Title

Information-Theoretic Approaches for Sensor Selection and Placement in Sensor Networks for Target Localization and Tracking

Authors

Hanbiao Wang
Kung Yao
Deborah Estrin

Publication Date

2005

Information-Theoretic Approaches for Sensor Selection and Placement in Sensor Networks for Target Localization and Tracking

Hanbiao Wang, Kung Yao, and Deborah Estrin

***Abstract*— In this paper, we describes the information-theoretic approaches to sensor selection and sensor placement in sensor networks for target localization and tracking. We have developed a sensor selection heuristic to activate the most informative candidate sensor for collaborative target localization and tracking. The fusion of the observation by the selected sensor with the prior target location distribution yields nearly the greatest reduction of the entropy of the expected posterior target location distribution. Our sensor selection heuristic is computationally less complex and thus more suitable to sensor networks with moderate computing power than the mutual information sensor selection criteria. We have also developed a method to compute the posterior target location distribution with the minimum entropy that could be achieved by the fusion of observations of the sensor network with a given deployment geometry. We have found that the covariance matrix of the posterior target location distribution with the minimum entropy is consistent with the Cramer-Rao lower bound (CRB) of the target location estimate. Using the minimum entropy of the posterior target location distribution, we have characterized the effect of the sensor placement geometry on the localization accuracy.**

***Index Terms*— information theory, sensor selection, sensor placement, sensor networks, target localization and tracking.**

I. Introduction

The emerging sensor networks could revolutionize a wide range of applications including target localization and tracking [1]. Multi-sensor data fusion is one of the key technologies to exploit the huge potential of sensor networks [2]. Information-theoretic concepts not only provide guidance to minimize the consumption of sensor resources for a given information gain requirement through selective sensor activation but also provide guidance to maximize the information gain of a given set of sensors through intelligent sensor configuration. Information-theoretic sensor management has been shown to be able to greatly improve the cost-effectiveness of multi-sensor data fusion [3], [4], [5], [6], [7], [8], [9], [10].

The existing information-theoretic sensor selection approaches are not optimized for computational complexity required by the moderate/low computational powers of sensor networks. In this paper, we describe a sensor select heuristic that is nearly as effective as the mutual information based sensor selection in the sense that the selected sensor observation results in the maximum average information gain. Our sensor selection heuristic is computationally much less complex and thus more suitable to sensor networks with moderate computing power than the mutual information sensor selection criteria. Much of the existing work on the information-theoretic sensor configuration is mostly about adaptive control of advanced sensors such as radars and cameras [11], [12]. In this paper, we describe an information-theoretic method to analyze the effect of the sensor placement geometry on the posterior target localization distribution that is produced by multi-sensor data fusion. An earlier version of our sensor selection heuristic has appeared in [9]. An earlier version of our sensor placement strategy has apeared in [10]. In this paper, we will discuss these two related problems in a coherent and unified framework based on Bayesian information fusion and information theory.

The rest of this paper is organized as follows. Sec. II reviews the recursive Bayesian estimation for target localization and tracking and discusses different measures of the estimation error of a target location distribution. Sec. III describes our sensor selection heuristic and compares it to the mutual information based sensor selection. Sec. IV describes our information-theoretic approach to analyze the effect of the sensor placement geometry on localization accuracy. Sec. V concludes this paper.

H. Wang and D. Estrin are with UCLA Computer Science Department, email: {hbwang,destrin}@cs.ucla.edu.

K. Yao is with UCLA Electrical Engineering Department, email: yao@ee.ucla.edu

II. Data Fusion for Localization

In this section, we review the recursive Bayesian estimation for target localization and tracking and discuss different measures of the target location estimation error.

In the recursive Bayesian estimation for target localization and tracking [13], [14], both the sought 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. Let $\boldsymbol{X}$ and $\boldsymbol{x}$ denote the target location random variable and its realization value respectively. Let $\boldsymbol{Z}_i$ and $\boldsymbol{z}_i$ denote the observation random variable of sensor i and its realization value respectively. The posterior target location distribution is incrementally updated by one sensor observation at a time,

$$\begin{aligned} &p(\boldsymbol{x}|\boldsymbol{z}_1,\ldots,\boldsymbol{z}_{i+1}) \\ =&Cp(\boldsymbol{z}_{i+1}|\boldsymbol{x},\boldsymbol{z}_1,\ldots,\boldsymbol{z}_i)p(\boldsymbol{x}|\boldsymbol{z}_1,\ldots,\boldsymbol{z}_i) \ , \end{aligned}$$

where C is a normalization constant. When $\boldsymbol{Z}_1,\ldots,\boldsymbol{Z}_{i+1}$ are conditionally independent with one another conditioned on $\boldsymbol{X}$, The above equation is simplified to

$$p(\boldsymbol{x}|\boldsymbol{z}_1,\ldots,\boldsymbol{z}_{i+1}) = Cp(\boldsymbol{z}_{i+1}|\boldsymbol{x})p(\boldsymbol{x}|\boldsymbol{z}_1,\ldots,\boldsymbol{z}_i) \ .$$

The incremental update of the target location distribution by a direction-of-arrival (DOA) sensor through the recursive Bayesian estimation is illustrated in Fig. 1. The left sub-figure of Fig. 1 shows the prior target location distribution $p(\boldsymbol{x}|\boldsymbol{z}_1,\ldots,\boldsymbol{z}_i)$ denoted by the oval image. The beam image originating from the DOA sensor is the target location distribution based only on this sensor's observation, $p(\boldsymbol{x}|\boldsymbol{z}_{i+1})$, which represents the new information provided by this sensor. We have assume a Gaussian DOA observation model with a standard deviation of 2 degrees. The right sub-figure of Fig. 1 shows the posterior target location distribution $p(\boldsymbol{x}|\boldsymbol{z}_1,\ldots,\boldsymbol{z}_i,\boldsymbol{z}_{i+1})$ denoted by the round image. The true target location is denoted by marker $+$. The posterior target location distribution has much smaller estimation error than the prior target location distribution.

One of the advantages of the recursive Bayesian estimation is that we can stop updating the posterior target location as soon as the estimation error is no larger than allowed. There are several different measures of the estimation error of the posterior target location distribution. One estimation error measure is the root-mean-square error (RMSE)

$$RMSE(\boldsymbol{X}) = \sqrt{E(\|\boldsymbol{x}-\boldsymbol{x}_t\|^2)} \ , \qquad (1)$$

where $\boldsymbol{x}_t$ is the true target location, $E(\cdot)$ is expectation w.r.t. the posterior target location distribution $p(\boldsymbol{x}|\ldots,\boldsymbol{z}_i,\ldots)$, $\|\cdot\|$ is the L_2 norm. In practice, the true target location $\boldsymbol{x}_t$ is usually unknown. In this paper, we assume the target location estimation is unbiased,

$$\boldsymbol{x}_t = E(\boldsymbol{x}) \ ,$$

where $E(\cdot)$ is expectation w.r.t. $p(\boldsymbol{x}|\ldots,\boldsymbol{z}_i,\ldots)$. Another estimation error measure is the covariance matrix

$$COV(\boldsymbol{X}) = E((\boldsymbol{x}-E(\boldsymbol{x}))^2) \ , \qquad (2)$$

where $E(\cdot)$ is expectation w.r.t. $p(\boldsymbol{x}|\ldots,\boldsymbol{z}_i,\ldots)$. Yet another estimation error measure is the Shannon entropy that measures the uncertainty of the posterior target location distribution, [15],

$$H(\boldsymbol{X}|\ldots,\boldsymbol{z}_i,\ldots) = -E(\ln p(\boldsymbol{x}|\ldots,\boldsymbol{z}_i,\ldots)) \ , \qquad (3)$$

where $E(\cdot)$ is expectation w.r.t. $p(\boldsymbol{x}|\ldots,\boldsymbol{z}_i,\ldots)$. A large entropy of the posterior target location distribution indicates a large estimation error of the target location.

To sort posterior target location distributions in the order of the estimation error, we need a scalar measure of the estimation error. Since the covariance matrix of the posterior target location distribution is a matrix and not a scalar, it is not a proper measure to sort the target location distributions. Both the RMSE and the Shannon entropy are scalar and thus can be used to sort posterior target location distributions. Because the Shannon entropy is a core component of the well-established information theory, we choose to use the Shannon entropy to quantify the uncertainty reduction (or information gain) of the target location distribution due to the additional sensor observation. To be brief, we will use the term entropy to denote the Shannon entropy from now on.

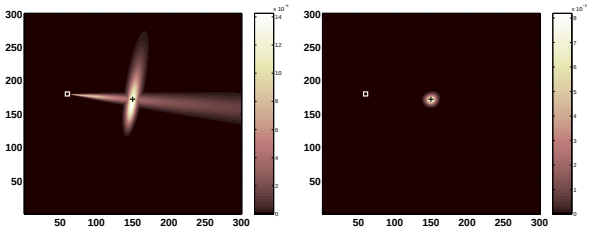

Fig. 1. Incremental update of the target location distribution by a DOA sensor denoted by the square through the recursive Bayesian estimation.

III. Sensor Selection Heuristic

In this section, we describes our sensor selection heuristic in detail. Subsec. III-A formulates the sensor selection problem in the sensor networks for target localization and tracking and reviews the mutual information based sensor selection. Subsec. III-B defines our sensor

selection heuristic. Subsec. III-C describes the relation between the entropy difference used in our sensor selection heuristic and the mutual information. Subsec. III-D validates our sensor selection heuristic using simulations. Subsec. III-E compares the computational complexity of our sensor selection heuristic to that of the mutual information based sensor selection. Subsec. III-F discusses the potential discrepancy in selection decision between our sensor selection heuristic and the mutual information based sensor selection.

A. Sensor Selection Problem

A greedy strategy has been used for sensor selection in sensor networks for target localization and tracking [7], [8]. This strategy selects the currently unused sensor whose observation is expected to result in the maximum entropy reduction of the posterior target location distribution. The observation of the selected sensor is incorporated into the target location distribution using recursive Bayesian estimation [13], [14]. The greedy sensor selection and the recursive information fusion repeat until the entropy of the posterior target location distribution is less than or equal to the desired level. Thus the entropy of the target location distribution is incrementally reduced to the desired level without consumption of more sensor resources than necessary. The core problem of the greedy sensor selection approach is how to efficiently evaluate the expected entropy reduction attributable to each candidate sensor without actually retrieving sensor data.

The sensor selection problem is formulated as follows. Given

1) the prior target location distribution: $p(\boldsymbol{x})$;
2) the set of candidate sensors for selection: $\mathcal{S}$;
3) the locations of candidate sensors: $\boldsymbol{x}_i, \forall i \in \mathcal{S}$;
4) the observation models of candidate sensors: $p(\boldsymbol{z}_i|\boldsymbol{x}), \forall i \in S$;

the objective is to find the sensor $\hat{i}$ whose observation $\boldsymbol{Z}_{\hat{i}}$ minimizes the expected conditional entropy of the posterior target location distribution,

$$\hat{i} = \arg\min_{i \in \mathcal{S}} H(\boldsymbol{X}|\boldsymbol{Z}_i) \ .$$

Equivalently, the observation of sensor $\hat{i}$ maximizes the expected reduction of the target location entropy,

$$\hat{i} = \arg\max_{i \in \mathcal{S}} (H(\boldsymbol{X}) - H(\boldsymbol{X}|\boldsymbol{Z}_i)) \ .$$

$H(\boldsymbol{X}) - H(\mathbf{X}|\boldsymbol{Z}_i)$ is one expression of $I(\boldsymbol{X}; \boldsymbol{Z}_i)$, the mutual information between the target location $\boldsymbol{X}$ and the predicted sensor observation $\boldsymbol{Z}_i$,

$$I(\boldsymbol{X}; \boldsymbol{Z}_i) = \int p(\boldsymbol{x}, \boldsymbol{z}_i) \ln \frac{p(\boldsymbol{x}, \boldsymbol{z}_i)}{p(\boldsymbol{x}) p(\boldsymbol{z}_i)} d\boldsymbol{x} d\boldsymbol{z}_i \ , \quad (4)$$

where $p(\boldsymbol{x}, \boldsymbol{z}_i) = p(\boldsymbol{z}_i|\boldsymbol{x}) p(\boldsymbol{x})$ and $p(\boldsymbol{z}_i) = \int p(\boldsymbol{x}, \boldsymbol{z}_i) d\boldsymbol{x}$. Thus the observation of sensor $\hat{i}$ maximizes the mutual information $I(\boldsymbol{X}; \boldsymbol{Z}_i)$,

$$\hat{i} = \arg\max_{i \in \mathcal{S}} I(\boldsymbol{X}; \boldsymbol{Z}_i) \ . \quad (5)$$

Sensor selection based on Eq. (5) is the maximum mutual information criterion described in [7], [8]. The target location $\boldsymbol{X}$ could be three-dimensional. The sensor observation $\boldsymbol{Z}_i$ could be two-dimensional (e.g. the direction to a target in a three-dimensional space is two-dimensional). Thus $I(\boldsymbol{X}; \boldsymbol{Z}_i)$ could be a complex integral in the joint state space $(\boldsymbol{X}, \boldsymbol{Z}_i)$ of five dimensions. The computational complexity of evaluating $I(\boldsymbol{X}; \boldsymbol{Z}_i)$ could be more than that of the capability of the low-end sensor nodes. If the observation $\boldsymbol{Z}_i$ is related to the target location $\boldsymbol{X}$ only through the sufficient statistics $\boldsymbol{Z}(\boldsymbol{X})$, then

$$I(\boldsymbol{X}; \boldsymbol{Z}_i) = I(\boldsymbol{Z}(\boldsymbol{X}); \boldsymbol{Z}_i) \ .$$

If $\boldsymbol{Z}(\boldsymbol{X})$ has fewer dimensions than $\boldsymbol{X}$, then $I(\boldsymbol{Z}(\boldsymbol{X}); \boldsymbol{Z}_i)$ is less complex to compute than $I(\boldsymbol{X}; \boldsymbol{Z}_i)$. In the above special scenario, $I(\boldsymbol{Z}(\boldsymbol{X}); \boldsymbol{Z}_i)$ has been proposed to replace $I(\boldsymbol{X}; \boldsymbol{Z}_i)$ to reduce the complexity of computing the mutual information in [7]. In this paper, we describe an alternative entropy based sensor selection heuristic. In general, the entropy based sensor selection heuristic is computationally much simpler than the mutual information based approaches. However, the observation of the sensor selected by the heuristic would still yield on average the greatest or nearly the greatest entropy reduction of the target location distribution as will be shown in Subsec. III-D.

B. Sensor Selection Heuristic

In our studies of sensor selection for localization, we have observed that the reduction of the localization uncertainty attributable to a sensor largely depends on the difference of two quantities, namely, the entropy of the noise-free sensor observation, and the entropy of that sensor observation model corresponding to the true target location. The noise-free sensor observation assumes that no error is introduced into the sensor observation. The sensor observation model corresponding to the true target location is the probability distribution of the sensor observation conditioned on the true target location. Loosely speaking, our sensor selection heuristic selects the candidate sensor with the maximum entropy difference described above.

Let $\boldsymbol{Z}_i^{\mathrm{v}}$ denote the noise-free observation of sensor i. Because $\boldsymbol{Z}_i^{\mathrm{v}}$ assumes no randomness in the process of observation regarding the target location, $\boldsymbol{Z}_i^{\mathrm{v}}$ is a

function of the target location $\boldsymbol{X}$ and the sensor location $\boldsymbol{x}_i$,

$$\boldsymbol{Z}_i^{\mathrm{v}} = f(\boldsymbol{X}, \boldsymbol{x}_i) \ . \tag{6}$$

In Eq. (6), because the target location $\boldsymbol{X}$ is a random variable, the noise-free sensor observation $\boldsymbol{Z}_i^{\mathrm{v}}$ is a random variable although the sensor location $\boldsymbol{x}_i$ is a deterministic quantity. Since the noise-free sensor observation $\boldsymbol{Z}_i^{\mathrm{v}}$ usually has less dimensions than the target location $\boldsymbol{X}$, the distribution of the noise-free sensor observation $\boldsymbol{Z}_i^{\mathrm{v}}$ is usually the geometric projection of the target location distribution $p(\boldsymbol{x})$ onto the observation perspective of sensor i,

$$P(\boldsymbol{Z}_i^{\mathrm{v}} \leq \boldsymbol{z}_i^{\mathrm{v}}) = \int\limits_{f(\boldsymbol{x},\boldsymbol{x}_i)\leq \boldsymbol{z}_i^{\mathrm{v}}} p(\boldsymbol{x}) d\boldsymbol{x} \ , \tag{7}$$

where the observation perspective of sensor i largely depends on the sensor location $\boldsymbol{x}_i$.

In practice, the subset of the state space of the target location $\boldsymbol{X}$ and the noise-free sensor observation $\boldsymbol{Z}_i^{\mathrm{v}}$ with the non-trivial probability density can be discretized into a grid for numerical analysis. Any probability density function value larger than a given threshold is considered as non-trivial. The discrete representation of $p(\boldsymbol{z}_i^{\mathrm{v}})$ can be computed as follows.

1) Let $\mathcal{X}$ be the set of the target location grid values with the non-trivial probability density;
2) Let $\mathcal{Z}$ be the set of the noise-free sensor observation grid values of the non-trivial probability density;
3) For each grid point $\boldsymbol{z}_i^{\mathrm{v}} \in \mathcal{Z}$, initialize $p(\boldsymbol{z}_i^{\mathrm{v}})$ to zero;
4) For each grid point $\boldsymbol{x} \in \mathcal{X}$, determine the corresponding grid point $\boldsymbol{z}_i^{\mathrm{v}} \in \mathcal{Z}$ using Eq. (6), and update its probability as $p(\boldsymbol{z}_i^{\mathrm{v}}) = p(\boldsymbol{z}_i^{\mathrm{v}}) + p(\boldsymbol{x})$;
5) Normalize $p(\boldsymbol{z}_i^{\mathrm{v}})$ to make the total probability of $\mathcal{Z}$ to be 1.

After the noise-free sensor observation distribution $p(\boldsymbol{z}_i^{\mathrm{v}})$ is computed, the noise-free sensor observation entropy $H(\boldsymbol{Z}_i^{\mathrm{v}})$ can be computed using Eq. (3).

The numerical computation of the noise-free observation distribution $p(\boldsymbol{z}_i^{\mathrm{v}})$ for a DOA sensor is illustrated in Fig. 2. In the left sub-figure of Fig. 2, the target location distribution is denoted by the image color, and the DOA sensor location is denoted by the square. The subset of the target location state space with the non-trivial probability density is discretized into a grid of 400×400. The true target location is denoted by marker $+$. The right sub-figure of Fig. 2 shows the discrete probability distribution of the DOA sensor's noise-free observation in the granularity of 2^{o}. Marker $\times$ denotes the probability of the noise-free DOA observation in the interval of $[36^{\mathrm{o}}, 38^{\mathrm{o}}]$, which is the summation of the probability of all target locations inside the sector delimited by the 36^{o} line and the 38^{o} line in the left sub-figure of Fig. 2.

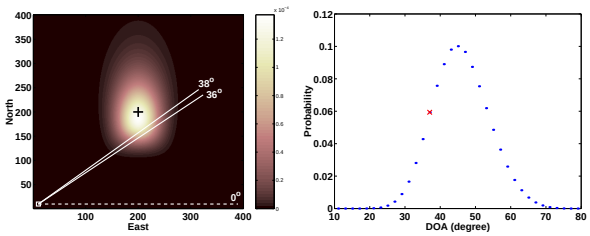

Fig. 2. A DOA sensor's noise-free observation about the target location.

The observation model of sensor i is $p(\boldsymbol{z}_i|\boldsymbol{x}^{\mathrm{t}})$ when the target is actually at $\boldsymbol{x}^{\mathrm{t}}$. The sensor observation model incorporates observation error from all sources, including the noise corruption to the signal used to observe the target, the signal modeling error in the estimation algorithm used by the sensor, the inaccuracy of the sensor hardware, and so on. The amount of uncertainty in the sensor observation model may depend on the target location. Since the true target location is unknown during the process of target localization and tracking, we have to use an estimated target location to approximate the true target location in order to determine the sensor observation model. For a single-modal target location distribution $p(\boldsymbol{x})$ that has a single peak, we can use the maximum likelihood estimate $\hat{\boldsymbol{x}}$ of the target location to approximate the true target location, and the entropy of the approximate sensor observation model is

$$H(\boldsymbol{Z}_i|\hat{\boldsymbol{x}}) = -\int p(\boldsymbol{z}_i|\hat{\boldsymbol{x}}) \ln p(\boldsymbol{z}_i|\hat{\boldsymbol{x}}) d\boldsymbol{z}_i \ . \tag{8}$$

For a multi-modal target location distribution $p(\boldsymbol{x})$ with more than one peaks, namely, $\hat{\boldsymbol{x}}^{(m)}$, $m = 1, \ldots, M$, the entropy of the observation model of sensor i can be approximated as a weighted average as follows

$$H(\boldsymbol{Z}_i|\hat{\boldsymbol{x}}) = \frac{\sum_{m=1}^{M} p(\hat{\boldsymbol{x}}^{(m)}) H(\boldsymbol{Z}_i|\hat{\boldsymbol{x}}^{(m)})}{\sum_{m=1}^{M} p(\hat{\boldsymbol{x}}^{(m)})} \ , \tag{9}$$

where $H(\boldsymbol{Z}_i|\hat{\boldsymbol{x}}^{(m)}) = -\int p(\boldsymbol{z}_i|\hat{\boldsymbol{x}}^{(m)}) \ln p(\boldsymbol{z}_i|\hat{\boldsymbol{x}}^{(m)}) d\boldsymbol{z}_i$.

We have repeatedly observed that the incorporation of the sensor observation with a larger entropy difference $H(\boldsymbol{Z}_i^{\mathrm{v}}) - H(\boldsymbol{Z}_i|\hat{\boldsymbol{x}})$ yields on average a larger reduction in the uncertainty of the posterior target location distribution. Thus, the entropy difference $H(\boldsymbol{Z}_i^{\mathrm{v}}) - H(\boldsymbol{Z}_i|\hat{\boldsymbol{x}})$ can sort candidate sensors into nearly the same order as the mutual information $I(\boldsymbol{X}; \boldsymbol{Z}_i)$. Specifically, the sensor with the maximum entropy difference $H(\boldsymbol{Z}_i^{\mathrm{v}}) - H(\boldsymbol{Z}_i|\hat{\boldsymbol{x}})$ also has nearly the maximum mutual information $I(\boldsymbol{X}; \boldsymbol{Z}_i)$. Hence we propose to use the entropy

difference $H(\boldsymbol{Z}_i^{\mathrm{v}}) - H(\boldsymbol{Z}_i|\hat{\boldsymbol{x}})$ as an alternative to the mutual information $I(\boldsymbol{X};\boldsymbol{Z}_i)$ for selecting the most informative sensor. Formally, the entropy based sensor selection heuristic is as follows.

1) compute the entropy difference $H(\boldsymbol{Z}_i^{\mathrm{v}}) - H(\boldsymbol{Z}_i|\hat{\boldsymbol{x}})$ for the set of candidate sensors $\mathcal{S}$;
2) select sensor $\hat{i}$ such that

$$\hat{i} = \arg\max_{i\in\mathcal{S}} (H(\boldsymbol{Z}_i^{\mathrm{v}}) - H(\boldsymbol{Z}_i|\hat{\boldsymbol{x}})) \ .$$

We will see that our sensor selection heuristic is computationally much simpler than the mutual information based sensor selection in Subsec. III-E

C. Relation to Mutual Information

In this subsection, mathematical analysis reveals that the entropy difference $H(\boldsymbol{Z}_i^{\mathrm{v}}) - H(\boldsymbol{Z}_i|\hat{\boldsymbol{x}})$ can reasonably approximate the mutual information $I(\boldsymbol{X};\boldsymbol{Z}_i)$. As a result, it is reasonably effective to use the entropy difference $H(\boldsymbol{Z}_i^{\mathrm{v}}) - H(\boldsymbol{Z}_i|\hat{\boldsymbol{x}})$ to select the sensor with the maximum mutual information $I(\boldsymbol{X};\boldsymbol{Z}_i)$. The mutual information $I(\boldsymbol{X};\boldsymbol{Z}_i)$ has another expression, namely, $H(\boldsymbol{Z}_i) - H(\boldsymbol{Z}_i|\boldsymbol{X})$. We will show that $H(\boldsymbol{Z}_i^{\mathrm{v}})$ and $H(\boldsymbol{Z}_i|\hat{\boldsymbol{x}})$ can reasonably approximate $H(\boldsymbol{Z}_i)$ and $H(\boldsymbol{Z}_i|\boldsymbol{X})$ respectively.

$H(\boldsymbol{Z}_i)$ is the entropy of the predicted sensor observation distribution, $p(\boldsymbol{z}_i) = p(\boldsymbol{z}_i|\boldsymbol{x})p(\boldsymbol{x})$. The predicted sensor observation distribution $p(\boldsymbol{z}_i)$ becomes the noise-free sensor observation distribution $p(\boldsymbol{z}_i^{\mathrm{v}})$ when the sensor observation model $p(\boldsymbol{z}_i|\boldsymbol{x})$ is deterministic without any uncertainty. The uncertainty in the sensor observation model $p(\boldsymbol{z}_i|\boldsymbol{x})$ makes the predicted sensor observation entropy $H(\boldsymbol{Z}_i)$ larger than the noise-free sensor observation entropy $H(\boldsymbol{Z}_i^{\mathrm{v}})$. When the sensor observation model $p(\boldsymbol{z}_i|\boldsymbol{x})$ has only a small amount of uncertainty, the predicted sensor observation entropy $H(\boldsymbol{Z}_i^{\mathrm{v}})$ closely approximates the noise-free sensor observation entropy $H(\boldsymbol{Z}_i)$.

$H(\boldsymbol{Z}_i|\boldsymbol{X})$ is actually the entropy of the sensor observation model averaged over all possible target locations,

$$\begin{aligned} H(\boldsymbol{Z}_i|\boldsymbol{X}) &= -\int p(\boldsymbol{x},\boldsymbol{z}_i)\ln p(\boldsymbol{z}_i|\boldsymbol{x})d\boldsymbol{x}d\boldsymbol{z}_i \\ &= \int p(\boldsymbol{x})\{-\int p(\boldsymbol{z}_i|\boldsymbol{x})\ln p(\boldsymbol{z}_i|\boldsymbol{x})d\boldsymbol{z}_i\}d\boldsymbol{x} \\ &= \int p(\boldsymbol{x})H(\boldsymbol{Z}_i|\boldsymbol{x})d\boldsymbol{x} \ . \end{aligned}$$

When $p(\boldsymbol{x})$ is a single-modal distribution, $H(\boldsymbol{Z}_i|\hat{\boldsymbol{x}})$ is defined in Eq. (8), which is the entropy of the sensor observation model for the most likely target location estimate $\hat{\boldsymbol{x}}$. When $p(\boldsymbol{x})$ is a multi-modal distribution, $H(\boldsymbol{Z}_i|\hat{\boldsymbol{x}})$ is defined in Eq. (9), which is the entropy of the sensor observation model averaged over all target locations with local maximum likelihood. When the entropy of the sensor observation model $H(\boldsymbol{Z}_i|\boldsymbol{x})$ changes slowly with the target location $\boldsymbol{x}$, $H(\boldsymbol{Z}_i|\hat{\boldsymbol{x}})$ can reasonably approximate $H(\boldsymbol{Z}_i|\boldsymbol{X})$.

Since $H(\boldsymbol{Z}_i^{\mathrm{v}})$ and $H(\boldsymbol{Z}_i|\hat{\boldsymbol{x}})$ can reasonably approximate $H(\boldsymbol{Z}_i)$ and $H(\boldsymbol{Z}_i|\boldsymbol{X})$ respectively, the entropy difference $H(\boldsymbol{Z}_i^{\mathrm{v}}) - H(\boldsymbol{Z}_i|\hat{\boldsymbol{x}})$ can reasonably approximate the mutual information $I(\boldsymbol{X};\boldsymbol{Z}_i) = H(\boldsymbol{Z}_i) - H(\boldsymbol{Z}_i|\boldsymbol{X})$. Such approximation is very close when $H(\boldsymbol{Z}_i|\hat{\boldsymbol{x}})$ is small relative to $H(\boldsymbol{Z}_i^{\mathrm{v}})$ and the entropy of the sensor observation model $H(\boldsymbol{Z}_i|\boldsymbol{x})$ changes slowly with the target location $\boldsymbol{x}$. Thus the entropy difference $H(\boldsymbol{Z}_i^{\mathrm{v}}) - H(\boldsymbol{Z}_i|\hat{\boldsymbol{x}})$ sorts sensors into approximately the order of the mutual information $I(\boldsymbol{X};\boldsymbol{Z}_i)$. As a result, the sensor with the maximum entropy difference $H(\boldsymbol{Z}_i^{\mathrm{v}}) - H(\boldsymbol{Z}_i|\hat{\boldsymbol{x}})$ probably also has the maximum mutual information $I(\boldsymbol{X};\boldsymbol{Z}_i)$. Thus the entropy difference $H(\boldsymbol{Z}_i^{\mathrm{v}}) - H(\boldsymbol{Z}_i|\hat{\boldsymbol{x}})$ is a reasonable alternative to the mutual information $I(\boldsymbol{X};\boldsymbol{Z}_i)$ for sensor selection. The correlation between the entropy difference $H(\boldsymbol{Z}_i^{\mathrm{v}}) - H(\boldsymbol{Z}_i|\hat{\boldsymbol{x}})$ and mutual information $I(\boldsymbol{X};\boldsymbol{Z}_i)$ will be further explored using simulations in Subsec. III-D.

D. Validation of Sensor Selection Heuristic

This subsection evaluates our sensor selection heuristic relative to the mutual information based sensor selection using simulations. The Gaussian noise model has been widely assumed for sensor observations in many localization and tracking algorithms, e.g., the Kalman filter [16]. As a starting point, we assume the Gaussian sensor observation models in the evaluative simulations for simplicity. The simple Gaussian sensor observation models assumed here are not accurate especially when sensors are very close to the target. To avoid the problem of the over-simplified sensor observation models in the simulations, we only analyze sensors with some middle distance range to the target. The heuristic will be evaluated further under more realistic sensor observation models in the future.

Four scenarios of sensor selection for localization have been studied. Three of them involve DOA sensors, range sensors, and time-difference-of-arrival (TDOA) sensors exclusively as shown in Fig. 3, Fig. 4, and Fig. 5 respectively. In each of these scenarios, 500 candidate sensors of different combination of location and observation standard deviation are considered. Another scenario involves all these three types of sensors mixed together as shown in Fig. 6. In this scenario, we have considered 100 candidate sensors with randomly assigned observation type, location, and observation standard deviation.

In every sensor selection scenario, both the entropy difference $H(\boldsymbol{Z}_i^{\mathrm{v}}) - H(\boldsymbol{Z}_i|\hat{\boldsymbol{x}})$ and the mutual information $I(\boldsymbol{X};\boldsymbol{Z}_i)$ are evaluated and compared for all candidate sensors.

In the left sub-figures of Fig. 3, Fig. 4, Fig. 5, and Fig. 6, the image color depicts the prior target location distribution $p(\boldsymbol{x})$. The subset of the state space of the target location $\boldsymbol{X}$ with the non-trivial probability density is enclosed by the solid rectangle. The true target location is denoted by marker $+$. Sensors are uniformly randomly placed outside the dotted rectangle. The squares, circles, and triangles denote DOA sensors, range sensors, and TDOA sensors respectively. All TDOA observations are relative to a common reference sensor denoted by marker $\times$. The size of the sensor marker in the left sub-figure of Fig. 6 indicates the observation standard deviation σ that is randomly chosen to be 2, 4, 8, 16, or 32. The right sub-figures of Fig. 3, Fig. 4, Fig. 5, and Fig. 6 show the plot of the mutual information $I(\boldsymbol{X};\boldsymbol{Z}_i)$ vs entropy difference $H(\boldsymbol{Z}_i^{\mathrm{v}}) - H(\boldsymbol{Z}_i|\hat{\boldsymbol{x}})$ of all candidate sensors. Each marker denotes $(H(\boldsymbol{Z}_i^{\mathrm{v}}) - H(\boldsymbol{Z}_i|\hat{\boldsymbol{x}}), I(\boldsymbol{X};\boldsymbol{Z}_i))$ pair evaluated for one candidate sensor. σ is the standard deviation of the Gaussian observation model assumed for candidate sensors.

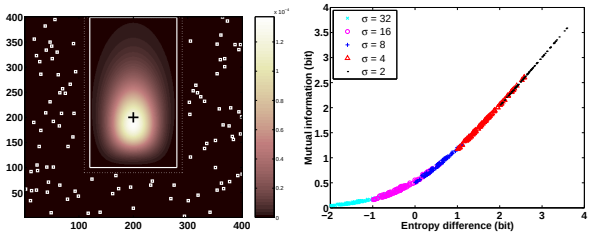

Fig. 3. Scenario of sensor selection for localization using DOA sensors exclusively.

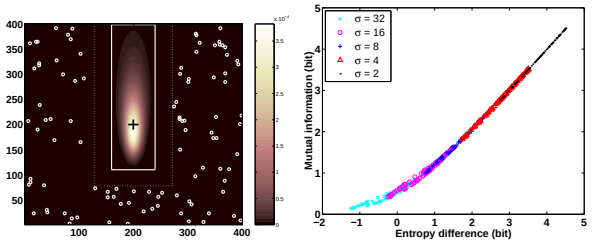

Fig. 4. Scenario of sensor selection for localization using range sensors exclusively.

In all sensor selection scenarios, the entropy difference $H(\boldsymbol{Z}_i^{\mathrm{v}}) - H(\boldsymbol{Z}_i|\hat{\boldsymbol{x}})$ correlates very well with the mutual information $I(\boldsymbol{X};\boldsymbol{Z}_i)$. Thus, the entropy difference $H(\boldsymbol{Z}_i^{\mathrm{v}}) - H(\boldsymbol{Z}_i|\hat{\boldsymbol{x}})$ can sort all candidate sensors into nearly the same order as the mutual information $I(\boldsymbol{X};\boldsymbol{Z}_i)$ does. The sensor with the maximal entropy

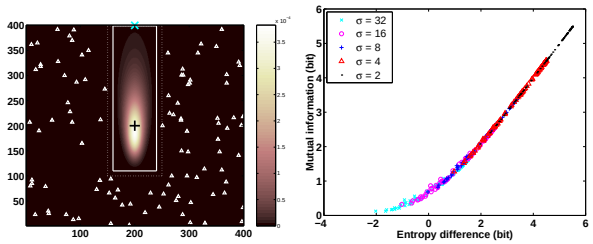

Fig. 5. Scenario of sensor selection for localization using TDOA sensors exclusively.

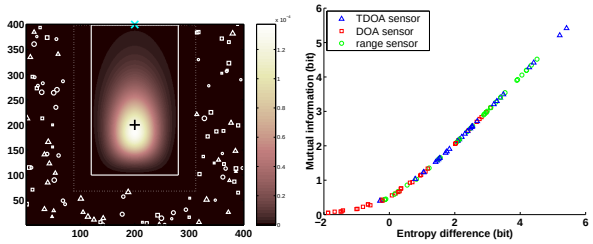

Fig. 6. Scenario of sensor selection for localization using DOA sensors, range sensors, and TDOA sensors together.

difference $H(\boldsymbol{Z}_i^{\mathrm{v}}) - H(\boldsymbol{Z}_i|\hat{\boldsymbol{x}})$ selected by the heuristic always has the maximum or nearly the maximal mutual information $I(\boldsymbol{X};\boldsymbol{Z}_i)$. The larger is the mutual information $I(\boldsymbol{X};\boldsymbol{Z}_i)$, the more consistent will be the decision between these two sensor selection criterion. Only when the mutual information $I(\boldsymbol{X};\boldsymbol{Z}_i)$ is very small, such correlation starts to show small dispersion as shown in Fig. 4 and Fig. 5. A sensor observation $\boldsymbol{Z}_i$ with very small mutual information with the target location $\boldsymbol{X}$ is expected to contribute very small amount of uncertainty reduction to the the target location distribution.

E. Complexity of Sensor Selection Heuristic

In this subsection, we analyze the computational complexity of our sensor selection heuristic and compare it to that of the mutual information based sensor selection. The computational complexity of these two sensor selection criterion depends on the number of dimensions of the target location $\boldsymbol{X}$ and the sensor observation $\boldsymbol{Z}_i$. We use the DOA sensor based three-dimensional target localization and tracking as an example to compare the computational complexity of these sensor selection criterion. The target location $\boldsymbol{X}$ is three-dimensional. Both the noise-free DOA observation $\boldsymbol{Z}_i^{\mathrm{v}}$ and the noisy DOA observation $\boldsymbol{Z}_i$ are two-dimensional. We assume that all random variables are discretized for numerical computation. Specifically, the three-dimensional target location subspace with non-trivial probability density is discretized into a grid of $n \times n \times n$. The scope of DOA observations with non-trivial probability density is also

discretized into a grid of $n \times n$. We assume there are K candidate sensors for selection. K is usually a small number relative to n.

Our sensor selection heuristic evaluates the entropy difference $H(\boldsymbol{Z}_i^{\mathrm{y}}) - H(\boldsymbol{Z}_i|\hat{\boldsymbol{x}})$ of all candidate sensors for selection and then selects the sensor with the maximum entropy difference. As shown in Subsec. III-B, $p(\boldsymbol{z}_i^{\mathrm{y}})$ can be computed from $p(\boldsymbol{x})$ with cost $O(n^3)$. As shown in Eq. (3), $H(\boldsymbol{Z}_i^{\mathrm{y}})$ can be computed from $p(\boldsymbol{z}_i^{\mathrm{y}})$ with cost $O(n^2)$. As shown in Eq. (8) and (9), $H(\boldsymbol{Z}_i|\hat{\boldsymbol{x}})$ can be computed from $p(\boldsymbol{z}_i|\hat{\boldsymbol{x}})$ with cost $O(n^2)$. Thus, the cost to compute the entropy difference $H(\boldsymbol{Z}_i^{\mathrm{y}}) - H(\boldsymbol{Z}_i|\hat{\boldsymbol{x}})$ for one candidate sensor is $O(n^3)$. Thus, the total cost for our heuristic to select one out of K candidate sensors is $O(n^3)$.

The mutual information based sensor selection evaluates the mutual information $I(\boldsymbol{X}; \boldsymbol{Z}_i)$ of all candidate sensors for selection and then select the one with the maximum mutual information. As shown in Eq. (4), the mutual information $I(\boldsymbol{X}; \boldsymbol{Z}_i)$ can be directly computed from $p(\boldsymbol{x})$ and $p(\boldsymbol{z}_i|\boldsymbol{x})$ with cost of $O(n^5)$. Thus, the total cost to select one out of K candidate sensors is $O(n^5)$. As we mentioned early in Subsec. III-A, the computational cost of mutual information $I(\boldsymbol{X}; \boldsymbol{Z}_i)$ could be reduced in some special scenarios. In general, however, our sensor selection heuristic is computationally much simpler than the mutual information based approaches.

F. Dispersion of Correlation with Mutual Information

As pointed out in Subsec. III-D, there is a little dispersion in the correlation between the entropy difference $H(\boldsymbol{Z}_i^{\mathrm{y}}) - H(\boldsymbol{Z}_i|\hat{\boldsymbol{x}})$ and the mutual information $I(\boldsymbol{X}; \boldsymbol{Z}_i)$ when the mutual information is very small. Such dispersion can be seen in the convex part of the plot of the entropy difference $H(\boldsymbol{Z}_i^{\mathrm{y}}) - H(\boldsymbol{Z}_i|\hat{\boldsymbol{x}})$ vs the mutual information $I(\boldsymbol{X}; \boldsymbol{Z}_i)$ in Fig. 4 and Fig. 5. Very small mutual information $I(\boldsymbol{X}; \boldsymbol{Z}_i)$ indicates that the sensor observation $\boldsymbol{Z}_i$ on average can only reduce very little uncertainty of the target location $\boldsymbol{X}$. Thus, there might be a discrepancy in selection decision between our sensor selection heuristic and the mutual information based sensor selection if and only if no candidate sensor is very informative. However, our simulations have shown that there is very little degradation in selection decision made by our sensor selection heuristic even if no candidate sensor is very informative.

We model the dispersion of the correlation between the entropy difference $H(\boldsymbol{Z}_i^{\mathrm{y}}) - H(\boldsymbol{Z}_i|\hat{\boldsymbol{x}})$ and the mutual information $I(\boldsymbol{X}; \boldsymbol{Z}_i)$ using a uniform distribution bounded by a parallelogram where a candidate sensor could assume any $(H(\boldsymbol{Z}_i^{\mathrm{y}}) - H(\boldsymbol{Z}_i|\hat{\boldsymbol{x}}), I(\boldsymbol{X}; \boldsymbol{Z}_i))$ pair within the parallelogram with uniform probability. As illustrated in Fig. 7, the geometry of the parallelogram is defined by three parameters, namely, a, b and c. Parameter a is the variation scope of the entropy difference $H(\boldsymbol{Z}_i^{\mathrm{y}}) - H(\boldsymbol{Z}_i|\hat{\boldsymbol{x}})$ of the candidate sensors considered in the current selection decision-making. Parameter c indicates the variation scope of the mutual information $I(\boldsymbol{X}; \boldsymbol{Z}_i)$ of the candidate sensors considered in the current selection decision-making. Parameter b describes the magnitude of dispersion of the correlation between the entropy difference $H(\boldsymbol{Z}_i^{\mathrm{y}}) - H(\boldsymbol{Z}_i|\hat{\boldsymbol{x}})$ and the mutual information $I(\boldsymbol{X}; \boldsymbol{Z}_i)$. We choose this dispersion model for simplicity. As the first order approximation, this dispersion model does capture the major features of the correlation dispersion revealed by simulations in Subsec. III-D, and help to reveal some major characteristics of the impact of the correlation dispersion on the performance of our sensor selection heuristic.

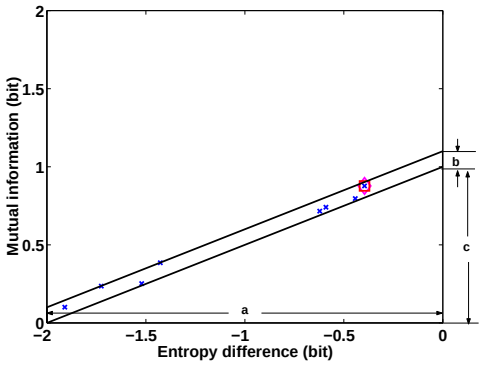

Fig. 7. Correlation dispersion between the entropy difference $H(\boldsymbol{Z}_i^{\mathrm{y}}) - H(\boldsymbol{Z}_i|\hat{\boldsymbol{x}})$ and the mutual information $I(\boldsymbol{X}; \boldsymbol{Z}_i)$ modeled by a uniform distribution bounded by a parallelogram.

Fig. 7 shows a typical dispersion scenario where no candidate sensor is very informative. The mutual information $I(\boldsymbol{X}; \boldsymbol{Z}_i)$ of the candidate sensors varies from 0 bit to 1 bit. Correspondingly, the entropy difference $H(\boldsymbol{Z}_i^{\mathrm{y}}) - H(\boldsymbol{Z}_i|\hat{\boldsymbol{x}})$ of the candidate sensors changes from -2 bit to 0 bit. The disperse of the correlation between $H(\boldsymbol{Z}_i^{\mathrm{y}}) - H(\boldsymbol{Z}_i|\hat{\boldsymbol{x}})$ and the mutual information $I(\boldsymbol{X}; \boldsymbol{Z}_i)$ is 0.1 bit. Given the above dispersion scenario, we run $10,000$ simulations. In each simulation, 8 candidate sensors randomly assume their $(H(\boldsymbol{Z}_i^{\mathrm{y}}) - H(\boldsymbol{Z}_i|\hat{\boldsymbol{x}}), I(\boldsymbol{X}; \boldsymbol{Z}_i))$ pairs within the specified dispersion range. In each simulation, we identify both the sensor with the maximum entropy difference and the sensor with the maximum mutual information. Fig. 7 also shows one particular realization of the simulations. Eight $\times$ markers are uniformly randomly distributed inside the parallelogram denote candidate sensors. Our sensor selection heuristic selects the rightmost sensor that is

enclosed by a square marker. The mutual information based approaches select the uppermost sensor that is enclosed by a diamond-shaped marker. The rightmost sensor happens also to be the uppermost sensor in this simulation.

For the dispersion shown in Fig. 7, with 87.8% chance, the sensor selected by our sensor selection heuristic also has the maximum mutual information. Even when our sensor selection heuristic fails to select the sensor of the maximum mutual information, the mutual information of the selected sensor is on average only about 0.026 bit less than the maximum mutual information. On average, the mutual information of the sensor selected by our sensor selection heuristic is about $0.026 \times (1-87.8\%) = 0.0032$ bit less than the maximum mutual information when there is dispersion in the correlation between the entropy difference $H(\boldsymbol{Z}_i^{\mathrm{v}}) - H(\boldsymbol{Z}_i|\hat{\boldsymbol{x}})$ and the mutual information $I(\boldsymbol{X}; \boldsymbol{Z}_i)$. Overall, our sensor selection heuristic introduces very little degradation to the quality of the sensor select decision even when no candidate sensor is very informative.

IV. Strategy for Sensor Placement

In Sec. III, we have described a computationally efficient strategy to select the most informative sensor from a given a sensor network deployment. In this section, we describe a strategy of sensor placement to minimize the localization uncertainty given the region where the target needs to be localized and tracked. Subsec. IV-A describes a method to compute the posterior target location distribution with the minimum entropy given a sensor placement geometry. Subsec. IV-B uses the minimum entropy of the posterior target location distribution to characterize the dependency of the localization uncertainty on the sensor placement geometry and the sensor observation type. Such dependency characteristics provide guidance to choose the optimal sensor placement geometry to minimize the localization uncertainty in a given region.

A. Min-Entropy Location Distribution

Given the deployment of N sensors to localize a target at location $\boldsymbol{x}^{\mathrm{t}}$, the estimation error in the posterior target location distribution $p(\boldsymbol{x}|\boldsymbol{z}_i, 1 \le i \le N)$ depends on the sensor observation values $\boldsymbol{z}_i, 1 \le i \le N$ as shown in Eq. (1), (2), and (3). In other words, given the same senor network deployment and the same true target location $\boldsymbol{x}^{\mathrm{t}}$, all three estimation error measures, namely the RMSE, the covariance, and the entropy of the posterior target location, can vary greatly with different realization of the sensor observation. One way to remove the randomness in the error measures of the posterior target location estimation is to use the lower bound of the posterior localization error to compare the localization capability of two sensor networks with different deployment geometry. The CRB is widely used in analysis of the lower bound of unbiased estimators [17], [18], [19]. The minimum covariance matrix of the posterior target location distribution is the CRB if the target location estimation is unbiased,

$$\begin{aligned} & COV(\boldsymbol{X}|\boldsymbol{z}_i, 1 \le i \le N) \\ \ge & \frac{1}{-E(\partial^2[\ln p(\boldsymbol{z}_i, 1 \le i \le N|\boldsymbol{x})]/\partial \boldsymbol{x}^2)} \ , \end{aligned}$$

where $E(\cdot)$ is expectation w.r.t. the sensor observation models $p(\boldsymbol{z}_i, 1 \le i \le N|\boldsymbol{x})$. As we pointed out in Sec. II, only scalar measures of the estimation error can be directly sorted into an order. Because the CRB is a matrix and not a scalar, the CRB can not be directly sorted into any order. The CRB of the target location estimation was converted into the RMSE of the target location estimation to compare the localization capability of multiple sensor networks of different deployment geometry in [20]. Because the covariance matrix does not fully describe a distribution, the conversion from the CRB to the RMSE can not be accurate when the distribution itself is unknown. We choose the minimum entropy of the posterior target location distribution over the lower bound of the RMSE of the posterior target location distribution to compare the localization capability of different sensor network deployment because the entropy has deep roots in the well-established information theory.

As shown in Eq. (3), the entropy of the posterior target location distribution $H(\boldsymbol{X}|\boldsymbol{z}_i, 1 \le i \le N)$ is a function of sensor observations $\boldsymbol{z}_i, 1 \le i \le N$. Formally,

$$H(\boldsymbol{X}|\boldsymbol{z}_i, 1 \le i \le N) = g(\boldsymbol{z}_1, \ldots, \boldsymbol{z}_N) \ , \qquad (10)$$

where $g(\cdot)$ is a complex multi-variate function. Let $H_{\min}$ be the minimum entropy of the posterior target location distribution. We can find out $H_{\min}$ by searching throughout the joint state space of sensor observations $\boldsymbol{z}_i, 1 \le i \le N$,

$$\begin{aligned} H_{\min} &= \min_{\boldsymbol{z}_i, 1 \le i \le N} g(\boldsymbol{z}_1, \ldots, \boldsymbol{z}_N) \\ &= g(\hat{\boldsymbol{z}}_1, \ldots, \hat{\boldsymbol{z}}_N) \ , \end{aligned} \qquad (11)$$

where $\hat{\boldsymbol{z}}_i, 1 \le i \le N$ is the sensor observation that minimizes entropy of the posterior target location distribution. If the partial derivatives of $g(\cdot)$ relative to $\boldsymbol{z}_i, 1 \le i \le N$ are well defined, Eq. (11) implies that

$$\partial g(\hat{\boldsymbol{z}}_1, \ldots, \hat{\boldsymbol{z}}_N)/\partial \boldsymbol{z}_i = 0, \ 1 \le i \le N \ . \qquad (12)$$

If the noise-free observation is a critical point of the sensor observation model $p(\boldsymbol{z}_1, \ldots, \boldsymbol{z}_N|\boldsymbol{x})$ and maximizes $p(\boldsymbol{z}_1, \ldots, \boldsymbol{z}_N|\boldsymbol{x})$, then

$$\partial p(\boldsymbol{z}_1^{\mathrm{v}}, \ldots, \boldsymbol{z}_N^{\mathrm{v}}|\boldsymbol{x})/\partial \boldsymbol{z}_i = 0,\ 1 \leq i \leq N,\ \ , \qquad (13)$$

We can prove that the min-entropy sensor observation is the noise-free sensor observation,

$$\hat{\boldsymbol{z}}_i = \boldsymbol{z}_i^{\mathrm{v}},\ 1 \leq i \leq N\ \ . \qquad (14)$$

if the noise-free observation is a critical point of the sensor observation model $p(\boldsymbol{z}_1, \ldots, \boldsymbol{z}_N|\boldsymbol{x})$ and maximizes $p(\boldsymbol{z}_1, \ldots, \boldsymbol{z}_N|\boldsymbol{x})$. Detail of the proof is in appendix I. Condition in Eq. (13) can be satisfied in most of the currently used sensor observation models.

Given the sensor network deployment geometry and the true target location $\boldsymbol{x}^{\mathrm{t}}$, the noise-free sensor observation $\boldsymbol{z}_i^{\mathrm{v}}, 1 \leq i \leq N$ can be computed according to Eq. (6). After the noise-free sensor observation is computed, we can compute the min-entropy posterior target location distribution as

$$p(\boldsymbol{x}|\boldsymbol{z}_1^{\mathrm{v}}, \ldots, \boldsymbol{z}_N^{\mathrm{v}}) = p(\boldsymbol{z}_1^{\mathrm{v}}, \ldots, \boldsymbol{z}_N^{\mathrm{v}}|\boldsymbol{x})p(\boldsymbol{x})\ \ ,$$

where $p(\boldsymbol{x})$ is the prior target location distribution. The minimum entropy $H_{\min}$ is simply the entropy of $p(\boldsymbol{x}|\boldsymbol{z}_1^{\mathrm{v}}, \ldots, \boldsymbol{z}_1^{\mathrm{v}})$. We can also compute the covariance matrix of the min-entropy posterior target location distribution,

$$\begin{aligned} &COV(\boldsymbol{X}|\boldsymbol{z}_1^{\mathrm{v}}, \ldots, \boldsymbol{z}_N^{\mathrm{v}}) \\ &= \int (\boldsymbol{x} - E(\boldsymbol{x}))(\boldsymbol{x} - E(\boldsymbol{x}))^T p(\boldsymbol{x}|\boldsymbol{z}_1^{\mathrm{v}}, \ldots, \boldsymbol{z}_N^{\mathrm{v}}) d\boldsymbol{x}\ \ , \end{aligned}$$

where T is the transpose operator, $E(\cdot)$ is expectation with respect to the min-entropy posterior target location distribution. We name such covariance matrix the Bayesian lower bound (BB) of the target location estimation.

We compare the BB to the CRB through two simulations of two-dimensional localization using TDOA sensors and range sensors as shown in Fig. 8 and Fig. 9 respectively. For simplicity, we assume the Gaussian distribution for all TDOA and range observations. The standard deviation $\sigma = 6$ time units is assumed for TDOA observations. The standard deviation $\sigma = 4$ distance units is assumed for range observations. In the left sub-figures of Fig. 8 and Fig. 9, the image color depicts 9 posterior target location distributions of the minimum entropy. Both TDOA sensors and range sensors are denoted by squares. The right sub-figures of Fig. 8 and Fig. 9 plot the elements of the BB matrix vs the corresponding element of the CRB matrix. In both simulations, the BB equals the CRB element to element. The consistency between the BB and the CRB indicates that our method to compute the min-entropy posterior target location distribution is valid.

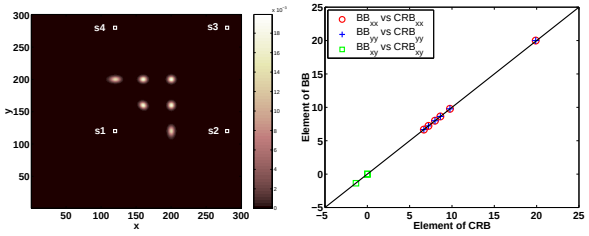

Fig. 8. Comparison of the BB with the CRB in localization using TDOA sensors.

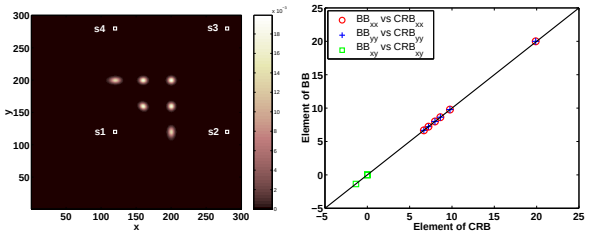

Fig. 9. Comparison of the BB with the CRB in localization using range sensors.

B. Effects of Sensor Placement Geometry

In this subsection, we use the minimum entropy of the posterior target location distribution to characterize the dependency of the localization uncertainty on the sensor network deployment geometry and the sensor observation type through simulations. We define the coverage of a sensor network for localization as the region where the target can be relatively accurately located by the sensor network. The localization uncertainty characteristics obtained in this section provides guidance to identify the coverage of sensor networks and to deploy sensor networks for the optimal localization accuracy in a given region. We have considered three types of information provided by sensor observations, including TDOA, the range to the target, and the direction-of-arrival (DOA) of the target signal.

In simulations as shown in Fig. 10, Fig. 11, Fig. 12, and Fig. 13, we consider two-dimensional localization using a small number of sensors with Gaussian observation uncertainty for simplicity. Given a sensor network deployment geometry, we consider the localization error lower bound at many different locations. The localization error lower bound is quantified using the minimum entropy of the posterior target location distribution. The spatial variation of the localization error lower bound indicates where localization is more accurate and where not. The left sub-figures of Fig. 10, Fig. 12, and Fig.

13, and both sub-figures of Fig. 11, are the map views of the spatial variation of the localization error lower bound. Sensors are denoted by square markers. The right sub-figures of Fig. 10, Fig. 12, and Fig. 13 show the spatial variation of the localization error lower bound in detail along profiles AB and CD that are defined in the corresponding left sub-figure.

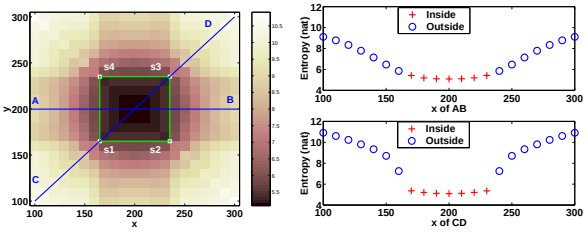

Fig. 10. Spatial variation of localization uncertainty lower bound of a TDOA sensor network.

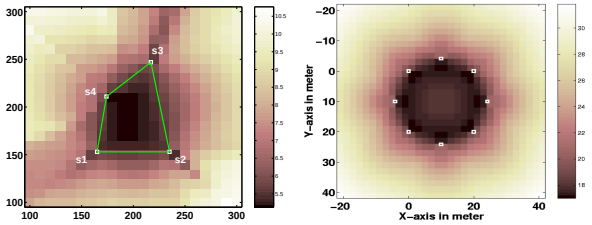

Fig. 11. Spatial variation of localization uncertainty lower bound of an unevenly placed TDOA sensors (left) and the AML based localization (right).

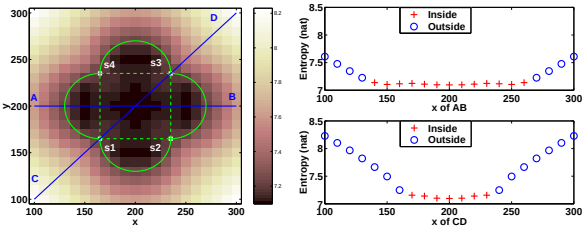

Fig. 12. Spatial variation of localization uncertainty lower bound of a range sensor network.

As shown in Fig. 10, if localization is essentially based on the TDOA information among all sensors, the coverage is the region inside the convex hull of all sensors used. In the right sub-figure of Fig. 10, marker $+$ denotes the relatively small lower bound of localization error inside the convex hull coverage. The convex hull coverage is true no matter whether TDOA sensors are evenly placed or not as shown in the left sub-figure of Fig. 11. The near-field AML algorithm only indirectly and partially relies on the time difference information between sensors for localization [21]. However, the convex hull coverage still holds even for the AML based

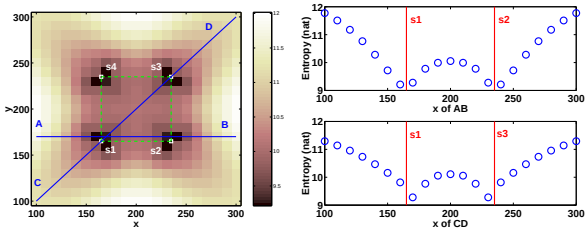

Fig. 13. Spatial variation of localization uncertainty lower bound of a DOA sensor network.

localization as shown in the right sub-figure of Fig. 11. The minimum entropy values in the right sub-figure of Fig. 11 are converted from the CRB of the AML algorithm[22]. Assuming the AML based posterior target location estimation is Gaussian, the conversion follows $H_{\min} = 1 + \ln(2\pi\sigma_a\sigma_b)$, where σ_a and σ_b are the square roots of the two eigen values of the CRB matrix. This result is consistent with the early findings of the convex hull characteristics of TDOA based localization in [23].

As shown in Fig. 12, in contrast to the coverage of the TDOA sensor networks, the coverage of the range sensor networks not only includes the area inside the convex hull of sensors, but also extends outward to the area enclosed by the arcs. These arcs have the convex hull edges as diameters. In the right sub-figure of Fig. 12, marker $+$ denotes the relatively small lower bound of localization error inside the coverage of the range senor networks. This result is consistent with the localization error characteristics of range sensors through the CRB analysis in [24]. When four range sensors are unevenly placed, our simulation indicates that the sensor network coverage is still enclosed by the arcs associated with the convex hull of sensors.

As shown in Fig. 13, the localization uncertainty characteristics using DOA information is very different from those using TDOA or range information. Although a target inside the convex hull of DOA sensors is still more accurately located than a target far from any sensor, the coverage of the DOA sensor networks is better described as the vicinity of individual DOA sensors. In the right sub-figure of Fig. 13, the DOA sensor locations are denoted by vertical bars. We can clearly see that the relatively small lower bound of localization error is near individual DOA sensors. In the simulation as shown in Fig. 13, the standard deviation $\sigma = 180/r + 0.2r$ degrees is assumed for all DOA observations, where r is the distance between the target and the DOA sensor. When σ changes with r differently, simulations indicate that the coverage of a DOA sensor network is still the vicinity of individual sensors, and is similar to that shown in Fig. 13.

V. Conclusion

In this paper, we have treated two related problems in sensor networks for target localization and tracking, namely, the sensor selection problem and the sensor placement problem in a coherent and unified framework based on Bayesian information fusion and information theory. We have described a sensor selection heuristic that approaches the quality of the sensor selection decision of the mutual information criteria but has much less computational complexity than the mutual information criteria. Our sensor selection heuristic is more suitable to sensor networks with moderate computing powers than the mutual information based sensor selections. We have also described a method to compute the posterior target location distribution with the minimum entropy. Using the minimum entropy of the posterior target location estimation, we have characterized the localization uncertainty of sensor networks with different placement geometry and observation types. Such localization uncertainty characteristics provide a strategy to optimize the sensor network deployment geometry in order to achieve the optimal localization accuracy in a given region.

Appendix I
Proof of Eq. (14)

According to Eq. (3) and (10),

$$g(\boldsymbol{z}_1,\ldots,\boldsymbol{z}_N) = -\int p(\boldsymbol{x}|\boldsymbol{z}_1,\ldots,\boldsymbol{z}_N)\ln p(\boldsymbol{x}|\boldsymbol{z}_1,\ldots,\boldsymbol{z}_N)d\boldsymbol{x} \ .$$

$$\begin{aligned}
&\partial g(\boldsymbol{z}_1,\ldots,\boldsymbol{z}_N)/\partial \boldsymbol{z}_i \\
=&-\partial[\int p(\boldsymbol{x}|\boldsymbol{z}_1,\ldots,\boldsymbol{z}_N) ln p(\boldsymbol{x}|\boldsymbol{z}_1,\ldots,\boldsymbol{z}_N)d\boldsymbol{x}]/\partial \boldsymbol{z}_i \\
=&-\int \partial[p(\boldsymbol{x}|\boldsymbol{z}_1,\ldots,\boldsymbol{z}_N)\ln p(\boldsymbol{x}|\boldsymbol{z}_1,\ldots,\boldsymbol{z}_N)]/\partial \boldsymbol{z}_i d\boldsymbol{x} \\
=&-\int \partial p(\boldsymbol{x}|\boldsymbol{z}_1,\ldots,\boldsymbol{z}_N)/\partial \boldsymbol{z}_i \ln p(\boldsymbol{x}|\boldsymbol{z}_1,\ldots,\boldsymbol{z}_N)d\boldsymbol{x} \\
&-\int \partial p(\boldsymbol{x}|\boldsymbol{z}_1,\ldots,\boldsymbol{z}_N)/\partial \boldsymbol{z}_i d\boldsymbol{x} \\
=&-\int \partial p(\boldsymbol{x}|\boldsymbol{z}_1,\ldots,\boldsymbol{z}_N)/\partial \boldsymbol{z}_i (1+\ln p(\boldsymbol{x}|\boldsymbol{z}_1,\ldots,\boldsymbol{z}_N))d\boldsymbol{x} \\
=&0 \ .
\end{aligned}$$

Then, according to Eq. (12),

$$\int \partial p(\boldsymbol{x}|\hat{\boldsymbol{z}}_1,\ldots,\hat{\boldsymbol{z}}_N)/\partial \boldsymbol{z}_i (1+\ln p(\boldsymbol{x}|\hat{\boldsymbol{z}}_1,\ldots,\hat{\boldsymbol{z}}_N))d\boldsymbol{x}$$
$$=0,\ 1\le i\le N \ .$$

One way to satisfy the above equation is

$$\partial p(\boldsymbol{x}|\hat{\boldsymbol{z}}_1,\ldots,\hat{\boldsymbol{z}}_N)/\partial \hat{\boldsymbol{z}}_i = 0,\ 1\le i\le N \ . \qquad (15)$$

Notice

$$p(\boldsymbol{x}|\boldsymbol{z}_1,\ldots,\boldsymbol{z}_N) = Cp(\boldsymbol{z}_1,\ldots,\boldsymbol{z}_N|\boldsymbol{x})p(\boldsymbol{x}), 1\le i\le N \ ,$$

where C is a normalization constant. Thus, Eq. (15) becomes

$$\partial p(\hat{\boldsymbol{z}}_1,\ldots,\hat{\boldsymbol{z}}_N|\boldsymbol{x})/\partial \boldsymbol{z}_i = 0,\ 1\le i\le N \ . \qquad (16)$$

If the noise-free observation is a critical point of the sensor observation model $p(\boldsymbol{z}_1,\ldots,\boldsymbol{z}_N|\boldsymbol{x})$ and maximizes $p(\boldsymbol{z}_1,\ldots,\boldsymbol{z}_N|\boldsymbol{x})$, then

$$\partial p(\boldsymbol{z}_1^{\mathrm{v}},\ldots,\boldsymbol{z}_N^{\mathrm{v}}|\boldsymbol{x})/\partial \boldsymbol{z}_i = 0,\ 1\le i\le N \ ,$$

then one solution to Eq. (16) is

$$\hat{\boldsymbol{z}}_i = \boldsymbol{z}_i^{\mathrm{v}},\ 1\le i\le N \ .$$

This is Eq. (14).
Proof is complete.

Acknowledgment

This work is partially supported by NSF CENS program under Cooperative Agreement CCR-0121778, NSF grant EF-0410438, AROD-MURI PSU Contract 50126, and UC-Discovery grant sponsored by ST Microelectronics.

References

[1] D. Culler, D. Estrin, and M. Srivastava, "Overview of sensor networks," *IEEE Computer*, vol. 37, no. 8, pp. 41–49, Aug. 2004.

[2] S. Kumar, D. Shepherd, and F. Zhao, "Collaborative signal and information processing in microsensor networks," *IEEE Signal Proc. Mag.*, vol. 19, no. 2, pp. 13–14, Mar 2002.

[3] K. Hintz and E. McVey, "A measure of the information gain attributable to cueing," *IEEE T. Syst. Man Cyb.*, vol. 21, no. 2, pp. 434–442, 1991.

[4] J. Manyika and H. Durrant-Whyte, *Data fusion and sensor management: a decentralized information-theoretic approach*. New York: Ellis Horwood, 1994.

[5] N. Xiong and P. Svensson, "Multi-sensor management for information fusion: issues and approaches," *Information Fusion*, vol. 3, no. 2, pp. 163–186, June 2002.

[6] J. Denzler and C. M. Brown, "Information theoretic sensor data selection for active object recognition and state estimation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 2, pp. 145–157, 2002.

[7] J. Liu, J. Reich, and F. Zhao, "Collaborative in-network processing for target tracking," *EURASIP JASP: Special Issues on Sensor Networks*, vol. 2003, no. 4, pp. 378–391, March 2003.

[8] E. Ertin, J. Fisher, and L. Potter, "Maximum mutual information principle for dynamic sensor query problems," in *Proc. IPSN'03*, Palo Alto, CA, April 2003.

[9] H. Wang, K. Yao, G. Pottie, and D. Estrin, "Entropy-based sensor selection heuristic for target localization," in *Proc. IPSN'04*, Berkeley, California, April 2004.

[10] H. Wang, L. Yip, K. Yao, and D. Estrin, "Lower bounds of localization uncertainty in sensor networks," in *Proc. ICASSP*, Montreal, Canada, May 2004.

[11] P. K. Varshney, *Distributed Detection and Data Fusion*. Secaucus, NJ: Springer-Verlag New York, Inc., 1996.
[12] B. Grocholsky, "Information-theoretic control of multiple sensor platforms," Ph.D. dissertation, The University of Sydney, 2002, available from http://www.acfr.usyd.edu.au.
[13] N. Bergman, "Recursive bayesian estimation: Navigation and tracking applications," Thesis of doctor of philosophy in electrical engineering, Departmentof Electrical Engineering, Linkoping University, Sweden, 1999.
[14] D. Fox, J. Hightower, L. Liao, D. Schulz, and G. Borriello, "Bayesian filtering for location estimation," *IEEE Pervasive Computing*, vol. 2, no. 3, pp. 24–33, July-September 2003.
[15] C. E. Shannon, "A mathematical theory of communication," *Bell Systems Technical Journal*, vol. 27, no. 6, pp. 379–423 and 623–656, 1948.
[16] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Trans. of the ASME–Journal of Basic Engineering*, vol. 82, no. Series D, pp. 35–45, 1960.
[17] H. Cramér, *Mathematical methods of statistics*. Princeton, NJ: Princeton University Press, 1946.
[18] C. R. Rao, "Information and the accuracy attainable in the estimation of statistical parameters," *Bull. Calcutta Math. Soc.*, vol. 37, pp. 81–91, 1945.
[19] R. McDonough and A. Whalen, *Detection of Signal in Noise, 2nd Edition*. San Diego, California: Academic Press, 1995.
[20] A. Savvides, W. Garber, R. Moses, and M. Srivastava, "An analysis of error inducing parameters in multihop sensor node localization," *IEEE Trans. Mobile Computing*, to appear.
[21] J. Chen, R. Hudson, and K. Yao, "Maximum-likelihood source localization and unknown sensor location estimation for wideband signals in the near-field," *IEEE T. Signal Proces.*, vol. 50, no. 8, pp. 1843–1854, Aug 2002.
[22] L. Yip, J. Chen, R. Hudson, and K. Yao, "Cramér-rao bound analysis of wideband source localization and doa estimation," in *Proc. SPIE*, vol. 4791, Dec 2002, pp. 304–316.
[23] K. Yao, R. Hudson, C. Reed, D. Chen, and F. Lorenzelli, "Blind beamforming source localization on a sensor array system," in *AWAIRS project presentation at UCLA*, Los Angeles, CA, December 1997.
[24] A. Savvides, W. Garber, S. Adlakha, R. Moses, and M. B. Srivastava, "On the error characteristics of multihop node localization in ad-hoc sensor networks," in *Proc. IPSN'03*, Palo Alto, CA, April 2003.