

Article

Sensor Selection via Maximizing Hybrid Bayesian Fisher Information and Mutual Information in Unreliable Sensor Networks

Qingli Yan ^{1,*} and Jianfeng Chen ^{2,*}

- School of Computer Science and Technology, Xi'an University of Posts & Telecommunications, Xi'an 710121, China
- ² School of Marine Science and Technology, Northwestern Polytechnical University, Xi'an 710012, China
- * Correspondence: gongchyy@163.com (Q.Y.); Chenjf@nwpu.edu.cn (J.C.)

Received: 22 January 2020; Accepted: 4 February 2020; Published: 7 February 2020



Abstract: The sensor selection problem is addressed for unreliable sensor networks. The Bayesian Fisher information (BFI) matrix, mutual information (MI) and their relationship are investigated under Gaussian mixture noise conditions. To overcome the flaw that the sensor selection methods based on either BFI matrix or MI could not provide coincident results, the multiple objective optimal (MOP) -based sensor selection approach is developed via minimizing the number of selected sensors while maximizing corresponding BFI matrix and MI. The variable weight decision making (VWDM) and technique for order of preference by similarity to ideal solution (TOPSIS) approaches are then proposed to find the candidate that can better trade off the cost and two performance metrics. Comparison results demonstrated that the proposed method can find a more informative sensor group, and ultimately, its overall localization performance outperforms the sensor selection methods based on BFI or MI.

Keywords: source localization; sensor selection; angle of arrival; multiple objective optimization; sensor networks

1. Introduction

Advances in sensor technology have made it possible to use a large number of sensors in various applications, such as environmental monitoring, battlefield surveillance, target localization and tracking [1,2], etc. Sensor selection is critical for saving energy to prolong the lifetime of sensor networks. A good sensor selection strategy needs to select most informative sensors to achieve a good balance between the localization accuracy and cost. In this paper, the sensor selection problem for angle of arrival (AOA)-based source localization is addressed, as AOA localization technology has been applied to many areas such as radar, sonar, wireless communications and indoor acoustic localization, to name but a few.

The sensor selection problem has been attracting much attention in the last decades [3–13]. Entropy and its variant mutual information (MI) are two popular performance metrics to design sensor selection methods. MI is a standard information quantity from the information theoretic point of a view. The MI between the predicted sensor observation and the current target location distribution was proposed to evaluate the expected information gain about the target location attributable to a sensor in [5,6]. The simple entropy-based heuristic for sensor selection is introduced in [7]. This method is computationally simpler than MI [5], but it works well only when the measurement noise is small. The maximum entropy fuzzy clustering was introduced to sensor selection for target tracking [8].

The Cramer-Rao lower bound (CRLB) (Bayesian CRLB if the priori distribution is known) provides a theoretical performance limit for an unbiased or asymptotically unbiased estimator, it thus is another



attractive metric to develop various sensor selection methods. For single target tracking, a subset of sensors was selected in a bearing-only sensor network to minimize the posteriori CRLB [9]. For time difference of arrival (TDOA)-based localization, the sensor selection method in non-line of sight (NLOS) condition was investigated in [10]. The global sensor selection method for AOA based localization was proposed via minimizing the trace of CRLB [11]. In [12,13], sensor selection methods for linear dynamical systems were proposed under correlated measurement noise condition and sensor selection approaches for non-linear measurement models were developed in [14,15].

All of the sensor selection works mentioned above are derived from either CRLB or MI. It has been demonstrated that they exhibit a good consistency when the estimation error of each individual sensor follows a Gaussian distribution. However, sensor failures, data loss, NLOS propagation or unexpected interference may impose uncertainty on sensor networks which result in the presence of unreliable measurements [16,17].

In the last decades, much attention has been devoted to using non-Gaussian noise model to model the sensor networks with uncertain observations. The Gaussian mixture noise has been applied to model the ambient noise in various applications (see e.g., [18,19] and the references therein). As illustrated in [20], the selection results based on MI and CRLB are different. MI is more easily influenced by uncertain probability of one sensor as its observation has insufficient information about the target; CRLB-based methods tend to select sensors which are close to the source even some of them having large uncertainties for received signal strength (RSS)-based target localization and tracking. In this work, we will investigate the selection results for AOA-based source localization.

This paper focuses on the sensor selection problem for AOA-based source localization in unreliable sensor networks. To select more informative sensors, we propose to incorporate both MI and Bayesian Fisher information (BFI) matrix, which is the inverse of Bayesian CRLB, into the selection scheme. In addition, as the number of selected sensors is usually unknown for practical applications, the best way is to select sensors that can trade off the localization performance and the cost. For this purpose, the number of selected sensors is also formulated as one objective to optimize. Thus, we have three objective functions to optimize: minimizing the number of selected sensors, maximizing the BFI and MI of the selected sensors. Obviously, the first one is conflict with the other two. The multi-objective evolutionary algorithm based on decomposition (MOEA/D) [21] is used to find the optimal solutions that can trade off these conflicting objective functions. Then, the decision-making method, variable weight decision making (VWDM) [22,23] and technique for order of preference by similarity to ideal solution (TOPSIS) [24,25], is proposed to find the final solution.

The rest of the paper is organized as follows. The measurement model is described in Section 2. The relationship between Fisher information (FI) and MI are derived in Section 3. The sensor selection method based on MOEA/D is proposed in Section 4. Section 5 provides simulation results and the conclusion are summarized in Section 6.

2. System Model

In this section, we review the recursive Bayesian estimation for target localization. The recursive Bayesian estimation is using expected posterior distribution to predict what the posterior distribution would look like if a simulated measurement of a new sensor is incorporated.

In the recursive Bayesian estimation for target localization and tracking [4], both the target location and the sensor observations are modeled as stochastic processes, and the posterior target location distribution conditioned on sensor observations is computed recursively from additional sensor observations step by step. Let **x** denote the target location random variable and its realization value, respectively. Let z_j denote the sensor observation. The posterior target location distribution is incrementally updated by one sensor observation at a time. When the recursive Bayesian estimation is applied to the target location, we can get that [4]:

$$f(\mathbf{x}|z_1,...,z_j) = Cf(z_j|\mathbf{x}, z_1,..., z_{j-1})f(\mathbf{x}|z_1,...,z_{j-1})$$
(1)

where *C* is normalization constant. When $z_1,..., z_j$ are conditionally independent with one another conditioned on **x**, the above equation is simplified to:

$$f(\mathbf{x}|z_1,\ldots,z_j) = Cf(z_j|\mathbf{x})f(\mathbf{x}|z_1,\ldots,z_{j-1})$$
(2)

For the AOA sensors, the observation is the estimated angle of each sensor which can be given by:

$$z_i = \theta_i(\mathbf{x}) + \eta_i \tag{3}$$

where $\theta_i(\mathbf{x}) = \tan^{-1}\left(\frac{y-y_i}{x-x_i}\right)$ represents the true angle, $\tan^{-1}(\mathbf{y})$ stands for the 4-quadrant arctangent. $\mathbf{x} = [x, y]^T$ is the target position, $\mathbf{s}_i = [x_i, y_i]^T$, i = 1, 2, ..., N, denotes the position of N sensors collecting angle measurements. η_i denotes the angle estimation error.

As mentioned above, the adverse environment may bring uncertainty to sensor networks. We follow [14,16] to model the uncertain scenario in which the probability density of function (PDF) of angle estimation error η_i is:

$$f(\eta_i) = pN(0,\sigma^2) + (1-p)N(\mu_o,\sigma_0^2)$$
(4)

where *p* is the reliable probability. $|\mu_0| >> 0$ and $\sigma_0^2 >> \sigma^2$.

3. Sensor Selection Criteria

3.1. BFI for Gaussian Mixture Noise

There are several different measures of the estimation error of the posterior target location distribution. One estimation error measure is the Bayesian CRLB of the target location which is the inverse of BFI.

In this section, the BFI is derived under Gaussian mixture noise. Let $\hat{\mathbf{x}}$ be an unbiased estimate of **x**, BFI satisfies the well-known inequality:

$$E\left\{\left[\hat{\mathbf{x}} - \mathbf{x}\right]\left[\hat{\mathbf{x}} - \mathbf{x}\right]^{\mathrm{T}}\right\} \ge \mathbf{J}^{-1}$$
(5)

where **J** is the BFI. It has been shown in [13] that, the BFI matrix consists of two parts: the information matrix obtained from the sensor measurements and the priori information matrix. Furthermore, under the assumption that the sensor measurements are conditionally independent on the given target information, the BFI matrix can be written as [13]:

$$\mathbf{J} = \sum_{i=1}^{N} \int \mathbf{J}_{i}^{s}(\mathbf{x}) f(\mathbf{x}) d\mathbf{x} + \mathbf{J}_{prior}$$

$$= \sum_{i=1}^{N} \mathbf{J}_{i} + \mathbf{J}_{prior}$$
(6)

where J_{prior} is the FI matrix of the priori information about target which typically comes from previous measurements or from other available measurements. Let $f(\mathbf{x})$ denotes the prior PDF of the target position distribution, J_{prior} can be expressed as:

$$\mathbf{J}_{prior} = -E \left\{ \frac{\partial}{\partial \mathbf{x}} \left(\frac{\partial \ln f(\mathbf{x})}{\partial \mathbf{x}} \right)^{\mathrm{T}} \right\}$$
(7)

Let J_i denote the standard FI of sensor s_i , it can be formulated as:

$$\mathbf{J}_i = \int \mathbf{J}_i^{\mathrm{s}}(\mathbf{x}) f(\mathbf{x}) d\mathbf{x}$$
(8)

where:

$$\mathbf{J}_{i}^{s}(\mathbf{x}) = \int \frac{1}{f(z_{i}|\mathbf{x})} \left(\frac{\partial f(z_{i}|\mathbf{x})}{\partial \mathbf{x}}\right) \left(\frac{\partial f(z_{i}|\mathbf{x})}{\partial \mathbf{x}}\right)^{\mathrm{T}} dz_{i}$$
(9)

$$\frac{\partial f(z_i|\mathbf{x})}{\partial \mathbf{x}} = \left\{ \frac{p}{\sqrt{2\pi\sigma}} \frac{z_i - \theta_i(\mathbf{x})}{\sigma^2} \exp\left(-\frac{(z_i - \theta_i(\mathbf{x}))^2}{2\sigma^2}\right) + \frac{1 - p}{\sqrt{2\pi\sigma_0}} \frac{z_i - \mu_0 - \theta_i(\mathbf{x})}{\sigma_0^2} \exp\left(-\frac{(z_i - \mu_0 - \theta_i(\mathbf{x}))^2}{2\sigma_0^2}\right) \right\} \frac{\partial \theta_i(\mathbf{x})}{\partial \mathbf{x}} \quad (10)$$

For AOA sensors, we have:

$$\frac{\partial \theta_i(\mathbf{x})}{\partial \mathbf{x}} = \begin{bmatrix} -\frac{\sin \theta_i(\mathbf{x})}{r_i} & \frac{\cos \theta_i(\mathbf{x})}{r_i} \end{bmatrix}^{\mathrm{T}}$$
(11)

where r_i is the distance between s_i and x. Let $p_1 = p$, $p_2 = 1 - p$, $\mu_1 = 0$, $\mu_2 = \mu_0$, $\sigma_1 = \sigma$, $\sigma_2^2 = \sigma_0^2$. Substituting Equations (10) and (11) into Equation (9), we can get:

$$\mathbf{J}_{i}^{s}(\mathbf{x}) = \begin{bmatrix} \frac{\sin^{2}\theta_{i}}{r_{i}^{2}} & -\frac{\sin\theta_{i}\cos\theta_{i}}{r_{i}^{2}} \\ -\frac{\sin\theta_{i}\cos\theta_{i}}{r_{i}^{2}} & \frac{\cos^{2}\theta_{i}}{r_{i}^{2}} \end{bmatrix} \int \rho(z_{i})dz_{i}$$
(12)

$$\rho(z_{i}) = \frac{\left\{\sum_{l=1}^{2} \frac{p_{l}}{\sqrt{2\pi\sigma_{l}}} \frac{z_{i}-\mu_{l}-\theta_{i}(\mathbf{x})}{\sigma_{l}^{2}} \exp\left(-\frac{(z_{i}-\mu_{l}-\theta_{i}(\mathbf{x}))^{2}}{2\sigma_{l}^{2}}\right)\right\}^{2}}{\sum_{l=1}^{2} \frac{p_{l}}{\sqrt{2\pi\sigma_{l}}} \exp\left(-\frac{(z_{i}-\mu_{l}-\theta_{i}(\mathbf{x}))^{2}}{2\sigma_{l}^{2}}\right)}$$
(13)

When ignoring prior information, it is well known that it is desirable for the nodes to be both close to the target and to provide good angular diversity by surrounding the target [11]. By inspection of Equation (12), it is clear that 2×2 BFI is positive semi-definite, and:

$$tr(\mathbf{J}) = \sum_{i=1}^{N} \mathbf{J}_{i}^{s}(\mathbf{x}) = \sum_{i=1}^{N} \frac{\int \rho(z_{i}) dz_{i}}{r_{i}^{2}}$$
(14)

Letting $\sum_{i=1}^{N} \frac{\int \rho(z_i) dz_i}{r_i^2} = M$, there exists two eigenvalues of J represented as:

$$\lambda_{\max} = \frac{1}{2}M(1+a), \lambda_{\min} = \frac{1}{2}M(1-a)$$
(15)

$$\det(\mathbf{J}) = \frac{1}{4}M^2(1-a^2) \le \frac{M^2}{4}$$
(16)

Equation (16) indicates that the range from source to sensors, angular diversity and bearing scaling factor determine BFIU. Due to the upper bound in Equation (16), it is expected that the selection approach will attempt to select nodes that both are close to the target and have a high sensing probability. In general, when prior information is available, the prior information is skewed to favor a certain direction; node selection methods will select sensors that reduce the error in the direction where it is high. For simplicity, we can use the maximum likelihood estimation of x as the actual target position to calculate the upper bound of Equation (16) to select nodes.

3.2. Mutual Information

Another estimation error measure is the Shannon entropy that measures the uncertainty of the posterior target location distribution. From the information theoretic point of view, sensors are tasked to observe the target in order to reduce the uncertainty about the target location distribution. One expression to denote the contribution of a sensor is MI. The greedy sensor selection method gradually reduces the uncertainty of the target location distribution by repeatedly selecting the currently

unused sensors with maximal MI. Given the distribution of the target state and the likelihood function of the sensor measurements, the MI of sensor can be written as [5]:

$$MI(z_i, \mathbf{x}) = H(z_i) - H(z_i|\mathbf{x})$$
(17)

$$H(z_i) = -\int f(z_i) \log f(z_i) dz_i$$

= $-\int_{\hat{\theta}_i} \left\{ \int_{\mathbf{x}} f(z_i | \mathbf{x}) f(\mathbf{x}) d\mathbf{x} \right\} \left\{ \log \left[\int_{\mathbf{x}} f(z_i | \mathbf{x}) f(\mathbf{x}) d\mathbf{x} \right] \right\} dz_i$ (18)

$$H(z_i|\mathbf{x}) = -\int f(\mathbf{x}) \left\{ \int f(z_i|\mathbf{x}) \log f(z_i|\mathbf{x}) dz_i \right\} d\mathbf{x}$$
(19)

3.3. The Relationship between Fisher Information and Mutual Information

In this section, we will demonstrate the relationship between FI and MI under Gaussian and non-Gaussian noise similar the work in [26]. The FI with respect to $\theta(\mathbf{x})$ is given as [20]:

$$J(\theta) = \int \left(\frac{\partial \ln f(z|\theta)}{\partial \theta}\right)^2 f(z|\theta) dz$$
(20)

Here, we assume the additive noise with density $q(\cdot)$, we can write $f(z|\theta) = q(z - \theta)(f(z|\mathbf{x}) = q(z - \theta(\mathbf{x})))$. In this case, $J(\theta)$ becomes independent of θ and thus can be rewritten as:

$$J[q] = \int \left(\frac{\partial \ln q(\eta)}{\partial \eta}\right)^2 q(\eta) d\eta$$
(21)

This quantity is referred to as FI of a random variable with respect to a scalar translation parameter and Equation (18) is a constant. Conceptually, the constant J[q] summarizes the total local dispersion of a distribution. Similarly, the Shannon entropy $H(z|\theta)(H(z|\mathbf{x}))$ is also independent of $\theta(\mathbf{x})$ and identical to the noise entropy:

$$H[q] = \int q(\eta) \ln q(\eta) d\eta$$
(22)

Stam's inequality specifies the relation between FI and Shannon entropy as following: for a given amount of FI, the Shannon entropy of a continuous random variable is minimized if and only if the variable is Gaussian distribution. When the variance of a Gaussian random variable is 1/J[q], Stam's inequality implies that:

$$H[q] \ge \frac{1}{2} \ln \left(\frac{2\pi e}{J[q]} \right) \tag{23}$$

Define:

$$D_0 = H[q] - \frac{1}{2} \ln \left(\frac{2\pi e}{J[q]}\right) \tag{24}$$

note that $D_0 > 0$ for distribution with lighter tails than a Gaussian, as well as for distributions that are asymmetric.

Because the transfer function is invertible, $MI(\mathbf{x}, z) = MI(\theta, z)$ even though $H(\mathbf{x}) \neq H(\theta)$ [26]. Thus we can write MI as:

$$MI(\theta, z) = H(z) - \int f(\theta)H(z|\theta)d\theta$$
(25)

As $H(z|\theta) = H[q]$, thus, Equation (25) can be written as:

$$MI(\theta, z) = H(z) - \int f(\theta) H[q] d\theta$$
(26)

From Equations (24) and (26), we can get that:

$$\mathrm{MI}(\theta, z) = \mathrm{H}(z) - \int f(\theta) \left(D_0 + \frac{1}{2} \ln \left(\frac{2\pi e}{\mathrm{J}[q]} \right) \right) d\theta \tag{27}$$

Because $J[q] = J(\theta)$, Equation (27) can be rewritten as:

$$MI(\theta, z) = H(z) - \int f(\theta) \left(D_0 + \frac{1}{2} \ln\left(\frac{2\pi e}{J(\theta)}\right) \right) d\theta$$

= $H(z) - \int f(\theta) \frac{1}{2} \ln\left(\frac{2\pi e}{J(\theta)}\right) d\theta - \int f(\theta) D_0 d\theta$ (28)
= $[H(z) - H(\theta)] + H(\theta) - \int f(\theta) \frac{1}{2} \ln\left(\frac{2\pi e}{J(\theta)}\right) d\theta - D_0$

Using the formulas for a change of variables, it is straightforward to verify that:

$$H(\theta) - \int f(\theta) \frac{1}{2} \ln\left(\frac{2\pi e}{J(\theta)}\right) d\theta = H(\mathbf{x}) - \int f(\mathbf{x}) \frac{1}{2} \ln\left(\frac{(2\pi e)^2}{\det(J(\mathbf{x}))}\right) d\mathbf{x}$$
(29)

Then, Equation (28) can be rewritten as:

$$MI(\theta, z) = MI(\mathbf{x}, z)$$

= $[H(z) - H(\theta)] + H(\mathbf{x}) - \int f(\mathbf{x}) \frac{1}{2} \ln\left(\frac{(2\pi e)^2}{\det(J(\mathbf{x}))}\right) d\mathbf{x} - D_0$ (30)
= $I_{Fisher} + C_0 - D_0$

where:

$$\mathbf{I}_{Fisher} = \mathbf{H}(\mathbf{x}) - \int f(\mathbf{x}) \frac{1}{2} \ln \left(\frac{(2\pi e)^2}{\det(\mathbf{J}(\mathbf{x}))} \right) d\mathbf{x}$$

 $C_0 = H(z) - H(\theta)$. It illustrate that the degree to which MI is well approximated by FI (I_{*Fisher*}) depends on the values of C_0 and D_0 . Both terms are nonnegative and quantify two very different aspects of the noise: $C_0 = H(\theta + \eta) - H(\theta)$ is monotonic in the magnitude of the noise. While D_0 represents the nongaussianity of the noise. From Equation (30), we can obtain the following conclusions:

- If the noise is Gaussian, $D_0 = 0$. And $C_0 \ge 0$ as the additive noise would increase the entropy, thus $MI(\mathbf{x}, z) \ge I_{Fisher}$. That is, if and only if the noise is Gaussian, I_{Fisher} is guaranteed to denote a lower bound on MI.
- As Stam's inequality tells us that $D_0 \ge 0$, we can get that $MI(\mathbf{x}, z) \le I_{Fisher} + C_0$. Specially, in the case of vanishing noise, $C_0 \rightarrow 0$, and it follows that $MI(\mathbf{x}, z) \le I_{Fisher}$. Thus, I_{Fisher} generally represents an upper bound on MI in the small noise regime.
- Only when the noise entropy goes to zero and the noise converges to a Gaussian at the same time, i.e., C₀ → 0, D₀ → 0, MI(**x**, z) = I_{*Fisher*}.

4. The Proposed Sensor Selection Method

Considering the problem mentioned above, we would like to take all performance evaluation metrics into account. In addition, the number of selected sensors is unknown in practice, which requires one to consider the localization performance and the cost for different applications. In this paper, we want to optimize multiple conflicting objectives at the same time: minimizing the number of selected sensors and maximizing the BFI matrix and MI of these sensors. In order to formulate the problem into a multiple objective optimal (MOP) framework, the first objective function is transformed into a maximization problem of the gap between the number of sensors to be selected and the number

of sensors. Let $\boldsymbol{\alpha} = [\alpha_1, \alpha_2, ..., \alpha_N]^T \in \{0, 1\}^N$ be the binary selection vector. The MOP scheme can be formulated as:

$$\max_{\boldsymbol{\alpha}} F(\boldsymbol{\alpha}) = \{f_1(\boldsymbol{\alpha}), f_2(\boldsymbol{\alpha}), f_3(\boldsymbol{\alpha})\}$$

s.t.
$$\boldsymbol{\alpha} \in \{0, 1\}^N$$
 (31)

where:

$$f_1(\boldsymbol{\alpha}) = \left(N - \sum_{i=1}^N \alpha_i\right) / N \tag{32}$$

$$f_2(\boldsymbol{\alpha}) = \frac{\sum_{i=1}^{N} \alpha_i M I_i}{\sum_{i=1}^{N} M I_i}$$
(33)

$$f_{3}(\boldsymbol{\alpha}) = \frac{\det\left(\sum_{i=1}^{N} \alpha_{i} \mathbf{J}_{i} + \mathbf{J}_{prior}\right)}{\det\left(\sum_{i=1}^{N} \mathbf{J}_{i} + \mathbf{J}_{prior}\right)}$$
(34)

where Equations (33) and (34) represent the normalized MI and BFI of selected sensors. As the sensor selection policy Equation (31) consists of three conflicting objective functions, any single solution cannot optimize them at the same time. However, MOP methods are proposed to find a set of solutions that can trade off the objectives.

Many multi-objective methods have been proposed in last decades, such as non-dominated sorting genetic algorithm (NSGA-II) [27], MOEA/D [21]. The MOEA/D method has lower computational complexity than NSGA-II, and it outperforms or performs similar to NSGA-II. In this paper, we will use MOEA/D to solve the optimization problem of Equation (31). The generalized MOP can be formulated as [21,27]:

$$\max_{\boldsymbol{\alpha}} F(\boldsymbol{\alpha}) = \{f_1(\boldsymbol{\alpha}), \dots, f_k(\boldsymbol{\alpha})\}$$

s.t. $\boldsymbol{\alpha} \in \Omega$ (35)

where Ω is the decision space. Assume α^1 and α^2 are two solutions of Equation (35), $F(\alpha^1)$ dominates $F(\alpha^2)$ if and only if $f_i(\alpha^1) \ge f_i(\alpha^2)$, $\forall i \in \{1, 2, 3\}$, and there exists one index j such that $f_j(\alpha^1) > f_j(\alpha^2)$. If there is no $\alpha \in \{0, 1\}^N$ such that $F(\alpha)$ dominates $F(\alpha^*)$, α^* is called a Pareto-optimal point and $F(\alpha^*)$ is a Pareto-optimal objective vector. That is to say, any improvement in one objective at a Pareto-optimal point must lead to deterioration in at least one other objective. The set of all Pareto-optimal points is called the Pareto set and the set of all Pareto-optimal objective vectors is the Pareto front (PF) [21–25].

MOEA/D decomposes the MOP problem into scalar optimization subproblems. It solves these subproblems in a collaborative way. Any decomposition approach developed in the area of mathematical programming can be incorporated into the framework of MOEA/D. In this paper, the scalar optimization subproblems based on classical Tchebycheff approach is given by [21]:

minimize
$$g^{te}(x|\lambda, z^*) = \max_{1 \le i \le m} \{\lambda_i | f_i(x) - z_i^* | \}$$

s.t. $x \in \Omega$ (36)

where $\lambda = [\lambda_1, \dots, \lambda_m]^T$ is a weight vector, and $\sum_{i=1}^m \lambda_i = 1, z^* = [z_1^*, \dots, z_m^*]^T$ is the reference point, i.e.,

 $z_i^* = \max\{f_i(x)|x \in \Omega\}$ for each i = 1, ..., m. For each Pareto optimal point x^* there exists a weight vector λ such that x^* is the optimal solution of Equation (31) and each optimal solution of Equation (36) is a Pareto optimal solution of Equation (31). Therefore, one is able to obtain different Pareto optimal solutions by altering the weight vector. The details about MOEA/D can be found in [21].

4.1. Performance Metrics for MOEA/D

The hypervolume indicator (I_H) [21,25] is used in our study to illustrate the efficiency of the MOEA/D in sensor selection problem. Let $y^* = (y_1^*, \ldots, y_m^*)$ be a point in the objective space which is dominated by any Pareto optimal vectors. Let P be the obtained approximation to the PF in the objective space. Then, the I_H value of P (with regard to y^*) is the volume of the region which is dominated by P and dominates y^* . The higher the hypervolume, the better the approximation. In our experiments, $y^* = (f_1^{\min}, f_2^{\min}, f_3^{\min})$ for the three objective ones, where f_i^{\min} indicates the minimum value of the ith objective in the obtained non-dominated set.

4.2. Select Solution from the Pareto-Optimal Solution

It is necessary to emphasize that the final aim of sensor selection is to obtain a single optimal solution. Since the optimization result of a MOP algorithm is a set of non-dominated solutions, the proper solution should be selected based on specific applications. There are many methods that one can employ in selecting a single solution from Pareto-optimal front. For the sensor selection problem proposed above, we need to evaluate three attributes to select the better candidate.

Let $\mathbf{w} = (w_1, w_2, ..., w_m)$ and $\mathbf{x} = (x_1, x_2, ..., x_m)$ be a constant weight vector and state value vector, a common decision making function is use $A = \sum w_j x_j$ to evaluate the alternatives. However, the constant weight vector is not work well in some cases. For example, if all factors are equality important, i.e., $\mathbf{w} = (w_1, w_2) = (1/3, 1/3, 1/3)$, Hence, the weighted average synthesis expression is $A = 1/3x_1 + 1/3x_2 + 1/3x_3$, Let $\mathbf{x}_1 = (0.6, 0.2, 0.7)$, $\mathbf{x}_2 = (0.6, 0.8, 0.1)$, $\mathbf{x}_3 = (0.6, 0.5, 0.4)$. Then we would expect $A(\mathbf{x}_1) \ll A(\mathbf{x}_3)$, $A(\mathbf{x}_2) \ll A(\mathbf{x}_3)$, however, known by the decision making function, we can find that $A(\mathbf{x}_1) = A(\mathbf{x}_2) = A(\mathbf{x}_3)$. This result contradicts with the expectation. Thus, the decision making based on constant weight vector has its limitations. To overcome this problem, Wang [22] proposed the variance weight method. Since then, a lot of work has been done for the variable weight decision making [23]. It emphasizes the weights should change with the state values of factors. According to the change trend of weight, the variable weight decision making mechanism can be divided punishment mechanism, incentive mechanism and mixed mechanism. The basic definition of variable weight theory is summarized as follows [22,23]:

Definition 1. A mapping $w = (w_1, w_2, ..., w_m)$ from $[0, 1]^n \to [0, 1]^n$, and $w_j : [0, 1]^n \to [0, 1]$, $(x_1, x_2, ..., x_m) \mapsto w_j(x_1, x_2, ..., x_m)$ is a variable weight vector with penalty for j = 1, 2, ..., n, if w satisfies the following properties [22]:

- (1) $\sum_{j=1}^{m} w_j(x_1, x_2, \dots, x_m) = 1.$
- (2) the function $w_i(x_1, x_2, ..., x_m)$ is continuous with respect to every variable x_i .
- (3) the function $w_j(x_1, x_2, ..., x_m)$ is monotonically decreasing (for punishment mechanism) or increasing (for incentive mechanism) with respect to the variable x_i .

Definition 2. A mapping $S = (S_1, S_2, ..., S_m)$ from $[0,1]^n \rightarrow [0,1]^n$, and $S_j : [0,1]^n \rightarrow [0,1]$, $(x_1, x_2, ..., x_m) \mapsto S_j(x_1, x_2, ..., x_m)$ is a state variable weight vector with penalty for j = 1, 2, ..., m, if S satisfies the following properties [22,23]:

- (1) the function $S_i(x_1, x_2, ..., x_m)$ is continuous with respect to every variable x_i .
- (2) for punishment mechanism $x_i \ge x_j \Rightarrow S_i(x_i, x_2, \dots, x_m) \le S_j(x_i, x_2, \dots, x_m)$; for incentive mechanism $x_i \ge x_j \Rightarrow S_i(x_i, x_2, \dots, x_m) \ge S_j(x_i, x_2, \dots, x_m)$
- (3) The mapping $w: [0,1]^n \to [0,1]^n$ given by

$$\mathbf{w}(w_1, w_2, \dots, w_m) = \frac{\mathbf{w} \cdot \mathbf{S}(x_1, x_2, \dots, x_m)}{\sum\limits_{j=1}^m w_j S_j(x_1, x_2, \dots, x_m)}$$
(37)

is a variable weight vector, where $\mathbf{w} = (w_1, w_2, \dots, w_m)$ is a constant weight vector, and $\mathbf{w} \cdot \mathbf{S}(x_1, x_2, \dots, x_m) = (w_1 S_1(x_1, x_2, \dots, x_m), \dots, w_m S_m(x_1, x_2, \dots, x_m)).$

Since the results of MOP are a set of Pareto solutions, several candidates which have the same value of the first state will be found. Thus the first step of finding the final solution is to select a sate vector from several candidates with the same first state value. The procedure can be summarized as follows [22,23]:

Step 1: Let the constant weight vector $\mathbf{w} = (w_1, w_2, ..., w_m)$, (m = 3 for sensor selection problem). Without any prior knowledge, one can assume $\mathbf{w} = (w_1, w_2, w_3) = (1/3, 1/3, 1/3)$.

Step 2: Construct the expression of state variable weight vector S. Analysis the meaning of three objective function, we can find that the better combination of selected sensors should have higher MI and BFI. We prefer to select the solution with large objective function values, but neglect the one with very small attribute value. Thus, weights need to make corresponding adjustments to the attribute values of indicators, punish the index weights with low attribute values, and encourage the index weights with high attribute values. The state variable weight vector can be expressed as:

$$S_j(\mathbf{x}_i) = \begin{cases} e^{-\alpha(x_{ij} - \overline{x}_j)^2}, & x_{ij} \le \overline{x}_j \\ e^{\alpha(x_{ij} - \overline{x}_j)^2}, & x_{ij} > \overline{x}_j \end{cases}$$
(38)

where $\alpha \ge 0$ denotes penalty factor, the bigger α is, the bigger the punishment range is. According to [23], α can be determined by:

$$A = \frac{1}{m} \sum_{j=1}^{m} w_j / [w_j + (1 - w_j)e^{-\alpha}]$$
(39)

 $A \in [0, 1)$ is the adjust level, usually, A is usually set to be 0.5. \overline{x}_j is the mean of *j*-th attribute, which denotes all the factors less than \overline{x}_j will be punished.

Step 3: Calculate the state variable weight matrix for all candidates:

$$\mathbf{w}_{i}(w_{1}, w_{2}, \dots, w_{m}) = \frac{\mathbf{w} \cdot \mathbf{S}_{i}(x_{1}, x_{2}, \dots, x_{m})}{\sum\limits_{j=1}^{m} w_{j} S_{i,j}(x_{1}, x_{2}, \dots, x_{m})}$$
(40)

4.3. TOPSIS

In this paper, the TOPSIS will be used to select the final solution. According to this technique, the chosen optimal solution should have the smallest Euclidean distance from the ideal solution and also the largest Euclidean distance from the negative-ideal solution. The ideal solution is a combination of the best value of each objective. In contrast, negative-ideal solution is a combination of the worst value of each objective.

Before introduce the method, common symbols are defined as follows: f_{ij} is the *i*th value of the *j*-th objective in the objective matrix, F_{ij} is the normalized value of f_{ij} , v_{ij} is the weighted value of F_{ij} and w_{ij} is the weight obtained from Equation (40). Assume *M* solutions have been found by MOP. The TOPSIS algorithm can be described as below [24,25]:

Step 1. Construct normalized objective matrix with *M* rows and three columns by:

$$F_{ij} = \frac{f_{ij}}{\sqrt{\sum_{i=1}^{m} f_{ij}^2}} \tag{41}$$

Step 2. Construct weighted normalized objective matrix by multiplying each column with its weight w_i :

$$v_{ij} = F_{ij} w_{ij} \tag{42}$$

Step 3. Calculate the Euclidean distance between each solution and the ideal and negative-ideal solution:

$$D_{i+} = \sqrt{\sum_{j=1}^{3} (v_{ij} - \max(\mathbf{v}_j)), i = 1, \dots, M}$$
(43)

$$D_{i-} = \sqrt{\sum_{j=1}^{3} (v_{ij} - \min(\mathbf{v}_j))}, i = 1, \dots, M$$
(44)

Step 4. Calculate the closeness of each optimal solution:

$$C_i = \frac{D_{i-}}{D_{i-} + D_{i+}}, i = 1, \dots, M$$
(45)

The optimal solution having the largest C_i is the final solution.

5. Simulations

In this section, we will use simulation results to illustrate the effectiveness of the proposed sensor selection method. N sensors are deployed in the interested area to estimate source location. Consider that 25 sensors with known reliable probability are uniformly deployed in the $100 \times 100 \text{ m}^2$ detection area as shown in Figure 1. In the current work, we assume that the sensing probabilities of the sensors are already known to the fusion center. In the literature [28], the estimation of the detection probabilities are studied. Also, as indicated in [29], the sensing probabilities can be derived from historical data. Since these probabilities are context and scenario dependent, we do not study their estimation specifically in this paper and leave it as a future research topic. Generally, if the sensors around the source have higher reliable probabilities compared to other sensors, it is highly likely that the algorithm will select those sensors owing to both larger estimated accuracy of angles and shorter distances between source and sensors. Our interest is in considering more challenging cases to test the performance of our algorithm. Similar with [20], we assume that the sensors around the source have relatively low reliable probabilities as shown in the Figure 1. We set $\sigma^2 = 1^\circ$, $\sigma_0^2 = 30^\circ$. The source randomly appears in the area. The priori state of the target follows a uniform distribution limited in a square area with length H.

As shown in [21], MOEA/D method runs much faster than NSGA-II under same conditions. In this section, we use the Hypervolume Indicator (I_H) to observe the convergence and distribution of PF. In MOEA/D, T is set to be 25. The population size N in both MOEA/D and NSGA-II is set to be 500. Both algorithms run 30 times independently, and each run stops after 500 generations. For both methods, the genetic operators are the one point crossover operator and the standard mutation operator, the crossover probability 1 while mutation operator probability is 1/N, respectively.

The evolution of the average I_H values with number of generations is plotted in Figure 2. It is evident that MOEA/D outperforms NSGA-II in both convergence speed and the quality of their final solution set.



Figure 1. The placement of 25 sensors. The black circles denote sensors, integer numbers on the right of the sensors are the index of sensors. Floating numbers below the sensor are reliable probabilities.



Figure 2. The evolution of hypervolume Indicator of MOEA/D and NSGA-II.

We compare the sensor selection results of different methods when nine sensors are selected. The convex optimization method proposed in [14] and the greedy heuristic approach developed in [5] are applied to select sensors. Figure 3 plots the selection results for two different source locations. It can be seen that the sensors selected by convex optimization method are closer to the target location even when they have low reliable probabilities, while the MI-based method always select sensors with large reliable probabilities. This is consistent with the result in [20] which is observed for RSS based target localization. The proposed method, however, can select sensors with relatively high reliable probabilities but not very far away from the source location.

100

80

60

40

20

0

100

80

y (m)





Figure 3. The sensor selection results using MI, BFI, and MOEA/D-SS. (**a**) Sensor selection results for the source (T1) close to the center. (**b**) Sensor selection results for the source (T2) close to the center.

To illustrate the advantage combing two metrics (BFI and MI), we will use MOEA/D to solve three objective optimal problem (MOP3) and two objective optimal problem (MOP2), respectively. The MOP2 can be formulated as:

$$\max_{\alpha} F(\alpha) = \{f_1(\alpha), f_2(\alpha)\}$$

s.t. $\alpha \in \{0, 1\}^N$ (46)

$$\max_{\alpha} F(\alpha) = \{f_1(\alpha), f_3(\alpha)\}$$

s.t. $\alpha \in \{0, 1\}^N$ (47)

From now on, we label Equations (31), (46) and (47) as MOP3-BFI, MOP2-BFI and MOP2-MI, respectively. And MOP3-BFIU denotes the multiple objective problems when using BFIU to replace BFI in Equation (31). The sequential importance resampling (SIR) particle filter [30,31] is then used to

achieve source localization. The initial Ns = 1000 particles are drawn from $f(x_0)$. The root mean square error (RMSE) of 1000 Monte Carlo runs is used to measure errors between the true source location and the estimations.

Figure 4 plots the RMSEs of the four compared methods as the number of selected sensors is increased. We can observe that MOP3-based sensor selection methods can improve localization performance compared to MOP2-based methods. That is to say, the proposed method using both BFI and MI as objectives has advantages over only one of them used. In addition, we can also observe that RMSE is decreased obviously with the number of sensors increasing in Figure 4a, while this phenomenon is not shown in Figure 4b. As we can see from Figure 4b, RMSE is decreased first, and then it goes up with the number of selected sensor increasing. This mainly because the second source is close to the edge of sensor networks, sensors with large distance and low reliable probabilities will have negative effect on localization accuracy.



Figure 4. RMSE versus the number of selected sensors using MOP2-MI, MOP2-BFI, MOP3-MI and MOP3-BFI. (a) RMSE vs. the number of sensors for T1. (b) RMSE vs. the number of selected sensors for T2.

As discussed in Section 2, σ_0 represents the interference which influences on the localization performance directly. Figure 5 plots the RMSE when nine sensors are selected as σ_0 changes. We can see that the proposed method can improve the localization performance when a fixed number of sensors are selected.



Figure 5. RMSE versus σ_0 when 9 sensors are selected. (a) RMSE vs. σ_0 for T1. (b) RMSE vs. σ_0 for T2.

As stated in Section 4.2, the constant weight of different objectives will determine the select result directly. In this section, we will illustrate that how the constant weight influences the selection results. We set the weight of the first objective function be the same for MOP3 and MOP2 based methods. For MOP3, the left two objective functions have the same weight.

As shown in Figure 6, the larger w_1 is, the fewer sensors are selected. This is because large w_1 means that the cost is more important than performance. Thus, the performance is sacrificed for saving cost. We can further observe that MOP3 based methods usually select more sensors than MOP2 based methods under the same weight. As a result, the localization error of MOP3 is lower than MOP2. Furthermore, MOP3-BFIU performs better than MOP3-BFI in localization accuracy.



Figure 6. RMSE and number of selected sensors versus w_1 . (a) The number of sensors vs constant weight w_1 for T1. (b) RMSE vs constant weight w_1 for T1. (c) The number of sensors vs constant weight w_1 for T2. (d) RMSE vs constant weight w_1 for T2.

For different source locations shown in Figure 7, Figure 8 plots the number of selected sensors and corresponding RMSE. We can observe that MOP3-BFI selects more sensors than MOP3-BFIU but having similar localization performance. For MOP2-BFI and MOP2-MI, it is hard to say which is one better to use as performance metrics to select sensors.



Figure 7. The placement of 25 sensors. The black circles denote sensors, integer numbers on the right of the sensors are the index of sensors. Floating numbers below the sensor are reliable probabilities. The diamonds represent positions of $[a_1, b_1]$.



Figure 8. The number of selected sensors and corresponding localization performance for different source locations. (a) The number of selected sensors vs different sources. (b) RMSE vs different sources.

6. Conclusions

In this paper, we propose a novel sensor selection scheme for AOA-based source localization in unreliable sensor networks. The relationship between FI and MI is investigated; it reveals that they have a good consistence only for Gaussian noise. By transforming multiple performance metrics into scale-equivariant functions, the MOEA/D method is proposed to find a set of Pareto optimal solutions. Then, VWDM and TOPSIS are proposed to select the final selection result. The simulation results show that the MOP3-based method can select more informative sensors which can provide better localization accuracy than MOP2-based methods; and different numbers of sensors can be selected by allocating different weight vectors based on our preference.

Author Contributions: Q.Y. and J.C. proposed the algorithm, and Q.Y. performed the simulations and wrote the paper. J.C. improved this paper. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by NSFC-Zhejiang Joint Fund for the Integration of Industrialization and Information with grant number U1609204.

Conflicts of Interest: The authors declare no conflict of interest.

References

- 1. Sayed, A.H.; Tarighat, A.; Khajehnouri, N. Network-Based wireless location: Challenges faced in developing techniques for accurate wireless location information. *IEEE Signal Process. Mag.* 2005, 22, 24–40. [CrossRef]
- Bakr, M.A.; Lee, S. A general framework for data fusion and outlier removal in distributed sensor networks. In Proceedings of the 2017 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI), Daegu, Korea, 16–18 November 2017.
- 3. Hintz, K.J. A measure of the information gain attributable to cueing. *IEEE Trans. Syst. Man Cybern.* **1991**, 21, 434–442. [CrossRef]
- 4. Zhao, F.; Shin, J.; Reich, J. Information-driven dynamic sensor collaboration. *IEEE Signal Process. Mag.* 2002, 19, 61–72. [CrossRef]
- 5. Hoffmann, G.M.; Tomlin, C.J. Mobile sensor network control using mutual information methods and particle filters. *IEEE Trans. Autom. Control.* **2010**, *55*, 32–47. [CrossRef]
- 6. Zhao, W.; Han, Y.; Wu, H. Weighted Distance Based Sensor Selection for Target Tracking in Wireless Sensor Networks. *IEEE Signal Process. Lett.* **2009**, *16*, 647–650. [CrossRef]
- Wang, H.; Yao, K.; Pottie, G.; Estrin, D. Entropy-based sensor selection heuristic for target localization. In Proceedings of the 3rd International Symposium on Information Processing in Sensor Networks, Berkeley, CA, USA, 26–27 April 2004; pp. 36–45.
- 8. Guo, J.; Yuan, X.; Han, C. Sensor selection based on maximum entropy fuzzy clustering for target tracking in large-scale sensor networks. *IET Signal Process.* **2017**, *11*, 613–621. [CrossRef]
- Zuo, L.; Niu, R.; Varshney, P.K. Posterior Crlb Based Sensor Selection for Target Tracking in Sensor Networks. In Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Honolulu, HI, USA, 15–20 April 2007.
- 10. Zhao, Y.; Li, Z.; Hao, B.; Jia, S. Sensor Selection for TDOA-Based Localization in Wireless Sensor Networks With Non-Line-of-Sight Condition. *IEEE Trans. Veh. Technol.* **2019**, *68*, 9935–9950. [CrossRef]
- 11. Kaplan, L.M. Global node selection for localization in a distributed sensor network. *IEEE Trans. Aerosp. Electron. Syst.* **2006**, *42*, 113–135. [CrossRef]
- 12. Zhang, H.; Ayoub, R.; Sundaram, S. Sensor selection for Kalman filtering of linear dynamical systems: Complexity, limitations and greedy algorithms. *Automatica* **2017**, *78*, 202–210. [CrossRef]
- 13. Liu, S.; Chepuri, S.P.; Fardad, M.; Maşazade, E.; Leus, G.; Varshney, P.K. Sensor Selection for Estimation with Correlated Measurement Noise. *IEEE Trans. Signal Process.* **2016**, *64*, 3509–3522. [CrossRef]
- 14. Chepuri, S.P.; Leus, G. Sparsity-promoting sensor selection for non-linear measurement models. *IEEE Trans. Signal Process* **2015**, *63*, 684–698. [CrossRef]
- 15. Wang, Z.; Shen, X.; Wang, P.; Zhu, Y. The Cramér–Rao Bounds and Sensor Selection for Nonlinear Systems with Uncertain Observations. *Sensors* **2018**, *18*, 1103. [CrossRef] [PubMed]
- 16. Yan, Q.; Chen, J.; Ottoy, G.; De Strycker, L. Robust AOA based acoustic source localization method with unreliable measurements. *Signal Process.* **2018**, *152*, 13–21. [CrossRef]
- 17. Yan, Q.; Chen, J.; Strycker, L.D. An Outlier Detection Method Based on Mahalanobis Distance for Source Localization. *Sensors* **2018**, *18*, 2186. [CrossRef] [PubMed]
- Kozick, R.J.; Blum, R.S.; Sadler, B.M. Signal processing in non-Gaussian noise using mixture distributions and the EM algorithm. In Proceedings of the Conference Record of the Thirty-First Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA, USA, 2–5 November 1997.
- 19. Madadi, Z.; Anand, G.V.; Premkumar, A.B. Three-dimensional localization of multiple acoustic sources in shallow ocean with non-Gaussian noise. *Digit. Signal Process.* **2014**, *32*, 85–99. [CrossRef]
- 20. Cao, N.; Choi, S.; Masazade, E.; Varshney, P.K. Sensor Selection for Target Tracking in Wireless Sensor Networks With Uncertainty. *IEEE Trans. Signal Process.* **2016**, *64*, 5191–5204. [CrossRef]
- 21. Zhang, Q.; Li, H. MOEA/D: A multiobjective evolutionary algorithm based on decomposition. *IEEE Trans. Evol. Comput.* **2007**, *11*, 712–731. [CrossRef]
- 22. Zeng, W.; Li, D.; Wang, P. Variable Weight Decision Making and Balance Function Analysis Based on Factor Space. *Int. J. Inf. Technol. Decis. Mak.* **2016**, *15*, 999–1014. [CrossRef]
- 23. Yu, G.F.; Fei, W.; Li, D.F. A Compromise-Typed Variable Weight Decision Method for Hybrid Multiattribute Decision Making. *IEEE Trans. Fuzzy Syst.* **2018**, *27*, 861–872. [CrossRef]

- 24. Behzadian, M.; Otaghsara, S.K.; Yazdani, M.; Ignatius, J. A state-of the-art survey of TOPSIS applications. *Expert Syst. Appl.* **2012**, *39*, 13051–13069. [CrossRef]
- 25. Lai, Y.J.; Liu, T.Y.; Hwang, C.L. Topsis for MODM. Eur. J. Oper. Res. 1994, 76, 486–500. [CrossRef]
- 26. Wei, X.X.; Stocker, A.A. Mutual Information, Fisher Information, and Efficient Coding. *Neural Comput.* **2015**, 28, 305–326. [CrossRef] [PubMed]
- 27. Deb, K.; Pratap, A.; Agarwal, S.; Meyarivan, T.A.M.T. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Trans. Evol. Comput.* **2002**, *6*, 182–197. [CrossRef]
- Kanungo, T.; Mount, D.M.; Netanyahu, N.S.; Piatko, C.D.; Silverman, R.; Wu, A.Y. An efficient k-means clustering algorithm: Analysis and implementation. *IEEE Trans. Pattern Anal. Mach. Intell.* 2002, 24, 880–892. [CrossRef]
- 29. Oymak, O. Sample Size Determination for Estimation of Sensor Detection Probabilities based on a Test Variable. Ph.D. Thesis, Naval Postgrad—Uate School, Monterey, CA, USA, 2007.
- 30. Nahi, N.E. Optimal recursive estimation with uncertain observation. *IEEE Trans. Inf. Theory* **1969**, *15*, 457–462. [CrossRef]
- 31. Arulampalam, M.S.; Maskell, S.; Gordon, N.; Tim, C. A tutorial on particle filters for online nonlinear/ non-gaussian Bayesian tracking. *IEEE Trans. Signal Process* **2002**, *50*, 174–188. [CrossRef]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).