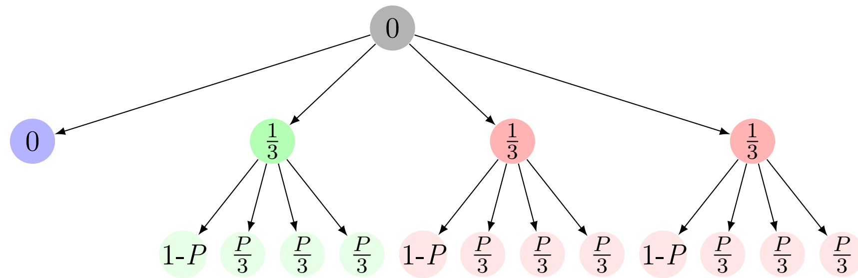


Figure 3.3: Posterior state distribution at time step k (middle row). Each state at time k could transition to one of its four children at time $k+1$. These possible $k+1$ states are shown in the bottom row.

negative entropy. The expected utility from an old measurement is, therefore,

$$\Phi^k = \frac{1}{3}H\left(1-P, \left\{\frac{P}{3}\right\}_3\right) + \frac{2}{3}H\left(\alpha\left\langle\{1-P\}_2, \left\{\frac{P}{3}\right\}_6\right\rangle\right), \tag{3.3}$$

where α is the normalizing factor required so that all probabilities inside the angle brackets sum to 1 and the subscripts outside the curly brackets indicate the number of possible states with that unnormalized probability.

We could alternatively select a current measurement from a sensor at time t_{k+1}. This is again computed using the expected utility defined in Equation 3.2. As shown in Figure 3.4, there are many more sensors to ask for a measurement at time $k+1$. Since $0 < P < 0.5$, the most likely nodes have probability of $\frac{1-P}{3}$. We arbitrarily select the green node in Figure 3.4; the probability of detection $P_D = \frac{1-P}{3}$. No propagation follows this selection because we are already at the current time (t_{k+1}).

We again use the negative entropy to quantify the information content of the distribution. Under the assumption that the target is at the green node at t_{k+1}, the green node is the only possible node in the posterior distribution. In this case its negative entropy is zero, indicating that there is no uncertainty in the distribution.

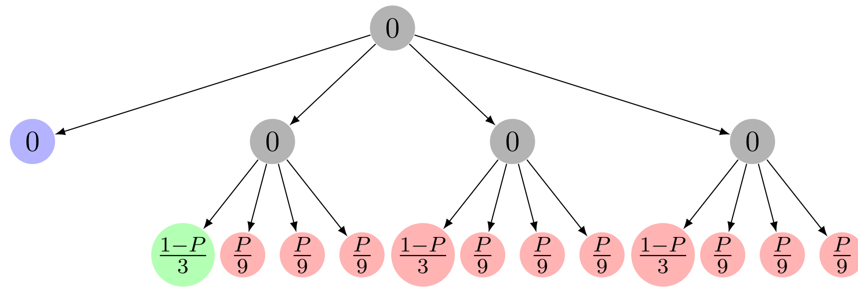


Figure 3.4: Suppose at time $k+1$ we select the green node shown at time $k+1$ (in the bottom row).

Therefore, the entire first term of the expected utility is zero. For the second term we assume that the target is not at the green node but one of the eleven red nodes. The likelihood of this occurring is $(1 - P_D) = \left(1 - \frac{1-P}{3}\right)$. The uncertainty is again quantified using negative entropy over the red nodes.

Accordingly, the expected utility of the most likely measurement at time t_{k+1} is:

$$\begin{aligned}\mathbf{\Phi}^{k+1} &= \frac{1-P}{3} H\left(\alpha\left\langle\frac{1-P}{3}\right\rangle\right) + \left[1 - \frac{1-P}{3}\right] H\left(\alpha\left\langle\left\{\frac{1-P}{3}\right\}_2, \left\{\frac{P}{9}\right\}_9\right\rangle\right) \\ &= \frac{2+P}{3} H\left(\alpha\left\langle\left\{\frac{1-P}{3}\right\}_2, \left\{\frac{P}{9}\right\}_9\right\rangle\right). \end{aligned} \tag{3.4}$$

The expected information of our posterior, described in Equations 3.3 and 3.4, is plotted in Figure 3.5. This figure shows the information gain, or the amount of information added to the estimate, assuming that we have no information to begin with. The y-axis quantities have little bearing on any physical interpretation since this specific problem is so abstracted.

On average, for all values of P in the allotted range, more information can be obtained by integrating an older measurement. It is important to note that these are expected values: it is certainly possible to have newer observations provide more

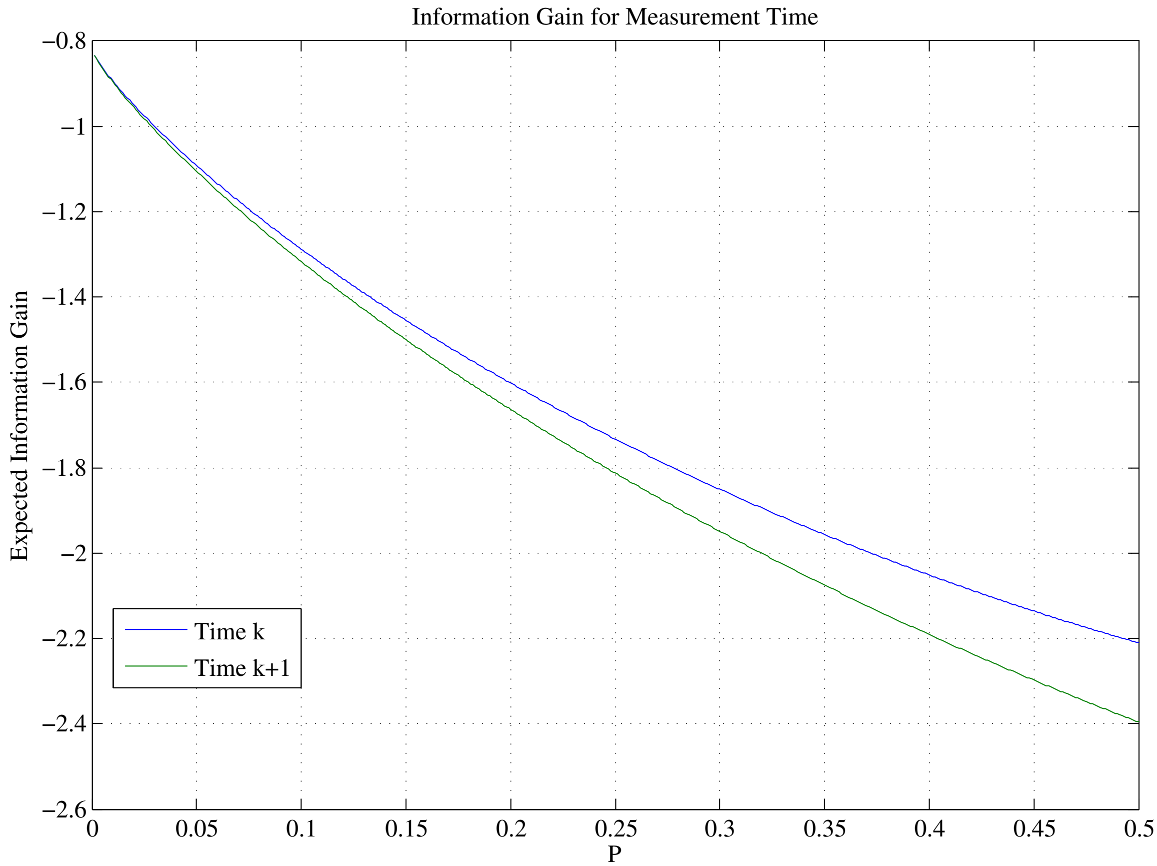


Figure 3.5: The expected utility of querying for a measurement at time t_k and t_{k+1}, and the change with respect to turning probability P.

information than older observations. Importantly, once we have lost the track of the target, it is better to query a sensor about an old observation regardless of the probability P. In this discrete state space, the degree of the benefit obtained is highly impacted by the number of possible states and the predictability of the target (here controlled by P). In the next chapter we consider the application of similar techniques to a more complex problem.

Chapter 4: Application

This chapter examines the application of the approach outlined in Chapter 3 to a realistic tracking scenario. The object dynamics and sensor models are identified, the tracking and sensor selection approach are described, and results are provided.

4.1 Problem Formulation

The system is characterized by a linear process model and a non-linear observation model. These are defined as follows.

4.1.1 Object Dynamics

The target being tracked is modeled in two dimensions with a position and a velocity in each dimension. Changes in velocity are provided by an additive velocity noise term in each dimension. The model also imposes discrete timing, where time step k corresponds to time t_k and $\Delta T = t_{k+1} - t_k$. The state of the system is given by

$$\mathbf{x}_k = \mathbf{F}_k \mathbf{x}_{k-1} + \mathbf{G}_k \mathbf{v}_k \tag{4.1}$$

where

- the $n \times 1$ state vector $\mathbf{x}_k = \begin{bmatrix} x_k & \dot{x}_k & y_k & \dot{y}_k \end{bmatrix}^\top$,

- the $n \times n$ state transition matrix $\mathbf{F}_k = \begin{bmatrix} 1 & \Delta T & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & \Delta T \\ 0 & 0 & 0 & 1 \end{bmatrix}$,

- the $n \times q$ process noise matrix $\mathbf{G}_k = \begin{bmatrix} \frac{(\Delta T)^2}{2} & 0 \\ \Delta T & 0 \\ 0 & \frac{(\Delta T)^2}{2} \\ 0 & \Delta T \end{bmatrix}$,

- the $q \times 1$ process noise vector $\mathbf{v}_k = \begin{bmatrix} v_x & v_y \end{bmatrix}^\top$, and

- the $q \times q$ process noise covariance matrix $\mathbf{Q}_k = \begin{bmatrix} \sigma_{vx}^2 & 0 \\ 0 & \sigma_{vy}^2 \end{bmatrix}$.

These equations comprise a matrix form of the standard kinematics equations of particle motion with an additional velocity noise term in both the $x-$ and $y-$ directions.

The process noise is modeled by a random process that is zero-mean, temporally uncorrelated (i.e. white), has known covariance, and is uncorrelated with the initial state. Formally, the following hold:

1. $\mathrm{E}\{\mathbf{v}_k\} = \bar{\mathbf{v}} = \mathbf{0}$

2. $\mathrm{E}\{\mathbf{v}_i \mathbf{v}_j^\top\} = 0$ for $i \neq j$ since the noise is temporally uncorrelated

3. $\mathrm{E}\{\mathbf{v}_k \mathbf{v}_k^\top\} = \mathbf{Q}_k$

4. $\mathrm{E}\{\mathbf{v}_k \mathbf{x}_0^\top\} = 0$ for all k.

4.1.2 Sensor Model

The sensor network uses homogeneous sensors that track the range and bearing of the target. Additionally, each sensor at each time step determines whether or not the target is detected or not. In many real sensor applications, a node may not receive an observation for a variety of reasons, including occlusions or communication delays.

This is modeled here by assuming that the target emits a beacon with some known amplitude and that this signal amplitude is attenuated with distance. Additionally, this signal amplitude is subject to additive noise. The received signal amplitude is compared to a threshold to determine whether or not the target was detected. If the sensor detects the target it produces an observation of range-bearing data; otherwise, no observation is produced. These two models are discussed here.

Observation Model

The sensors take noisy observations of the range and bearing from the sensor to the target. At time step k we model the observation of the target state $\mathbf{x}_k = \begin{bmatrix} x_k & \dot{x}_k & y_k & \dot{y}_k \end{bmatrix}^\top$ by some sensor i with location $\mathbf{s}_i = \begin{bmatrix} X_{i,k} & Y_{i,k} \end{bmatrix}^\top$ for $i \in \{1, \ldots, S\}$ with the following equation:

$$\mathbf{z}_k = \mathbf{h}(\mathbf{x}_k, \mathbf{w}_k), \tag{4.2}$$

where

- the $m \times 1$ observation vector $\mathbf{z}_k = \begin{bmatrix} z_r & z_\theta \end{bmatrix}^\top$,

- the $m \times 1$ non-linear observation model vector function

$$\mathbf{h}(\mathbf{x}_k, \mathbf{w}_k) = \begin{bmatrix} \sqrt{(x_k - X_{i,k})^2 + (y_k - Y_{i,k})^2} \\ \operatorname{atan2}\left(\dfrac{y_k - Y_{i,k}}{x_k - X_{i,k}}\right) \end{bmatrix} + \mathbf{w}_k,$$

- the $r \times 1$ observation noise vector $\mathbf{w}_k = \begin{bmatrix} w_r & w_\theta \end{bmatrix}^\top$, and

- the $r \times r$ observation noise covariance matrix $\mathbf{R}_k = \begin{bmatrix} \sigma_r^2 & 0 \\ 0 & \sigma_\theta^2 \end{bmatrix}$.

The observation noise is modeled by a random process that is zero-mean, temporally uncorrelated (i.e. white), has known covariance, and is uncorrelated with the initial state. Symbolically, the following hold.

1. $\mathrm{E}\{\mathbf{w}_k\} = \bar{\mathbf{w}} = \mathbf{0}$

2. $\mathrm{E}\{\mathbf{w}_i\mathbf{w}_j^\top\} = 0$ for $i \neq j$ since the noise is temporally uncorrelated

3. $\mathrm{E}\{\mathbf{w}_k\mathbf{w}_k^\top\} = \mathbf{R}_k$

4. $\mathrm{E}\{\mathbf{w}_k\mathbf{x}_0^\top\} = 0$ for all k

Detection Model

The sensors collect radius and bearing measurements $\mathbf{z}_k$, specified as above. The sensors also collect an amplitude z_k that is used to determine whether or not the target was actually detected. This amplitude is attenuated with distance and subject to additive noise $n \sim \mathcal{N}(0, \sigma_n)$.

Sensor i receives the amplitude $z_i = s(r_i) + n$, where the k subscript has been dropped for convenience of notation. The received signal amplitude is given by

$$s(r_i) = \frac{a}{\|\mathbf{x} - \mathbf{s}_i\|_2^2 + b} = \frac{a}{(x_k - X_{i,k})^2 + (y_k - Y_{i,k})^2 + b}.$$

Without loss of generality, we assume that $a = 1$ and $b = 1$ for simplicity.

We define a cutoff τ that allows us to differentiate between two hypotheses: H_0, in which we cannot detect the target and believe the signal is dominated by noise, and H_1, in which the target is visible over the noise. We write this as

$$z_i \underset{H_0}{\overset{H_1}{\gtrless}} \tau,$$

where for H_0 we believe that $z_i = n$ and for H_1 we believe $z_i = s(r_i) + n$. The end product of the detection model is a binary guess as to whether the target was or was not detected.

We can set this threshold τ to tune performance as specified by

$$
\begin{aligned}
P_{FA}(\tau) &= \Pr\{\text{declare } H_1 \mid H_0 \text{ is true}\} \\
&= \Pr\{z_i > \tau \mid H_0 \text{ is true}\} \\
&= \Pr\{n > \tau\} \\
&= \Phi\left(\frac{\tau}{\sigma_n}\right)
\end{aligned} \tag{4.3}
$$

and

$$
\begin{aligned}
P_D(\tau) &= \Pr\{\text{declare } H_1 \mid H_1 \text{ is true}\} \\
&= \Pr\{z_i > \tau \mid H_1 \text{ is true}\} \\
&= \Pr\{n > \tau - s(r_i)\} \\
&= \Phi\left(\frac{\tau - s(r_i)}{\sigma_n}\right)
\end{aligned} \tag{4.4}
$$

where the function $\Phi(x)$ is the standard normal CDF. The tradeoff between the false alarm probability P_{FA} and the detection probability P_D is shown in Figure 4.1, the sensor's receiver operating characteristics (ROC) curve. We set the noise level and threshold to achieve a sensor detection radius of $R = 50$ meters and a false alarm probability of $P_{FA} = 0.001$. With a sensor detection radius R and zero noise contribution assume, when the target is a distance R away from the sensor its likelihood of detection is 0.5. Targets closer than R to the sensor will be declared H_1 and targets further from the sensor will be declared H_0, assuming zero noise contribution. The threshold and noise level are then set as

$$
\tau = \frac{1}{R^2 + 1} \qquad \text{and} \qquad \sigma_n = \frac{\tau}{\Phi^{-1}(P_{FA})},
$$

respectively. This provides the sensor response curve shown in Figure 4.2.

4.2 Filtering Techniques

The target is tracked as it moves through a field of sensors as described above. The Extended Information filter (EIF), a variant of the Kalman filter, is used. The

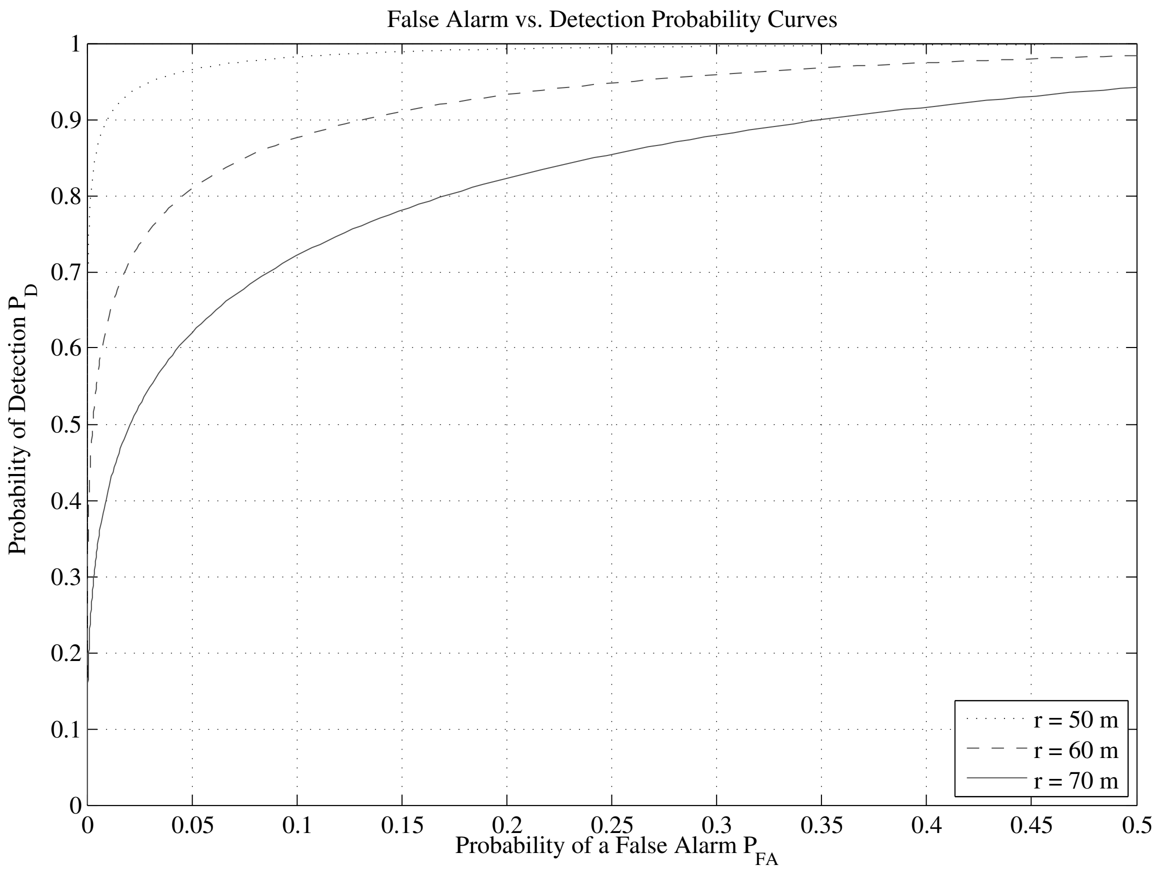


Figure 4.1: Plots of detection probability vs. the probability of a false alarm for several distances.

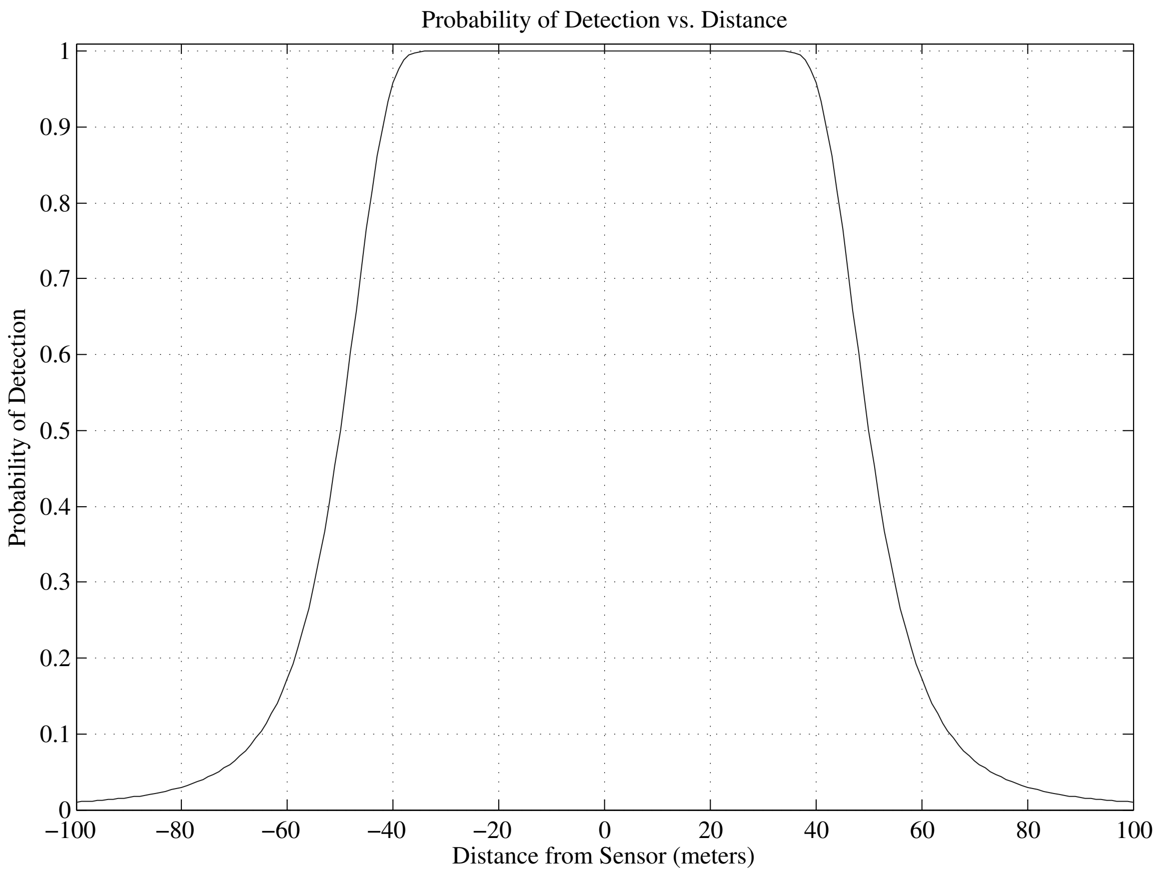


Figure 4.2: Plot of detection probability vs. distance. (Here, $R = 50$ meters).

Kalman filter tracks the state by iteratively refining a state estimate $\hat{\mathbf{x}}$ and a state estimate covariance matrix $\mathbf{P}$. The covariance matrix $\mathbf{P}$ retains information about the uncertainty of the estimate. The track is then given as a series of multivariate Gaussians distributions $\mathcal{N}(\hat{\mathbf{x}}_{k|k}, \mathbf{P}_{k|k})$. The information filter instead uses the inverse of the covariance matrix:

$$\mathbf{Y}_{i|j} = \mathbf{P}_{i|j}^{-1}.$$

We also define the information state vector as an analog to the state estimate:

$$\hat{\mathbf{y}}_{i|j} = \mathbf{Y}_{i|j}\hat{\mathbf{x}}_{i|j} = \mathbf{P}_{i|j}^{-1}\hat{\mathbf{x}}_{i|j}.$$

It is stressed that the Kalman filter and the information filter are mathematically equivalent.

The Information filter described has a linear process model and a nonlinear observation model. The filter proceeds by repeating a sequence of two steps: a propagation step and an update step.

- **Propagate:**

$$\hat{\mathbf{y}}_{k|k-1} = [\mathbf{1} - \mathbf{\Omega}_k \mathbf{G}_k^\top]\mathbf{F}_k^{-\top}\hat{\mathbf{y}}_{k-1|k-1} + \mathbf{Y}_{k|k-1}\mathbf{B}_k\mathbf{u}_k \tag{4.5}$$

$$\mathbf{Y}_{k|k-1} = \mathbf{M}_k - \mathbf{\Omega}_k \mathbf{\Sigma}_k \mathbf{\Omega}_k^\top \tag{4.6}$$

  where

  - the propagated information

$$\mathbf{M}_k = \mathbf{F}_k^{-\top}\mathbf{Y}_{k-1|k-1}\mathbf{F}_k^{-1}$$

  - the 'information propagation gain matrix'

$$\mathbf{\Omega}_k = \mathbf{M}_k \mathbf{G}_k \mathbf{\Sigma}_k^{-1}$$

– the 'information innovation covariance matrix'

$$\mathbf{\Sigma}_k = \mathbf{G}_k^\top \mathbf{M}_k \mathbf{G}_k + \mathbf{Q}_k^{-1}$$

- **Update:**

$$\hat{\mathbf{y}}_{k|k} = \hat{\mathbf{y}}_{k|k-1} + \mathbf{i}_k \tag{4.7}$$

$$\mathbf{Y}_{k|k} = \mathbf{Y}_{k|k-1} + \mathbf{I}_k \tag{4.8}$$

where

– the information state contribution

$$\mathbf{i}_k = \nabla \mathbf{h}_{\mathbf{x},k}^\top \mathbf{R}_k^{-1} \left[\nu_k + \nabla \mathbf{h}_{\mathbf{x},k} \hat{\mathbf{x}}_{k|k-1} \right]$$

– the information matrix contribution

$$\mathbf{I}_k = \nabla \mathbf{h}_{\mathbf{x},k}^\top \mathbf{R}_k^{-1} \nabla \mathbf{h}_{\mathbf{x},k}$$

– the innovation

$$\nu_k = \mathbf{z}_k - \mathbf{h}\left(\hat{\mathbf{x}}_{k|k-1}\right)$$

– the predicted observation

$$\mathbf{h}(\hat{\mathbf{x}}_{k|k-1}) = \begin{bmatrix} \sqrt{(\hat{x}_{k|k-1} - X_{i,k})^2 + (\hat{y}_{k|k-1} - Y_{i,k})^2} \\ \text{atan2}\left(\dfrac{\hat{y}_{k|k-1} - Y_{i,k}}{\hat{x}_{k|k-1} - X_{i,k}} \right) \end{bmatrix}$$

– the Jacobian of the observation model vector function with respect to the predicted state $\hat{\mathbf{x}}_{k|k-1}$

$$\nabla \mathbf{h}_{\mathbf{x},k} = \begin{bmatrix} \dfrac{\hat{x}_{k|k-1} - X_{i,k}}{\hat{r}} & 0 & \dfrac{\hat{y}_{k|k-1} - Y_{i,k}}{\hat{r}} & 0 \\ -\dfrac{\hat{y}_{k|k-1} - Y_{i,k}}{\hat{r}^2} & 0 & \dfrac{\hat{x}_{k|k-1} - X_{i,k}}{\hat{r}^2} & 0 \end{bmatrix}$$

where $\hat{r} = \sqrt{(\hat{x}_{k|k-1} - X_{i,k})^2 + (\hat{y}_{k|k-1} - Y_{i,k})^2}$ is the distance from the i-th sensor to the target. For delayed observations measured at time $k-d$, we linearize around $\hat{\mathbf{x}}_{k-d|k-d}$.

Using the inverse covariance provides several benefits. First, it allows the filter to be initialized with no information. The equivalent covariance matrix, by comparison, would have to be initialized with infinite uncertainty. Many sensor networks used in tracking also use the information matrix because it significantly simplifies the state update procedure. Granted, this comes at the cost of increased complexity in the state prediction, but it is much more common to update the filter with multiple observations and only update once.

Conceptualizing the estimate uncertainty with the information matrix also make information-based metrics more intuitive. We can directly state that the maximization of mutual information between sensor output and the target state can be achieved by maximizing the determinant of the information matrix for a Gaussian posterior.

4.3 Sensor Selection

Sensor selection is determined by the expected utility condition outlined in Chapter 3. At each time step the tracker can integrate only one observation. This observation can be from time k or time $k-1$. Recall that the lead node maintains the recent estimate $\mathbf{y}_{k-1|k-1}$ and $\mathbf{Y}_{k-1|k-1}$. Further, each sensor knows all of the sensor locations and retains its own two most recent observations: one from time step k and one from time step $k-1$.

At each time step the lead node computes either $2M$ or $2M-1$ expected utilities are computed, assuming M possible sensors with two cached observations at each. $2M$ are computed if at time step $k-1$ we selected a prior observation; $2M-1$ are computed if at time step k we selected a current observation. The sensor with the maximal value becomes the next lead node, regardless of whether the highest expected utility was from time step k or time step $k-1$.

The expected utility calculation is performed according to Equation 3.2, reproduced here:

$$\mathbf{\Phi}\left(\mathbf{x}_k \mid \mathcal{Z}_{k-1} \cup \mathbf{z}^j_{k-\gamma}\right) = P_D \varphi\left(\mathbf{x}_k \mid \mathcal{Z}_{k-1} \cup \mathbf{z}^{j\,(+)}_{k-\gamma}\right) + (1 - P_D)\varphi\left(\mathbf{x}_k \mid \mathcal{Z}_{k-1} \cup \mathbf{z}^{j\,(-)}_{k-\gamma}\right).$$

In this context, P_D and $(1 - P_D)$ are the probability of detection, $\Pr\{z_i > \tau\}$, and the probability of a non-detection, $\Pr\{z_i > \tau\}$, for some sensor i. These probabilities are computed using a Monte Carlo simulation.

The information content of a multivariate Gaussian distribution is given by the determinant of the inverse covariance matrix, or equivalently, the determinant of the information matrix $\mathbf{Y}$. The information matrix of a detection can be directly computed from sensor locations and target estimate. As such, the information matrix of the posterior after integrating a measurement from sensor i is $\varphi\left(\mathbf{x}_k \mid \mathcal{Z}_{k-1} \cup \mathbf{z}^{j\,(+)}_{k-\gamma}\right) = \det\left(\mathbf{Y}_{k|k-1} + \mathbf{I}_{k-d}\right)$. For a non-detection, we estimate the posterior covariance matrix by Monte Carlo sampling and computing the covariance $\mathbf{\Sigma}$ of the sample population. Inverting this provides the approximate posterior information matrix $\mathbf{Y}^{(-)}_k$ that is used to compute $\varphi\left(\mathbf{x}_k \mid \mathcal{Z}_{k-1} \cup \mathbf{z}^{j\,(-)}_{k-\gamma}\right) = \det\left(\mathbf{Y}^{(-)}_k\right)$. Similar techniques are used to actually incorporate a non-detection event into the current estimate.

To estimate the utility of a current observation the information contribution $\mathbf{I}$ of the sensor is added to the predicted information matrix $\mathbf{Y}_{k|k-1}$ (for a detection event) or by computing the non-detection posterior (for a non-detection event). For an old measurement, the observation's information contribution is added to the previous updated information matrix to producing an enhanced $\mathbf{Y}_{k-1|k-1}$ (for a detection event) or by computing the non-detection posterior. This is propagated forward to the current time $k|k-1$ and designated the estimate at time k without incorporating a current observation.

4.4 Experimental Results

The implemented algorithm was used to track moving targets over many different sensor distributions. The target trajectories are driven by Brownian velocity and heading, which are used to compute the trajectory in x and y with velocities $\dot{x}$ and $\dot{y}$.

Two types of sensor distributions were used. The first type of distribution places sensors in a uniform grid, separated by the sensor detection radius. The second type of distribution is a random field, with positions generated by a uniform random sensor distribution.

The performance of three different types of trackers are compared. The first type follows the approach discussed above and incorporates delayed measurements and non-detection events. The second type of tracker works similarly but excludes delayed measurements from consideration. The third type of tracker, used as a baseline, selects sensors using expected utility but does not incorporate delayed measurements or non-detection events.

Table 4.1 provides the sum-of-squares errors for each of the different sensor distribution and tracker configurations. The columns each correspond to one of the three tracker types. The data is provided for both types of sensor distributions. The mean sum-of-squared error of each configuration is computed from all of the trials in which the track was never lost.

Table 4.2 shows the track retention ability of each of the different tracking algorithms. We consider a track to have been retained if it is never lost, or if the information content of the estimate never below a certain value η:

$$\det(\mathbf{Y}_k) \leq \eta.$$

Table 4.1: Sum of squared errors performance indicators.

Sensor Distr.	Current & Old	Current Only	Standard
Grid	2,400.7	2,439.65	4,632.3
Random	2,442.3	3,029.5	3,713.6

Table 4.2: Number of tracks retained out of the total.

Sensor Distr.	Current & Old	Current Only	Standard
Grid	19/50	18/50	15/50
Random	17/50	19/50	16/50

Figures 4.3-4.5 show an example of tracker performances over one specific target trajectory and a grid-like sensor distribution. Figures 4.6 and 4.7 show an example of the inclusion of old observation leading to the loss of a target that would have otherwise been tracked successfully.

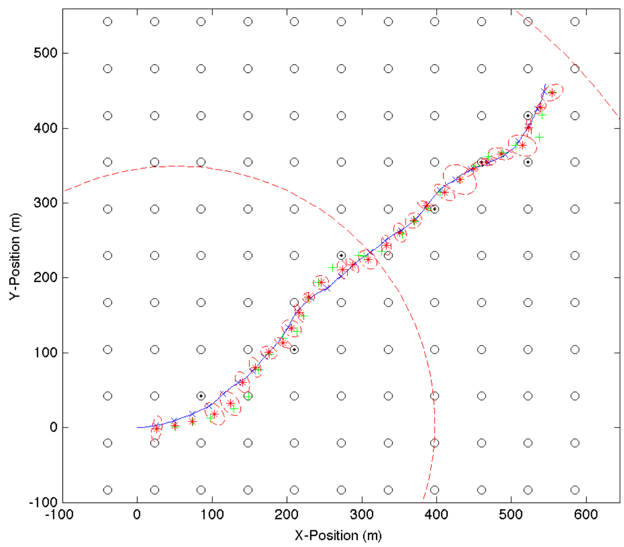


Figure 4.3: Target tracking with delayed observations permitted. Sensors arranged as a grid.

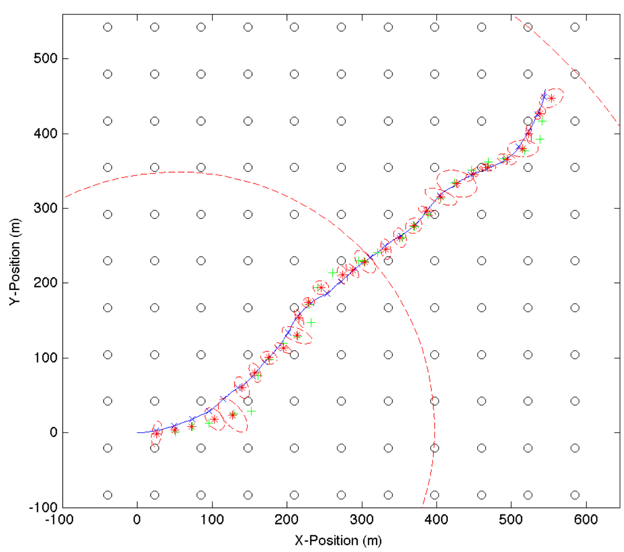


Figure 4.4: Target tracking with delayed observations prohibited. Sensors arranged as a grid.

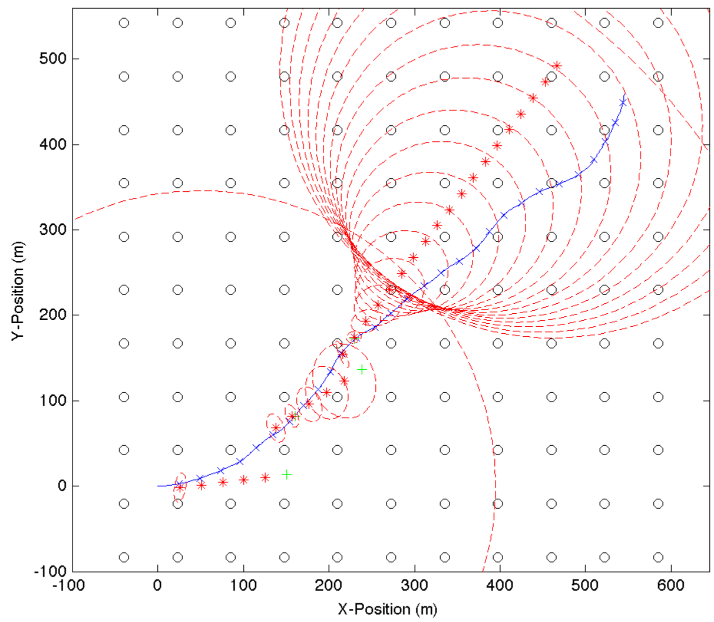


Figure 4.5: Target tracking with standard EIF approach. Sensors arranged as a grid.

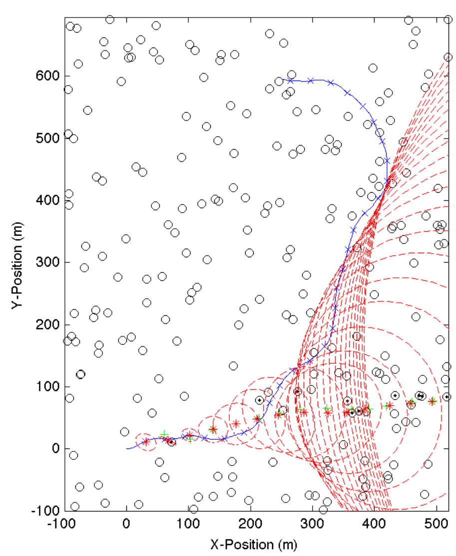


Figure 4.6: Target tracking considering missed detections and both old and new observations.

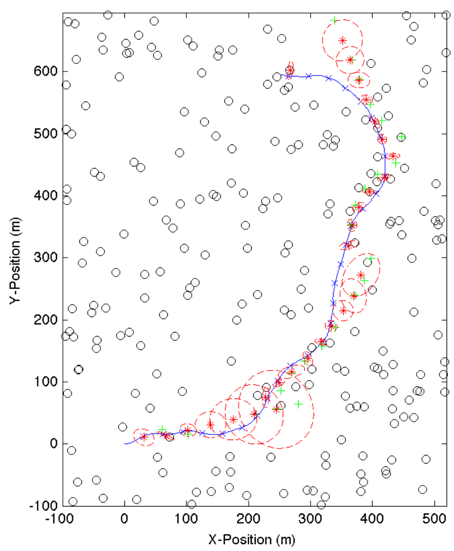


Figure 4.7: Target tracking considering missed detections and only new observations.

Chapter 5: Discussion

The expected utility metric, Φ was applied to problems in both discrete and real state spaces. We now establish the conditions leading to finding value in delayed measurements and consider how useful delayed information actually is.

5.1 Conditions for Integrating Delayed Information

According to our sensor selection criteria a delayed observation will be selected when

$$\Phi^{k-1} > \Phi^{k}.$$

Intuitively, the information value of detecting the target is much greater than the value of not detecting it. A detection contributes some amount of information that is inversely related to the observation noise covariance matrix. Lower observation noise covariances provide greater information contributions or, equivalently, provide a more drastic reduction in the estimate's uncertainty. A non-detection event generally contributes less information because it rarely provides a drastic reduction in the estimate's uncertainty.

If the target is out of the sensor's range or it is within range with a sufficiently large covariance, then the posterior is very close to that of the prior distribution. Instead, if the target covariance is small and the target is within range of the sensor, then the posterior reflects the masking. It will usually have a mean outside of the main sensor response area and a slightly smaller covariance ellipse. In these cases, where there is an increase in information, the weighting factor $1 - P_D \approx 0$ because the target is within close range of the sensor.

By this intuition, the above criteria will produce similar selections to

$$P_D^{k-1} \varphi \left(\mathbf{x}_k \mid \mathcal{Z}_{k-1} \cup \mathbf{z}_{k-1}^{(+)} \right) > P_D^k \varphi \left(\mathbf{x}_k \mid \mathcal{Z}_{k-1} \cup \mathbf{z}_k^{(+)} \right).$$

Using a simplified notation, the above is rewritten as

$$P_D^{k-1} \varphi^{k-1} > P_D^k \varphi^k. \tag{5.1}$$

This highlights the two most important factors governing the selection of an old observation: the loss of information during the propagation step and the sensor distribution.

Propagation generally decreases our information content. As time passes, an old observation has a decreasing relevance to the current state. This is evident from the Kalman filter propagation equation:

$$\mathbf{P}_{k|k-1} = \mathbf{F}_k \mathbf{P}_{k-1|k-1} \mathbf{F}_k^\top + \mathbf{G}_k \mathbf{Q}_k \mathbf{G}_k^\top \tag{5.2}$$

where $\mathbf{P}$ is the state estimate covariance, $\mathbf{F}_k$ is the state transition matrix, $\mathbf{G}_k$ is the process noise propagation matrix, and $\mathbf{Q}_k$ is the process noise covariance matrix. For the values used in Chapter 4, this grows quickly: $\mathbf{F}_k$ grows on the order of ΔT, so $\mathbf{F}_k \mathbf{P}_{k-1|k-1} \mathbf{F}_k^\top$ grows as ΔT^2; $\mathbf{G}_k$ grows on the order of ΔT^2, so $\mathbf{G}_k \mathbf{Q}_k \mathbf{G}_k^\top$ grows on the order of ΔT^4! The uncertainty of an estimate grows quickly as it ages, although the order of growth is governed by the relative magnitudes of $\mathbf{P}_{k-1|k-1}$ and $\mathbf{Q}_k$. Conversely, the information utility of an estimate will be reduced quickly as it ages. Therefore, the following relationship generally holds:

$$\varphi^{k-1} < \varphi^{k}. \tag{5.3}$$

We will only select an old observation when $P_D^{k-1}\varphi^{k-1} > P_D^k\varphi^k$. Given that Equation 5.3 holds, for an old observation to be selected,

$$P_D^{k-1} \gg P_D^k. \tag{5.4}$$

More accurately, since we select the sensor with the maximum expected utility, we rewrite this condition for some sensor j as

$$P_{D,j}^{k-1} \gg P_{D,i}^k \qquad \forall i \in \{1, \ldots, S\}. \tag{5.5}$$

This specifically happens when the target transitions from an area of sensor coverage into an area with no sensor coverage. At the same time, sufficient coverage in the preceding area is needed so that there exists another sensor whose observation we want to integrate.

Although it is generally true that $\varphi^{k-1} < \varphi^k$, delayed information is more likely to be selected when φ^{k-1} is very close to φ^k as possible. This happens when an old measurement exists with a high information content, generally because the target was observed by multiple close sensors at the previous time. We also observe φ^{k-1} being very close to φ^k when the propagation penalty is low. This penalty, obtained by rearranging and inverting the factors of Equation 5.2, has both a multiplicative component and an additive component. By definition, the magnitudes of $\mathbf{F}_k$ and $\mathbf{G}_k$ increase with the sample period ΔT. Thus, having a higher sampling rate can reduce the propagation penalty. Decreasing the magnitude of the process noise covariance matrix can also lower the propagation penalty.

Satisfying these conditions increases the likelihood that a delayed observation will have an information utility comparable to that of many current observations.

5.2 The Value of Delayed Information

Above we identified several conditions that can result in the tracker incorporating delayed measurements into the state estimate. Empirical evidence for the benefits of using delayed measurements in general tracking algorithms have also been acquired.

Table 4.1 shows that the tracker including both current and old observations has, on average, lower error rates than the others. The tracker that does not include delayed information but still estimates the posteriors from non-detection events also does better, on average, compared to the standard method.

Table 4.2 shows the track retention capabilities of each sensor. For the sensors arranged on a uniform grid, the tracker with delayed information achieves the highest overall retention rate. It is expected that with more trials, a similar pattern would occur for the random sensor distribution. In general, it was observed that sparser sensor distributions resulted in higher values of delayed information but that valuable delayed information was also harder to obtain.

It is important to note that the described approach is a 'greedy' algorithm. Specifically, this sensor selection approach guaranteed to always select the observation that maximizes the short-term information gain by maximizing estimated by the expected utility function. Repeated application of this sensor selection criteria is not, however, guaranteed to maximize long-term information content.

Under all three approaches considered here, incorporating an observation that maximizes expected utility is maximizing the immediate information content of the posterior distribution. As a side effect, the observation may introduce some bias to the estimate that, over time, results in degraded tracking performance. This is especially the case for targets with relatively unpredictable trajectories: the more dynamic a trajectory, the greater the risk that this bias will result in the loss of some useful observation. Occasionally this will cause the tracker to lose the target completely, as shown in Figure 4.6 and 4.7.

Although it may not be suitable for all situations, delayed information appears particularly well-suited for scenarios where targets are more predictable or have a restricted state space, like tracking vehicles on a road.

Chapter 6: Conclusion

This thesis set out to find low-complexity methods for estimating the value of delayed information. Through the development of the metric of Expected Information Utility, delayed information was shown to be valuable in several circumstances. These circumstances are often identifiable through certain attributes of the sensor distribution or the target dynamics. Delayed information has high value in situations where the sensor distribution is sparse or the target motion is relatively predictable. However, it is also noted that the expected utility metric established here is a 'greedy' constraint and is only guaranteed to maximize short-term information gain. In certain situations, incorporating a measurement that is optimal according to the constraint might result in worsening tracker performance, up to and including track loss.

Future work might seek to quantify the intuitions developed in this thesis. The expected utility metric could be extended to differently-configured systems with more explicit communication costs or asynchronous observations. Additionally, dynamic programming could be applied to differentiate between cases where delayed information will improve versus degrading track quality over a long-term horizon.

Bibliography

[1] Y. Bar-Shalom. Update with out-of-sequence measurements in tracking: exact solution. *Aerospace and Electronic Systems, IEEE Transactions on*, 38(3):769 – 777, jul 2002.

[2] Arnaud Doucet and Adam M. Johansen. A tutorial on particle filtering and smoothing: fifteen years later, 2011.

[3] Hugh Durrant-Whyte. *Introduction to Decentralized Data Fusion*. Australian Centre for Field Robotics, 2002.

[4] Emre Ertin, John W. Fisher, and Lee C. Potter. Maximum mutual information principle for dynamic sensor query problems. In *Proc. IPSN 2003*, pages 405–416. Springer-Verlag, April 2003.

[5] A.O. Hero and D. Cochran. Sensor management: Past, present, and future. *Sensors Journal, IEEE*, 11(12):3064 –3075, dec. 2011.

[6] M.P. Kolba and L.M. Collins. Information-based sensor management in the presence of uncertainty. *Signal Processing, IEEE Transactions on*, 55(6):2731 –2735, june 2007.

[7] Chris M. Kreucher, Keith D. Kastella, and Alfred O. Hero III. Information-based sensor management for multitarget tracking. pages 480–489, 2003.

[8] Juan Liu, D. Petrovic, and Feng Zhao. Multi-step information-directed sensor querying in distributed sensor networks. In *Acoustics, Speech, and Signal Processing, 2003. Proceedings. (ICASSP '03). 2003 IEEE International Conference on*, volume 5, pages V – 145–8 vol.5, april 2003.

[9] James M. Manyika and Hugh F. Durrant-Whyte. Information-theoretic approach to management in decentralized data fusion. *SPIE*, 1828 Sensor Fusion V:202–213, 1992.

[10] Wayne W. Schmaedeke. Information-based sensor management. pages 156–164, 1993.

[11] Hanbiao Wang, Kung Yao, and Deborah Estrin. Information-theoretic approaches for sensor selection and placement for target localization and tracking in sensor networks. *journal of Communications and Networks*, pages 438–229, 2005.

[12] Hanbiao Wang, Kung Yao, Greg Pottie, and Deborah Estrin. Entropy-based sensor selection heuristic for target localization. In *Proceedings of the 3rd international symposium on Information processing in sensor networks*, IPSN '04, pages 36–45, New York, NY, USA, 2004. ACM.

[13] J.L. Williams, III Fisher, J.W., and A.S. Willsky. An approximate dynamic programming approach to a communication constrained sensor management problem. In *Information Fusion, 2005 8th International Conference on*, volume 1, page 8 pp., july 2005.

[14] Shuo Zhang and Y. Bar-Shalom. Particle filter processing of out-of-sequence measurements: Exact bayesian solution. In *Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP), 2011 4th IEEE International Workshop on*, pages 401 –404, dec. 2011.

[15] Feng Zhao, Jie Liu, Juan Liu, L. Guibas, and J. Reich. Collaborative signal and information processing: an information-directed approach. *Proceedings of the IEEE*, 91(8):1199 – 1209, aug. 2003.

[16] Feng Zhao, Jaewon Shin, and J. Reich. Information-driven dynamic sensor collaboration. *Signal Processing Magazine, IEEE*, 19(2):61 –72, mar 2002.