Proceedings of the
47th IEEE Conference on Decision and Control
Cancun, Mexico, Dec. 9-11, 2008

TuB14.6

Dynamic Ping Optimization for Surveillance in Multistatic Sonar Buoy Networks with Energy Constraints

Anshu Saksena and I-Jeng Wang

Abstract— In this paper we study the problem of dynamic optimization of ping schedule in an active sonar buoy network deployed to provide persistent surveillance of a littoral area through multistatic detection. The goal of ping scheduling is to dynamically determine when to ping and which ping source to engage in order to achieve the desirable detection performance. For applications where persistent surveillance is needed for an extended period of time, it is expected that the energy available at each ping source is limited relative to the required system lifetime. Hence efficient management of power consumption for pinging is important to support the required lifetime of the network while maintaining acceptable detection performance. Our approach to ping optimization is based on the application of approximate Partially Observable Markov Decision Process (POMDP) techniques such as the rollout algorithm. To enable a practical implementation of the policy rollout, we apply sampling-based techniques based on a simplified model that approximates the detailed multistatic model. Using high fidelity sonar simulations, we evaluate the performance of the proposed approach and compare it with the greedy technique in terms of detection performance and system lifetime.

I. INTRODUCTION

In this paper we derive a formulation of the dynamic ping optimization problem in multistatic sonar buoy networks using the framework of Partially Observable Markov Decision Processes (POMDP). The basic operating concept of a multistatic sonar buoy network for underwater surveillance is to proactively ping from an acoustic source and correlate the echo returns (specifically time of arrivals and bearings) across receivers in the field to detect, localize, and track targets of interest. In order to provide necessary coverage for the surveillance area and sufficient diversity in multistatic geometry for target localization, multiple receivers and acoustic sources are deployed and distributed throughout the field. The ping optimization problem considered here is a specific example of sensor management in sensor networks, which has received significant interest in recent years in generic sensor network settings (see, for example, [1], [2], [3]). Our work focuses on the practical application of existing algorithms (in particular, the approximate dynamic programming techniques proposed in [3], [4]) to multistatic sonar networks where sensing and detection models are more complicated than in settings considered in the literature.

The benefit of optimizing ping sequences to tracking performance for multistatic sonar systems has been demonstrated through simulations in [5], where a greedy approach is presented to optimize the expected detection probability during the search phase when no detections are confirmed and to improve the track with worst quality after tracks are established. In this paper we focus on the aspects of ping source selection with additional energy constraint during the *surveillance* phase. The surveillance phase is defined as the stage where no specific track with sufficient confidence has been established based on the detection results from the network of receivers. In contrast to [5], where complete absence of confirmed detections is assumed, we permit the existence of initial detections in our formulation. The rationale for this unique feature is two-fold:

- In practical applications, sparse detections are expected before sufficient measurements are available to establish a track (the track initiation problem). Furthermore, an approach assuming no detection has limited applicability since false alarms are inevitable for most practical scenarios.
- The separation of surveillance and tactical (after tracks are established) phases postulated by existing literature is somewhat artificial. In reality, both situations exist simultaneously since continuing search of new targets is desirable even after tracks are established. A formulation incorporating detections during the surveillance phase (such as the one proposed here) can likely be extended to address requirements for both search and tracking simultaneously.

In Section II, we formally define all the necessary attributes of a POMDP for the ping optimization problem. Based on the formulation, we discuss in Section III how the policy rollout technique can be applied to sequentially solve for ping source selections by taking a nonmyopic view of performance over a time horizon into the near future. This approach serves to trade-off the detection performance versus energy management requirement. We present simulation results to demonstrate the performance of proposed algorithms in Section IV. Throughout the discussion, we assume that there are N_s sources and N_r receivers deployed in the field and their positions are stationary.

II. MATHEMATICAL FORMULATION OF PING OPTIMIZATION PROBLEMS

We assume a discrete time model and use $k = 0, 1, 2, \ldots$ to denote the kth stage with 0th stage denoting the starting time. We will use time k to denote the beginning of the
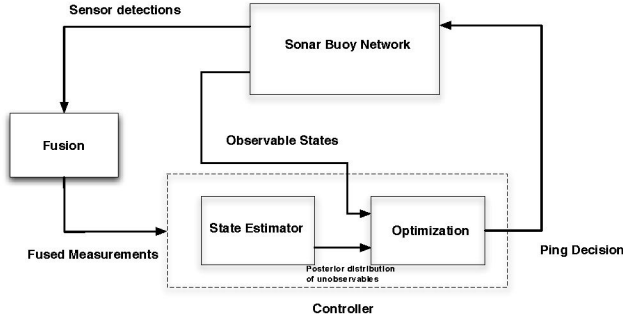
Fig. 1. **An illustration of the POMDP for ping optimization.**

kth stage. Following the standard MDP framework, we will define the states $X(k)$, the action $a(k)$, the observation $Z(k)$ and its dependency on states and actions, the state transition, the immediate reward $r(X(k), a(k))$, the horizon H, and the terminal cost $r_T(X(H))$ if needed. Given the initial state $X(0)$ (or its distribution), the optimization problem defined by the MDP is given by

$$V_H^*(X(0)) = \max_{a(0),\dots,a(H-1)} E[r(X(0), a(0)) + \cdots$$
$$+ r(X(H-1), a(H-1)) + r_T(X(H))], \quad (1)$$

where $a(k) \in \{0,\dots,N_s\}$ is the ping source selection at time k, with $a(k) = 0$ corresponding to the action of not pinging from any source, and the maximization is taken over all feasible sequences of actions. Bellman's equation provides a characterization of the performance of the optimal decisions

$$V_H^*(X) = \max_a Q(X, a),$$
$$Q(X, a) \triangleq r(X, a) + E_{X'}[V_{H-1}^*(X')],$$

where X' is the next state after taking the action a at state X and $Q(X, a)$ is typically referred to as the Q-value of action a at state X.

Given that the states are not fully observable, we cast the problem as a POMDP using the concept of the belief state or the conditional probability. The basic idea of the POMDP is illustrated in Figure 1.

A. States and Observations

The states of the ping optimization can be classified into two types of states: the target states that capture the dynamic behaviors of targets; and the sensor states that characterize the status of the sensor systems.

1) Target states and observation: For the surveillance case where no credible tracks have been established, it is difficult to model all the uncertainty, including the number of targets in the field. Instead of directly modelling the target states, we characterize the potential target presence across the field. We further discretize the field to arrive at a more tractable formulation. Specifically, we assume that the field is discretized into a grid $\{(x_i, y_i), i = 1,\dots, N_g\}$, where N_g

is the size of the grid. The surveillance target state at time k is defined by

$$X(k) = \left[t_1(k),\dots, t_{N_g}(k)\right]^T,$$

where

$$t_i(k) = \begin{cases} 1 & \text{if there is at least 1 target at } (x_i, y_i) \text{ at time } k, \\ 0 & \text{otherwise.} \end{cases}$$

Note that in this model we have made two approximating assumptions: (1) target localizations are mapped to the grid; and (2) we make no differentiation in our models between the cases of a single target and multiple targets being at a grid point.

Since no specific target tracking information has been established during the surveillance phase, a randomized model derived from a stochastic spatial diffusion process determines the random movement and arrival of potential targets. We define the state transition probability by assuming that the target dynamics are independent of each other and not affected by the action taken:

$$P(X(k+1)|X(k), a(k)) = \prod_i P(t_i(k+1)|X(k)),$$

where the individual target transition model is given by a Gaussian distribution.

The target states are not directly observable but can be estimated from the observation:

$$Z(k) = \left[z_1(k),\dots, z_{N_g}(k)\right]^T,$$

where

$$z_i(k) = \begin{cases} 1 & \text{if a detection occurs at } (x_i, y_i) \text{ after } a(k), \\ 0 & \text{otherwise.} \end{cases}$$

The observation model is specified by the conditional probability $P(Z(k)|X(k), a(k))$. If we assume that the detection result at the grid point (x_i, y_i) with action $a(k)$ is independent of both detection $z_j(k)$ and target presence $t_j(k)$ at other grid points (x_j, y_j), with $j \neq i$, at time k, then

$$P(Z(k)|X(k), a(k)) = \prod_i P(z_i(k)|t_i(k), a(k)), \text{ and}$$

$$P(z_i(k) = 1|t_i(k) = 1, a(k)) = p_d(x_i, y_i, a(k)), \quad (2)$$
$$P(z_i(k) = 1|t_i(k) = 0, a(k)) = p_{fa}(x_i, y_i, a(k)). \quad (3)$$

2) Sensor states: In the initial formulation, we assume that all the sensors are stationary and focus on modelling the energy levels at the sources.

We assume a simple model for the energy states:

$$E(k) = [e_1(k),\dots, e_{N_s}(k)]^T, \qquad e_i(k) \in \{0\} \cup \mathbb{N},$$

where $e_i(k)$ is the remaining number of pings the source i can support in the beginning of the time k *before* an action

$a(k)$ is taken. Assuming that the initial energy state $E(0)$ is given, the state transition is defined by

$$e_i(k+1) = \begin{cases} e_i(k) - 1 & \text{if } a(k) = i \text{ and } e_i(k) \geq 1, \\ e_i(k) & \text{otherwise.} \end{cases}$$

In this simple model we assume that the energy consumption for pinging is deterministic. Furthermore, we assume that the energy state $E(k)$ is directly observable (that is, can be measured without errors).

B. Belief states and estimation

To apply the POMDP technique, we define the belief states and derive a recursion that enables us to estimate the belief states sequentially based on the observations $Z(k)$.

The belief state is defined as the conditional probability of the states given the information state (all prior observations and actions), that is,

$$P\{X(k)|Z(0), Z(1), \ldots, Z(k-1), a(0), \ldots, a(k-1)\}.$$

We denote the belief state by

$$P_T(k) = \left[P_T(x_1, y_1; k), P_T(x_2, y_2; k), \ldots, P_T(x_{N_g}, y_{N_g}; k)\right]^T,$$

where $P_T(x_i, y_i; k)$ denotes the conditional probability of target presence at (x_i, y_i) at time k given all prior observations and actions.

The belief state $P_T(k)$ can be estimated and propagated sequentially by a two-step recursion. Given the current belief state $P_T(k)$:

1) **Bayes update:** Given an action $a(k)$ and the resulting observation $Z(k)$, compute the *a posteriori* belief state $P_{T|Z}$ using Bayes rule.

$$P_{T|Z(k)}(x_i, y_i; k) =$$
$$\begin{cases} \frac{p_d P_T}{p_d P_T + p_{fa}[1 - P_T]} & \text{if } z_i(k) = 1, \\ \frac{(1 - p_d) P_T}{[1 - p_d] P_T + [1 - p_{fa}][1 - P_T]} & \text{otherwise.} \end{cases}$$

where p_d, p_{fa}, and P_T have been written instead of $p_d(x_i, y_i, a(k))$, $p_{fa}(x_i, y_i, a(k))$, and $P_T(x_i, y_i; k)$, respectively, for readability.

2) **Diffusion:** Propagate the *a posteriori* belief state according to the transition model

$$P_T(x_i, y_i; k+1) = 1 - \prod_j \left(1 - P_{T|Z(k)}(x_j, y_j; k)P_{ij}\right)$$

The following diagram illustrates the propagation of belief state:

$$P_T(k) \xrightarrow{a(k)} Z(k) \xrightarrow{\text{Bayes update}} P_{T|Z(k)}(k) \xrightarrow{\text{diffusion}} P_T(k+1)$$

C. Feasible actions and policies

An action that meets all the constraints defined in the problem is called a feasible action. An obvious constraint for the ping optimization problem is that the energy level at a source is sufficient to ping. That is, a_i is feasible only if $e_i \geq 1$ in our formulation. Note that the feasibility of an action will depend on the current state $\{P_T(k), E(k)\}$.
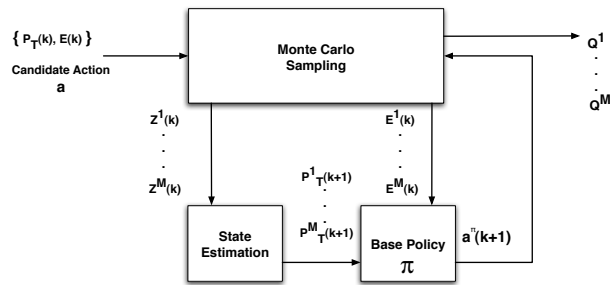

Fig. 2. An illustration of the policy rollout.

We denote the set of feasible actions at time k by $\mathcal{A}_k \subset \{0, \ldots, N_s\}$.

A (Markov) policy π is a mapping from the states $\{P_T(k), E(k)\}$ to an action $a(k)$. A policy is feasible if $\pi(P_T(k), E(k)) \in \mathcal{A}_k$ for all $k = 0, \ldots, H - 1$. Two typical examples of policy for the ping optimization problem are:

- Fixed sequence policies: One can define a feasible policy by repeating the same sequence of ping decisions as long as the remaining energy level at a source is sufficient to ping. The round-robin algorithm falls under this class of policies.
- Greedy policies: The greedy policy is a policy that selects a decision among the feasible actions by maximizing the immediate reward. That is

$$\pi(P_T(k), E(k)) = \text{argmax}_{a \in \mathcal{A}_k} E\left[r(P_T(k), E(k), a)\right].$$

III. OPTIMIZATION CRITERIA AND POLICY ROLLOUT

A. Rewards and dynamic programming formulation

For detection performance we define the reward function using the average detection probability across the field:

$$r(P_T, a) = \frac{1}{N_g} \sum_{i=1}^{N_g} P_T^\alpha(x_i, y_i) p_d(x_i, y_i, a), \qquad (4)$$

where $\alpha \geq 1$ is a parameter that allows us to control the search strategy.

To model the energy constraint, we assume a simplified model where each ping source has a pre-specified number of pings throughout its lifetime. Even though we do not explicitly model energy related considerations as part of the reward in POMDP, it is expected that the energy constraint will play a role in the performance of POMDP if the selected horizon is long enough relative to the available ping resources.

With the reward and constraint defined above, we formulate ping optimization as the following stochastic dynamic programming problem:

$$\max_{a(0), \ldots, a(H-1); a(k) \in \mathcal{A}_k} E\left[\sum_{k=0}^{H-1} r(P_T(k), a(k))\right], \quad (5)$$

where r is defined in (4). For our experiments, we use the formulation (5) in a receding horizon setting. The receding

horizon setting can potentially reduce the impact of model inaccuracy on the prediction of future detection performance in dynamic programming.

B. Policy Rollout Technique

To address the complexity of the dynamic programming formulated for the ping optimization problem, we apply the policy rollout technique [9]. The basic idea of the policy rollout algorithm is to use the performance resulting from a base policy π to estimate the Q-value for making the decision. Given the complexity of the underlying models, it is difficult to directly estimate the performance of a base policy. Hence, we apply a sampling-based approach similar to the ones used in [3], [4] to this estimation problem.

The sampling-based approach relies on the outcome of multiple Monte Carlo simulations with the base policy to estimate the Q-value of each feasible action. Consider the traces of states from M independent Monte Carlo runs by taking an action a and following a feasible base policy π afterward:

$$\{P_T(k), E(k)\} \xrightarrow{a} Z^i(k) \rightarrow \{P_T^i(k+1), E^i(k+1)\}$$
$$\xrightarrow{a^\pi(k+1)} \{P_T^i(k+2), E^i(k+2)\}$$
$$\vdots$$
$$\xrightarrow{a^\pi(H-1)} \{P_T^i(H), E^i(H)\},$$

$i = 1, \ldots, M$. Note that the specific actions taken following the base policy will generally differ among the traces since they depend on the specific belief and energy state trajectory, $\{P_T^i, E^i\}$, obtained in each Monte Carlo run.

From each sample path, we obtain a sample of the Q-value for the base policy π as

$$Q^i(P_T(k), a) = r(P_T(k), a) + r(P_T^i(k+1), a^\pi(k+1))$$
$$+ \cdots + r(P_T^i(H-1), a^\pi(H-1)).$$

An estimate of the Q-value for the action a at the belief state $P_T(k)$, denoted by $\bar{Q}(P_T(k), a)$, can be obtained as the empirical average over these Monte Carlo runs. Following the policy rollout strategy, the action at time k is chosen based on

$$a^*(k) = \text{argmax}_{a \in \mathcal{A}_k} \bar{Q}(P_T(k), a).$$

In this paper, we consider the following approximation model for simulations of $Z(k)$ in policy rollout:

$$z_i(k) = 1 \quad \text{with probability } p_d P_T + p_{fa}(1 - P_T),$$

where P_T, p_d, and p_{fa} are again used as shorthand for $P_T(x_i, y_i; k)$, $p_d(x_i, y_i, a(k))$, and $p_{fa}(x_i, y_i, a(k))$, respectively, for the sake of readability.

IV. HIGH FIDELITY SIMULATION RESULTS

A. Sonar Models

In order to test performance over a range of environmental settings and conditions, numerical modeling was used to simulate the motion of a target through a distributed

TABLE I

TOTAL PROBABILITY OF DETECTION FOR TARGET 1

	Mean	Std dev	Num scenarios
Round-robin	293.9	87.3	11
Greedy	333.4	92.3	11
Rollout H=10	358.8	102.7	7
Rollout H=15	339.2	103.0	8
Rollout H=20	328.0	93.5	10

TABLE II

TOTAL PROBABILITY OF DETECTION FOR TARGET 2

	Mean	Std dev	Num scenarios
Round-robin	214.2	56.8	11
Greedy	240.1	65.2	11
Rollout H=10	215.0	55.2	7
Rollout H=15	226.4	63.1	8
Rollout H=20	228.3	65.4	10

sensor network. The signal excess for echoes associated with source-receiver pairs was determined for each event in a ping schedule using a physics-based model. Clutter was established statistically from a measured characteristic distribution.

Transmission loss calculations are performed using APLNM, a range-dependent coupled normal mode code. Sound velocity profiles, bathymetry, and bottom properties are extracted from standard databases.

Background interference was calculated using a power sum of an assumed spectral ambient noise level (NL) and a calculated spectral reverberation level (RL). The RL was calculated by assuming a constant sound speed, with an equitime ellipse of reverberant patches (RVB) constrained by the signal bandwidth. The receiver beam pattern is included directly in establishing RL, with a single directivity index applied to the NL. Both NL and RVB are determined in units of intensity, and are assumed constant over the bandwidth of the signal.

Potential clutter events are established using a uniform density of events for each receiver beam (e.g. one event per second). These events are drawn randomly from a generalized gamma distribution having mean, standard deviation, and skewness parameters fit to represent measured clutter statistics from at-sea experiments.

B. Evaluation with Simulated Data

Our simulations are based on a sonobuoy field consisting of 16 receivers arranged in a 4x4 grid and 4 active sonar sources located within the grid of receivers as shown in Figure 3. Each source begins with enough energy to ping

TABLE III

SYSTEM LIFETIME

	Mean	Min	Max	Num scenarios
Round-robin	600.0	600	600	11
Greedy	601.0	601	601	11
Rollout H=10	638.0	601	840	7
Rollout H=15	635.8	603	840	8
Rollout H=20	631.7	603	840	10

150 times. The first target enters the field at stage 15, and the second enters at stage 250.

We compare detection performance of three source selection policies: the round-robin policy, the greedy policy, and policy rollout using the greedy policy as the base policy. Policy rollout optimized the decision at each stage based on the average of up to 50 Monte Carlo simulations of the observations. At each stage, there are 5 possible actions: ping from one of the 4 sources, or do not ping from any source.

We measure surveillance performance by summing the probability of detecting each target over each simulation run. A number of different target track scenarios were run using round-robin, greedy, and policy rollout strategies, with their results summarized in Table I for the first target entering the field and in Table II for the second. Policy rollout was run with horizon lengths of 10, 15, and 20 stages with the results for each listed in the table separately.

Since target 1 is tracked with sufficient energy reserves, prediction accuracy over the horizon is governed by the uninformed diffusion target motion model – the longer horizons seem to have overextended the relevance of those predictions. Since target 2 arrived later, the end game of limited resources becomes more of a factor. Longer horizons may be beneficial in this case since prediction of energy depletion is accurate even over long horizons. While these trends seem reasonable, the statistics of these results preclude us from concluding that any of these policies performs better or worse than any other.

The system lifetime, defined as the time after which all sources have exhausted their energy, is shown in Table III. Since we do not allow pings to occur simultaneously, 600 is a lower bound on the lifetime from any algorithm. Since the round-robin policy pings at every stage, it produces lifetimes of 600 for all scenarios. With the exception of the first stage when there are no targets believed to be in the area and our implementation selects not to ping, the greedy policy also pings at every stage, until source energies are exhausted at

Fig. 4. Plot of the energy reserves in the number of pings remaining on each source as a function of time under the greedy source selection policy.
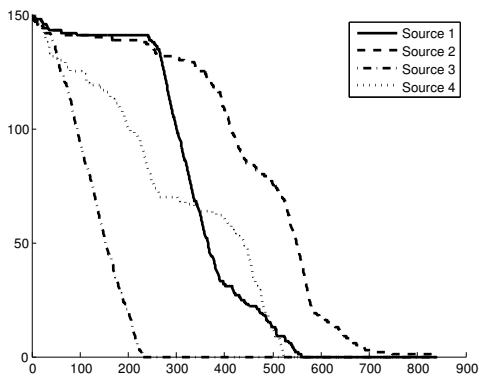
Fig. 5. Plot of the energy reserves in the number of pings remaining on each source as a function of time under the rollout source selection policy with horizon length of 20.

stage 601.

However, the table shows that rollout produces longer lifetimes, as would be expected by the algorithm's anticipation of the scarcity and subsequent rationing of energy. For some scenarios, there were sources that still had energy when the simulation was terminated at the 840th stage. Since persistent surveillance benefits from detection coverage in time as well as space, this is an important result, and the fact that this is accomplished without a significant impact on detection performance over that lifetime demonstrates the promise of this type of approach.

We now briefly present more detailed data from one scenario to illustrate the performance of policy rollout with horizon length 20 versus the greedy algorithm. Figure 3 shows an example of the simulated tracks that was used to generate the plots presented here. Figure 4 shows the energy reserves of each source as a function of time for the greedy policy, and Figure 5 shows this for policy rollout, giving an idea of how judiciously each policy makes use of each source.
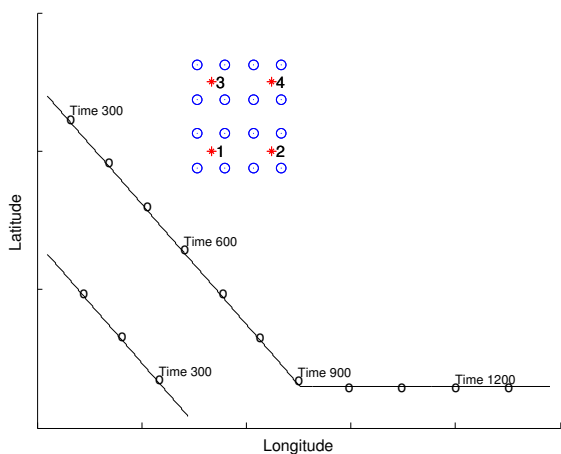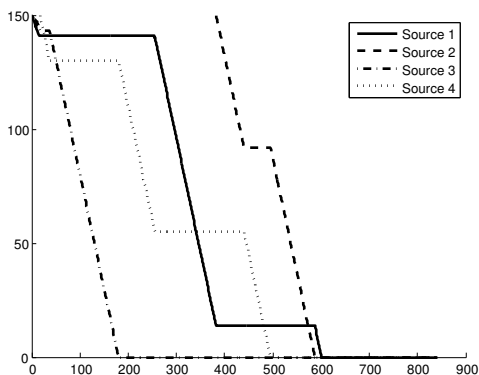

Fig. 3. Simulated scenario: asterisks indicate acoustic sources, circles receivers, lines show target tracks, and circles indicate the time at which the target is at that point of the track. Targets enter the field from the Northwest, turn East, and leave.

V. CONCLUSIONS AND FUTURE WORK

In this paper, we present a mathematical formulation of the dynamic ping optimization problem for sonar buoy

networks with energy constraints. We develop an implementation of the sampling-based policy rollout to optimize the average detection performance. Even with a much simplified model used in policy rollout, the approximate dynamic programming approach leads to a longer system lifetime while achieving detection performance similar to the greedy approach, as illustrated by the experimental results using high-fidelity sonar simulations. As an immediate future work, we will consider applying state aggregation techniques to further improve the efficiency of policy rollout.

We address the energy constraint by defining a horizon long enough in stochastic dynamic programming. One might consider explicitly modelling the energy consideration to enable a receding horizon formulation with shorter horizon (and hence reducing the computational costs of policy rollout). Theoretically, the most appropriate metric for energy management is the lifetime of the system. However, directly using the lifetime as the reward will lead to a difficult POMDP. A good alternative is to use a load-balancing type of metric to capture the fairness of energy consumption across the field. Given the energy levels $E = [e_1, \ldots, e_{N_s}]^T$, define the deviation of energy level at source i as

$$\sigma_i^e(E) = \frac{|e_i - \bar{e}|}{\bar{e}},$$

where $\bar{e}$ is the average among $\{e_i\}$. We can then use the energy deviations to either define the reward

$$r(P_T, E, a) = -\sum_{i=1}^{N_s} \sigma_i^e(E'), \qquad (6)$$

or define a constraint

$$\sigma_i^e(E') \leq \rho(E) < 1, \text{for all } i = 1, \ldots, N_s, \qquad (7)$$

where E' is the resulting energy levels *after* the action a is taken given current energy levels E, and $\rho(E)$ is defined to ensure the existence of a feasible action while maintaining a desirable level of load balancing.

Finally, since we consider the detections during the surveillance in our formulation, it is straightforward to incorporate the outcome of tracking algorithms into our formulation with an appropriate expansion of the state space to include the current target course and speed, etc.. In essence, the state estimation process we discussed in Section II-B is a simple Bayesian tracking algorithm with the probabilistic diffusion as its underlying dynamic model. Among the existing tracking approaches, the Probability Hypothesis Density (PHD) filter [10], [11] is perhaps the most natural one to consider. The PHD filter recursions propagate the intensity of the target states in a way that is very similiar to the propagation of the belief state considered in our formulation. A key challenge in incorporating tracking outcomes in ping optimization is to devise a performance metric that adequately addresses the operational priority regarding the trade-off between continuing tracking of confirmed targets and persistent surveillance of the entire field. One possible approach is to experiment with the choice of parameter α in the average detection reward defined in (4) to control the

preference of future detection on locations with higher target presence belief where established tracks would be.

VI. ACKNOWLEDGEMENTS

REFERENCES

[1] V. Krishnamurthy, "Algorithms for optimal scheduling of hidden Markov model sensors," *IEEE Transactions on Signal Processing*, vol. 50, no. 6, pp. 1382–1397, 2002.

[2] C. M. Kreucher, A. O. Hero, III, K. D. Kastella, and M. R. Morelande, "An information-based approach to sensor management in large dynamic networks," *Proceedings of the IEEE*, vol. 95, no. 5, May 2007.

[3] Y. Li, L. W. Krakow, E. K. P. Chong, and K. N. Groom, "Approximate stochastic dynamic programming for sensor scheduling to track multiple targets," *Digital Signal Processing*, to appear.

[4] E. K. P. Chong, C. Kreucher, and A. O. Hero III, "POMDP approximation methods based on heuristics and simulation," Chapter 8 in Foundations and Applications of Sensor Management, A. O. Hero, D. Castanon, D. Cochran, and K. Kastella, Eds., Chapter 5, Springer, 2008, ISBN: 978-0-387-27892-6, pp. 95-120.

[5] D. W. Krout, M. A. El-Sharkawi, W. J. L. Fox, and M. U. Hazen, "Intelligent ping sequencing for multistatic sonar systems," *2006 9th International Conference on Information Fusion*, 2006.

[6] W. A. Albersheim, "A Closed-Form Approximation to Robertson's Detection Characteristics," *Proceedings of the IEEE*, Vol. 69, No. 7, July 1981.

[7] A. Saksena, L. Benmohamed, J. Dunne, D. Lucarelli, I-J. Wang, "Improving System-wide Detection Performance for Sonar Buoy Networks using In-Network Fusion," *MILCOM 2007 Unclassified Proceedings*, Orlando, FL.

[8] A. C. Coon, "Spatial coorelation of detections for impulsive echo ranging sonar," *Johns Hopkins APL Technical Digest*, vol. 18, no.1, pp. 105-112, 1997.

[9] D. P. Bertsekas and D. A. Castañon, "Rollout algorithms for stochastic scheduling problems," *Journal of Heuristics*, vol. 5, pp. 89–108, 1999.

[10] R. Mahler, "Multi-target Bayes filtering via first-order multi-target moments," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 39, no. 4, pp.1152–1178, 2003.

[11] B.-N. Vo and W.-K. Ma, "The Gaussian mixture probability hypothesis denisty filter," *IEEE Transactions on Signal Processing*, vol. 54, no.11, pp. 4091–4104, 2006.