Data compression in emitter location systems via sensor pairing and selection
Xi Hu and Mark L. Fowler*
Dept. of Electrical and Computer Eng., Binghamton University, Binghamton, NY, USA 13902

ABSTRACT

Data compression ideas can be extended to assess the data quality across multiple sensors to manage the network of sensors to optimize the location accuracy subject to communication constraints. From an unconstrained-resources viewpoint it is desirable to use the complete set of deployed sensors; however, that generally results in an excessive data volume. We have previously presented here results on selecting pre-paired sensors. We have now extended our results to enable optimal joint pairing/selection of sensors.

Pairing and selecting sensors to participate in sensing is crucial to satisfying trade-offs between accuracy and time-line requirements. We propose two methods that use Fisher information to determine sensor pairing/selection. The first method optimally determines pairings as well as selections of pairs but with the constraint that no sensors are shared between pairs. The second method allows sensors to be shared between pairs. In the first method, it is simple to evaluate the Fisher information but is challenging to make the optimal selections of sensors. However, the opposite is true in the second method: it is more challenging to evaluate the Fisher information but is simple to make the optimal selections of sensors.

Keywords: Sensor Selection, Sensor Management, Data Compression, Sensor Networks, TDOA/FDOA Emitter Location, Fisher Information

1. INTRODUCTION

Our general interest is in achieving network-wide optimization over a large number of simultaneously deployed airborne sensors to enable more efficient and effective cooperation within the network of sensors. To provide a concrete perspective we consider the specific scenario of using the sensors to locate a non-cooperative RF emitter using TDOA/FDOA-based methods\(^1\), here TDOA is “Time-Difference-Of-Arrival” and FDOA is “Frequency-Difference-Of-Arrival”, which can be jointly estimated by cross-correlating signals from a pair of the sensors\(^2\). The accuracy of the TDOA/FDOA estimates depend on the signal SNR and the time-frequency structure of the intercepted signal\(^3\); however, the accuracy of the location estimate depends also on the emitter/sensor geometry through the so-called “geometric dilution of precision” or GDOP\(^4\). The goal of our work is to optimize over the set of all sensor assets, under the constraint of limited network communication resources.

The sensors simultaneously intercept an RF emitter’s signal data and then cooperatively share the signal data between paired sensors to estimate the TDOA/FDOA for each sensor pair, which are then used to locate the RF emitter. After data collection at the sensors, they send a small amount of data to a central node where it is possible to determine a rough estimate of the emitter location – accurate enough to assess the impact of the relative emitter-sensor geometry on the location processing task, thus allowing subsequent processing to be optimized with respect to the geometry and error sources. (An alternative to this is when the network is cued with a rough location from some cueing sensor system.) Thus, the central node then uses knowledge of the current positions and trajectories of the remaining sensors to further reduce the participating subset based on the quality and the error sensitivity of their data sets. For example, one sensor may have high-quality data but its position and trajectory give it a poor GDOP, whereas another sensor could have low-quality data but have good GDOP. By eliminating sensors that have negligible usefulness to the final outcome of the task it is possible to significantly reduce the amount of network communication needed to accomplish the task with little degradation of the location accuracy. Further reduction in the needed communication resources is then achieved through

* m Fowler@binghamton.edu; 607-777-6973; http://www.ws.binghamton.edu/fowler/
This work was supported in part by the Air Force Office of Scientific Research through Grant number FA9550-06-1-0249.
location-optimized compression when this final subset of participating sensors shares its data to support the location-
estimation tasks.

We propose various approaches to this problem and discuss trade-offs between them. The first method assumes that the
sensors have pre-paired and share their data between these pairs; sensor selection then consists of selecting pairs to
optimize performance while meeting constraints on number of pairs selected. The second method consists of optimally
determining pairings as well as selections of pairs but with the constraint that no sensors are shared between pairs. The
third method consists of allowing sensors to be shared between pairs.

We discuss several aspects of these three methods. The first method is simple to solve but clearly the pre-pairing
requirement makes this method clumsy and very sub-optimal. In the second method, it is simple to evaluate the Fisher
information but is challenging to make the optimal selections of sensors. However, in the third method things are
reversed in that it is more challenging to evaluate the Fisher information but is simple to make the optimal selections of
sensors.

2. PROBLEM DESCRIPTION

For simplicity we consider only the 2-D geometrical scenario. In the scenario we consider a rough estimate of emitter
location has already been made (either by our system or by a cueing system). As shown in Figure 1, we wish to find the
location of a stationary emitter, denoted by \( \mathbf{u} = [x_e, y_e]^T \), using signals intercepted at \( N \) unmanned aerial vehicle (UAV)
sensors denoted \( S_1 \) to \( S_N \), whose positions are \( \mathbf{x}_i = [x_i, y_i]^T \) and speeds are \( \mathbf{x}_i = [\dot{x}_i, \dot{y}_i]^T \), for \( i = 1, 2, \ldots, N \).

\[ r_i = |\mathbf{x}_i - \mathbf{u}| = \sqrt{(x_i - x_e)^2 + (y_i - y_e)^2} \quad (1) \]

To compute the TDOA/FDOA measurements the sensors must be paired. We consider three types of pairings within the
network of sensors, as shown in Figure 2:
- Type-I: No Sensor Sharing (two pairs that do not share a sensor are said to be “independent pairs”);
- Type-II: De-Centralized Sensor Sharing (i.e., sensors are shared between pairs but no sensor is part of more
  than two pairs);
- Type-III: Centralized Sensor Sharing (i.e., a common reference sensor is used).

For the \( i^{th} \) pair of sensors the TDOA \( \tau_i \) and FDOA \( \omega_i \) between the signals received at the two sensors in the pair are
given by
\[
\tau_i = \frac{1}{c}(r_{i,1} - r_{i,2})
\]
\[
\omega_i = \frac{f_c}{c}(u_{i,1}^T \cdot \dot{x}_{i,1} - u_{i,2}^T \cdot \dot{x}_{i,2}),
\]

(2)

where \( u_{i,k} \) is the unit vector pointing from the \( k^{th} \) sensor in the \( i^{th} \) pair to the emitter, for \( k = 1, 2 \), and \( f_c \) is the transmitted frequency of the transmitter (assumed estimated in advance).

Assume there are \( M \) pairs totally. Let \( \mathbf{\theta}_m = [\tau_{k_m,j_m}, \omega_{k_m,j_m}]^T \) be the parameter vector to be estimated by the \( m^{th} \) pair of sensors, which is paired by \( (k_m)^{th} \) and \( (j_m)^{th} \) sensors, where \( m = 1, 2, \ldots, M \); and \( k_m \neq j_m \). Let \( \hat{\tau}_{k_m,j_m} \) and \( \hat{\omega}_{k_m,j_m} \) be the estimates, \( \Delta \tau_{k_m,j_m} \) and \( \Delta \omega_{k_m,j_m} \) be the estimation errors, then

\[
\hat{\tau}_{k_m,j_m} = \tau_{k_m,j_m} + \Delta \tau_{k_m,j_m}
\]
\[
\hat{\omega}_{k_m,j_m} = \omega_{k_m,j_m} + \Delta \omega_{k_m,j_m}
\]

(3)

Because the estimate \( \hat{\mathbf{\theta}}_m \) is obtained by maximum likelihood (ML) estimator[3], the asymptotic properties of ML estimators[5] gives that the PDF of it is Gaussian with covariance matrix that is the inverse of the Fisher information matrix (FIM), so

\[
\begin{bmatrix}
\Delta \tau_{k_m,j_m} \\
\Delta \omega_{k_m,j_m}
\end{bmatrix}
\sim N(0, F^{-1}_m)
\]

(4)
As we know $\mathbf{F}_m$ depends only on the sensors received signals according to

$$
\mathbf{F}_m = 2 \Re \left[ \frac{\partial \mathbf{s}_m^T}{\partial \theta_m} \mathbf{C}_m^{-1} \frac{\partial \mathbf{s}_m}{\partial \theta_m} \right],
$$

(5)

where $\mathbf{s}_m$ is the vector of received signals and $\mathbf{C}_m$ is the covariance of the AWGN at the $m$th sensor pair. The FIM of $\theta = [\theta_1^T, \theta_2^T, \ldots, \theta_M^T]^T$ has a block structure as

$$
\mathbf{F}_\theta = \begin{bmatrix}
\mathbf{I}_{11} & \mathbf{I}_{12} & \cdots & \mathbf{I}_{1M} \\
\mathbf{I}_{21} & \mathbf{I}_{22} & \cdots & \mathbf{I}_{2M} \\
\vdots & \vdots & \ddots & \vdots \\
\mathbf{I}_{M1} & \mathbf{I}_{M2} & \cdots & \mathbf{I}_{MM}
\end{bmatrix},
$$

(6)

where $\mathbf{I}_{m,k}$ is the cross term FIM between $m$th and $k$th pairs, which is evaluated in the Appendix.

The TDOA/FDOA estimates are then used by the sensor system to estimate the location of the emitter. Because of the asymptotic properties of the ML estimator of TDOA/FDOA we can take the TDOA/FDOA estimates as Gaussian so that the FIM of the estimate of the geo-location is given by

$$
\mathbf{J}_{\text{geo}} = \left[ \mathbf{G}_1^T, \ldots, \mathbf{G}_m^T, \ldots, \mathbf{G}_M^T \right] \mathbf{F}_\theta \left[ \mathbf{G}_1, \ldots, \mathbf{G}_M \right]^T,
$$

(7)

where $\mathbf{G}_m$ is the Jacobian matrix of the $m$th pair of sensors, defined by $\mathbf{G}_m = \frac{\partial \theta_m(u)}{\partial u}$ and calculated by

$$
\mathbf{G}_m = \begin{bmatrix}
(x_{xa} - u)^T / r_{xa} - (x_{ja} - u)^T / r_{ja} \\
(x_{xa} - u)^T \tilde{r}_{xa} / r_{xa}^2 - (x_{ja} - u)^T \tilde{r}_{ja} / r_{ja}^2 + \tilde{x}_{xa}^T / r_{xa} + \tilde{x}_{ja}^T / r_{ja}
\end{bmatrix}.
$$

(8)

Our objective is to select an optimal subset of sensors and pair them as well. The criterion we used to make the decision is the trace of FIM of geo-location as

$$
\max_{\text{all possible subset solutions}} \{ \text{trace}(\mathbf{J}_{\text{geo}}(\text{subset})) \}
$$

(9)

In the following sections, we discuss sensor selection algorithms for the three network types.

### 3. ALGORITHMS

#### 3.1 Pre-paired sensors

When sensors are pre-paired, we simply select pairs instead of sensors. The FIM $\mathbf{F}_m$ and cross-FIM $\mathbf{I}_{m,k}$ are evaluated based on the paring and sensor sharing.
Type-I: No Sensor Sharing—When no sensor is shared the cross-FIMs $I_{m,k}$ are zero. The $F_{I}$ are evaluated individually for each pair. Then $F_{\theta}$ will have block diagonal structure as

$$F_{\theta} = \begin{bmatrix} F_{I_1} & 0 & \ldots & 0 \\ 0 & F_{I_2} & & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \ldots & F_{I_M} \end{bmatrix}$$ (10)

The problem of selecting $K$ sensor pairs from $N$ pairs is specified by

$$\max_{p_1, \ldots, p_N} \{ \text{trace}(p_1 (G_1^T F_{I_1} G_1) + \cdots + p_N (G_N^T F_{I_N} G_N)) \}$$

s.t. $p_1 + \cdots + p_N = K < N$, $p_i \in \{0,1\}$

(11)

The solution of this was discussed in [9]: we simply select the $K$ pre-paired sensor pairs that have the largest values of

$$\text{trace}(J_{geo,k}) = \text{trace}(G_k^T F_{I_k} G_k)$$

(12)

Type-II: De-Centralized Sensor Sharing—Here we treat sensors by sensor sets, where a sensor set is defined as a group of sensors which have no connections to sensors outside the group and do not have any independent pairs inside the group. For the sensor network in Figure 2 (b), the sets are defined as in Figure 3.

The geo-location FIM of each sensor set is computed; for example, the evaluation of set-1 is

$$J_{geo, set-1} = G_1^T F_{I_1} G_1 + G_2^T F_{I_2} G_2 + 2G_1^T F_{I_1,2} G_2$$

(13)

Then the problem of selecting $K$ sensors from $M$ sets is specified by

$$\max_{p_1, \ldots, p_M} \{ \text{trace}(p_1 \cdot J_{geo, set-1} + \cdots + p_M \cdot J_{geo, set-M}) \}$$

s.t. $p_1 \cdot n_1 + \cdots + p_M \cdot n_M = K < N$, $p_i \in \{0,1\}$

$n_i$ is the number of sensors in set-i

(14)
and can be easily solved. For example, if we are asked to select 5 sensors, we can check the set which has 5 sensors, or the two sets which have 2 sensors and 3 sensors respectively, and add the trace of the two sets up, compare it with the one with 5 sensors and choose the larger one.

Type-III: Centralized Sensor Sharing—For the pre-paired case, the central sensor is already specified and the remaining \( N - 1 \) sensors pair with it to form \( N - 1 \) centralized pairs. There are \( C_{N-1}^K \) possible ways to select \( K \) pairs. The FIM of this set will have the following structure

\[
J_{\text{geo},k} = \sum_{k=1}^{K} G_k^T F_{I_k} G_k + \sum_{m=1,k=m+1}^{m-K-1,k+K} 2G_{m,k}^T I_{m,k} G_{k}
\]  

(15)

If the \( r^{th} \) sensor is the reference sensor then the Appendix states that \( F_{I_k} = F_k + F_r \) and \( I_{m,k} = F_r \). In this case, we have to evaluate the trace of all the FIMs of geo-location of the \( C_{N-1}^K \) possible combinations, and choose the largest one:

\[
\max_{\text{set}, k \in \{C_{N-1}^K\}} \{\text{trace}(J_{\text{set},k})\}
\]

\[
\{C_{N-1}^K\} \text{ is all the possible combination set}
\]  

(16)

3.2 Non-pre-paired sensors

We are given a set of sensors and asked to optimally choose a subset and the optimal pairings as well. In this case the pairing provides more flexibility to enable better performance but it introduces additional complexity as well.

Type-I Pairing of Sensors: No Sensor Sharing—For \( N \) sensors, there could be \( N/2 \) independent pairs. To choose \( K \) pairs is a time-consuming work if we enumerated all the possible solutions. For example, \( N = 10 \), there are \( C_{10}^2 = 45 \) possible pairs, and \( (N-1) \cdot (N-3) \cdots 3 \cdot 1 = 945 \) possible ways to make 5 pairs as a subset. Fortunately, since there is no sensor sharing and we select sensors pair by pair, the selection of the next pair will not affect the selection of the previous one. This yields a tree structure and allows use of integer dynamic programming method. For this paper we used the “Branch and Bound” method to choose a pair at each step. The objective function is

\[
\max_{\text{all feasible solutions}} \left\{ \sum_{k=1}^{K} \text{trace}(J_{\text{geo},k}^{\text{pair in the solution}}) \right\}
\]

(17)

A “feasible solution” means any selection/pairing of sensors where no sensors are shared and the selected number of sensors is as required.

Type-II Pairing of Sensors: De-Centralized Sensor Sharing—For \( N \) sensors, there could be \( C_{N}^2 \) possible pairs. To choose \( K \) pairs, there are \( C_{N}^K \) possible ways to pair and then select \( K \) pairs. For example, \( N = 10, K = 5 \), the number of ways is 122,1759, which is quite large and nonconductive to listing all of them. But fortunately, among all this large number of ways to pair and select, only a small number of them are unique. We have established the following theorem.

Theorem: For \( M \) sensors, at most independent \( M - 1 \) pairs can be used as a “sensor set”; and different pairing methods of the \( M \) sensors to make \( M - 1 \) independent pairs will result in the same CRLB of geo-location.

We can exploit this result to simplify the optimal selection and pairing for this case. When we are given \( N \) sensors and asked to make \( K \) pairs, there are many solutions for this network. We can use at least \( K+1 \) sensors to make it or at
most $2K$. Since the main advantage to share sensors is to save some sensor energies, we would like to use the number of sensors as less as possible. So here we only choose $K + 1$ sensors to make $K$ pairs.

For example, for given $N = 7$ and $K = 3$ pairs needed, compute the FIM of geo-location of all $C_7^4 = 35$ solutions, and find the one with the largest trace. Inside each solution, sensors are “paired by sequence.” For example, as in Figure 4, the solution set is $\{S_4, S_1, S_2, S_3\}$.

![Figure 4 An example of pairing by sequence](image)

### 4. SIMULATION RESULTS

To demonstrate the capability of the sensor selection methods we present some simulation results for the case of locating an emitter with a random lay-down of 14 sensors. The sensor selection proceeds as follows. Each sensor intercepts the emitter signal data at SNRs in the range of 10–15dB (where the SNR variation is assumed to depend quadratically on the range to the emitter). The full set of sensors share a very small amount of data to obtain a rough estimate of the emitter location; alternatively, we could consider the case where the system is cued by some other sensor system that provides a rough location that is to be improved using our sensors.

Figure 5 shows the performance of sensor selections without sensor sharing. We select 6 to 14 sensors to make 3 to 7 pairs, shown on the horizontal axis. The vertical axis shows the standard deviation of the geo-location error versus the number of sensors/pairs selected. The upper curve (-Δ-) shows the performance for the pre-paired sensor case without sharing; the lower curve (-O-) shows the performance when using the selection and pairing method discussed above for the case of no sensor sharing. Not surprisingly, the ability to select the pairing on the basis of the sensor geometry and the rough emitter location enables better performance than using pre-paired sensors.

![Figure 5 Performance of sensor selection w/o sharing](image)
Figure 6 shows the time consumption used in pairing sensors for the non-sharing case versus the number of sensors/pairs selected. The upper line (-Δ-) shows the time required for the enumeration-based method, the lower one (-O-) shows the time required for our selection and pairing method. These time results are for matlab-based implementations.

Figure 7 shows the performance of sensor selections allowing sensor sharing. We select 5 to 11 sensors from 12, to make 4 to 10 pairs. It also shows the standard deviation of the geo-location error versus the number of sensors/pairs selected. The upper curve (-Δ-) shows the performance for the pre-paired sensor case with sharing; the lower curve (-O-) shows the performance using our selection and pairing method with sharing that is based on the Theorem in Section 3.2.

5. DISCUSSION

The results above show that it is possible to select and pair an optimal subset of sensors while significantly retaining performance levels. The sensor selection optimization problem was based on the fact that the geometry property and data quality of sensors play important roles in the emitter location estimation. We have used Fisher information to capture this inter-play between data quality and geometry. We have discussed different situations: (i) pre-paired sensors vs. optimally pairing the sensors, and (ii) allowing shared sensors or not. Following are some general conclusions made from this work.
From the point of view of data compression, by reducing the number of sensors needed to achieve a desired accuracy we have reduced the total amount of data that needs to be communicated within the network. Coupled with compression algorithms tailored for emitter location\cite{10} this significantly increases the reduction of data to be communicated. For example, if the intersensor communication employs a compression ratio of 5:1 (a reasonable amount given our previous results\cite{10}) and we reduce the number of total sensors by a factor of two (reasonable given the results in Section 4) then the total network-wide compression ratio is 10:1, which is a significant reduction.

**Conclusions: Without Sensor Sharing**

- FIM of Geo-Location is easy to calculate, since each pair is independent;
- However, the pairing method is more complicated, since we need to consider all the possible pairing ways;
- From a system point of view, the communication among different pairs can be done simultaneously;
- The number of pairs needed is small; beyond a certain point the accuracy improves slowly as more pairs are selected to participate.

**Conclusions: With Sensor Sharing**

- For a total of \(N\) sensors we can have as many as \(N - 1\) pairs, the more the higher accuracy of location estimation;
- Fortunately, FIM of all the possible independent sets are the same, so we do not need to consider about the pairing method. One simple way is to pair the sensors in nature order. This is the main result of this work and leads to a major reduction in the optimization processing required.
- However, since not all the pairs are uncoupled, there are cross terms in the TDOA/FDOA FIM. This complicates the computation required to support the optimization processing.
- Some sensors work in more than one pair, the communication among them needs to be considered carefully to avoid collision. This will be the focus of future work.

**APPENDIX: EVALUATION OF FIM CROSS-TERM**

Consider the case where two pairs share one sensor, as shown in [Figure 8](#).

The three received signals at the sensors are
where \( s[nT] \) is the sampled transmitted signal, and \( \omega_i[n], i = 1, 2, 3 \) is the AWGN received by sensor.

Let

\[
\mathbf{s} = \begin{bmatrix} s_1 \\ s_2 \\ s_3 \end{bmatrix} \quad \text{and} \quad \tau_{13} = \tau_1 - \tau_3; \quad \tau_{23} = \tau_2 - \tau_3
\]

(19)

The cross term FIM between pair-1 and pair-2 can be evaluated as

\[
\mathbf{I}_{1,2} = \left( \frac{\partial \mathbf{s}}{\partial \tau_{1,3}} \right)^H \cdot \left( \frac{\partial \mathbf{s}}{\partial \tau_{2,3}} \right)
\]

(20)

Since

\[
\frac{\partial \mathbf{s}}{\partial \tau_{1,3}} = \begin{bmatrix} \frac{\partial s_1}{\partial \tau_1} \\ \frac{\partial s_2}{\partial \tau_1} \\ \frac{\partial s_3}{\partial \tau_1} \end{bmatrix} = \begin{bmatrix} 0 \\ \frac{\partial s_2}{\partial \tau_2} \\ \frac{\partial s_3}{\partial \tau_2} \end{bmatrix} = \begin{bmatrix} 0 \\ \frac{\partial s_2}{\partial \tau_3} \\ \frac{\partial s_3}{\partial \tau_3} \end{bmatrix}
\]

(21)

also

\[
\frac{\partial \mathbf{s}}{\partial \tau_{2,3}} = \begin{bmatrix} \frac{\partial s_1}{\partial \tau_2} \\ \frac{\partial s_2}{\partial \tau_2} \\ \frac{\partial s_3}{\partial \tau_2} \end{bmatrix} = \begin{bmatrix} 0 \\ \frac{\partial s_2}{\partial \tau_2} \\ \frac{\partial s_3}{\partial \tau_2} \end{bmatrix} = \begin{bmatrix} 0 \\ \frac{\partial s_2}{\partial \tau_3} \\ \frac{\partial s_3}{\partial \tau_3} \end{bmatrix}
\]

(22)

Substitute (21) and (22) into (20), we get

\[
\mathbf{I}_{1,2}^{(1,2)} = \left( \frac{\partial s_1}{\partial \tau_3} \right)^H \cdot \left( \frac{\partial s_2}{\partial \tau_3} \right)
\]

(23)
This is exactly the FI of TDOA of sensor $S_1$’s received signal. Following the same rule we get

$$I_{1,2} = \left( \frac{\partial s}{\partial \tau_{1,3}} \right)^H \cdot \left( \frac{\partial s}{\partial v_{1,3}} \right)^H \cdot \left( \frac{\partial s}{\partial \tau_{3}} \right)$$  \hspace{1cm} (24)$$

$$I_{1,2} = \left( \frac{\partial s}{\partial \tau_{1,3}} \right)^H \cdot \left( \frac{\partial s}{\partial v_{2,3}} \right)^H \cdot \left( \frac{\partial s}{\partial \tau_{3}} \right)$$  \hspace{1cm} (25)$$

$$I_{1,2} = \left( \frac{\partial s}{\partial v_{1,3}} \right)^H \cdot \left( \frac{\partial s}{\partial \tau_{2,3}} \right)^H \cdot \left( \frac{\partial s}{\partial \tau_{3}} \right)$$  \hspace{1cm} (26)$$

Therefore the FIM cross-term between pairs is just the FIM of the shared sensor itself. The FIM of

$$\theta = [\tau_{1,3}, v_{1,3}, \tau_{2,3}, v_{2,3}]^T$$

is

$$J(\theta) = \begin{bmatrix} F_1 & I_{1,2} \\ I_{1,2} & F_2 \end{bmatrix} = \begin{bmatrix} F_1 + F_3 \\ F_3 \end{bmatrix}$$  \hspace{1cm} (27)$$

where $F_i$ is the FIM of TDOA/FDOA of $i^{th}$ sensor.

REFERENCES