Mutual Information Based Resource Management Applied to Road Constrained Target Tracking

David Stein, Steve Theophanis, Walter Kuklinski, James Witkoskie, and Mike Otero

The MITRE Corporation
M/S E050
202 Burlington Road
Bedford, MA 01730-1420 USA
dstein@mitre.org

ABSTRACT

Netted sensors offer advantages for many surveillance applications. Target tracking and identification may be enhanced by jointly exploiting a variety of data sources, and certain surveillance applications may be more readily accomplished with jointly operated small in situ sensors than with large standoff sensors. Efficiently utilizing sensor-network data requires reliable sensor fusion and resource management algorithms. Resource management is particularly important when a limited number of tasks can be performed either because sensors may be used in one of several modes at any given time, or various resources, e.g. energy, computational capabilities and communication bandwidth are limited. In this paper tracking is viewed as a parameter estimation problem. Parameters are values in a state space and inference about the parameters is based on sensor measurements. The utility of sensor measurements is assessed using the mutual information between the parameters and the measurements. Resource management is achieved by minimizing average expected entropy subject to constraints. This approach is applied to a random set tracking algorithm that is based on Gaussian mixture models. Quadratic mutual information, which in this context is computable in closed form, is used as a substitute for mutual information when comparing the utility of sets of sensors of the same cardinality. The Mobius transformation is utilized to reduce the computational requirements of the optimization process. The tracking and resource management algorithms are demonstrated using a simulation capability. Four acoustic arrays, that measure angle of arrival and two radars, that measure range, monitor a triangular road network. For the example shown, two vehicles traversing the network, the tracker and resource manager are able to maintain the approximate quality of the estimate, as measured by average entropy of the distribution of the state space parameters, using, on average, less than 2.5 of the six sensors.

1.0 INTRODUCTION

Sensor networks are utilized when a single sensor can not adequately provide all desired information, e.g., target identification and localization may require multiple sensing modalities. Furthermore, kinematic or other variables may be better estimated using multiple sensors. Efficiently employing sensor networks requires methods to jointly utilize the available information to achieve the sensing objectives subject to constraints on energy usage, communications bandwidth, computational capability, or sensor modes of operation [5].

The present paper considers the problem of sensor selection for tracking a varying number of vehicles that are moving along roads that are under surveillance by a network of sensors. A random set tracker has been developed [11,12] to perform the localization using multiple sensor data. In applications, acoustic sensors provide angle of arrival and radars provide range estimates. Processed data from all tasked sensors are available at a central processor which uses the incoming data to update state parameter estimates that include number of targets and their location and velocity. Future work will also include vehicle identity. The resource management problem is to select the sensors to maintain the quality of the estimates subject to resource constraints.

In [12] the restless bandit approach is used to solve the sensor selection problem for this tracker. To apply this technique a Markov decision process is defined: monitoring road segments are the tasks; states are defined as knowledge of vehicular occupancy of task segments; actions are probing the task segments with sensors; rewards are generated for transitioning from one state to another and depend upon whether sensors were or were not employed. The optimal action given the current state provides the maximum expected infinite time horizon discounted reward, and it is found by solving a linear programming problem [12]. Resource optimization depends only on the current state, and as there are finitely many, the optimal action for a given state can be precomputed and stored in a lookup table.

Mutual information is the basis for sensor selection in earlier work [13,14]. These papers consider a single target moving through a two dimensional sensor field. In this work, at each time instant, the state space probability density function (pdf) is handed off to that sensor, the new leader, which optimizes the mutual information. The calculations are performed by discretizing the state and measurement spaces, which have a combined dimension of three. Dimension issues preclude applying a grid based approach to the problem considered herein of tracking a varying number of targets using potentially multiple simultaneous measurements. Information theoretic control using particle filters is described in [1].

The present paper equates the resource optimization problem with minimizing the average entropy of the state space probability density function subject to constraints. This approach does not require a Markov decision structure. However, resource optimization computations must be done in real time and not precomputed. The approach is described in section two, and various approximations and simplifications are introduced. In section three, simulation results using the random set tracker demonstrate the efficacy of the approach. For the problem of tracking two vehicles moving through a road network, little degradation in entropy occurs using on average less than 2.5 out of six sensors.

2.0 ENTROPY BASED RESOURCE MANAGEMENT

Tracking is viewed, herein, as a parameter estimation problem in which certain parameters of interest, e.g. position and velocity, may be time varying and other attributes, such as shape and color that may be used to identify the object, are fixed. The parameter estimation problem is then cast in the form of a state-space model in which inference regarding the parameters of interest is conditioned on measurements. The present approach assumes that the inference process produces probability density functions of the state space variables conditioned on past and/or current measurements and of the measurement space conditioned on past measurements. The state space model may have the general form

$$X^t = \left(\overbrace{\kappa_t}^{\substack{number\\ of\\ targets}}, \overbrace{\nu_1^{\,t},\ldots,\nu_{k_t}^{\,t}}^{\substack{kinematic\\ variables}}, \overbrace{\delta_1^{\,t},\ldots,\delta_{k_t}^{\,t}}^{\substack{target\\ identifiers}}, \overbrace{\omega_1^{\,t},\ldots,\omega_{c_t}^{\,t}}^{\substack{discreet\\ markov\\ chain}} \right). \tag{1}$$

Assume that sensors are employed and that Y^t denotes the measurement vector at time t. The state space model includes a measurement model, $g_t(Y^t \mid X^t)$,where g_t is the probability density on the measurements conditioned on the state space variate. Note that the measurement density further depends on the target-sensor relations, $\alpha \in \mathrm{A}$, i.e.,

$$g_t(Y^t \mid X^t) = \sum_{\alpha \in \mathrm{A}} g_t(Y^t \mid X^t, \alpha) p(\alpha) \tag{2}$$

The state space model is completed by specifying a state space propagation model, i.e. the probability distribution function on state space at time *t+1* conditioned on the variate at time *t*. $f_{t+1}(X^{t+1} \mid X^t)$
This formulation includes time series models of the form

$$X^{t+1} = \varphi(X^t, V^{t+1}, u^{t+1}) \quad \text{and} \quad Y^t = \psi(X^t, W^t),$$

where V^t and W^t are noise processes and u^t is a control parameter.

Witkoskie et al. develop a tracker in this context. Reference [12] allows for a variable number of targets and kinematic parameters, and reference [11], additionally, allows for target identifiers. In this work, the probability density functions are mixtures of Gaussians. Sensor fusion takes place through the estimation of state variables using multiple sensor data. As described below and in [12], the methods have been applied to tracking vehicles confined to roads using range-radars and angle of arrival estimates obtained from acoustic sensors.

In this context optimal control is the selection of sensors to optimize an objective function subject to resource constraints. For the present work sensors are chosen to minimize the expected entropy of the state space distribution subject to a resource constraint on each sensor or a total resource constraint.

Let

$$H(X^t \mid Y^{0:t}) = \int f(X^t \mid Y^{0:t}) \ln(f(X^t \mid Y^{0:t})) dX^t \tag{3}$$

be the entropy in the state space distribution conditioned on the measurements up to time t. Also define

$$A(X^t \mid Y^{0:t-1}) = \int f(X^t \mid Y^{0:t-1}) \ln(f(X^t \mid Y^{0:t-1})) dX^t - H(X^{t-1} \mid Y^{0:t-1}), \tag{4}$$

and

$$R(X^t \mid Y^{0:t}) = \int f(X^t \mid Y^{0:t-1}) \ln(f(X^t \mid Y^{0:t-1})) dX^t - H(X^t \mid Y^{0:t}). \tag{5}$$

Then

$$H(X^t) = H(X^{t-1}) + A(X^t) - R(X^t). \tag{6}$$

Define a cost function

$$C(T) = \int_{t_0}^{t_0+T} H(X^t) dt. \tag{7}$$

In discreet time this becomes

$$C=\sum_{j=0}^{N}H\left(X^{t_j}\right)=(N+1)H\left(X^{t_0}\right)+\sum_{j=1}^{N}(N-j+1)A\left(X^{t_j}\right)-\sum_{j=1}^{N}(N-j+1)R\left(X^{t_j}\right). \tag{8}$$

Assume that s sensors are available and that $R\left(X^{t_j}\right)=R_j\left(l_1^{\,j},\ldots,l_s^{\,j}\right)$ where $l_i^{\,j}$ is the resource committed to sensor i at time j. The resource optimization problem may then be expressed as

$$\text{minimize } C \text{ subject to the constraints } \sum_{j=1}^{N}l_i^{\,j}\le L_i. \tag{9}$$

If the sensors share a common resource then the resource optimization problem may be expressed as

$$\text{minimize } C \text{ subject to the constraints } \sum_{j=1}^{N}\sum_{i=1}^{s}l_i^{\,j}\le L. \tag{10}$$

Using Lagrange multipliers the solutions to problems (9) and (10) are found to be, assuming that the increment in entropy, $A\left(X^t\right)$, is independent of past measurements.

$$(N-j+1)\frac{\partial R_j}{\partial l_i^{\,j}}\left(l_1^{\,j},\ldots,l_s^{\,j}\right)=\lambda_i \tag{11}$$

and

$$(N-j+1)\sum_{i=1}^{s}\frac{\partial R_j}{\partial l_i^{\,j}}\left(l_1^{\,j},\ldots,l_s^{\,j}\right)=\lambda, \tag{12}$$

respectively.

In practice, R in Equations (5)—(12) is replaced with its expected value, which is equal to the mutual information, defined in (28) below, between the state variable conditioned on past measurements and current measurements, i.e.,

$$E(R)=I\left(X^t\mid Y^{0:t-1};Y^t\right). \tag{13}$$

The PDf in state-space is represented as a Gaussian mixture

$$f\left(X^{t+1}\mid Y^{0:t}\right)=\sum_{j=1}^{n_{t+1}}w_j^{\,t+1}N\left(\mu_j^{\,t+1|t},\Gamma_j^{\,t+1|t};X^{t+1}\right). \tag{14}$$

For each set, S, of sensors and each component, j, of the state space mixture define a set of target–sensor relations, $\mathrm{A}(S,j)$. The probability, $p(j,\alpha)$, of any target-sensor relation, $\alpha\in\mathrm{A}(S,j)$, is determined by the target-sensor geometry and characteristics of the background and sensors. Using a linearization of the measurement model about the mean of each mixture component, the measurement space density may be expressed as

$$g_S\left(Y^{t+1}\mid Y^{0:t}\right)=\sum_{j=1}^{n_{t+1}}\sum_{\alpha\in\mathrm{A}(S,j)}p(j,\alpha)N\left(\hat{\mu}_{j\alpha}^{\,t+1|t},\hat{\Gamma}_{j\alpha}^{\,t+1|t};Y^{t+1}\right), \tag{15}$$

where $\hat{\mu}_{j\alpha}^{\,t+1|t}$ and $\hat{\Gamma}_{j\alpha}^{\,t+1|t}$ are obtained by linearizing the measurement model about $\mu_j^{\,t+1|t}$. The density in state space conditioned on the sensor set S and measurements Y^{t+1} is approximated by

$$f_S\left(X^{t+1} \mid Y^{0:t}, Y^{t+1}\right) = \sum_{j=1}^{n_{t+1}} \sum_{\alpha \in \mathrm{A}(S,j)} p_{j\alpha}\left(Y^{t+1}\right) N\left(\tilde{\mu}_{j\alpha}{}^{t+1|t}(Y^{t+1}), \tilde{\Gamma}_{j\alpha}{}^{t+1|t}; X^{t+1}\right) \quad (16))$$

$$\approx \sum_{j=1}^{n_{t+1}} \sum_{\alpha \in (S,j)} p_{j\alpha}\left(Y^{t+1}\right) N\left(\mu_j{}^{t+1|t}, \tilde{\Gamma}_{j\alpha}{}^{t+1|t}; X^{t+1}\right) \quad (17)$$

$$\approx \sum_{j=1}^{n_{t+1}} p_j\left(Y^{t+1}\right) N\left(\mu_j{}^{t+1|t}, \overline{\Gamma}_j{}^{t+1|t}; X^{t+1}\right) \quad (18)$$

$$\equiv \tilde{f}\left(X^{t+1}; Y^{0:t}, S, Y^{t+1}\right). \quad (19)$$

where $\tilde{\mu}_{j\alpha}{}^{t+1|t}$ and $\tilde{\Gamma}_{j\alpha}{}^{t+1|t}$ are the Kalman filter updates of $\mu_j{}^{t+1|t}$, and $\Gamma_j{}^{t+1|t}$, respectively [2]. Note that $\tilde{\Gamma}_{j\alpha}{}^{t+1|t}$ is independent of the measurements. In going from (17) to (18), $\tilde{\mu}_{j\alpha}{}^{t+1|t}(Y^{t+1})$ is replaced with $\mu_j{}^{t+1|t}$, and on going from (18) to (19) the mixture terms for a given j are combined and approximated by the given normal.

The entropy of a normal mixture distribution is replaced with the following upper bound on the entropy. If $f(x) = \sum_{j=1}^{n} \omega_j N\left(\mu_j, \Gamma_j; x\right)$ is a normal mixture density function, then

$$\begin{aligned}
H(x) &= -\int f(x)\ln\left(f(x)\right)dx \\
&\le -\sum_{j=1}^{n} \omega_j \ln\left(\omega_j\right) - \sum_{j=1}^{n} \int N\left(\mu_j, \Gamma_j; x\right)\ln\left(N\left(\mu_j, \Gamma_j; x\right)\right)dx \\
&= -\sum_{j=1}^{n} \omega_j \ln\left(\omega_j\right) + \sum_{j=1}^{n} \frac{d_j}{2}\left(1 + \ln(2\pi)\right) + \sum_{j=1}^{n} \frac{\omega_j}{2}\ln\left(\left|\Gamma_j\right|\right) \qquad (20) \\
&= \tilde{H}(x), \qquad (21)
\end{aligned}$$

where d_j is the dimension of class *j*.

$\tilde{H}(X)$ is the entropy of the class weights, ω_j, plus the weighted entropy of the classes. Note that

$$\tilde{H}(X) = H\left((X, j)\right), \quad (22)$$

which might be called the complete entropy. Overlapping classes are combined as described in [12], and the more distinct the classes, the tighter the bound.

The mutual information is (see (28))

$$I\left(X^{t+1}, Y_S{}^{t+1}\right) = H\left(X^{t+1} \mid Y^{0:t}\right) - E_{Y_S{}^{t+1}}\left(H(X^{t+1} \mid Y^{0:t}, Y_S{}^{t+1})\right), \quad (23)$$

and for computational purposes it is replaced with

$$\tilde{I}\left(X^{t+1},Y_S^{t+1}\right)=\tilde{H}\left(X^{t+1}\mid Y_S^{0:t}\right)-\left(-\sum_{j=1}^{n_{t+1}}\int p\left(j\mid Y_S^{t}\right)\ln\left(p\left(j\mid Y_S^{t}\right)\right)p\left(Y_S^{t}\right)dY_S^{t}+\sum_{j=1}^{n_{t+1}}w_j^{t+1}H\left(N\left(\mu_j^{t+1|t},\overline{\Gamma}_j^{t+1|t}\right)\right)\right)$$

$$=\tilde{H}\left(X^{t+1}\mid Y_S^{0:t}\right)-\left(-\sum_{j=1}^{n}w_j^{t+1}\ln\left(w_j^{t+1}\right)-\sum_{j=1}^{n_t}w_j^{t+1}\int p\left(Y_S^{t}\mid j\right)\ln\left(\frac{p\left(Y_S^{t}\mid j\right)}{p\left(Y_S^{t}\right)}\right)dY_S^{t}\right.$$

$$\left.+\sum_{j=1}^{n}w_j^{t+1}H\left(N\left(\mu_j^{t+1:t},\overline{\Gamma}_j^{t+1:t}\right)\right)\right)$$

$$=\sum_{j=1}^{n}w_j^{t+1}H\left(N\left(\mu_j^{t+1:t},\Gamma_j^{t+1:t}\right)\right)-\sum_{j=1}^{n}w_j^{t+1}H\left(N\left(\mu_j^{t+1:t},\overline{\Gamma}_j^{t+1:t}\right)\right)$$

$$+\sum_{j=1}^{n_t}w_j^{t+1}\int p\left(Y_S^{t}\mid j\right)\ln\left(\frac{p\left(Y_S^{t}\mid j\right)}{p\left(Y_S^{t}\right)}\right)dY_S. \qquad (24)$$

Note that the difference of weighted normal entropies in (24) captures the reduction in the class covariance due to the measurements and the remaining term captures the reduction in the class ambiguities. Presently, Monte Carlo integration is used to approximate the integral in (24).

The search described above over all subsets of sensors grows exponentially with the number of sensors, $|S,|$ available. However, the Mobius transformation [3,8] may be invoked to reduce the search to one that is of order $|S|^k$ for small *k*. Following [3], suppose that Θ is a finite set and that f and g are functions on 2^{Θ}, the set of subsets of Θ. Then

$$f(A)=\sum_{B\subset A}g(B)\text{ for all }A\subset\Theta\text{ if and only if }g(A)=\sum_{B\subset A}(-1)^{|A-B|}f(B)\text{ for all }A\subset\Theta. \qquad (25)$$

If (25) holds, then *g* is the Mobius transform of *f*. To reduce the computational complexity of optimizing a function over a discreet set, select $k>0$ and define

$$\tilde{g}(A)=\begin{cases}g(A)\text{ if }|A|\le k\\ 0,\text{ otherwise.}\end{cases} \qquad (26)$$

Define

$$\tilde{f}(A)=\sum_{B\subset A}\tilde{g}(A). \qquad (27)$$

In applications to parameter estimation, *k* is chosen so that any parameter is well localized with k or fewer data. For the example described below $k=2$.

For certain applications, quadratic mutual information, I_T, may be computed in closed form, whereas Shannon's mutual information (I) often requires Monte Carlo techniques [6,7]. Shannon's mutual information is defined as follows and may be expressed using the Kullback-Leibler metric, *KL,* between the joint density and the product density. QMI is defined using the quadratic divergence between the joint density and the product density [10]. I_T provides a lower bound to I [10].

$$
\begin{aligned}
I(X,Y) &= H(X) - H(X \mid Y) \\
&= H(X) + \iint p(X \mid Y)\ln\left(p(X \mid Y)\right) dX p(Y) dY \\
&= -\int p(X)\ln\left(p(X)\right) dX + \iint \frac{p(X,Y)}{p(Y)} \ln\left(\frac{p(X,Y)}{p(Y)}\right) dX p(Y) dY \\
&= -\int p(X)\ln\left(p(X)\right) dX - \int p(Y)\ln\left(p(Y)\right) dY + \iint p(X,Y)\ln\left(p(X,Y)\right) dX dY \\
&= \iint p(X,Y)\ln\left(\frac{p(X,Y)}{p(X)p(Y)}\right) dX dY \\
&= KL\left(p_{XY}, p_X p_Y\right).
\end{aligned} \tag{28}
$$

QMI substitutes the quadratic divergence for the KL metric [10]. The quadratic divergence between two probability densities *f* and *g* is [10]

$$D(f,g) = \int \left(f(x) - g(x)\right)^2 dx, \tag{29}$$

and the QMI is [10]

$$I_T(X,Y) = D\left(p_{XY}, p_X p_Y\right). \tag{30}$$

Figure 1 illustrates the sensor subset selection procedure. The bottom row illustrates the propagation of the state space density in the absence of additional measurements. The top row illustrates for each selected subset the calculation of the expected information gain from measurements obtained with these sensors and the search over subsets of sensors. Subset selection takes place by applying one or more thresholds in accordance with (11) or (12).

3.0 APPLICATION

MITRE's Netted Sensor Program has deployed networks of radar, acoustic and imaging sensors and developed a simulation capability that is used to demonstrate the capabilities of sensors and algorithms. The methods described above were applied to the problem of tracking simulated vehicles traveling through a triangular mesh of roads as depicted in Figure 2(b). Two range radars and four acoustic arrays that provide bearing estimates [12] were utilized, and they were positioned as shown in Figure 2(b). The sensors in this study draw equal power, and are assumed to be connected to a common power supply, so that total energy consumption is the only constraint imposed.

Each trial of the simulation has duration 40 seconds. For this study, a target enters the mesh at location A at time 5 seconds and takes the direct path to location C where it exits; a second target enters the network from location C a few seconds later and takes the direct path to location B. The FISST tracker, as described in [11,12] is used to estimate the location of the vehicles as a function of time, and more generally to compute the probability density function on the state space given the radar and acoustic measurements, which in this case are simulated.

Subsets of sensors are selected using the following procedure. At every fifth time instant, i.e., 1.25 seconds, the entropy based resource manager described above computes the quadratic mutual information (30) for each subset of cardinality 1 and 2, sets the Mobius transform of sets with cardinality greater than two to zero, and uses (27) to approximate the quadratic mutual information of each possible subset of sensors. For each cardinality from one to the number of sensors, the maximal-qmi set is selected, and Shannon's mutual information (SMI) is computed (24). The derivative of SMI with respect to total energy consumption is computed, and a threshold is applied to select the sensor subset (12) as shown in Figure 2 (a). In (12) the factor $(N - j + 1)$ is ignored. Figure 2(b) shows the sensors selected for a given configuration of targets and a threshold, as well as the target locations, estimated positions and 95% confidence ellipses. Selected sensor subsets at higher thresholds are noted in the caption.

The resource manager was run at a range of thresholds to determine the relationship between the entropy of the system and the number of sensors utilized. Results are shown in Figures 3(a) and 3(b). For each threshold setting, the simulation was run 10 times and the average entropy at each pulse for each threshold is plotted in Figure 3(a). The average number of sensors utilized for each threshold was also determined and the plots in this figure are labeled by the corresponding average number of sensors utilized. The entropy plots are only degraded minimally if as few as 2.4 sensors, on average, are selected. Figure 3(b) compares the average entropy across trials and across the duration of each run for each of the thresholds shown in Figure 3(a). The average lifetime extension factor is the maximal number of sensors divided by the average number of sensors employed for the given threshold setting, and the average relative entropy is the ratio of the average entropy obtained using the resource manager to the average entropy if all available sensors are utilized. One sees that the lifetime of the system can be extended by approximately a factor of three with minimal average degradation as measured by entropy.

4.0 SUMMARY AND CONCLUSIONS

A resource management algorithm, based on minimizing average entropy of the state space probability density function subject to constraints, that is applicable to tracking with a network of heterogeneous sensors multiple targets moving through a road mesh was developed and demonstrated. For the problem considered, using on average 2.5 sensors out of a total of six resulted in little degradation in average entropy of the state space probability density function. This work extends earlier work on entropy based resource management in that it applies to a varying number of targets tracked with heterogeneous sensors and allows for multiple sensor data to contribute at any time instant to the update of estimates of state space parameters. Computational simplifications using quadratic mutual information and the Mobius transform were introduced and shown to be effective. In particular, using the Mobius transform reduces the computational complexity from one that is exponential in the number of sensors to one that is polynomial in the number of sensors. Future work will address remaining computational issues, incorporate identification along with kinematic variables, incorporate longer time horizons and address the relationships between resource allocation based on mutual information and Markov decision processes.

REFERENCES

[1] Christophe Andrieu, Arnaud Doucet**,** Sumeetpal S. Singh, and Vladislav B. Tadic. "Particle Methods for Change Detection, System Identification and Control," *Proceedings of the IEEE*, vol. 92, No 3, March 2004, pp. 423—438.

[2] Yaakov Bar-Shalom, X. Rong Li, and Thiagalingam Kirubarajan, *Estimation with Applications to Tracking and Navigation*, John Wiley and Sons, New York, 2001, Chapter 5.

[3] Alain Chateauneuf and Jean-Yves Jaffray, "Some Characterizations of Lower Probabilities and other Monotone Capacities Through the Use of Mobius Inversion," *Mathematical Social Sciences* Vol. 17 (1989) pp. 263—283.

[4] Tommy W. S. Chow and D. Huang, “Estimating Optimal Feature Subsets Using Efficient Estimation of High-Dimensional Mutual Information,” *IEEE Transactions on Neural Netw*orks, vol. 16 No. 1, January 2005, pp. 213—224.

[5] Deborah Estrin, Ramesh Govindan, John Heidemann, and Satish Kumar, “Next Century Challenges: Scalable Coordination in Sensor Networks,” *Mobicom ’99*, Seattle, Washington USA, pp. 263—270.

[6] Peter Hall and Sally Morton, “On the Estimation of Entropy,” *Ann. Inst. Statist. Math*., Vol. 45 No. 1, 1993, pp. 69—88.

[7] Harry Joe, “Estimation of Entropy and Other Functionals of a Multivariate Density,” *Ann. Inst. Statist. Math.*, Vol. 41, No. 4, 1989, pp. 683—697.

[8] Ivan Kojadinovic, “Relevance Measures for Subset Variable Selection in Regression Problems Based On K-Additive Mutual Information,” *Computational Statistics and Data Analysis,*” Vol. 49, 2005, pp. 1205—1227.

[9] Juan Liu, Feng Zhao, and Dragan Petrovic, “Information-Directed Routing in Ad Hoc Sensor Networks,” *IEEE Journal on Selected Areas in Communications,* Vol. 23 Vol. 4, 2005, pp. 851—861.

[10] Kari Torkkola, “Feature Extraction by Non-parametric Mutual Information Maximization,” *Journal of Machine Learning Research,* Vol. 3, 2003, pp. 1415—1438.

[11] James Witkoskie, Walter Kuklinski, Stephen Theophanis, David Stein and Michael Otero, “Feature-aided Random Set Tracking on a Road Constrained Network,” MSS National Symposium on Sensor and Data Fusion, McLean Virginia, 6—8 June 2006.

[12] James Witkoskie, Walter Kuklinski, Stephen Theophanis and Michael Otero, “Random Set Tracker Experiment of a Road Constrained Network with Resource Management,” to appear, Fusion 2006, 9th International Conference on Information Fusion, Florence Italy, 10—13 July 2006.

[13] Feng Zhao, Juan Liu, Juan Liu, Leonidas Guibas, and James Reich, “Collaborative Signal and Information Processing: An Information-Directed Approach,” *Proceedings of the IEEE*, Vol. 91. No. 8, August 2003, pp. 1199—1209.

[14] Juan Liu, James Reich, and Feng Zhao, “Collaborative In-Network Processing for Target Tracking,” EURASIP Journal on Applied Signal Processing 2003: 4, 378—391.

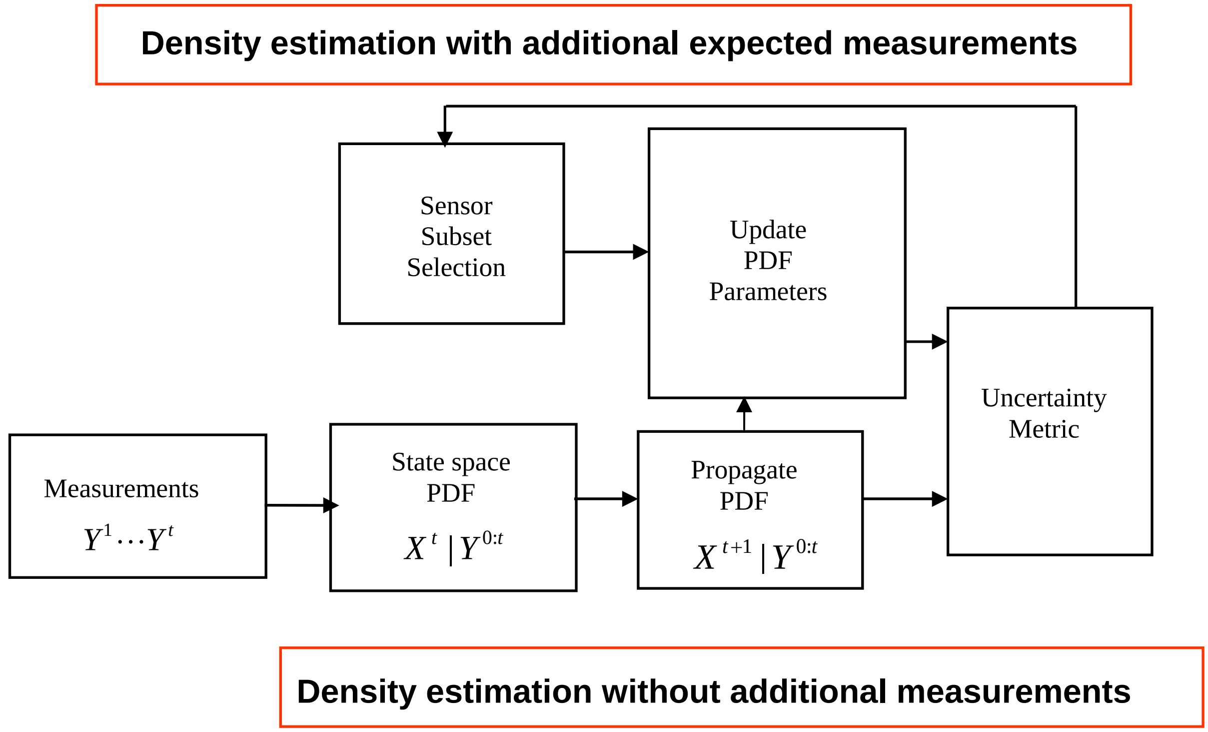
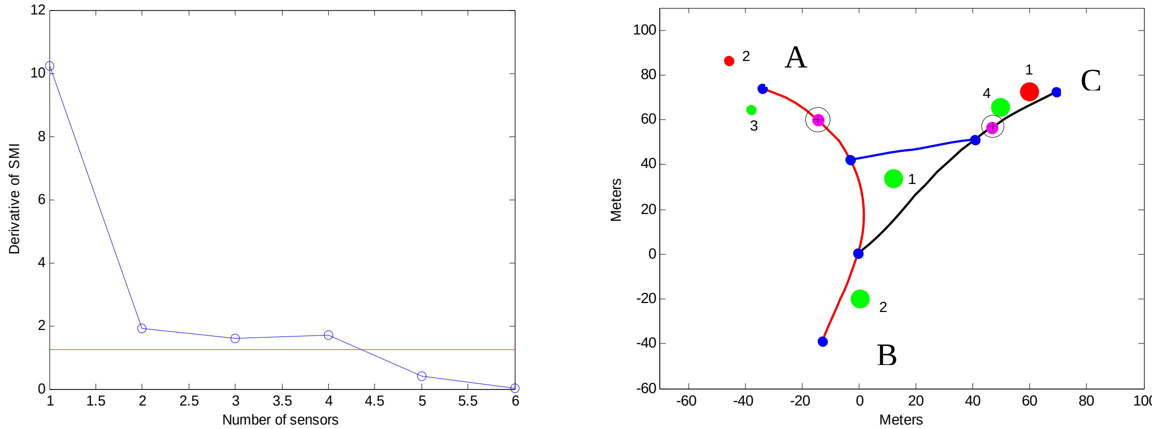
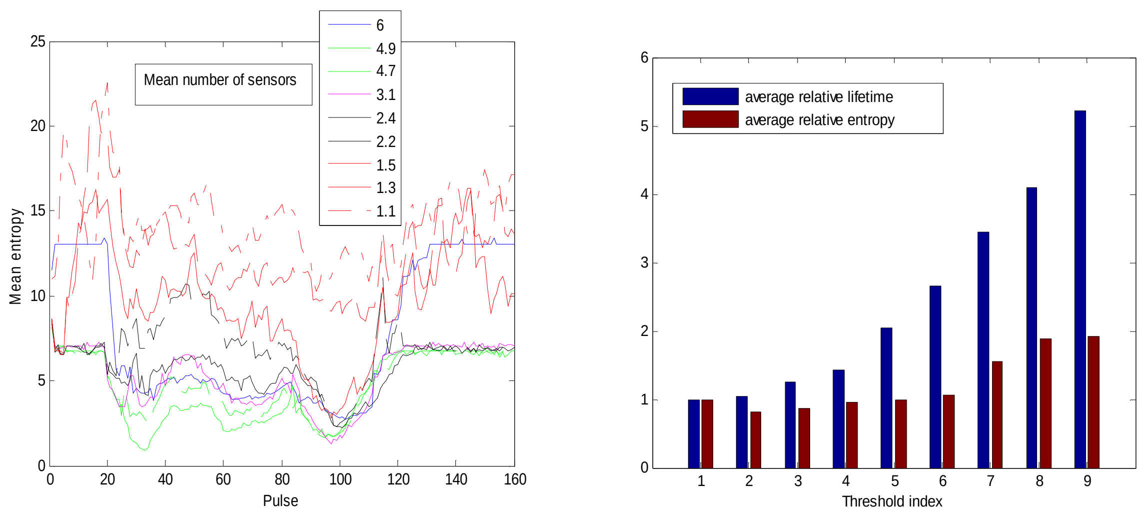


Figure 1. Diagram of the resource management algorithm.



(a) (b)

Figure 2. (a) The derivative of the mutual information curve with respect to energy used to select subsets of sensors. (b) A snapshot of the tracking scenario showing the targets (magenta), 95% confidence ellipses, active arrays (larger green disks), active radar (larger red disk), inactive array (smaller green disk), inactive radar (smaller red disk). Thus, arrays 1,2, and 4 and radar 1 are on, while radar 2 (R-2) and A-3 are off. If the threshold had been set higher so that two sensors were selected they would have been R-1 and A-3, and if three sensors had been selected they would have been R-1, A-3, and A-4.



(a) (b)

Figure 3. (a) Average entropy curves for a range of threshold values are shown for each position in a run of the simulation. The first target enters the system at 20 pulses and both targets have left by 120 pulses. Each graph corresponds to a particular threshold, and the labels show the correspondence between threshold and average number of sensors selected. (b) The curves in 3(a) are further averaged and compared with the average entropy (red) measured if all of the sensors are utilized all of the time. Additionally, the average life time extension factor in comparison with using all sensors all of the time is shown in blue.