A Multiobjective Optimization Approach to Obtain Decision Thresholds for Distributed Detection in Wireless Sensor Networks

Engin Masazade, *Student Member, IEEE*, Ramesh Rajagopalan, *Student Member, IEEE*, Pramod K. Varshney, *Fellow, IEEE*, Chilukuri K. Mohan, *Senior Member, IEEE*, Gullu Kiziltas Sendur, *Member, IEEE*, and Mehmet Keskinoz, *Member, IEEE*

***Abstract*—For distributed detection in a wireless sensor network, sensors arrive at decisions about a specific event that are then sent to a central fusion center that makes global inference about the event. For such systems, the determination of the decision thresholds for local sensors is an essential task. In this paper, we study the distributed detection problem and evaluate the sensor thresholds by formulating and solving a multiobjective optimization problem, where the objectives are to minimize the probability of error and the total energy consumption of the network. The problem is investigated and solved for two types of fusion schemes: 1) parallel decision fusion and 2) serial decision fusion. The Pareto optimal solutions are obtained using two different multiobjective optimization techniques. The normal boundary intersection (NBI) method converts the multiobjective problem into a number of single objective-constrained subproblems, where each subproblem can be solved with appropriate optimization methods and nondominating sorting genetic algorithm-II (NSGA-II), which is a multiobjective evolutionary algorithm. In our simulations, NBI yielded better and evenly distributed Pareto optimal solutions in a shorter time as compared with NSGA-II. The simulation results show that, instead of only minimizing the probability of error, multiobjective optimization provides a number of design alternatives, which achieve significant energy savings at the cost of slightly increasing the best achievable decision error probability. The simulation results also show that the parallel fusion model achieves better error probability, but the serial fusion model is more efficient in terms of energy consumption.**

***Index Terms*—Distributed detection, multiobjective optimization, wireless sensor networks (WSNs).**

Manuscript received September 15, 2008; revised February 15, 2009 and April 29, 2009. First published August 11, 2009; current version published March 17, 2010. The work of E. Masazade was supported by The Scientific and Technological Research Council of Turkey (TUBITAK) under the Research Abroad Support Scheme. The work of M. Keskinoz was supported by TUBITAK under Grant 105E161. The work of R. Rajagopalan, P. K. Varshney, and C. K. Mohan was supported in part by the Air Force Office of Scientific Research under Grant FA-9550-06-1-0277. This paper was presented in part at the Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, CA, October 26–29, 2008. This paper was recommended by Associate Editor S. X. Yang.

E. Masazade, G. K. Sendur, and M. Keskinoz are with the Faculty of Engineering and Natural Sciences, Sabanci University, Istanbul 34956, Turkey (e-mail: enginm@su.sabanciuniv.edu; gkiziltas@sabanciuniv.edu; keskinoz@sabanciuniv.edu).

R. Rajagopalan is with the Department of Electrical and Computer Engineering, Florida State University, Panama City, FL 32405 USA (e-mail: rxrajagopalan@fsu.edu).

P. K. Varshney and C. K. Mohan are with the Department of Electrical Engineering and Computer Science, Syracuse University, Syracuse, NY 13244 USA (e-mail:varshney@syr.edu; ckmohan@syr.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSMCB.2009.2026633

I. Introduction

WIRELESS sensor networks (WSNs) typically consist of a large number of densely deployed sensor nodes that cooperatively monitor parameters or events of interest, such as the temperature or velocity of an object or a break in by an intruder. Sensors transmit raw measurements or their processed versions to a central fusion center that performs a final inference about the underlying parameters or events. WSNs are currently used in a wide range of application areas, such as battlefield security or surveillance, environment monitoring, health monitoring, and disaster relief operations. In this paper, we study the detection problem, where the objective of the WSN is to distinguish between two hypotheses, such as the absence (Hypothesis 0) or presence (Hypothesis 1) of a certain event. Such detection ability of a WSN is crucial for various applications. As an example, in a surveillance scenario, the presence or absence of a target is usually determined before attributes, such as its position or velocity, are estimated [1].

Since the sensors assumed here are tiny battery-powered devices, they suffer from several constraints, such as severe energy, computation, and storage limitations. The transmission of raw measurements to the fusion center incurs excessive energy and bandwidth consumption. In distributed detection, by taking advantage of the limited onboard signal processing capabilities of sensors, the measurements are first preprocessed, and a quantized version of the decision statistic is sent to the fusion center. For binary quantization and under different performance criteria [Bayes, Neyman–Pearson (NP)], the design of the optimal fusion rule is relatively straightforward, but the evaluation of the decision thresholds at peripheral sensors is more complicated as a result of the distributed nature of the WSN. Therefore, obtaining local sensor decision rules is a major issue in the distributed detection problem [2].

For a given number of sensors and under the assumption of conditionally independent observations, the optimal decision rule at each sensor reduces to a likelihood ratio test (LRT) [2] for both Bayesian and NP criteria and for different decision fusion topologies, such as parallel or serial. In parallel decision fusion, each sensor directly sends its decision to the fusion center. In serial decision fusion, all the sensors are connected in series. The routing path defines how these sensors are interconnected, and in this paper, we assume that it is known in advance. In the serial case, each sensor generates its decision by combining the decision coming from its predecessor with its own measurement. Then, the decision of the last sensor on

the path is accepted as the final inference. Under both decision fusion schemes, the LRTs at each sensor are coupled with other sensor decisions and the fusion rule. Optimal values of the local sensor thresholds are typically found using person-by-person optimization (PBPO) [2], where each sensor threshold is iteratively optimized by assuming a fixed fusion rule and decision rules at the other sensors. In the asymptotic regime, where the number of sensors is very large, an identical decision rule for all the sensors is asymptotically optimal [3]. This result considerably simplifies the design of decision rules.

In this paper, we assume ideal channels between the sensors and the fusion center (for recent work involving nonideal channels, see [4]–[6]). Under the NP criterion and considering fading channels between sensors and the fusion center, an exhaustive search has been employed in [5] over all threshold selections to determine their optimal values. The computational complexity of such an approach exponentially increases with the number of local sensors, and this approach for finding the optimal sensor thresholds is only practical with relatively few sensors. In this paper, we assume ideal channels, where the fusion center receives the sensor decisions without any error. This requires that each sensor decision be transmitted with sufficient energy, which is a function of the distance between the sensor and the fusion center [7]. Each sensor arrives at a binary decision about the event by comparing its decision statistic with a threshold. We assume that if the sensor positively decides about the presence of the event, it transmits one bit; otherwise, it stays silent. Therefore, the thresholds of the local sensors determine not only the network's probability of error but also the total energy consumption. A recent work [8] considers the design of local sensor decision rules that minimize the probability of error subject to a transmission rate constraint for each sensor. Under conditionally independent observations, a constrained minimization problem is defined, and the optimal thresholds are obtained using the well-known PBPO procedure. Although conditional independence assumption simplifies the derivation of decision rules, it may not be valid in many realistic cases. In this paper, we consider the case where the event has an isotropic signal emission with path loss [9], [10]. Then, in the presence of the event, each sensor's received measurement from the event depends on the distance between the sensor and the event location. Each noisy sensor measurement then follows the same probability distribution with different means as long as the measurement noise is independent and identically distributed across sensors. The sensors in proximity of the event are more likely to decide on the presence of the event. In other words, an isotropic signal source for the event implies a high degree of spatial correlation. A related work [11] proposes a collaborative detection scheme, where a sensor close to the signal source requests collaboration and receives the decisions of the $K_{\max}$ sensors within the neighborhood. The authors showed that increasing $K_{\max}$, namely, including more sensors to the collaboration that are located far from the event, considerably degrades the detection performance. Moreover, if the location of the event can only be described in terms of its probability density function (pdf), then the received sensor decisions are no longer conditionally independent because of the unknown event location. Then, the optimality of LRTs for local sensor decision-making fails, and the derivation of optimal sensor decision rules becomes complicated.

Sensor network design usually involves simultaneous consideration of multiple conflicting objectives [5], [12], such as maximizing the lifetime of the network or maximizing the detection capability, while minimizing the transmission costs. In a conventional WSN setting, one of the desired objectives is optimized while treating others as constraints of the problem or the problem is converted into a single objective problem by assigning weights to each objective function. In the constrained minimization case, one single solution is obtained based on available resource limitations and the solution has to be reevaluated each time when the amount of resource is changed. In the weighted sum approach, relative weights of the objectives are usually not known or difficult to determine. These drawbacks can be overcome via multiobjective optimization methods [14]–[20] which optimize all the objectives simultaneously and generate a set of solutions at the same time reflecting different tradeoffs between the objectives. Multiobjective optimization has recently been introduced for WSN design [21] where the mobile agent routing and sensor placement problems and the tradeoff solutions between the desired objectives were determined through the use of multiobjective evolutionary algorithms.

In this paper, we study the event detection problem for sensor networks under isotropic signal emission, and the event location is only known in terms of its pdf. In addition, we assume that the sensors employ the on–off keying scheme, where they send one bit of data to the fusion center only if they decide on the presence of the event. Then, sensor decision thresholds determine not only the probability of error but also the total energy consumption of the network. Therefore, instead of having a single solution that minimizes the probability of error of the network, by using the multiobjective optimization approach, we seek several sensor threshold sets that deliver a significant energy saving as compared with the energy consumption of the minimum probability of error solution without sacrificing the probability of error too much. Thus, we are able to obtain a set of solutions that provide tradeoffs between energy consumption and probability of error performance.

Thus, we formulate a multiobjective optimization problem (MOP) with two objectives, minimizing the probability of error at the fusion center (global probability of error) P_e and minimizing the total network energy consumption (global energy consumption) E_T, where the sensor decision thresholds are selected as the variables of the MOP. We solve the MOP and generate the Pareto optimal solutions between these two conflicting objectives through normal boundary intersection (NBI) [14] and nondominating sorting genetic algorithm-II (NSGA-II) [18]. In our prior work [22], we only studied the problem for parallel decision fusion, where each sensor performed binary quantization by comparing its measurement with its threshold. The sensor decisions were then directly sent to the fusion center. In this paper, we extend the work reported in [22] and compare the results of parallel decision fusion with serial decision fusion. In the serial case, it is hard to evaluate the optimal decision rule of each sensor since the event location is only known in terms of its pdf. The simulation results show that when each sensor makes its decision based on the decision of its predecessor and its own observation, the performance is poor if the sensor is very far away from the event location. For this reason, motivated by the counting rule considered in [13], we use a heuristic decision rule at each sensor. Our

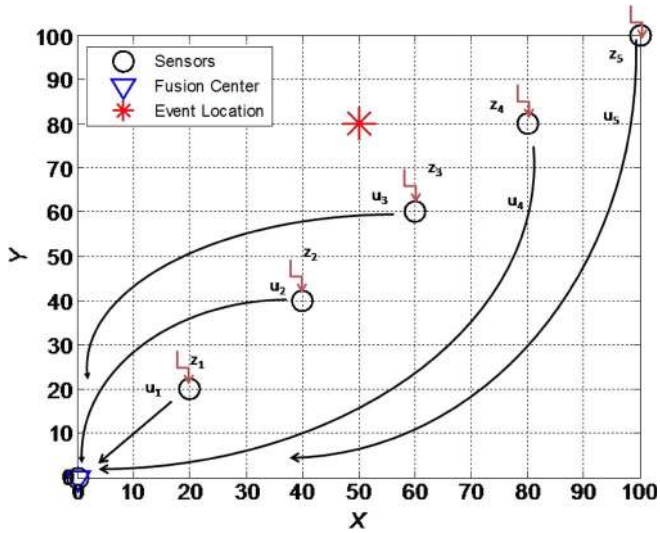

Fig. 1. WSN model with parallel decision fusion.

decision statistic used for the serial case is the aggregation of sensor decisions from all the previous sensors and its own observation. In this paper, we also compare the multiobjective optimization methods NBI and NSGA-II in detail by using different performance metrics [19]. Finally, we compare the performance of the network both for different and identical sensor thresholds employed at each sensor.

The rest of this paper is organized as follows: In Section II, we state the WSN assumptions and describe each objective function under both parallel and serial decision fusion schemes. In Section III, we briefly review the fundamentals of MOP and describe the NBI and NSGA-II methods. In Section IV, we present our simulation results, and finally, we devote Section V to the conclusions.

II. Problem Definition

In this section, we first state the WSN assumptions. Then, we define the mathematical models for both objective functions for parallel and serial decision fusion topologies.

A. WSN Model and Statement of the MOP

A representative WSN consisting of N sensors $\{s_i,\ i = 1, 2, \ldots, N\}$ with parallel decision fusion is shown in Fig. 1. The distances between s_i and the fusion center and the event location (x, y) are denoted as $d_{f,i}$ and d_i, respectively. We assume the event location to be a random variable with an associated prior pdf, and therefore, d_i is a random variable.

Throughout this paper, we denote the probability mass function of discrete variables by $P(.)$ and the pdf of continuous variables by $p(.)$ or $p(.,.)$, depending on the number of random variables.

Specifically, we assume that the location of the event is uniformly distributed with joint pdf as

$$p(x, y) = \frac{1}{A \times B}, \qquad 0 \le x \le A, \quad 0 \le y \le B \tag{1}$$

where the region of interest (ROI) is an area of size $A \times B$. Other pdf's can be employed in a similar manner. The average distance of s_i located at (x_i, y_i) to the event location (x, y) is then expressed as

$$\bar{d}_i = \int_0^A \int_0^B \sqrt{(x - x_i)^2 + (y - y_i)^2}\, p(x, y)\, dy\, dx. \tag{2}$$

Suppose that a signal that follows the power attenuation model, such as an acoustic signal that is radiated from an event source with energy K_0 [9], and sensors $s_i,\ i = 1, 2, \ldots, N$, are deployed at positions $(x_i, y_i),\ i = 1, 2, \ldots, N$. Then, the received energy (e_i) observed at s_i is

$$e_i(x_i, y_i, x, y) = \frac{K_0}{1 + a\left(\sqrt{(x - x_i)^2 + (y - y_i)^2}\right)^n} \tag{3}$$

where n is the signal decay exponent, and a is an adjustable constant [9]. When $n = 2$, the energy of the event decays at a rate inversely proportional to the square of the distance $d_i = \sqrt{(x - x_i)^2 + (y - y_i)^2}$. Then, under each hypothesis, the received measurement of each sensor (z_i) can be expressed as

$$\begin{aligned} z_i &= n_i, \text{ under } H_0 \\ z_i &= \sqrt{e_i(x_i, y_i, x, y)} + n_i, \text{ under } H_1 \end{aligned} \tag{4}$$

where n_i is the measurement noise that follows the normal distribution at each sensor, and it is assumed to be independent across the sensors. z_i then follows a normal distribution with parameters

$$z_i \sim \begin{cases} N(0, \sigma^2), & \text{under } H_0 \\ N\left(\sqrt{e_i(x_i, y_i, x, y)}, \sigma^2\right), & \text{under } H_1 \end{cases}. \tag{5}$$

Throughout this paper, we assume that the noise variance $\sigma^2 = 1$. When z_i exceeds a certain threshold denoted as t_i, sensor s_i transmits a 1-bit decision $(u_i = 1)$ to the fusion center. Otherwise, it does not transmit anything.

The functions global probability of error P_e and global energy consumption E_T are functions of the local sensor thresholds t_i and constitute the objective functions of the MOP. The MOP considered here is formulated as

$$\min_{t_1, t_2, \ldots, t_N} \{P_e(t_1, t_2, \ldots, t_N), E_T(t_1, t_2, \ldots, t_N)\}, \quad t_{\min} \le t_i \le t_{\max}, \quad i \in \{1, 2, \ldots, N\}. \tag{6}$$

We first solve the above problem for N nonidentical decision thresholds $\{t_1, t_2, \ldots, t_N\}$ employed at each sensor. We also compare the performance of nonidentical decision thresholds with identical decision threshold at each sensor $\{t = t_1 = \cdots = t_N\}$ via simulation.

In the next sections, we derive the objective functions for the global probability of error and the global energy consumption under parallel and serial decision fusion models.

B. Parallel Decision Fusion

In this section, we derive mathematical expressions for the two objectives, namely, the global probability of error and the global energy consumption for parallel decision fusion.

1) Global Probability of Error: Let u_0 be the global decision at the fusion center about the presence or absence of an event, and let P_0 and P_1 be the *a priori* probabilities of H_0 and H_1, respectively. The global probability of error is given by [2]

$$P_e = P_0 P_F + P_1(1 - P_D) \tag{7}$$

where $P_F = P(u_0 = 1|H_0)$ denotes the global probability of false alarm, and $P_D = P(u_0 = 1|H_1)$ denotes the global probability of detection. Given the vector of local sensor decisions of size $1 \times N$, $\mathbf{u} = [\, u_1 \quad u_2 \quad \cdots \quad u_N \,]$, and $u_i \in \{0, 1\}$, the probability of error is expressed as

$$\begin{aligned} P_e &= P_0 P(u_0 = 1|H_0) + P_1 \left(1 - P(u_0 = 1|H_1)\right) \\ P_e &= P_1 + P(u_0 = 1|\mathbf{u}) \left[P_0 P(\mathbf{u}|H_0) - P_1 P(\mathbf{u}|H_1)\right]. \end{aligned} \tag{8}$$

P_e is minimized if

$$\begin{aligned} P(u_0 = 1|\mathbf{u}) &= 0 \text{ when } \left[P_0 P(\mathbf{u}|H_0) - P_1 P(\mathbf{u}|H_1)\right] > 0 \\ P(u_0 = 1|\mathbf{u}) &= 1 \text{ when } \left[P_0 P(\mathbf{u}|H_0) - P_1 P(\mathbf{u}|H_1)\right] < 0. \end{aligned} \tag{9}$$

The above property leads to the following LRT at the fusion center [2]:

$$\frac{P(\mathbf{u}|H_1)}{P(\mathbf{u}|H_0)} \underset{u_0=0}{\overset{u_0=1}{\gtrless}} \frac{P_0}{P_1}. \tag{10}$$

By conditioning P_F over each possible incoming vector of decisions $\mathbf{u}$ and then averaging over $\mathbf{u}$, P_F is expressed as

$$P_F = P(u_0 = 1|H_0) = \sum_{all\, \mathbf{u}} P(u_0 = 1|\mathbf{u}) P(\mathbf{u}|H_0) \tag{11}$$

where, according to the received decision vector $\mathbf{u}$, $P(u_0 = 1|\mathbf{u})$ is either 0 or 1 based on the fusion rule expressed in (10). Since the noise samples are assumed to be independent and identically distributed, we have

$$P(\mathbf{u}|H_0) = \prod_{i=1}^{N} P(u_i|H_0) \tag{12}$$

where the false alarm probability of an individual sensor $P_{F,i}$ is

$$P_{F,i} = P(u_i = 1|H_0) = Q(t_i) \tag{13}$$

and $P(u_i = 0|H_0) = 1 - P(u_i = 1|H_0)$, where $Q(.)$ is the complementary distribution function of the Gaussian defined as

$$Q(t_i) = \int_{t_i}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx. \tag{14}$$

Since the event location is random, $P(\mathbf{u}|H_1)$ cannot directly be written as the product of individual decisions, as in (12). Instead, the global probability of detection needs to first be conditioned on the location of the event and then be averaged over its pdf. For a given event location (x, y), the conditional global probability of detection CP_D is

$$\begin{aligned} CP_D &= P(u_0 = 1|x, y, H_1) \\ &= \sum_{all\, \mathbf{u}} P(u_0 = 1|\mathbf{u}) P(\mathbf{u}|x, y, H_1) \end{aligned} \tag{15}$$

and since noise distribution is independent across sensors as

$$P(\mathbf{u}|x, y, H_1) = \prod_{i=1}^{N} P(u_i|x, y, H_1) \tag{16}$$

where the conditional probability of detection of an individual sensor $CP_{D,i}$ under given event location (x, y) is expressed as

$$\begin{aligned} CP_{D,i} &= P(u_i = 1|x, y, H_1) \\ &= Q\left(t_i - \sqrt{e_i(x_i, y_i, x, y)}\right) \end{aligned} \tag{17}$$

and $P(u_i = 0|x, y, H_1) = 1 - P(u_i = 1|x, y, H_1)$. In addition, the error probability of an *individual sensor* $P_{\mathrm{ind},i}(t_i)$ as a function of its decision threshold t_i can be expressed as

$$\begin{aligned} P_{\mathrm{ind},i}(t_i) &= P_0 P(u_i = 1|H_0) \\ &\quad + P_1 \int_0^A \int_0^B P(u_i = 0|x, y, H_1) dy dx. \end{aligned} \tag{18}$$

The global detection probability P_D is found by averaging $CP_{D,i}$ over the pdf of the event location as

$$P_D = \int_0^A \int_0^B P(u_0 = 1|x, y, H_1) p(x, y) dy dx. \tag{19}$$

Our first objective function, i.e., the probability of error, is given by

$$\begin{aligned} &P_e(t_1, t_2, \ldots, t_N) \\ &= P_0 \sum_{all\, \mathbf{u}} P(u_0 = 1|\mathbf{u}) P(\mathbf{u}|H_0) \\ &+ P_1 \left[\int_0^A \int_0^B \sum_{all\, \mathbf{u}} P(u_0 = 0|\mathbf{u}) P(\mathbf{u}|x, y, H_1) p(x, y) dy dx\right]. \end{aligned} \tag{20}$$

2) Global Energy Consumption: In this paper, we employ an energy-efficient on–off keying scheme, where only the sensors that detect the event transmit their decision to the fusion center. We also assume that the transmitted local decisions are delivered to the fusion center without any error. Then, the energy consumption at sensor s_i for perfectly transmitting m bits to the fusion center over distance $d_{f,i}$ is defined as [7]

$$E_{\mathrm{TX}}(m, d_{f,i}) = E_{\mathrm{elec}} \times m + \epsilon_{\mathrm{amp}} \times m \times d_{f,i}^2 \text{ [in joules]}. \tag{21}$$

According to this model, a sensor dissipates $E_{\mathrm{elec}} = 50$ nJ/bit to run the transmitter circuitry and $\epsilon_{\mathrm{amp}} = 100$ pJ/bit/m^2 for the transmitter amplifier.

The energy consumption of the network is the total transmission energy of all single-bit decisions transmitted to the fusion center. In other words, in (21), m becomes 1 if $u_i = 1$, and m is 0 (no transmission) if $u_i = 0$. An *individual sensor*'s energy consumption can be expressed as

$$E_{\mathrm{ind},i}(t_i) = E(1, d_{f,i}) \left[P(u_i = 1|H_0) P_0 + P(u_i = 1|H_1) P_1\right]. \tag{22}$$

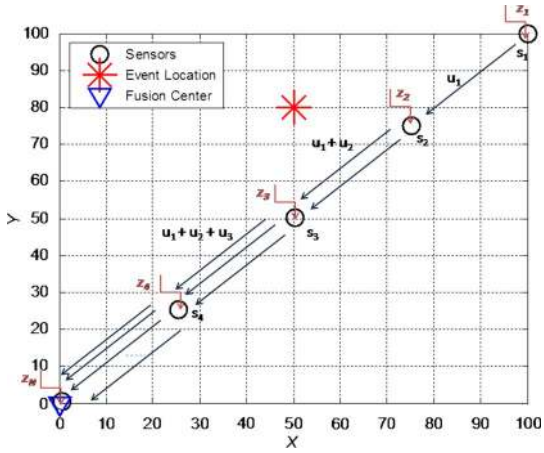

Fig. 2. WSN model with serial decision fusion.

The transmission of the decision vector $\mathbf{u}$ to the fusion center then requires

$$E_C(\mathbf{u}) = \sum_{i=1}^{N} E_{\mathrm{TX}}(u_i, d_{f,i}). \tag{23}$$

The energy consumption E_T of the network is then found by conditioning $E_C(\mathbf{u})$ on all possible vector of decisions as

$$\begin{aligned} &E_T(t_1, t_2, \ldots, t_N) \\ &\quad = \sum_{all\,\mathbf{u}} E_C(\mathbf{u}) P(\mathbf{u}) \\ &\quad = \sum_{all\,\mathbf{u}} E_C(\mathbf{u}) \left(P(\mathbf{u}|H_0) P_0 + P(\mathbf{u}|H_1) P_1 \right). \end{aligned} \tag{24}$$

Using the relation

$$P(\mathbf{u}|H_1) = \int_x \int_y \left[\prod_{i=1}^{N} P(u_i|x, y, H_1) \right] p(x, y) dy dx \tag{25}$$

together with (12) and (24), our second objective, i.e., global energy consumption, is obtained as

$$\begin{aligned} &E_T(t_1, t_2, \ldots, t_N) \\ &= \sum_{all\,\mathbf{u}} E_C(\mathbf{u}) \left[P_0 \prod_{i=1}^{N} P(u_i|H_0) \right. \\ &\qquad \left. + P_1 \int_x \int_y \left[\prod_{i=1}^{N} P(u_i|x, y, H_1) \right] p(x, y) dy dx \right]. \end{aligned} \tag{26}$$

C. Serial Decision Fusion

In this section, we derive mathematical expressions for the probability of error and the total energy consumption for serial decision fusion. In the serial fusion scheme, as described earlier and shown in Fig. 2, the decision of s_i is a function of its own measurement from the event z_i and the decisions of its predecessors $\mathbf{u}_{i-1} = [u_1, \ldots, u_{i-1}]$. The aggregate decision u_i is then forwarded to the successor sensor together with the past decisions $\mathbf{u}_{i-1}$. The last sensor in the serial configuration takes the final decision, which is a binary value that represents either of the two hypotheses. We assume that the routing path is known to all the sensors.

1) Probability of Error: In serial topology, the decision of the Nth sensor u_N is the decision of the entire WSN. Therefore, the probability of error is expressed as

$$\begin{aligned} P_e = P_0 P(u_N = 1|H_0) \\ + P_1 \int_x \int_y P(u_N = 0|x, y, H_1) p(x, y) dy dx \end{aligned} \tag{27}$$

where $P_F = P(u_N = 1|H_0)$ denotes the probability of false alarm, and $P_D = 1 - P(u_N = 0|H_1)$ denotes the probability of detection. To calculate these two quantities, the decision of s_N should be conditioned on both the received measurement z_N and the decisions of all its predecessors $\mathbf{u}_{N-1}$ as

$$\begin{aligned} &P_{F,N} \\ &= \sum_{all\,\mathbf{u}_{N-1}} \int_{z_N} P(u_N = 1|\mathbf{u}_{N-1}, z_N, H_0) P(\mathbf{u}_{N-1}, z_N|H_0) dz_N \\ &P_{D,N} \\ &= \int_x \int_y \left(\sum_{all\,\mathbf{u}_{N-1}} \int_{z_N} P(u_N = 1|\mathbf{u}_{N-1}, z_N, x, y, H_1) \right. \\ &\qquad \left. \times P(\mathbf{u}_{N-1}, z_N|x, y, H_1) dz_N \right) p(x, y) dy dx. \end{aligned} \tag{28}$$

At each sensor, we assume that the measurement is independent of the received incoming decisions, so their joint probabilities can be expressed according to

$$\begin{aligned} P(\mathbf{u}_{N-1}, z_N|H_0) &= P(\mathbf{u}_{N-1}|H_0) p(z_N|H_0) \\ P(\mathbf{u}_{N-1}, z_N|x, y, H_1) &= P(\mathbf{u}_{N-1}|x, y, H_1) p(z_N|x, y, H_1). \end{aligned} \tag{29}$$

For simplicity, we only show the derivation of the probability of false alarm. The calculation of the probability of detection is then quite straightforward except for an outer integration on the event location. Plugging (29) into (28) and using the fact that the sensor decision is independent of the underlying hypothesis, we obtain

$$\begin{aligned} P_{F,N} = \sum_{all\,\mathbf{u}_{N-1}} \left[\int_{z_N} P(u_N = 1|\mathbf{u}_{N-1}, z_N) p(z_N|H_0) dz_N \right] \\ \times P(\mathbf{u}_{N-1}|H_0). \end{aligned} \tag{30}$$

Given the event location (x, y) and the independent and identically distributed noise at each sensor, the optimum decision rule at s_i is an LRT that uses the decisions of the previous sensors $\mathbf{u}_{i-1}$ together with its own observation z_i. It is expressed as [2]

$$\begin{aligned} &P(u_i = 1|\mathbf{u}_{i-1}, z_i) = 0, \\ &\qquad \text{if } \frac{P(\mathbf{u}_{i-1}|H_1, x, y)}{P(\mathbf{u}_{i-1}|H_0)} \frac{p(z_i|H_1, x, y)}{p(z_i|H_0)} \le t_i \\ &P(u_i = 1|\mathbf{u}_{i-1}, z_i) = 1, \\ &\qquad \text{if } \frac{P(\mathbf{u}_{i-1}|H_1, x, y)}{P(\mathbf{u}_{i-1}|H_0)} \frac{p(z_i|H_1, x, y)}{p(z_i|H_0)} > t_i. \end{aligned} \tag{31}$$

When the location of the event is a random variable, the sensor measurements from the event become correlated, the

LRT shown in (31) is not necessarily optimal at the local sensors, and the derivation of their optimal rules becomes a very hard problem. For this reason, motivated by the counting rule considered in [13], we use a heuristic decision statistic δ_i at each sensor in the following form:

$$P(u_i=1|\mathbf{u}_{i-1},z_N)=0, \quad \text{if } \delta_i=z_i+\sum_{k=1}^{i-1}u_k\le t_i$$
$$P(u_i=1|\mathbf{u}_{i-1},z_i)=1, \quad \text{if } \delta_i=z_i+\sum_{k=1}^{i-1}u_k> t_i. \quad (32)$$

Basically, each sensor computes its decision statistic δ_i by summing the number of 1s received from its predecessor sensors together with its own measurement. Then, this decision statistic is compared with a certain threshold t_i. This heuristic rule works even when there is no prior information available in the network, such as the location of sensors or the location of the event.

The inner integration term in (30) can be written as

$$\begin{aligned}&\int_{z_N} P(u_N=1|\mathbf{u}_{N-1},z_N)p(z_N|H_0)dz_N\\ &\quad=\int_{t_N-\left(\sum_{i=1}^{N-1}u_i\right)}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{z_N^2}{2}}dz_N\\ &\quad=Q\left(t_N-\left(\sum_{i=1}^{N-1}u_i\right)\right). \end{aligned}\quad (33)$$

In addition, in (30), the probability mass function of the received decisions $P(u_{N-1},\ldots,u_1|H_0)$ needs to iteratively be conditioned on sensor decisions as

$$\begin{aligned}&P(u_{N-1},\ldots,u_1|H_0)\\ &\quad=\int_{z_{N-1}} P(u_{N-1}|\mathbf{u}_{N-2},z_{N-1})p(z_{N-1}|H_0)dz_{N-1},\ldots\\ &\qquad\times\int_{z_1}P(u_1|z_1)p(z_1|H_0)dz_1. \end{aligned}\quad (34)$$

In (34), depending on the local sensor decisions, each inner integral is replaced by an appropriate $Q(.)$ or $1-Q(.)$ function, as defined in (33) based on the decision and the threshold of each local sensor.

Finally, the probability of error is found by averaging over all possible decisions, as shown in (35), located at the bottom of the page.

2) Energy Consumption: In serial decision fusion, a sensor's energy consumption depends not only on the distance between the source and destination but also on the number of bits received and the distance from its predecessors. Each sensor receives $m-1$ bits from its predecessors and transmits m bits to its next successor, including its own decision. We define the distance between s_i and s_{i+1} as $d_{i,i+1}$; then, $E(i)$, which is the energy consumption of s_i, is the sum of energy used for receiving $m-1$ bits from its predecessors $E_{\mathrm{RX},i}$ and transmitting m bits to its successor over distance $d_{i,i+1}E_{\mathrm{TX},i}$ [7] as

$$\begin{aligned}E_{\mathrm{RX},i}(m-1)&=E_{\mathrm{elec}}\times(m-1)\\ E_{\mathrm{TX},i}(m,d_{i,i+1})&=E_{\mathrm{elec}}\times m+\epsilon_{\mathrm{amp}}\times m\times d_{i,i+1}^2\\ E(i)&=E_{\mathrm{RX},i}+E_{\mathrm{TX},i}\end{aligned}\quad (36)$$

for $i=\{2,3,\ldots,N-1\}$, $E(1)=E_{\mathrm{TX},1}(u_1,d_{1,2})$, and $E(N)=E_{\mathrm{RX}}(\sum_{i=1}^{N-1}u_i)$ joules.

Since the decision of the Nth sensor u_N is the final inference, u_N does not contribute to the energy consumption. Given the vector of past sensor decisions $\mathbf{u}_{N-1}$, the energy consumption in the network is expressed as

$$\begin{aligned}&E_C(\mathbf{u}_{N-1})\\ &\quad=E_{\mathrm{TX}}(u_1,d_{1,2})\\ &\qquad+\sum_{i=2}^{N-1}\left[E_{\mathrm{TX},i}\left(\sum_{q=1}^{i}u_i,d_{i,i+1}\right)+E_{\mathrm{RX},i}\left(\sum_{q=1}^{i-1}u_i\right)\right]\\ &\qquad+E_{\mathrm{RX},N}\left(\sum_{i=1}^{N-1}u_i\right).\end{aligned}\quad (37)$$

Finally, conditioning on all possible vector of decisions, the energy consumption E_T of the network is found according to

$$\begin{aligned}E_T&=\sum_{\text{all }\mathbf{u}_{N-1}}E_C(\mathbf{u}_{N-1})P(\mathbf{u}_{N-1})\\ &=\sum_{\text{all }\mathbf{u}_{N-1}}E_C(\mathbf{u}_{N-1})\big(P(\mathbf{u}_{N-1}|H_0)P(H_0)\\ &\qquad+P(\mathbf{u}_{N-1}|H_1)P(H_1)\big)\end{aligned}\quad (38)$$

where $P(\mathbf{u}_{N-1}|H_0)=P(u_{N-1},\ldots,u_1|H_0)$ and $P(\mathbf{u}_{N-1}|H_1)=P(u_{N-1},\ldots,u_1|H_1)$ are calculated as described in (34) and (35), respectively.

$$\begin{aligned}P_e=&P_0\left(\int_{z_N}P(u_N=1|\mathbf{u}_{N-1},z_N)p(z_N|H_0)dz_N,\ldots,\int_{z_1}\sum_{u_1}P(u_1|z_1)p(z_1|H_0)dz_1\right)\\ &+P_1\left(1-\int_x\int_y\left[\int_{z_N}P(u_N=1|\mathbf{u}_{N-1},z_N,x,y)p(z_N|x,y,H_1)dz_N,\ldots\right.\right.\\ &\qquad\left.\left.\times\int_{z_1}\sum_{u_1}P(u_1|z_1,x,y)p(z_1|x,y,H_1)dz_1\right]p(x,y)dydx\right)\end{aligned}\quad (35)$$

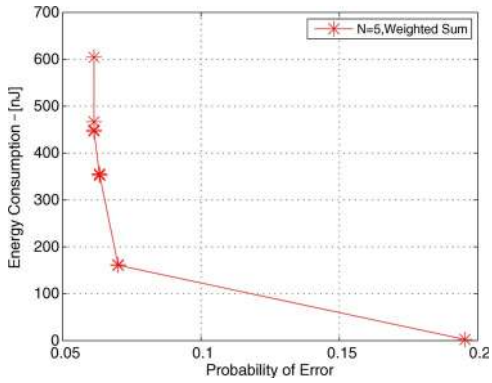

Fig. 3. Performance when the weighted sum of the objective functions is minimized for $N = 5$.

III. Multiobjective Optimization

In this section, we briefly review the fundamentals of multiobjective optimization. Then, we summarize NBI and NSGA-II, which are efficient methods to solve MOPs. The mathematical description of multiobjective optimization can be given as

$$\min_{\chi \in C} \quad [\, f_1(\chi) \quad f_2(\chi) \quad \cdots \quad f_n(\chi)\,]^T \tag{39}$$

where χ is a solution to the MOP. The number of objectives $n \geq 2$ and the feasible set C

$$C : \{\chi : h(\chi) = 0, g(\chi) \leq 0, a \leq \chi \leq b\} \tag{40}$$

is subject to the equality and inequality constraints denoted as $h(\chi)$ and $g(\chi)$, respectively, and explicit variable bounds $[a, b]$. In a minimization problem, a solution χ_1 dominates another solution χ_2 ($\chi_1 \gg \chi_2$) if and only if

$$\begin{aligned} f_u(\chi_1) &\leq f_u(\chi_2) \qquad \forall\, u \in \{1, 2, \ldots, n\} \\ f_v(\chi_1) &< f_v(\chi_2) \qquad \exists v \in \{1, 2, \ldots, n\} \end{aligned} \tag{41}$$

and a solution χ^* is the Pareto optimal solution for the MOP if and only if there is no $\chi \in C$ that dominates χ^*. Pareto optimal points are also known as nondominated points. A well-known technique for solving MOPs is to minimize a weighted sum of the objectives. Before describing the other MOP methods, the performance where the weighted sum of the objectives P_e and E_T for parallel decision fusion with five sensors is presented in Fig. 3. The weight pairs for the 11 points shown in Fig. 3 are $\{(0, 1), (0.1, 0.9), \ldots, (0.9, 0.1), (1, 0)\}$. In this problem, each solution to the MOP represents the set of local sensor thresholds $\chi = [t_1, t_2, \ldots, t_N]$. As seen from the figure, minimizing the weighted sum of the objectives suffers from several drawbacks [14]. First of all, a uniform spread of weights rarely produces a uniform spread of points on the Pareto front. Some of the optimal design solutions are closely spaced, which reduce the number of design alternatives. Second, if the Pareto optimal curve is not a convex function, the Pareto points on the concave parts of the actual Pareto optimal curve will be missed. Moreover, since it is up to the user to choose appropriate weights, the decision on the preferences may not be clear to the user until the solution is generated. Similarly, when compared with other existing MOP algorithms, such as Timmel's population-based method and Schaffler's stochastic method [20], the NBI method chosen in this paper and explained below is computationally efficient in locating Pareto optimal points. Therefore, we first consider the application of NBI to solve our MOP problem, which is briefly described as follows.

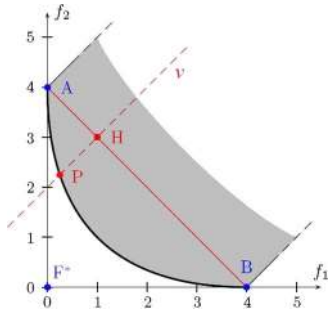

Fig. 4. Point P is the solution of the single-objective constrained NBI subproblem outlined with the dashed line v [15].

A. NBI

The NBI method [14] reduces the MOP to multiple numbers of single-objective constrained problems called NBI subproblems. This method starts with separately finding the optimizers of each objective function. For the two-objective example illustrated in Fig. 4, the shaded area represents the region of feasible design, and the curve at the lower boundary is the Pareto optimal front. The convex hull of individual minima (CHIM) is defined as the line segment AB. Any NBI problem is then specified by a reference point on the CHIM, such as the point H. Let χ_j^* be the minimizer of the jth objective and $F_j^* = F(\chi_j^*) = [f_1(\chi_j^*), \ldots, f_n(\chi_j^*)]^T$, and the payoff matrix Φ is an $n \times n$ matrix whose jth column is $F_j^* - F^*$. $\Phi\beta$ then denotes the reference point H, and each NBI subproblem is defined as

$$\begin{aligned} \max_{\chi, \tau} \quad & \tau \\ s.t. \quad & \Phi\beta + \tau v = F(\chi) \\ & h(\chi) = 0, g(\chi) \leq 0, a \leq \chi \leq b. \end{aligned} \tag{42}$$

The length of the line segment HP τ represents the new variable introduced by the NBI subproblem. The new constraint given the NBI subproblem ensures that the point lies inside the feasible set C. The number of NBI subproblems determines the resolution of the Pareto front. Clearly, larger values for this parameter imply a better resolution of the Pareto front. If the Pareto set is disconnected, then it is concluded that some of the subproblems have no solution [14]. Each NBI subproblem can be solved with any appropriate optimization method.

Figs. 5 and 6 show the contour plots of the global probability of error and the global energy consumption for parallel and serial configurations. Since a closed-form expression for either objective function (for the most general case of N sensors) is not available, the Hessian matrix consisting of the second derivatives of the objective functions with respect to sensor thresholds needed in a formal proof of unimodality cannot analytically be determined. Hence, any attempt to prove unimodality with the available information will only be approximate. For this reason, we choose to present numerical examples for the objective function's behavior rather than a formal proof. Simulation results show that, even for a two-sensor network, the global probability of error is not a unimodal function of

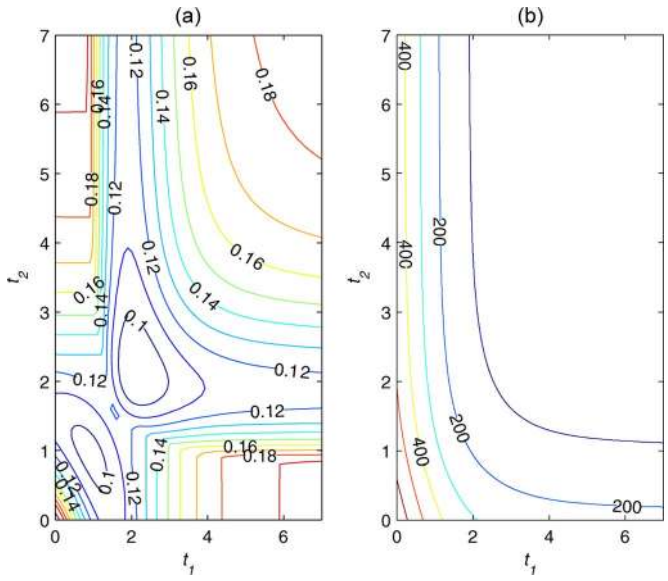

Fig. 5. Contour lines of the objective functions with $N = 2$ sensors, parallel configuration. (a) Probability of error. (b) Energy consumption.

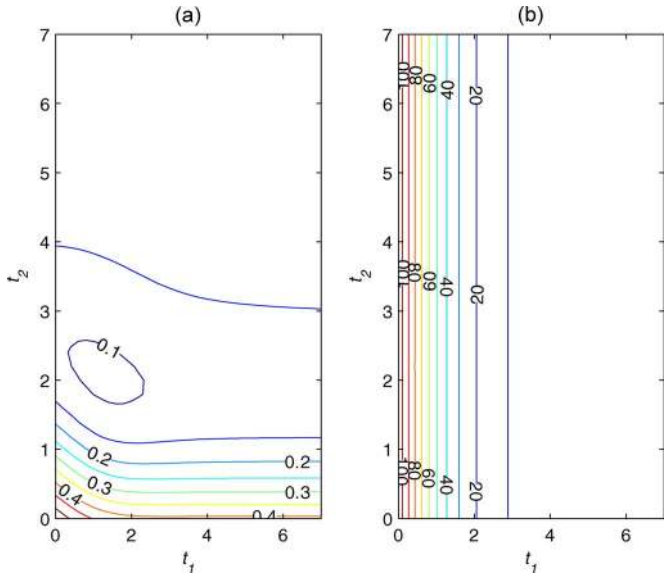

Fig. 6. Contour lines of the objective functions with $N = 2$ sensors, serial configuration. (a) Probability of error. (b) Energy consumption.

sensor thresholds. For such cases, gradient-based approaches may yield a local optimum instead of a global optimum. Hence, the obtained results need to be compared with other global techniques, such as an exhaustive search or a genetic algorithm, to ensure that the solutions really converge to the global optimum solution. For this purpose, next, we describe an evolutionary algorithm for MOPs called NSGA-II.

B. NSGA-II

NSGA-II [18] is a state-of-the-art multiobjective evolutionary algorithm that simultaneously obtains M Pareto optimal solutions in n-dimensional objective space. A solution in the population is represented as a sequence of decision variables, namely, the sensor thresholds. Unlike NBI, using NSGA-II, the Pareto front (tradeoff curve) is directly found, so there is no need to separately calculate the individual minimizers of each objective function. NSGA-II is an elitist algorithm, where good solutions are preserved in the population.

NSGA-II is based on nondomination in each front. Each solution in the population is assigned a fitness and crowding distance value. The solutions with the same fitness are then resorted based on their crowding distance, which is a closure measure of each solution to its neighbors. For each generation of the algorithm, the computational complexity $O(n \times M^2)$ is governed by this nondominated sorting operation (see [18] for details). The mating population is subsequently generated by using binary tournament selection. If both of the solutions have the same fitness, then the solution with larger crowding distance is selected. We use a real-parameter recombination operator called simulated binary crossover (SBX), which is commonly used in the evolutionary algorithm literature [18], [23]. The SBX has a parameter distribution index η_C, whose value determines the closeness of the offspring to their parents. Let p_1 and p_2 be two individual solutions obtained from binary tournament selection. In SBX, offspring solutions c_1 and c_2 are obtained from parent solutions p_1 and p_2 according to

$$\begin{aligned} c_1 &= \frac{1}{2}\left[(1-\zeta)p_1 + (1+\zeta)p_2\right] \\ c_2 &= \frac{1}{2}\left[(1+\zeta)p_1 + (1-\zeta)p_2\right] \end{aligned} \tag{43}$$

where $\zeta \geq 0$ is a random number with pdf [24]

$$\begin{aligned} p(\zeta) &= \frac{1}{2}(\eta_C + 1)\zeta^{\eta_C}, \qquad 0 \leq \zeta \leq 1 \\ p(\zeta) &= \frac{1}{2}(\eta_C + 1)\frac{1}{\zeta^{\eta_C+2}}, \qquad \zeta > 1. \end{aligned} \tag{44}$$

Along with the SBX, we use polynomial mutation that also makes use of a parameter distribution index η_M. In polynomial mutation, the offspring solution c_l is obtained from the parent solution p_l according to

$$c_l = p_l + (t_{\max} - t_{\min})\delta \tag{45}$$

where δ is a small variation calculated from the density function [24]

$$\begin{aligned} p(\delta) &= (2q)^{\frac{1}{\eta_M+1}} - 1, \qquad q < 0.5 \\ p(\delta) &= 1 - (2(1-q))^{\frac{1}{\eta_M+1}}, \qquad q \geq 0.5 \end{aligned} \tag{46}$$

where q is a random number with uniform distribution between (0, 1).

The population is then updated by selecting the solutions starting from the first front. If the number of solutions in the last allowable front is larger than the available places in the population, then the solutions with larger crowding distance are selected first. After several iterations, the entire population only contains the solutions near or at the Pareto optimal front.

C. Performance Metrics

For performance comparison between the solutions found with NBI and NSGA-II, we use three metrics: 1) generational distance (GD); 2) domination (Dom) metric; and 3) spacing metric [19], [26].

TABLE I
GD BETWEEN NBI AND NSGA-II, SPREAD METRIC, AND MEAN EXECUTION TIMES (E.T.) FOR NSGA-II AND NBI

	GD (Mean)	GD (Std. Dev.)	S (Mean)	S (Std. Dev.)	Mean E.T. (seconds)
NSGA II: G=20	603.9421	805.1765	19.2959	22.7605	0.2476e4
NSGA II: G=50	253.0000	784.6029	4.5374	9.7953	0.6449e4
NSGA II: G=100	5.0322	0.4002	1.6428	0.1827	1.2350e4
NSGA II: G=200	5.1058	0.4269	1.7263	0.3049	2.3105e4
NSGA II: G=500	5.0539	0.4498	1.9158	0.2094	5.5232e4
NBI: resolution 100			1.5594		1.7118e4

The GD [19]

$$GD(A,B)=\sqrt{\sum_{i=1}^{M} g_i^2} \tag{47}$$

measures the distance between the nondominated solutions obtained by algorithms A and B. g_i is the Euclidean distance between the solution $i \in A$ and the nearest solution in B.

The Dom metric [19] is based on the number of solutions (obtained by one algorithm) dominated by each solution obtained by the other algorithm. The Dom metric is defined as

$$Dom(A,B)=\frac{d(A,B)}{d(A,B)+d(B,A)}$$
$$d(X,Y)=\sum_{x}\left|\{y\in Y | x \gg y\}\right|. \tag{48}$$

If each solution of algorithm A dominates every solution of algorithm B, then $Dom(A,B)=1$ and $Dom(B,A)=0$, where $Dom(B,A)=1-Dom(A,B)$.

The spacing metric [26] measures the uniformity of the solutions obtained in the Pareto optimal front. The S metric is defined as

$$S(A)=\sqrt{\frac{1}{M-1}\sum_{i=1}^{M}(r_i-\bar{r})^2} \tag{49}$$

where M is the number of nondominated solutions in the archive, and r_i is the sum of the differences in objective function values between solution i and its two nearest neighbors for each objective. The spacing metric approaches zero when the Pareto optimal solutions are near uniformly spaced.

IV. SIMULATION RESULTS

In this section, we first describe the simulation settings and then present the Pareto fronts obtained from NBI and NSGA-II algorithms and discuss the effects of nondominated solutions on WSN performance.

A. Simulation Settings

In our simulations, we use the WSN configuration described in Section II-A. The solution of the MOP is illustrated with deterministic sensor placements, where the sensors are equidistantly placed on the $y=x$ line in the ROI $A \times B = 100$ m $\times$ 100 m, as shown in Fig. 1. As an example, boundary or pipeline surveillance requires placing the sensors on a straight line. The proposed MOP can be applied to any configuration as long as the sensor placements and the characteristics of the event of interest are known. The fusion center is located at the origin. According to our objective functions shown in (20) and (26), adding an additional sensor doubles the number of possible vectors of received decisions. Therefore, the search space of both objectives exponentially increases with N, i.e., it is 2^N. For this reason, we illustrate the proposed MOP with relatively few sensors. The *a priori* probabilities for H_0 and H_1 are selected as $P_0=0.8$ and $P_1=0.2$, respectively. The parameters of the event detection model are set as $K_0=10^6$ J, $a=200$, and $n=2$. The standard deviation of the measurement noise σ is set to 1. The minimum $t_{\min}$ and maximum $t_{\max}$ values for the thresholds are taken as 0 and 10, respectively. For NBI, individual minimizers of each objective function and each NBI subproblem are determined by using MATLAB's `fmincon` routine. For the `fmincon` routine, all sensor thresholds are initialized at $t_i^0=8$, where $P_e^0 \approx 0.2$ and $E_T^0 \approx 0$, and the algorithm termination tolerances of `fmincon` routine are all set to 10^{-7}. The resolution of the Pareto optimal front is selected as $R_{NBI}=10$. For NSGA-II, we use a population of size $M=100$. The crossover and mutation probabilities are set at 0.9 and 0.1, respectively [27]. The parameter distribution indexes of SBX and polynomial mutation are set to $\eta_C=20$ and $\eta_M=20$, respectively. We observed that slight changes in these parameters do not significantly change the results. NSGA-II and NBI methods are implemented via available public codes in [27] and [28], respectively. All simulations are performed on a computer with a 3.2-GHz Pentium processor.

B. Performance Comparison of NBI and NSGA-II

For performance comparison between NBI and NSGA-II, we select both the Pareto front resolution of NBI and the population size of NSGA-II as 100 with $N=4$ variables with parallel decision fusion. The GD, the Dom metric, and the spacing metric (S) are averaged over ten different NSGA-II trials. In Table I, we vary the number of generations (G) and measure the GD between 100 solutions of NSGA-II and NBI. The simulation results show that the average GD between NBI and NSGA-II is small after $G=100$ generations. Moreover, according to Table I, after $G=100$ generations, the spacing metric of NSGA-II converges with a small standard deviation. On the other hand, the solutions corresponding to NBI are more evenly spaced as compared with solutions of NSGA-II, since NBI yields a smaller S-metric. Table II reveals that the solutions obtained by NBI dominate the solutions of NSGA-II. Note that the resolution of the Pareto optimal front is independent from the convergence of the NBI. Therefore, ten Pareto optimal solutions found with NBI clearly dominate every solution found with NSGA-II. Therefore, by solving just a few number of subproblems, the same Pareto optimal front can be achieved, and evenly distributed trade off solutions can be obtained in a very short time as compared with NSGA-II. It should be pointed out that NBI is known to be better for two objective problems,

TABLE II
DOM METRIC BETWEEN NBI AND NSGA-II

Dom(A,B) A:NBI, B:NSGA-II	B: G=20	B: G=50	B: G=100	B: G=200	B: G=500
A:res. 100,E.T.:1.7e4 s.	1	1	1	1	0.997
A:res. 20,E.T.:4.9e3 s.	1	1	1	1	1
A:res. 10,E.T.:2.3e3 s.	1	1	1	1	1

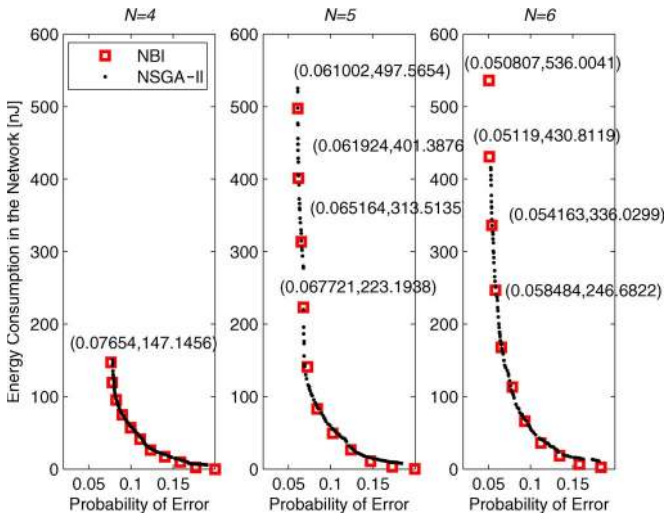

Fig. 7. Pareto optimal solutions generated via NBI and NSGA-II methods for parallel fusion and nonidentical decision thresholds at each sensor.

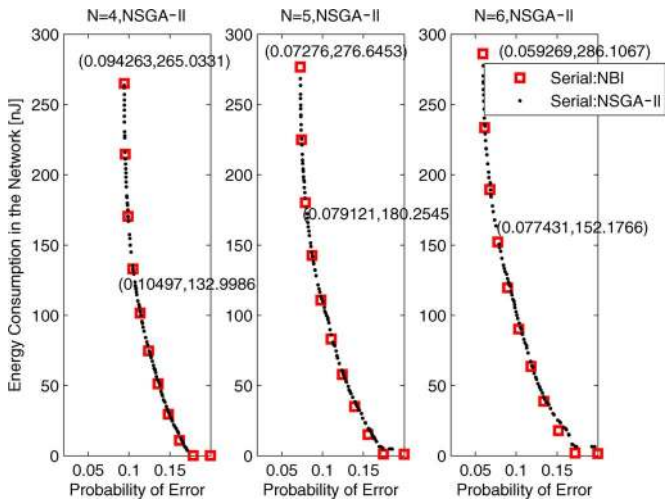

Fig. 8. Pareto optimal solutions obtained by NBI and NSGA-II methods for serial fusion and nonidentical decision thresholds at each sensor.

but for problems with a larger number of objectives, NSGA-II may be better [25].

C. Optimal Pareto Fronts

In this section, we present the Pareto optimal solutions for parallel and serial decision fusion for the case of nonidentical decision thresholds employed at each of the N sensors.

1) Pareto Optimal Fronts Under Parallel Fusion: For parallel decision fusion and the WSN configuration, as described in Section III-A, Fig. 7 shows the Pareto optimal fronts generated with NBI and NSGA-II, where each solution is shown in terms of objective function pairs $[P_e, E_T]$. The number of decision variables is selected as $N = 4, 5, 6$. NSGA-II is executed with population size $M = 100$ and number of generations $G = 100$, and NBI is executed with a resolution of 10. The simulation results show that NBI and NSGA-II yield Pareto optimal fronts that are fairly close to each other. Adding more sensors to the network decreases the error probability, and NBI results in nearly equidistant points on the Pareto front. For $N = 5$ sensors, if we only minimize P_e, the best achievable global probability of error is 0.061, which consumes 497.5654 nJ. By using the solution for the MOP, instead of selecting this solution, we may accept the neighboring solution on the Pareto optimal front with a global error probability of 0.0619 and a global energy consumption of 401.38 nJ. Therefore, a 1.5% increase in the global probability of error delivers 23% saving in global energy consumption. Similarly, for the $N = 6$ case, instead of operating on the minimum probability error solution [0.05, 536 nJ], selecting the solution [0.058, 246 nJ] yields 53.9% energy saving in exchange for a 15.11% increase in the global probability of error.

2) Pareto Optimal Fronts Under Serial Fusion: Fig. 8 shows the Pareto optimal solutions obtained with NBI and NSGA-II for the serial decision fusion case for $N = 4, 5, 6$ sensors in the network according to the WSN configuration, as shown in Fig. 2. The heuristic decision rule of each sensor proposed in (30) yields the global probability of error that is slightly worse than the parallel configuration. As an example for $N = 5$ and $N = 6$ sensors, the minimum achievable error probabilities for parallel decision fusion are 0.061 and 0.05, whereas the serial case yields the minimum error probabilities 0.072 and 0.059, respectively. The global energy consumption of the serial configuration is determined by the distance between neighboring sensors and the number of received and transmitted bits of each sensor. For the $N = 4$ case, the distance between two neighboring sensors is relatively large. The minimum error solution for the serial fusion consumes 265 nJ, whereas the parallel configuration consumes 147 nJ. Although increasing the number of sensors increases the number of bits for reception and transmission, the distance between sensors significantly decreases. Since the energy consumption of the network is determined by the square of the intersensor distance, increasing the number of sensors decreases the network's total energy consumption as compared with the parallel case. As an example, for $N = 6$ sensors, under parallel network configuration, the minimum achievable global probability of error is about 0.050 with an energy consumption of 536 nJ, whereas under serial configuration, the minimum achievable error probability is 0.059 with an energy consumption of 286 nJ. In other words, deploying the network serially increases the probability of error by 18% but decreases the energy consumption of the network by 46%.

D. Performance of WSN

In this section, we analyze the performance of WSNs based on the selected Pareto solution with N decision variables

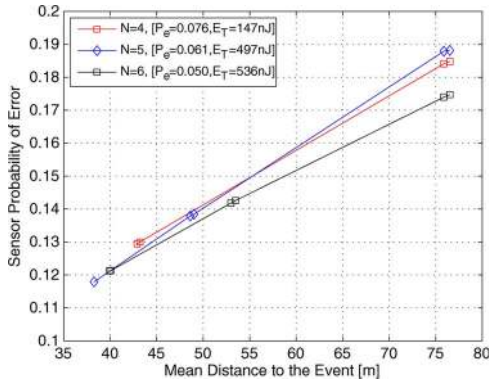

Fig. 9. Parallel decision fusion. Local sensor error probability as a function of its mean distance to the event location.

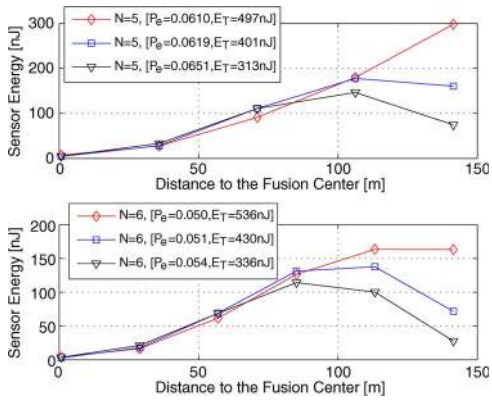

Fig. 10. Parallel decision fusion. Local sensor energy consumption as a function of its distance to the fusion center.

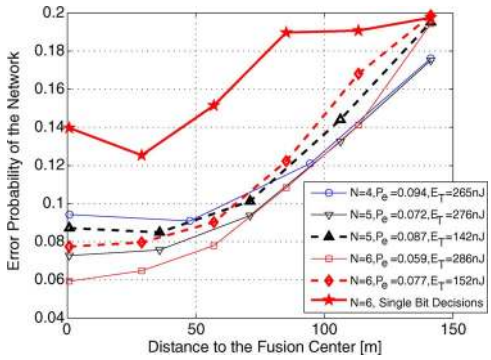

Fig. 11. Serial decision fusion. Global error probability as a function of the hop count on the routing path.

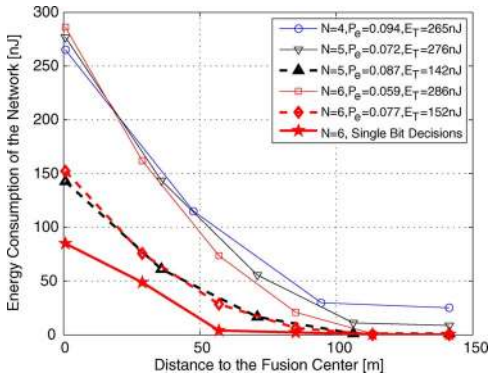

Fig. 12. Serial decision fusion. Global energy consumption as a function of the hop count on the routing path.

presented in the previous section. Under parallel decision fusion, we first determine the error probability of an individual sensor $P_{\mathrm{ind}}(t_i)$, as given in (18), as a function of its mean distance to the event location $\bar{d}_i$. We then calculate an individual sensor's energy consumption $E_{\mathrm{ind}}(t_i)$, as given in (22), as a function of its distance to the fusion center $d_{f,i}$. For serial decision fusion, we calculate the global probability of error and the global energy consumption of the network for a given number of sensor on the routing path.

1) WSN Performance Under Parallel Decision Fusion: For the minimum global probability of error solutions, Fig. 9 shows that the local sensor thresholds are assigned in such a way that the individual sensor error probability increases with the mean sensor distance to the event location. Due to this, a sensor more frequently transmits if it is close to the event and does not transmit that frequently if it is far. Then, the error probability of a sensor far away from the mean event location is close to the prior probability P_1. In terms of energy consumption, Fig. 10 shows that the energy consumption of a sensor increases with its distance to the fusion center. This is an expected result since the energy consumption of a sensor increases with the square of the distance to the fusion center. Fig. 10 also shows that for the consecutive Pareto optimal solutions with increased global probability of error and decreased global energy consumption, i.e., $N = 5 : [0.0619, 401\ \mathrm{nJ}]$, $N = 5 : [0.0651, 313\ \mathrm{nJ}]$, $N = 6 : [0.051, 430\ \mathrm{nJ}]$, and $N = 6 : [0.054, 336\ \mathrm{nJ}]$, the energy consumption of the sensors that are far away from the fusion center decreases as their thresholds increase. Since these sensors are also relatively far from the expected event location, a decrease in their transmission rate only makes a slight difference in the minimum achievable global probability of error. On the other hand, since delivering their decisions to the fusion center has much energy cost, a decrease in their transmission rate provides significant savings in global energy consumption.

2) WSN Performance Under Serial Decision Fusion: Figs. 11 and 12 show the global probability of error and the global energy consumption as a function of number of sensors (hops) on the routing path, respectively. In these figures, sensor 1 (s_1) in Fig. 2 that is farthest away from the fusion center generates the first decision and transmits to s_2. At each sensor, we calculate the global probability of error and the global energy consumption. As a benchmark, we compare the performance of the proposed decision rule given in (32) with a simple rule where each sensor decision u_i is only the aggregation of the decision of the previous sensor and its measurement, that is, $u_i = z_i + u_{i-1}$. In Figs. 11 and 12, since s_1 has the farthest distance to the average event location, a higher threshold is assigned to this sensor, and it is operating at a probability of error close to the prior probability P_1. For the benchmark case, over consecutive sensors, the global probability of error decreases, since the sensors become much closer to the mean event location. On the other hand, the sensors near the fusion center are far away from the mean event location, so the measurements of these sensors add uncertainty to the received decisions. That is why the global probability of error increases again (see the topmost curve in Fig. 11). Our proposed rule considers not only the decision of the previous sensor but also the decisions of all predecessors. Therefore, increasing the

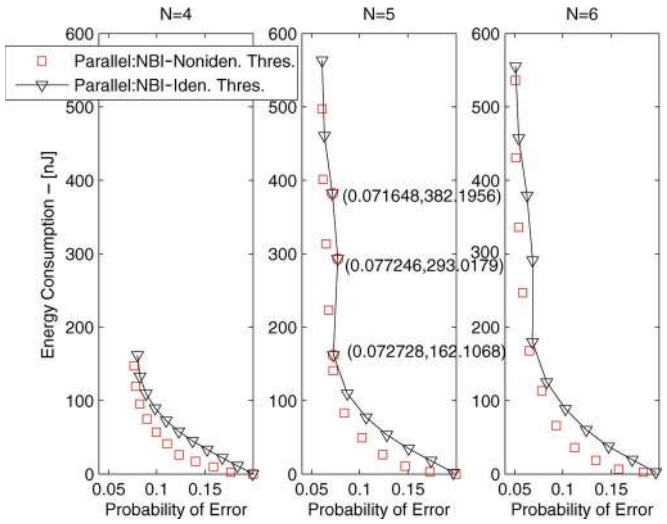

Fig. 13. Parallel decision fusion. Pareto optimal solutions for identical and nonidentical sensor thresholds.

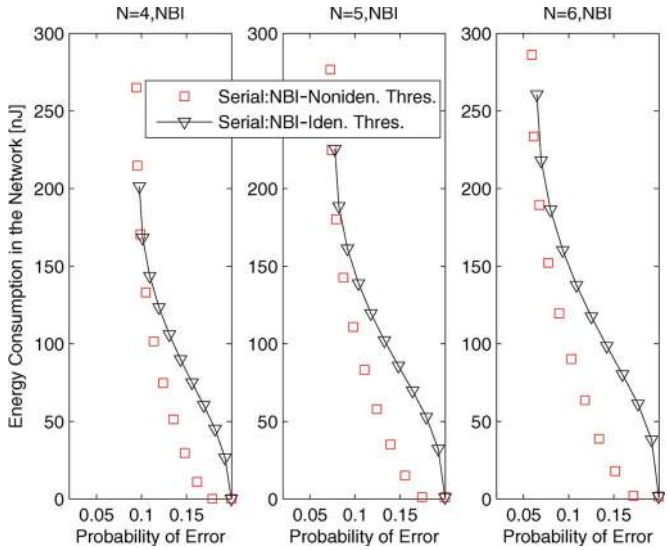

Fig. 14. Serial decision fusion. Pareto optimal solutions for identical and nonidentical sensor thresholds.

number of sensors on the routing path increases the number of available decisions about the event, and the global probability of error successively decreases at each sensor. As shown in Fig. 12, the global energy consumption of the network increases at each sensor as a result of the increased number of transmitted and received decisions at each sensor.

E. Identical Decision Thresholds

Since the search space of both objectives exponentially increases with N, adding an additional sensor roughly doubles the computation time. To simplify the problem, we may constrain the decision rules to be identical of all the sensors. For parallel decision fusion, Fig. 13 shows the optimal Pareto fronts for the case of identical and nonidentical decision thresholds. For $N = 4, 5, 6$ sensors, identical decision thresholds for all the sensors yield the objective function pair values with minimum global probability of error as [0.080, 161 nJ], [0.061, 563 nJ], and [0.051, 555 nJ], whereas the nonidentical threshold selection gives the objective function pair values [0.076, 147 nJ], [0.061, 497 nJ], and [0.0508, 536 nJ], respectively. Simulation results show that, as the number of sensors in the network increases, an identical decision threshold for all the sensors achieves nearly the same error probability as compared with nonidentical threshold selection. In [3], it is shown that as the number of sensors grows to infinity, the probability of error goes to zero for any reasonable set of decision thresholds. Therefore, fine adjustment of decision thresholds at each sensor becomes unnecessary if the number of sensors is large. Therefore, particularly for a large number of sensors, an identical decision threshold at all the sensors simplifies the problem and yet gives near-optimal results. Note that the energy consumption of the network is slightly higher for the identical threshold case. This is due to the fact that, under nonidentical threshold assignment, the sensors that are far from the event are assigned a higher threshold that reduces the transmission rate of these sensors. On the other hand, for identical threshold selection, these sensors are assigned a lower threshold value and transmit more frequently. In addition, the simulation results show that for identical threshold selection, the Pareto optimal curve between E_T and P_e is not convex, which implies that some of the candidate solutions on the Pareto front are still dominated. As an example, for $N = 5$ case, the two solutions [0.077246, 293 nJ] and [0.072728, 162 nJ] are on the Pareto front, but [0.077246, 293 nJ] is already dominated by [0.072728, 162 nJ]. On the other hand, nonidentical sensor threshold selection yields a convex Pareto front, where all the solutions are nondominated, which provides more alternatives to the designer. For serial decision fusion, Fig. 14 shows the Pareto optimal solutions for the case of identical and nonidentical decision thresholds. The simulation results show that the best achievable error probability with identical threshold selection is slightly worse than the nonidentical threshold selection. Similar to the parallel decision fusion system, the identical threshold scheme has a nonconvex tradeoff curve, and for a given global probability of error, the energy consumption of identical threshold selection is higher than nonidentical threshold selection.

V. Conclusion

In this paper, we have studied the binary distributed detection problem. The event signal is represented by an isotropic emission model, and the location of the event is only known in terms of its pdf, where the decision thresholds that minimize the probability of error cannot be determined using existing methods, such as PBPO. We formulated and solved a MOP with two conflicting objectives: 1) global probability of error and 2) global energy consumption of the network, where each solution of this problem corresponds to placing a different emphasis on the two objectives. The proposed MOP is solved by two different methods. NBI and NSGA-II yield Pareto optimal fronts that are very close to each other. The simulation results show that, for our problem, NBI provides better and more uniformly distributed solutions in a shorter time as compared with NSGA-II.

Rather than only minimizing the global probability of error, an MOP approach provides a number of alternative solutions that provide significant energy savings as compared with the minimum error solution at the cost of slightly increasing the

best achievable global probability of error. Under parallel decision fusion, the consecutive Pareto optimal solutions significantly decrease the global energy consumption by allowing a slight increase in the minimum achievable probability of error. Under serial decision fusion, increasing the number of sensors on the routing path successively increases the amount of information about the event and the probability of error at each sensor decreases. We have also shown that an identical decision threshold for all the sensors achieves nearly the same error probability as compared with nonidentical threshold selections at each sensor as the number of sensors in the network increases. Therefore, particularly for a large number of sensors, an identical decision rule for all sensors can be employed to achieve nearly the best probability of error performance.

In this paper, we have shown the utility of the MOP approach in providing different alternatives to WSN designers. The proposed MOP can easily be extended to multiobjective problems with more than two objectives as well as under specified constraints. Future work will include adapting the proposed framework to a larger number of sensors and more than two objectives, as well as the development of computationally efficient approaches. An extension of our methodology for a general network topology with multiple events occurring at the same time will also be addressed. For this purpose, the earlier work of Alhakeem and Varshney [29] could be used.

ACKNOWLEDGMENT

The authors would like to thank all the reviewers and Dr. R. Niu and Dr. H. Chen for their invaluable comments.

REFERENCES

[1] J.-F. Chamberland and V. V. Veeravalli, "Wireless sensors in distributed detection applications," *IEEE Signal Process. Mag.*, vol. 24, no. 3, pp. 16–25, May 2007.

[2] P. K. Varshney, *Distributed Detection and Data Fusion*. Berlin, Germany: Springer-Verlag, 1996.

[3] J. N. Tsitsiklis, "Decentralized detection by a large number of sensors," *Math. Control, Signals, Syst.*, vol. 1, no. 2, pp. 167–182, Jun. 1988.

[4] B. Chen, L. Tong, and P. K. Varshney, "Channel-aware distributed detection in wireless sensor networks," *IEEE Signal Process. Mag.*, vol. 23, no. 4, pp. 16–26, Jul. 2006.

[5] I. Bahceci, "Multiple-input multiple-output wireless systems: Coding, distributed detection and antenna selection," Ph.D. dissertation, Georgia Inst. Technol., Atlanta, GA, 2005.

[6] X. Zhang, H. V. Poor, and M. Chiang, "Optimal power allocation for distributed detection over MIMO channels in wireless sensor networks," *IEEE Trans. Signal Process.*, vol. 56, no. 9, pp. 4124–4140, Sep. 2008.

[7] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy-efficient communication protocol for wireless microsensor networks," in *Proc. 33rd Int. Conf. Syst. Sci.*, Jan. 2000, p. 8020.

[8] R. Jiang, Y. Lin, B. Chen, and B. Suter, "Distributed sensor censoring for detection in sensor networks under communication constraints," in *Proc. IEEE Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, Nov. 2005, pp. 946–950.

[9] R. Niu and P. K. Varshney, "Distributed detection and fusion in a large wireless sensor network of random size," *EURASIP J. Wireless Commun. Netw.*, vol. 5, no. 4, pp. 462–472, Sep. 2005.

[10] D. Li, K. D. Wong, Y. H. Hu, and A. M. Sayeed, "Detection, classification, and tracking of targets," *IEEE Signal Process. Mag.*, vol. 19, no. 2, pp. 17–29, Mar. 2002.

[11] A. Nasipuri and K. Li, "Multisensor collaboration in wireless sensor networks for detection of spatially correlated signals," *Int. J. Mob. Netw. Design Innov.*, vol. 1, no. 3/4, pp. 215–223, Jan. 2006.

[12] J.-J. Xiao, S. Cui, Z.-Q. Luo, and A. J. Goldsmith, "Power scheduling of universal decentralized estimation in sensor networks," *IEEE Trans. Signal Process.*, vol. 54, no. 2, pp. 413–422, Feb. 2006.

[13] R. Niu and P. K. Varshney, "Joint detection and localization in sensor networks based on local decisions," in *Proc. IEEE Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, Nov. 2006, pp. 525–529.

[14] I. Das and J. Dennis, "Normal-boundary intersection: A new method for generating the Pareto surface in nonlinear multicriteria optimization problems," *SIAM J. Optim.*, vol. 8, no. 3, pp. 631–657, 1998.

[15] E. Rigoni and S. Poles, *NBI and MOGA-II, Two Complementary Algorithms for Multi-Objective Optimizations*. [Online]. Available: http://drops.dagstuhl.de/opus/volltexte/2005/272/pdf/04461.PolesSilvia.Paper.272.pdf

[16] K. Deb, *Multi-Objective Optimization using Evolutionary Algorithms*. New York: Wiley, 2001.

[17] E. Zitzler, "Evolutionary algorithms for multiobjective optimization: Methods and applications," Ph.D. dissertation, Swiss Federal Inst. Technol., Zurich, Switzerland, 1999.

[18] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," *IEEE Trans. Evol. Comput.*, vol. 6, no. 2, pp. 182–197, Apr. 2002.

[19] R. Rajagopalan, C. K. Mohan, K. G. Mehrotra, and P. K. Varshney, "EMOCA: An evolutionary multi-objective crowding algorithm," *J. Intell. Syst.*, vol. 17, no. 1/3, pp. 107–123, 2008.

[20] P. K. Shukla and K. Deb, *On Finding Multiple Pareto-Optimal Solutions Using Classical and Evolutionary Generating Methods*. [Online]. Available: http://www.iitk.ac.in/kangal/reports.shtml

[21] L. T. Bui and S. Alam, *Multi-Objective Optimization in Computational Intelligence: Theory and Practice*. Calgary, AB, Canada: Idea Group Inc., 2008, ch. 8, pp. 208–238.

[22] E. Masazade, R. Rajagopalan, P. K. Varshney, G. K. Sendur, and M. Keskinoz, "Evaluation of local decision thresholds for distributed detection in wireless sensor networks using multiobjective optimization," in *Proc. IEEE Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, Oct. 2008.

[23] K. Deb and B. Agrawal, "Simulated binary crossover for continuous search space," *Complex Syst.*, vol. 9, no. 2, pp. 115–148, Apr. 1995.

[24] M. M. Raghuwanshi and O. G. Kakde, "Survey on multiobjective evolutionary and real coded genetic algorithms," *Complex. Int.*, vol. 11, pp. 150–161, 2005.

[25] P. K. Shukla, K. Deb, and S. Tiwari, *Comparing Classical Generating Methods With an Evolutionary Multi-Objective Optimization Method*, vol. 3410/2005. Berlin, Germany: Springer-Verlag, 2005, pp. 311–325.

[26] J. R. Schott, "Fault tolerant design using single and multi-criteria genetic algorithm optimization," M.S. thesis, Dept. Aeronaut. Astronaut., MIT, Cambridge, MA, May 1995.

[27] A. Seshadri, *Kanpur Genetic Algorithms Laboratory*. [Online]. Available: http://www.iitk.ac.in/kangal/codes.shtml

[28] I. Das, *The Normal-Boundary Intersection Home Page*. [Online]. Available: http://softlib.rice.edu/NBI.html

[29] S. Alhakeem and P. K. Varshney, "A unified approach to the design of decentralized detection systems," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 31, no. 1, pp. 9–20, Jan. 1995.

Engin Masazade (S'03) was born in Istanbul, Turkey, on July 19, 1980. He received the B.S. degree in 2003 from Istanbul Technical University, Istanbul, and the M.S. degree in 2006 from Sabanci University, Istanbul, where he is currently working toward the Ph.D. degree.

During his M.S. and Ph.D. studies, he was with the Communication Theory and Technologies Group under the supervision of Dr. M. Keskinoz. His research interests include bit error rate estimation and cross-layer design for multiband OFDM systems, distributed detection, estimation, and localization for wireless sensor networks.

Mr. Masazade has been awarded with a research abroad support scheme by the Scientific and Technological Research Council of Turkey (TUBITAK) for his studies at Syracuse University, Syracuse, NY, under the supervision of Prof. P. K. Varshney, in 2008.

Ramesh Rajagopalan (S'05) received the Masters (Honors) degree in physics and the Bachelor's in Engineering (Honors) degree in electrical and electronics from the Birla Institute of Technology and Science, Pilani, India, in 2002 and the Ph.D. degree in electrical engineering from Syracuse University, Syracuse, NY, in 2008.

He is currently a Faculty member in electrical and computer engineering with Florida State University, Panama City, where he has been teaching since January 2008. His research interests include design and analysis of algorithms, evolutionary computation, wireless communications and networks, distributed sensor networks, and multiobjective optimization. He has authored/coauthored conference and journal papers in the areas of sensor networks and multiobjective optimization.

Dr. Rajagopalan was the recipient of the Graduate School Masters Prize and the Doctoral Prize at Syracuse University in recognition of his research and scholarship. He received the Best Poster Award for his work on path planning algorithms at Nunan Research Day, Syracuse University.

Pramod K. Varshney (S'72–M'77–SM'82–F'97) was born in Allahabad, India, on July 1, 1952. He received the B.S. degree in electrical engineering and computer science (with highest honors) and the M.S. and Ph.D. degrees in electrical engineering from the University of Illinois at Urbana-Champaign, Urbana, in 1972, 1974, and 1976, respectively.

During 1972–1976, he held teaching and research assistantships with the University of Illinois at Urbana-Champaign. Since 1976, he has been with Syracuse University, Syracuse, NY, where he is currently a Distinguished Professor of electrical engineering and computer science and the Director of the Center for Advanced Systems and Engineering (CASE). He served as the Associate Chair of the department during 1993–1996. He is also an Adjunct Professor of radiology with Upstate Medical University, Syracuse, NY. He has extensively published. He is the author of *Distributed Detection and Data Fusion* (Springer-Verlag, 1997). He has served as a consultant to several major companies. He is a member of Tau Beta Pi His current research interests are in distributed sensor networks and data fusion, detection and estimation theory, wireless communications, image processing, radar signal processing, and remote sensing.

Dr. Varshney was a James Scholar, while at the University of Illinois, a Bronze Tablet Senior, and a Fellow. He is the recipient of the 1981 ASEE Dow Outstanding Young Faculty Award. He was elected to the grade of Fellow of the IEEE in 1997 for his contributions in the area of distributed detection and data fusion. He was the Guest Editor of the special issue on data fusion of the PROCEEDINGS OF THE IEEE, January 1997. In 2000, he received the Third Millennium Medal from the IEEE and Chancellor's Citation for exceptional academic achievement at Syracuse University. He serves as a distinguished lecturer for the Audio Engineering Society of the IEEE. He is on the editorial board of the *International Journal of Distributed Sensor Networks* and the IEEE TRANSACTIONS ON SIGNAL PROCESSING. He was the President of International Society of Information Fusion during 2001.

Chilukuri K. Mohan (SM'94) received the B.Tech. degree in computer science from the Indian Institute of Technology, Kanpur, India, in 1983 and the Ph.D. degree in computer science from the State University of New York, Stony Brook, in 1988.

He is currently a Professor (and the Chair) with the Department of Electrical Engineering and Computer Science, Syracuse University, Syracuse, NY, where he has been teaching since 1988. He has coauthored *Elements of Artificial Neural Networks* (MIT Press, 1997) and authored *Frontiers of Expert Systems: Reasoning with Limited Knowledge* (Kluwer, 2000). He has also authored/coauthored over 140 research papers in various areas of artificial intelligence.

Dr. Mohan is a member of an IEEE Task Force on Swarm Intelligence and serves on several international conference committees.

Gullu Kiziltas Sendur (S'03–M'05) was born in Turkey in 1972. She received the B.Sc. and M.Sc. degrees in mechanical engineering from the Middle East Technical University, Ankara, Turkey, in 1995 and 1998, respectively, and the Ph.D. degree in mechanical engineering from the University of Michigan, Ann Arbor, in 2003.

During her doctoral work, she focused on extending topology optimization design methods to high-frequency electromagnetic applications. She was a Postdoctoral Researcher with both the Electro-Science Laboratory, Ohio State University, Columbus, and the University of Michigan, from May 2003 to September 2005. She also coordinated with the Ceramic Research Group, Material Science and Engineering Department, University of Michigan, on the advanced fabrication of dielectric composites. She is currently an Assistant Professor with the Mechatronics Program, Sabanci University, Istanbul, Turkey. Her current research interests include the design, analysis, and fabrication of complex engineering systems, such as miniaturized electromagnetic, electromechanical, and biomedical devices and multidisciplinary design optimization.

Mehmet Keskinoz (M'98) received the M.S. and Ph.D. degrees from Carnegie Mellon University, Pittsburgh, PA, in 1997 and 2001, respectively.

In 2001, he joined the Electronics Engineering Program, Sabanci University Istanbul, Turkey, where he is currently an Associate Professor. His research interests include signal processing for wired and wireless communications, UWB communications, multiband OFDM UWB systems, wireless mesh networks, magnetic and optical data storage systems, distributed detection and data fusion for wireless sensor networks, Turbo and LDPC coding, synchronization, and digital watermarking.

Dr. Keskinoz is a member of the IEEE Communication Society, the IEEE Signal Processing Society, and the Optical Society of America. He was the recipient of a Turkish National Science Foundation research grant on distributed detection in wireless sensor networks and a Career Award on wireless mesh networks in August 2005. He is a co-Guest Editor of the IEEE COMMUNICATIONS MAGAZINE, January 2009 Special Issue on Advances in Signal Processing for Wireless and Wired Communications.