Assessing Performance Tradeoffs in Undersea Distributed Sensor Networks

Thomas A. Wettergren, Russell Costa, John G. Baylog, and Sandie P. Grage
Naval Undersea Warfare Center
1176 Howell Street
Newport, RI 02841
Email: t.a.wettergren@ieee.org

***Abstract*— We consider the problem of determining the correct set of sensors to employ in the design of large area undersea surveillance sensor networks. As sensor technologies evolve, such networks are becoming increasingly practical. In turn, optimal selection of the number and type of sensors to deploy becomes an increasingly nontrivial process. Choices of field level detection and false alarm performance, as well as cost, all enter into this tradeoff decision space. In particular, the multiobjective nature of the problem leads to families of "optimal" solutions that each correspond to different tradeoffs between these often conflicting objectives. In this paper, we address these tradeoffs using a simple model of multi-sensor search performance and show the tradeoffs as Pareto efficient sets of solutions that satisfy system constraints. We also provide a means to determine the specific characteristics of the systems that lead to different design choices and explain how these designs perform as comprehensive search systems.**

I. Introduction

The general goal in deploying large distributed sensor networks is to detect and/or track a target that has entered the surveillance region [1]. Design choices often begin with deployment of as many sensors as is feasible within the surveillance region [2]. However, in the undersea domain, sensors are necessarily complicated, finely-tuned engineered devices. If these are sensors of the large gain variety, they become individually costly and a highly populated field deployment becomes an expensive venture [3]. Unfortunately, cost concerns are often evaluated as an afterthought to a design in the form of an ancillary objective or a limiting constraint. We thus explicitly include cost regulation as an a priori design objective. To avoid the high costs of large range sensors (such as with large acoustic arrays), an alternative is to deploy more numerous but less expensive shorter range sensors. However, such sensor characteristics can often lead to unnecessarily large numbers of false alarms; the regulation of which is often another initially unstated design objective. We thus include field level false alarm performance as an explicit system design objective in our sensor field construction. Managing the tradeoff between detection performance, false alarm performance, and cost is fundamental to the evaluation of appropriate tradeoffs between sensor characteristics and numbers. The mapping of this tradeoff space is a multiobjective design optimization problem [4] whose solution yields a Pareto tradeoff curve.

To perform the multiobjective optimization of sensor characteristics, we employ analytical models of system performance that connect the design parameters of sensor numbers, sensor detection threshold, and fusion strategy to the design objectives of maximal detection performance, minimal false alarms, and minimal cost. The field level performance of distributed sensor networks involves more than the concatenation of numerous individual sensor detection decisions, specifically, it involves the examination of multiple sensor detections that all originate from the same target over a fixed interval of time. We refer to this process as track-before-detect (see [5] for a description), since the final determination of a target presence is not made until multiple sensor detections occur and are kinematically consistent with target motion (the track). Thus, the process is more of a process of search (searching for the combination of detections that are consistent with a target track) rather than mere detection, and therefore we refer to the field level detection performance by the terms *probability of successful search* P_{SS} and *probability of false search* P_{FS}. Successful search is an important objective since it is the tactical purpose for deploying the sensor network. Unfortunately, false search often becomes an equally important objective since false searches lead to the expenditure of valuable assets in prosecuting each false search result. Furthermore, cost is always an objective, and becomes especially important in cases where the sensor network is deployed and not retrieved.

In the next section, we review the derivation of expressions for probability of successful search and probability of false search for modeling the track-before-detect process. We then describe a numerical multiobjective optimizer and use it to identify tradeoff surfaces that show how different choices of system parameters lead to different implied preferences amongst the objectives. In all cases, we restrict our attention to homogeneous fields of sensors that are spatially distributed in a uniform random manner (randomly sampled from a uniform distribution).

II. Sensor Network Performance Models

We model the process of track-before-detect by considering the interaction of m sensors with a single target traveling within the bounded search region $\mathcal{S} \subset R^2$. We assume the sensor processing and target characteristics are known to an extent such that the target detection process for a fixed sensor-target geometry is well-understood. In order to have a sensor detect a target traveling at speed v and constant

1-4244-0115-1/06/$20.00 ©2006 IEEE

heading over a period of time Δt, the target must come within the detection radius R_d of the sensor and stay within that range for a period of time δt (corresponding to the sensor's detection processing). This creates an effective sensor detection range of $r_d^2 = \max\{0,\ R_d^2 - (v\,\delta t/2)^2\}$ which is the instantaneous range required between the sensor and target for a detection opportunity to occur. Even when a detection opportunity occurs, there is still a non-zero probability that the target will not be detected. We model this probabilistically by a probability of detection p_d (with $0 < p_d < 1$) defined as the conditional probability that a sensor that falls within instantaneous range r_d reports a detection.

Assume there are m identical sensors which are deployed uniformly (uniform random) over the search region $\mathcal{S}$ of area a_0 (such that $\int_{\mathcal{S}} d\mathbf{x} = a_0$). The probability of having k of the m sensors report detections over the target time Δt is modeled as the result of a spatial Poisson process [6] with density $p_d\phi$. Based on the Poisson model, the probability of a successful search (defined as the occurrence of *at least* k detections) is given by

$$P_{SS} = 1 - \exp(-p_d m\phi) \sum_{j=0}^{k-1} \frac{(p_d m\phi)^j}{j!} \tag{1}$$

where

$$\phi = \frac{2 r_d v \Delta t + \pi r_d^2}{a_0} \tag{2}$$

represents the fractional measure of the region of detectability of a target relative to the entire search region $\mathcal{S}$. This detection performance model (1) explicitly depends on the target time interval Δt (via ϕ) to account for scenarios when the sensor detection regions do not overlap (i.e. when there are holes in the instantaneous coverage) yet the target motion allows for the multiple sensor detections to occur over the extended time interval.

We similarly consider the occurrence of *at least* k false alarms that are kinematically consistent with target motion as a false search. We do not consider the case of mixtures of real and false detections providing a false search. In this context, that implies the false alarms must occur within a restricted geometric region that is consistent with the detectable region of the target, that is, a region of size defined by the area fraction ϕ. For false alarms that occur independently at rate FAR, the probability of any sensor reporting a false alarm during an arbitrary target time interval of length Δt is given by $p_{fa} = 1 - \exp(-FAR \cdot \Delta t) \approx FAR \cdot \Delta t$ where the approximation holds for small arguments of the exponential. These sensor false alarm events occur as a set of independent probabilistic events with a similar spatial density as in the successful search expression, however, the location of the region of target detectability is now of arbitrary orientation, leading to a probability of false search of

$$P_{FS} = 1 - \exp(-\pi p_{fa} m) \left[\sum_{j=0}^{k-1} \frac{(p_{fa} m\phi)^j}{j!} \right]^{\pi/\phi} \tag{3}$$

which has been validated with comparisons to Monte Carlo simulations. Equations (1) and (3) represent the probability of successful search and probability of false search associated with uniformly distributed sets of sensors in a multi-sensor target detection strategy.

Design goals of increasing successful search while decreasing false search can obviously be met by increasing numbers of sensors m and/or detection range R_d while decreasing the false alarm rate FAR. The reason why this strategy is not employed in practice is due to the practical limitation of cost. In [3], a simple cost model of passive acoustic sensor nodes for undersea sensor networks was developed for cost-effective coverage studies. We recreate that cost here for providing constraints for tradeoff studies. Specifically, each sensor node is modeled as a small coherent array (for fixed false alarm rate FAR and detection probability p_d) with cost given by a fixed (overhead) cost and a per-channel cost, specifically

$$C_{NODE}(R_d) = F_C + C_{CH} M_{CH}(R_d) \tag{4}$$

In this model, $M_{CH}(R_d) = \alpha R_d^2$ is the cost factor for achieving a certain range limit R_d (assume each sensor node is short range) leading to the cost model of

$$C_{NODE}(R_d) = c_0 + c_1 R_d^2 \tag{5}$$

for fixed coefficients c_0 and c_1. Values of $c_0/c_1 \approx 20{,}000\,\mathrm{m}^2$ have been found [3] to be consistent with deployable autonomous passive acoustic sensor nodes. Given this nodal cost, the cost of the field of m sensors is given by

$$\begin{aligned} C_{FIELD}(m, r_d) &= m \cdot C_{NODE}(R_d) \\ &= m\,c_0 + m\,c_1 \left(r_d^2 + (v\,\delta t/2)^2\right) \end{aligned} \tag{6}$$

where we have written the expression in terms of the effective sensor range r_d in favor of the original R_d as described above. The objective functions provided in equations (1), (3) and (6) provide analytical representations of the general field design goals. Specifically, these are functions of the fixed scenario parameters $\{a_0,\ v,\ \delta t,\ \Delta t\}$ and the adjustable sensor field parameters $\{p_d,\ p_{fa},\ r_d,\ m,\ k\}$ which must be chosen to optimize the various objectives.

III. Numerical Determination of Tradeoffs

To numerically determine the tradeoffs between objectives, we treat the problem as one of constrained multiobjective optimization. For instance, we may want to achieve simultaneously some or all of the following simultaneous goals: maximize P_{SS}, minimize P_{FS}, or minimize C_{FIELD}. Such problems are historically treated as cost/benefit analysis problems, where the mathematical goal is to find the "knee" in a performance tradeoff curve between conflicting objectives. The tradeoff is represented mathematically by the notion of Pareto optimality or non-dominance of the objective vector. Specificaly, a Pareto optimal solution is one where no single objective can be improved without deteriorating at least one of the others. Thus, the set of all Pareto optimal points (or Pareto set) provides an accurate representation of the tradeoff among the conflicting objectives.

A. Pareto Set Generation Method

Many different approaches to generating numerical approximations to Pareto sets are available (see [4] for an overview). Most approaches fall into two categories: gradient-based methods that rely on solving many single-objective problems based on different objective combinations and evolutionary methods that iteratively modify a group of designs to create a "better" group of designs which approximate the Pareto set. The former methods suffer from issues of local convergence (due to the complexities of the combined objectives), where the latter suffer from sparse representations of the Pareto set. Either issue makes it difficult to generate results from which to infer design and employment guidance. In order to overcome the limitations of prior methods, we developed a new method of multiobjective optimization, named Genetic Algorithm-based Normal Boundary Intersection (GANBI) [8]. GANBI employs features of gradient based methods to combine the objectives but does so in an evolutionary manner that attempts to iteratively modify a group of designs to approach the Pareto set.

The approach that the GANBI method takes to multiobjective optimization is based on a genetic modification of the normal boundary intersection method of Das and Dennis [9]. The algorithm functions as follows. First, we optimize each objective independent of the others, but still under the overall system constraints. Let X represent a choice of parameter values, and let $f_j(X)$ represent the j-th objective evaluated at those parameter values. Assume, without loss of generality, that $f_j(X)$ is a minimization objective (if not, replace $f_j(X) \mapsto -f_j(X)$). Then define f_j^* as the optimal value of the j-th objective when the optimization is performed independent of the other objectives, and let X_j^* represent the corresponding set of parameter values; therefore $f_j^* = f_j(X_j^*)$. The convex hull formed by combining all of these individual objective optimizations (ie: a line segment in 2D, a triangle in 3D, etc.) is a very crude approximation to the Pareto surface. This convex hull is mathematically represented as

$$CH = \Phi B + [f_1^*,\ f_2^*,\ \ldots]^T \tag{7}$$

where Φ is the matrix whose (i,j)-th component is given by $\Phi_{i,j} = f_i(X_j^*) - f_i(X_i^*)$ and

$$B = [b_1,\ b_2,\ \ldots]^T,\ \sum b_i = 1,\ b_i \geq 0 \tag{8}$$

This set CH represents all the possible convex combinations of the individual minima.

We form a set of normal lines to this convex hull, and seek to find "designs" (that is, sets of parameters) that translate to points in the objective space along these normal lines. Of the designs that lie on (or very close to) the normal lines, we seek those that lie furthest from the convex hull of individual minima while still meeting all of the constraints. This concept is illustrated graphically in Fig. 1 for a two-objective minimization problem. Here the arrows are the normals and the goal is to move as far along each arrow as possible. Mathematically, for a design point X with multiobjective value

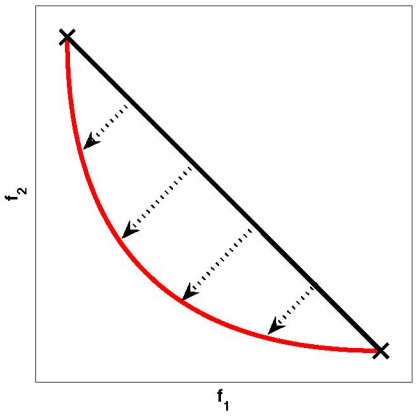

Fig. 1. Graphical description of the GANBI solution technique showing Pareto set (red), indvidual minima (x), convex hull (black), and normal lines (dotted arrows)

$F = [f_1(X),\ f_2(X),\ \ldots]^T$, we form the distance to the normal lines as

$$d_i = \left| \left(I - A(A^TA)^{-1}A^T\right)\left(F^* - F - \Phi B_i\right) \right| \tag{9}$$

where $A = \Phi U$ (with U as a vector of all ones) and $|\cdot|$ represents the L_2 norm. Note that B_i is the specific B-value that corresponds to the normal line under consideration. The distance from CH measured along the normal is correspondingly given by

$$z_i = (A^TA)^{-1}A^T\left(F^* - F - \Phi B_i\right) \tag{10}$$

and these two measures are combined according to $h_i = d_i - 2z_i$. This gives a new objective h_i for each design point X that measures how far its objective values are away from the convex hull combined with distance to the i-th normal line (the weighting value of 2 in the h_i definition was determined empirically to speed convergence). All of the h_i objectives have the goal of minimization (effectively making d_i small and z_i large), leading to a multiobjective minimization problem. The solution with this new set of objectives $\{h_i(X)\}$ tends to spread points along the Pareto set approximation because the normal lines are naturally spread along the convex hull.

We simultaneously apply each design to all of the GANBI normals using a multiobjective genetic approach. In this manner, each of the arrows (indexed by i) in Fig. 1 represents a new objective h_i in some abstract objective space. The objective space $\{h_i\}$ is then explored in a multiobjective genetic sense (using existing solution techniques) to arrive at a set of solutions that span the Pareto front in this abstract objective space, which become (by default) very well-spread in the true objective space. In this manner, the GANBI method behaves as a preprocessor for other multiobjective solvers. We have examined the benefits and computational costs of this new method with many genetic multiobjective approaches, and we have concluded that Deb's NSGA method [7] provides the best performance [8].

When applying GANBI to a particular problem, a set of parameters X refers to a design. At each optimization iteration, a set of designs is formed and each of their locations in the GANBI objective space $\{h_i\}$ is obtained. These objective and parameter values are fed to the NSGA multiobjective optimization solver, and a new set of designs is generated according to the rules of the NSGA algorithm. Ideally, the iterates would repeat until convergence; in practice, a large number of iterations is run until very minor changes occur between iterations. When constraints are included in the problem description, we use a penalty method [10] to artificially increase the objective values $\{h_i\}$ of designs that violate the constraints. The magnitude of this penalty is tied to the degree of constraint violation.

B. Numerical Results

For the distributed sensor design problem, there are three objectives of interest: maximize P_{SS}, minimize P_{FS}, and minimize C_{FIELD}. While numerically tractable, it is generally difficult to visualize the resulting Pareto set for three-objective problems like this. In particular, it is difficult to read-off specific values of the individual objectives at points along the surface that defines the Pareto set. For that reason, we effectively remove the cost objective by replacing it with a cost constraint of the form $C_{FIELD} \leq C_{max}$ and make repeated two-objective solutions for various chosen values of C_{max}. This process is similar to the ε-constraint method [4] of nonlinear multiobjective optimization; the difference here is that we only apply the constraint to a single objective and are still left with a Pareto tradeoff problem in two-objective space. This Pareto tradeoff problem is restated in minimization form as follows:

$$\min_{X} \begin{bmatrix} 1 - P_{SS}(X) \\ P_{FS}(X) \end{bmatrix}, \text{ s.t. } C_{FIELD}(X) \leq C_{max} \quad (11)$$

where X is the set of adjustable design parameters.

In addition to design parameters, there are fixed scenario parameters that define the specific application of the sensor network. If design robustness to scenario variations is desired (such that a single sensor choice performs well across the set of expected scenarios), then the entire objective evaluation may be marginalized with respect to a statistical prior distributed for these values. This assessment of robust designs is a subject of future study; for this example, we consider a single fixed scenario.

To avoid dealing with the complexities associated with a variety of processing choices, we also fix the values of the parameters $\{p_d, p_{fa}\}$ and assume the resulting r_d values are appropriately scaled to match the fixed values. For our example, we use the following fixed parameter values: $v =$ 5kts, $\delta t =$ 40sec, $\Delta t =$ 30min, $a_0 =$ 100nmi^2, $c_0 =$ \$120, $c_1 =$ \$0.006/m^2, $p_d = 0.9$, and $p_{fa} = 0.044$. Note that the value of p_{fa} corresponds to a probability of false alarm of 10^{-3} over the sensor integration time δt. The set of adjustable design parameters consists of sensor range r_d, number of sensors m, and number of required detections k. The achievable values of these adjustable design parameters are bounded according to $0 \leq r_d \leq 10^4$m, $1 \leq m \leq 10^5$, and $k \in \{1, 2, 3, 4\}$. Such bounds are applied to avoid impractical designs (such as having too many sensors to deploy).

We first consider the design of a sensor field to cover the search area of size a_0 subject to a cost constraint of $C_{max} =$ \$1,000,000. To develop the tradeoff between maximizing P_{SS} and minimizing P_{FS}, problem (11) was solved using the GANBI approach for the parameters $X = \{r_d, m, k\}$. We ran 200 iterations of the genetic algorithm with 4 GANBI normals using 100 individual designs in the population at each iteration. The total number of designs investigated (2×10^4) represents a very small sampling of the design space. With the parameter set X represented in the genetic algorithm by a 32-bit binary string, the fraction of the design space sampled was approximately 4×10^{-6}. Thus, the GANBI method provided a very efficient sampling of the design space.

Figure 2 shows the resulting approximate Pareto set computed for this case. Rather than showing the minimization objectives of equation (11) directly, the plot shows performance of the sensor field in a manner similar to a ROC curve. The vertical axis represents the logarithm of the probability of false search converted to probability of occurrence per day. Specifically, it is given by

$$\log_{10}\left(1 - (1 - P_{FS})^{86400/\Delta t}\right) \quad (12)$$

which is a monotonic function of P_{FS}, thus the goal of minimizing P_{FS} corresponds to values lower on the plot. The horizontal axis represents the probability of successful search, and the objective of minimizing $(1 - P_{SS})$ corresponds to values on the right of the plot. Thus, the multiobjective design goal corresponds to the lower right of the plot. The point on the upper left is included since many trivial designs lead to regular false searches with no successful searches (i.e. use a sparse set of few short-range sensors with $k = 1$). The results show a clear optimal tradeoff between designs with good search effectiveness and many false searches (upper right of plot) and designs with poor search effectiveness but less false searches (lower left of plot). In table I, the specific parameter values of the three labeled points along the tradeoff are identified. From these values, it is clear that the designs corresponding to the Pareto set all have approximately the same sensor range ($r_d \approx$ 300m), and the number of sensors used provides the variation along the curve (more sensors leads to higher P_{SS} and P_{FS} whereas fewer sensors leads to lower P_{SS} and P_{FS}). This trend shows that an optimal sensor can be designed and the number employed in the field can be chosen later to meet the desired performance tradeoff. The number of required detections k varies somewhat sporadically along the tradeoff curve because of the sparsity of the resulting sensor fields.

We next consider a scenario whose only difference with the previous example is in the cost constraint. Specifically, we relax the cost restriction to $C_{max} =$ \$5,000,000 in order to examine the impact of more densely populated fields.

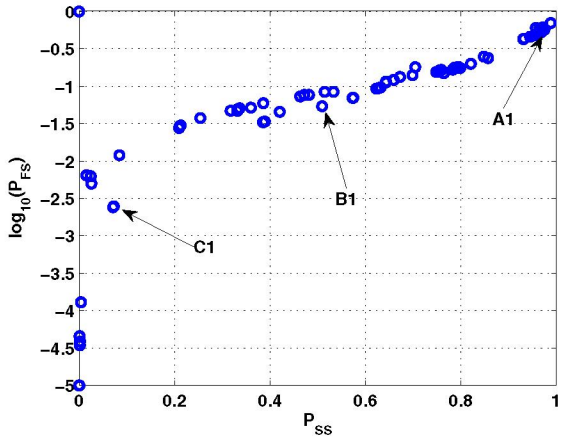

Fig. 2. Approximate Pareto set for $C_{max} = \$1,000,000$.

TABLE I
DESIGN VALUES FOR POINTS CALLED OUT IN FIG. 2.

Point	k	m	r_d
A1	2	985	313 m
B1	1	396	299 m
C1	3	131	314 m

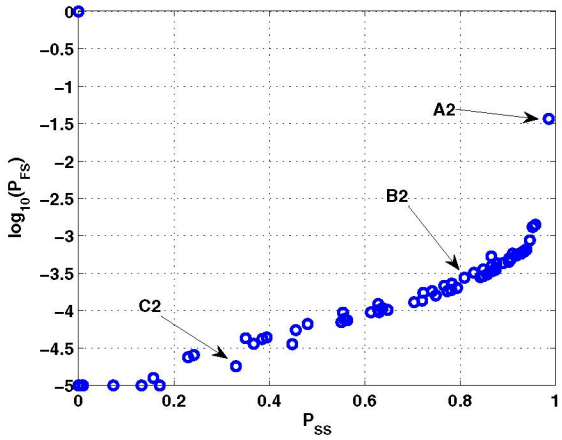

Fig. 3. Approximate Pareto set for $C_{max} = \$5,000,000$.

TABLE II
DESIGN VALUES FOR POINTS CALLED OUT IN FIG. 3.

Point	k	m	r_d
A2	3	241	1312 m
B2	3	8	9529 m
C2	3	3	9914 m

The resulting Pareto set approximation is shown in Fig. 3. Obviously, the relaxed cost constraint allows the optimal tradeoff to be pushed closer to the desired goal of the bottom right. However, once again the results show a clear optimal tradeoff between designs with good search effectiveness and many false searches (upper right of plot) and designs with poor search effectiveness but less false searches (lower left of plot). Table II gives the parameter values corresponding to the three labeled points on the tradeoff. The nearly perfect search effectiveness of point A2 is achieved by using many sensors of shorter range (albeit a longer range than found in the more cost-restrictive case), leading to necessarily larger probability of false search. As the sensor range increases (point B2), fewer sensors are needed – leading to less chance of a false search at a cost of slightly lower search effectiveness. Finally, the point C2 shows a very low search effectiveness obtained by reducing the number of sensors m to the number of detections required k, although the false search almost never occurs. We note that this point occurs at the edge of the design space (recall that r_d is limited to $r_d \leq 10^4$m. Once again, the variation of the number of required detections k has little impact on the resulting tradeoff, although larger values of k are required for these densely populated sensor fields.

In general, these numerical results show a trend that has been an intuitive notion in sensor field design: larger numbers of shorter range sensors used to obtain large P_{SS} with large P_{FS} versus smaller numbers of larger range sensors used to obtain lower P_{FS} with low P_{SS}. In these examples, we showed that this trend is an optimal tradeoff only when the field cost constraint is not too restrictive. For very restrictive cost constraints (and hence, very sparse fields), the optimal tradeoff is achieved by simply varying the number of sensors employed with each sensor reflecting a single optimal detection range. Thus, the specific value of the constraint employed must be taken into consideration when making pre-employment design choices.

IV. CONCLUSION

We examined the impact of system design parameters on multiple objective design tradeoffs for undersea distributed sensor networks. Specifically, we examined how the scale of the field (given by range of each sensor and number of sensors employed) as well as the number of multi-sensor detections to use in track-before-detect affect the performance of a sensor network. Performance was measured in terms of search effectiveness, false searches, and cost. In general, optimal tradeoffs for a fixed cost constraint lead to design choices that span from good search effectiveness at the expense of more false searches to poorer search effectiveness with minimal false searches. The decision of where to operate on this tradeoff curve is left to the operator and/or decision-maker.

If cost constraints are very strict, it was shown that an optimal sensor size may be determined that is independent of the search performance tradeoff goal, whereas for less-restrictive cost constraints, the optimal size depends on desired search/false-search effectiveness. This result informs the intuitive notion of trading off between many short-range sensors and few long-range sensors by showing how overall cost constraints may prohibit such tradeoffs from being optimal.

The tradeoffs resulting from these analyses may be used to provide insight into both design and employment decision-making. Furthermore, adaptive employment decisions based on observed performance become more informed when they are based on optimal tradeoffs between design goals. Future extensions of this work include the addition of the complexities associated with nonuniform sensor placement as well as those due to nonuniform environmental characteristics.

ACKNOWLEDGMENT

This work was supported by the Office of Naval Research Code 321MS and by the In-House Laboratory Independent Research Program of the Naval Undersea Warfare Center.

REFERENCES

[1] D. Culler, D. Estrin and M. Srivastava, "Overview of Sensor Networks," *IEEE Computer Magazine*, pp. 41–49, August 2004.

[2] Y. Zou and K. Chakrabarty, "Uncertainty-Aware Sensor Deployment Algorithms for Surveillance Applications," *Proc. IEEE GLOBECOM*, pp. 2972–2976, 2003.

[3] C. M. Traweek and T. A. Wettergren, "Efficient Sensor Characteristic Selection for Cost-Effective Distributed Sensor Networks," *IEEE Journal of Oceanic Engineering*, to appear.

[4] K. M. Miettinen, *Nonlinear Multiobjective Optimization*, Kluwer Academic Publishers, Norwell, MA, 1998.

[5] H. M. Shertukde and Y. Bar-Shalom, "Detection and Estimation for Multiple Targets with Two Omnidirectional Sensors in the Presence of False Measurements," *IEEE Trans. ASSP*, vol. 38, no. 5, pp. 749–763, 1990.

[6] A. Papoulis, *Probability, Random Variables, and Stochastic Processes, Third Edition*, McGraw-Hill, Boston, MA, 1991.

[7] N. Srinivas and K. Deb, "Multiobjective Function Optimization Using Nondominated Sorting Genetic Algorithms," *Evolutionary Computation*, vol. 2, no. 3, pp. 221–248, 1995.

[8] T. A. Wettergren, "The Genetic-Algorithm-Based Normal Boundary Intersection (GANBI) Method: An Efficient Approach to Pareto Multiobjective Optimization for Engineering Design, " NUWC-NPT Technical Report 11,741, May 2006.

[9] I. Das and J. E. Dennis, "Normal-Boundary Intersection: A New Method for Generating the Pareto Surface in Nonlinear Multicriteria Optimization Problems," *SIAM Journal of Optimization*, vol. 8, no. 3, pp. 631–657, 1998.

[10] D. E. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning*, Addison-Wesley, Boston, MA, 1989.