

# Energy Based Acoustic Source Localization

Xiaohong Sheng and Yu-Hen Hu\*

University of Wisconsin – Madison,  
Department of Electrical and Computer Engineering  
Madison WI 53706, USA,  
sheng@ece.wisc.edu, hu@engr.wisc.edu,  
<http://www.ece.wisc.edu/~sensit>

**Abstract.** A novel source localization approach using acoustic energy measurements from the individual sensors in the sensor field is presented. This new approach is based on the acoustic energy decay model that acoustic energy decays inverse of distance square under the conditions that the sound propagates in the free and homogenous space and the targets are pre-detected to be in a certain region of the sensor field. This new approach is power efficient and needs low communication bandwidth and therefore, is suitable for the source localization in the distributed sensor network system. Maximum Likelihood (*ML*) estimation with Expectation Maximization (*EM*) solution and *projection* solution are proposed to solve this energy based source location (*EBL*) problem. Cramer-Rao Bound (*CRB*) is derived and used for the sensor deployment analysis. Experiments and simulations are conducted to evaluate *ML* algorithm with different solutions and to compare it with the Non-linear Least Square (*NLS*) algorithm using energy ratio function that we proposed previously. Results show that energy based acoustic source localization algorithms are accurate and robust.

## 1 Introduction

Efficient collaborative signal processing algorithms that consume less energy for computation and communication are important in wireless distributed sensor network communication system [1]. An important collaborative signal processing task is source localization. The objective is to estimate the positions of the moving targets within a sensor field that is monitored by a sensor network. In this paper, our focus will be on collaborative source localization based on acoustic signatures.

Most localization methods depend on three types of physical variables measured by or derived from sensor readings for localization: time delay of arrival (*TDOA*), direction of arrival (*DOA*) and received sensor signal strength or power. *DOA* can be estimated by exploiting the phase difference measured at receiving sensors [2], [3], [4], [5],[6] and is applicable in the case of a coherent, narrow band source. *TDOA* is suitable for broadband source and has been extensively

---

\* This project is supported by DARPA under grant no. F 30602-00-2-0555

investigated [7], [8] [9], [10], [11], [12]. In practice, *DOA* measurements typically require costly antenna array on each node. The *TDOA* techniques require a high demand on the accurate measurement or estimation of time delay. In contrast, received sensor signal strength is comparatively much easier and less costly to obtain from the time series recordings from each sensor.

In [13], we show that, for single target at noiseless situation, each energy ratio dictates that the potential target location must be on a hyper-sphere within the sensor field. With noise taken into account, the target location is solved as the position that is closest to all the hyper-spheres formed by all energy ratios in the least square sense. Using this energy ratio function, the source energy is eliminated, the task of source localization estimation can be simplified by solving a Nonlinear Least Square (*NLS*) problem. Yet, this method can only be used for single target localization.

In this paper, we presented a novel approach to estimate the source location based on acoustic energy measured at individual sensors in the sensor field. We set up an acoustic energy decay model in the free and homogenous (no gusty wind) space under the conditions that the acoustic sources are not far away from the sensors and also not too close to the sensors so that they can be treated as omni-directional points.

Based on this acoustic energy decay model, we presented a new approach using acoustic energy measurements from the individual sensors in the sensor field to locate the targets in the region. Maximum Likelihood (*ML*) estimation with Expectation Maximization (*EM*) solution and project solution using *Exhaustive Search (ES)* and *Multi-Resolution (MR) search* are proposed to solve this *EBL* problem. *Cramer-Rao Bound (CRB)* is derived and used to analyze the sensor deployment to improve the performance of *EBL* algorithms.

Experiments and simulations were conducted to verify the energy decay model and to evaluate different algorithms and solutions of this *EBL* problem. Results show that energy based localization with *ML* estimation using *projection* solution outperforms other method by the cost of heavy computation burden. *Projection* solution with *MR* search reduces the computation burden a lot with the cost of reducing the performance a little bit.

Performance variation of *ML* estimation when targets have different energy intensity is also analyzed. It shows that, when the two targets have significant difference of energy source, the target with smaller energy becomes more ambiguous, and therefore, we get less accurate localization estimation for that target.

This paper is organized as follows: In section2, we formulate the acoustic energy decay model in the sensor network system when certain conditions are satisfied. Based on this energy decay model, we derive *ML* estimation and its *EM* solution and *projection* solution with *MR* and *exhaustive* search to solve this *EBL* problem. In section 3, we derive the *CRB* for this *EBL* problem. The effect of sensor deployment to the *CRB* is also analyzed in this section. Experiments and simulations are provided in section 4 to evaluate and compare different algorithms and solutions for solving the *EBL* problems. Conclusion and future work are given in section 5.

## 2 Energy-Based Source Localization

Energy-based source localization is motivated by a simple observation that the sound level decreases when the distance between sound source and the listener becomes large. By modelling the relation between sound level (energy) and distance from the sound source, one may estimate the source location using multiple energy reading at different, known sensor locations.

### 2.1 An Acoustic Energy Decay Model

The way sound propagates with the distance from the source is dependent on the size and shape of the source, the surrounding environment, prevailing air currents and the frequencies of the propagating sound. Other factors that may affect sound propagation may include wind direction, strength of wind, vegetation such as forest and other obstructions.

To simplify the problem, we make some assumptions in developing the energy decay model for energy based source localization in the wireless sensor network system. These assumptions are normally satisfied in certain sensor network systems. The assumptions we made are:

1. Sound propagates in the free air,
2. Target is pre-detected to be in a particular region of a sensor field. The region size is not very big so that targets are not far from the sensors. (For example, in our experiment, the region size is about  $100 * 100M^2$ )
3. Sound source can still be treated as an omni-directional point. (We can assume the dimension of the engine of the vehicle is relative small compared with the distance between the sensor and the vehicle).
4. The propagation medium (air) is roughly homogenous ( i.e. no gusty wind) and there is no sound reverberation.

In such environment, the acoustic intensity attenuated at a rate that is inversely proportional to the distance between source and the sensor [14]. Since sound waveform is additive, the acoustic wave intensity signature received by each sensor is:

$$x_i(n) = s_i(n) + \nu_i(n) \quad (1)$$

Where:  $s_i(n) = \gamma_i \sum_{k=1}^K \frac{a_k(n-t_{ki})}{\|\boldsymbol{\rho}_k(n-t_{ki})-\mathbf{r}_i\|}$  and  $i = 1, 2, \dots, N$

In the above equation,  $x_i(n)$  is the  $n^{th}$  acoustic signature sampled on the  $i^{th}$  sensor over a time interval  $[1/f_s]$  by a matched filter,  $f_s$  is the sampling frequency;  $\nu_i(n)$  is the zero-mean additive white Gaussian noise (AWGN) on the  $n^{th}$  time interval;  $K$  is the number of targets;  $N$  is the number of the sensors of the particular region in the sensor field;  $a_k(n-t_{ki})$  is a scalar denoting the acoustic source intensity emitted by the  $k^{th}$  target;  $t_{ki}$  is the propagation delay from the  $k^{th}$  source to the  $i^{th}$  sensor;  $\boldsymbol{\rho}_k$  is the  $p \times 1$  vector denoting the coordinates of the  $k^{th}$  target;  $\mathbf{r}_i$  is a  $p \times 1$  vector denoting the Cartesian coordinates of the  $i^{th}$  stationary sensor;  $p$  is the dimension of location;  $\gamma_i$  is a scaling factor corresponding to the sensor gain of the  $i^{th}$  acoustic sensor.

Assume  $s_i(n)$  and  $v_i(n)$  are uncorrelated,  $a_{k_1}$  and  $a_{k_2}$  are uncorrelated for  $k_1 \neq k_2$ ,  $E[v_i(n)] = 0$ , and  $E[a_k(n)] = 0$ , we get:

$$E [x_i^2(n)] = E [s_i^2(n)] + E [\nu_i^2(n)] \tag{2}$$

where  $E [s_i^2(n)] = g_i \sum_{k=1}^K \frac{S_k(n-t_{ki})}{\|\boldsymbol{\rho}_k(n-t_{ki})-\mathbf{r}_i\|^2}$

In above,  $g_i = \gamma_i^2$  and  $S_k(n-t_{ki}) = E [a_k^2(n-t_{ki})]$ .

Since we assume that the targets are pre-detected to be in a particular region of a sensor field and the region size is not big, we can assume that the targets are not far from the sensors in the region. Therefore, the propagation delay  $t_{ki}$  is small enough that we can assume:  $a(n-t_{ki}) \approx a(n)$  and  $\boldsymbol{\rho}(n-t_{ki}) \approx \boldsymbol{\rho}(n)$ .

The expectation of energy is calculated by averaging over a time window  $T = M/f_s$ , where  $M$  is the number of sample points we used for averaging the energy,  $f_s$  is the sampling frequency. Denote  $E [x_i^2(n)]$  as  $y_i(t)$ ,  $E [a_k^2(n)]$  as  $y_s(t)$  and  $E [\nu_i^2(n)]$  as  $\varepsilon_i(t)$ , we get the energy decay model as:

$$y_i(t) = y_s(t) + \varepsilon_i(t) = g_i \sum_{j=1}^K \frac{S_j(t)}{\|\boldsymbol{\rho}_j(t) - \mathbf{r}_i\|^2} + \varepsilon_i(t) \tag{3}$$

Where  $t = \frac{T}{2}, \frac{3T}{2}, \frac{5T}{2}, \dots$

Background noise  $\nu_i(n)$  is independent zero mean AWGN with variance  $\sigma_n^2$ ,  $\nu_i^2(n)$  is independent  $\chi^2$  distributed with mean  $\sigma_n^2$  and variance  $\frac{2\sigma_n^4}{M}$ . If  $M$  is sufficiently large (practically  $M > 30$ ), by central limit theorem,  $\varepsilon_i$  is approximately normal:  $\varepsilon \sim N(\sigma_n^2, \frac{2\sigma_n^4}{M})$ .

### 2.2 Maximum Likelihood Estimation for EBL Problem

To simplify our notation, in the following parts, we will not denote time t explicitly in our equation. All parameters refer to the same time window automatically, i.e., we denote  $y_i$  for  $y_i(t)$ .

Following we will introduce the ML estimation with different solutions to estimate the source location. Note that the estimation is based on single frame of energy readings from different individual sensors. Estimation based on sequential energy readings are under developing.

Define

$$\mathbf{Z} = \left[ \frac{y_1-\mu_1}{\sigma_1} \quad \frac{y_2-\mu_2}{\sigma_2} \quad \dots \quad \frac{y_N-\mu_N}{\sigma_N} \right]^T \tag{4}$$

Equation (3) can be simplified as:

$$\mathbf{Z} = \mathbf{GDS} + \boldsymbol{\xi} = \mathbf{HS} + \boldsymbol{\xi} \tag{5}$$

Where:

$$\mathbf{S} = [S_1 \ S_2 \ \dots \ S_K]^T \tag{6}$$

$$\mathbf{H} = \mathbf{GD} \tag{7}$$

$$\mathbf{G} = \text{diag} \left[ \frac{g_1}{\sigma_1} \frac{g_2}{\sigma_2} \dots \frac{g_n}{\sigma_n} \right] \tag{8}$$

$$\mathbf{D} = \begin{bmatrix} \frac{1}{d_{11}^2} & \frac{1}{d_{12}^2} & \dots & \frac{1}{d_{1K}^2} \\ \frac{1}{d_{21}^2} & \frac{1}{d_{22}^2} & \dots & \frac{1}{d_{2K}^2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{d_{n1}^2} & \frac{1}{d_{n2}^2} & \dots & \frac{1}{d_{nK}^2} \end{bmatrix} \tag{9}$$

$d_{ij} = |\boldsymbol{\rho}_j - \mathbf{r}_i|$  is the Euclidean distance between the  $i^{th}$  sensor and the  $j^{th}$  source.

Then,  $\xi_i = \frac{(\varepsilon_i - \mu_i)}{\sigma_i} \sim N(0, 1)$ ,  $z_i = \frac{y_i - \mu_i}{\sigma_i} \sim N\left(\frac{g_i}{\sigma_i} \sum_{j=1}^K \frac{S_j}{d_{ij}^2}, 1\right)$

The unknown parameters  $\boldsymbol{\theta}$  in the above function is:

$$\boldsymbol{\theta} = [\boldsymbol{\rho}_1^T \ \boldsymbol{\rho}_2^T \ \dots \ \boldsymbol{\rho}_K^T \ S_1 \ S_2 \ \dots \ S_K]^T$$

The log-likelihood function of equation (5) is:

$$\ell(\boldsymbol{\theta}) \sim \frac{-1}{2} \sum_{i=1}^N \left\| z_i - \frac{g_i}{\sigma_i} \sum_{j=1}^K \frac{S_j}{d_{ij}^2} \right\|^2 = \frac{-1}{2} \|\mathbf{Z} - \mathbf{GDS}\|^2 \tag{10}$$

*ML* estimate of the parameters  $\boldsymbol{\theta}$  is the values that maximize  $\ell(\boldsymbol{\theta})$ , or equivalently, minimize

$$\mathbf{L}(\boldsymbol{\theta}) = \|\mathbf{Z} - \mathbf{GDS}\|^2 \tag{11}$$

Equation (11) has  $K(p + 1)$  unknown parameters, there must be at least  $K(p + 1)$  or more sensors reporting acoustic energy readings to yield an unique solution to this nonlinear least square problem.

**Expectation Maximization Solution.** Define pseudoinverse of  $\mathbf{H}$  as  $\mathbf{H}^\dagger$ , perform reduced SVD of  $\mathbf{H}$ , and setting  $\frac{\partial \mathbf{L}}{\partial \mathbf{S}}$  to be zero, we get:

$$\mathbf{S} = \mathbf{H}^\dagger \mathbf{Z} \tag{12}$$

Where:

$$\mathbf{H} = \mathbf{GD} = \mathbf{U}_H \boldsymbol{\Sigma}_H \mathbf{V}_H^T \tag{13}$$

$$\mathbf{H}^\dagger = (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T \tag{14}$$

Now, set the gradient of  $L$  with respect to  $\boldsymbol{\rho}_j$  to zero, we get:

$$\nabla_{\boldsymbol{\rho}_j} L = 2s_j \sum_{i=1}^N \frac{g_i}{\sigma_i} \left( \frac{\boldsymbol{\rho}_j - \mathbf{r}_i}{d_{ij}^4} \right) \left( z_i - \frac{g_i}{\sigma_i} \sum_{m=1}^K \frac{s_m}{d_{im}^2} \right) = 0 \tag{15}$$

Where the relation

$$\nabla \boldsymbol{\rho}_j \left( \frac{1}{d_{im}^2} \right) = \begin{cases} \frac{\boldsymbol{\rho}_m - \mathbf{r}_i}{-d_{im}^4}, & j = m \\ 0, & j \neq m \end{cases}$$

is used. Solving equation (15) for  $\mathbf{j}$ , we have

$$\boldsymbol{\rho}_j = \frac{\sum_{i=1}^N \frac{g_i}{\sigma_i} \left( \frac{1}{d_{ij}^4} \right) \left( z_i - \frac{g_i}{\sigma_i} \sum_{m=1}^K \frac{s_m}{d_{im}^2} \right) \mathbf{r}_i}{\sum_{i=1}^N \frac{g_i}{\sigma_i} \left( \frac{1}{d_{ij}^4} \right) \left( z_i - \frac{g_i}{\sigma_i} \sum_{m=1}^K \frac{s_m}{d_{im}^2} \right)} \quad (16)$$

Equation (16) represents  $Kp$  nonlinear constraints on the target location coordinates  $\{\boldsymbol{\rho}_j, 1 \leq j \leq K\}$ . Note that  $\boldsymbol{\rho}_j$  appears on both sides of equation (16) because  $d_{ij}$  contains  $\boldsymbol{\rho}_j$  implicitly.  $\boldsymbol{\rho}_j$  can't be solved explicitly. Yet, we can solve it by iterative procedure, a special case of Expectation Maximization (*EM*) algorithm [15]. In [16], it is proved that such special case of *EM* algorithm is guaranteed to be convergence (in fact, all *EM* algorithms are guaranteed to be convergence). Besides, it avoids the complexities of non-linear optimization algorithms. The procedure of this algorithm is as follows:

*EM Algorithm*

*Initialization:* Initial estimates of  $\{\boldsymbol{\rho}_j, 1 \leq j \leq K\}$

*Repeat until convergence*

*Expectation Step.* Estimate  $\mathbf{S}$  using equation (12), update  $\mathbf{S}$ .

*Maximization Step.* Substitute  $\mathbf{S}$  into equation (16), update  $\boldsymbol{\rho}_j$

**Projection Solution with (MR) Search.** Insert (12) into the cost function (11), we get modified cost function as follows:

$$\begin{aligned} \arg_{\{\rho_1, \rho_2, \dots, \rho_k\}} \min L &= \arg_{\{\rho_1, \rho_2, \dots, \rho_k\}} \min (\mathbf{Z}^T (\mathbf{I} - \mathbf{P}_H)^T (\mathbf{I} - \mathbf{P}_H) \mathbf{Z}) \\ &= \arg_{\{\rho_1, \rho_2, \dots, \rho_k\}} \max (\mathbf{Z}^T \mathbf{P}_H^T \mathbf{Z}) = \arg_{\{\rho_1, \rho_2, \dots, \rho_k\}} \max \mathbf{Z}^T \mathbf{U}_H \mathbf{U}_H^T \mathbf{Z} \end{aligned} \quad (17)$$

Where

$$\mathbf{P}_H = \mathbf{H}(\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T = \mathbf{U}_H \mathbf{U}_H^T \quad (18)$$

is the projection matrix of  $\mathbf{H}$ .

For single source,  $j = 1$ ,

$$\mathbf{H} = \left[ \frac{g_1}{\sigma_1 d_1^2}, \frac{g_2}{\sigma_2 d_2^2}, \dots, \frac{g_n}{\sigma_n d_n^2} \right]^T, \quad \mathbf{U}_H = \frac{\mathbf{H}}{\|\mathbf{H}\|}$$

Exhaustive search can be used to get the source location to maximize function (17). However, the computation complexity is very high. For example, suppose our detected search region is  $128 \times 128$ , if we use *exhaustive* search using the grid size of  $8 \times 8$ , we need  $256^K$  times of search for every estimation point, where  $K$  is the number of the targets. Rather, we can use *MR* search to reduce the number of search times. For example, we can use the initial search grid size  $16 \times 16$  followed by the fine search grid size  $8 \times 8$ , then, the number of search times is

reduced to  $64^K + 4^K$ . For two targets, it needs 4112 search times using *MR* search with this search strategy and  $256^2$  search times using *exhaustive* search to get one estimation. We can further reduce the number of search times by reducing our search region based on the previous location estimation, the time interval between two localization operation, possible vehicle speed and estimation error. In our simulation, all these conditions are used. The search area we used for the *projection* solution is  $(x_i - 32, x_i + 32) \times (y_i - 32, y_i + 32)$ , where  $(x_i, y_i)$  is the previous estimation location of the  $i^{\text{th}}$  target. Therefore, for single target, we need only 20 search; for two targets, we need 272 search for every localization estimation, which is feasible for our distributed wireless networking system.

### 3 Cramer-Rao Bounds and Sensor Deployment Analysis

Cramer-Rao Bound (*CRB*) is a theoretical lower bound of the variance that we can reach for the unbiased estimation. It is useful to indicate the performance bounds of a particular algorithm. *CRB* also facilitates analysis of factors that impact most on the performance of an algorithm. *CRB* is defined as the inverse of the *Fisher Matrix*:

$$\mathbf{J} = -E \left( \frac{\partial}{\partial \boldsymbol{\theta}} \left[ \frac{\partial}{\partial \boldsymbol{\theta}} \ln f_{\boldsymbol{\theta}}(\mathbf{Z}) \right] \right)$$

For the problem with log-likelihood function described as equation (10), *Fisher matrix* is:

$$\mathbf{J} = \frac{\partial (\mathbf{DS})^T}{\partial \boldsymbol{\theta}} \mathbf{G}^T \mathbf{G} \frac{\partial (\mathbf{DS})}{\partial \boldsymbol{\theta}^T} \quad (19)$$

$$\frac{\partial \mathbf{DS}}{\partial \boldsymbol{\theta}^T} = \left[ \frac{\partial \mathbf{DS}}{\partial \rho_1^T} \quad \frac{\partial \mathbf{DS}}{\partial \rho_2^T} \quad \cdots \quad \frac{\partial \mathbf{DS}}{\partial \rho_K^T} \quad \frac{\partial \mathbf{DS}}{\partial \mathbf{S}^T} \right] \quad (20)$$

$$\frac{\partial \mathbf{DS}}{\partial \mathbf{S}^T} = \mathbf{D} \quad (21)$$

$$\mathbf{B}_j = \frac{\partial \mathbf{DS}^T}{\partial \rho_j} = \left[ \frac{-2S_j}{d_{1j}^3} \mathbf{b}_{1j} \quad \frac{-2S_j}{d_{2j}^3} \mathbf{b}_{2j} \quad \cdots \quad \frac{-2S_j}{d_{Nj}^3} \mathbf{b}_{Nj} \right]^T \quad (22)$$

In above equation,  $\mathbf{b}_{ij}$  is the unit vector from source  $j$  to sensor  $i$ , which can be expressed as:

$$\mathbf{b}_{ij} = \frac{\partial d_{ij}}{\partial \rho_j} = \frac{\boldsymbol{\rho}_j - \mathbf{r}_i}{d_{ij}}$$

Define:

$$\mathbf{B} = [\mathbf{B}_1 \quad \mathbf{B}_2 \quad \cdots \quad \mathbf{B}_K] \quad (23)$$

We get the *Fisher Matrix*  $\mathbf{J}$  as follows:

$$\mathbf{J} = \begin{bmatrix} \mathbf{B}^T \\ \mathbf{D}^T \end{bmatrix} \mathbf{G}^T \mathbf{G} \begin{bmatrix} \mathbf{B} & \mathbf{D} \end{bmatrix} \tag{24}$$

Note that in the above equations,  $\mathbf{J}$  is  $(p + 1)K \times (p + 1)K$  matrix,  $\mathbf{B}_j$  is  $N \times p$  matrix,  $\mathbf{B}$  is  $n \times Kp$ ,  $\mathbf{D}$  is  $N \times K$  matrix,  $\mathbf{G}$  is  $N \times N$  matrix and  $\mathbf{b}_{ij}$  is  $p \times 1$  vector

The CRB is:

$$\mathbf{J}^{-1} = \left( \begin{bmatrix} \mathbf{B}^T \\ \mathbf{D}^T \end{bmatrix} \mathbf{G}^T \mathbf{G} \begin{bmatrix} \mathbf{B} & \mathbf{D} \end{bmatrix} \right)^{-1} \tag{25}$$

For single target, the formula is reduced to:

$$\mathbf{J} = \begin{bmatrix} \mathbf{J}_{11} & \mathbf{J}_{12} \\ \mathbf{J}_{21} & J_{22} \end{bmatrix} \tag{26}$$

Where:

$$\mathbf{J}_{11} = \sum_{i=1}^n \frac{4s^2 g_i^2}{\sigma_i^2 d_i^6} \mathbf{b}_i \mathbf{b}_i^T \tag{27}$$

$$\mathbf{J}_{21}^T = \mathbf{J}_{12} = -2s \sum_{i=1}^n \frac{g_i^2}{\sigma_i^2 d_i^5} \mathbf{b}_i \tag{28}$$

$$J_{22} = \sum_{i=1}^n \frac{g_i^2}{\sigma_i^2 d_i^4} \tag{29}$$

The variance of the unknown parameter estimation is bounded by the CRB, i.e.

$$\text{var}(\widehat{\rho}_{ij}) \geq (\mathbf{J}^{-1})_{(i-1)p+j, (i-1)p+j} \quad \{i = 1 \cdots K, j = 1 \cdots p\}$$

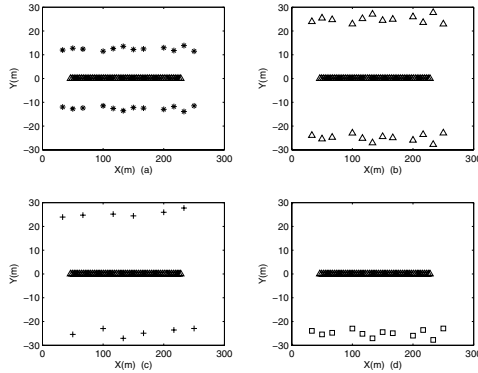
Where  $\text{var}(\widehat{\rho}_{ij})$  is the variance of the estimation location for  $i^{th}$  source  $(\rho_i)$ , at  $j^{th}$  coordinate direction.

From above *CRB* equation, we know that *CRB* is sensitive to the overall weighted  $\left(\frac{g_i}{\sigma_i}\right)$  distance from the targets to the sensors. The longer the distance, the bigger *CRB* could be. When the sensors are deployed close to the road, the weighted overall distance from the targets to the sensors is reduced. Therefore, we can get smaller *CRB*. When the sensors are dense, there are more sensors close to the targets, therefore, smaller *CRB* we can approach. By *Chebyshev's* inequality, we know that the probability of estimation error is less than the ratio of the variance of that random variable and the square of that estimation error, i.e.

$$P(|X - E(X)| \geq a) \leq \frac{\text{Var}(X)}{a^2}$$

Since the variance is lower bounded by the  $CRB$  and for  $ML$  estimation, variance asymptotically approaches its  $CRB$ , the bigger the  $CRB$ , the higher probability of the estimation error we might get. Using dense sensors that are close to the road gives smaller  $CRB$ , and therefore, improves the performance in the sense that,  $\{\forall a > 0, P(|\hat{X} - X| > a)\}$  is smaller.

Simulations of different sensor deployment with two targets producing similar acoustic energy moving in opposite direction are conducted to check the relation between  $CRB$  and the sensor deployment. The results are shown in Fig.1 and Fig.2. These results are consistent to our theoretical analysis. Note that when two targets are close to each other, we have more ambiguity, and therefore, the  $CRB$  increases abruptly at the middle part in Fig. 2.

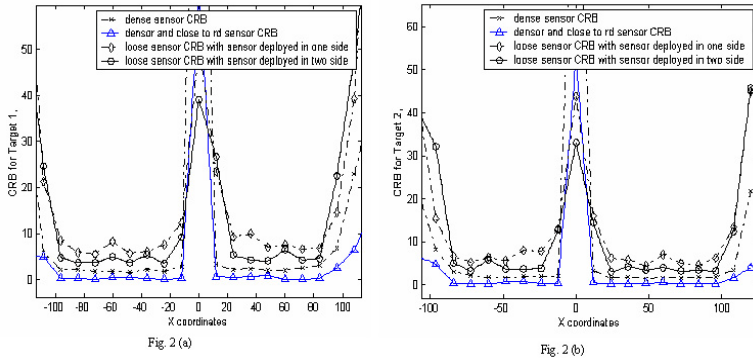


**Fig. 1.** Sensor Deployment, (a) dense sensor, (b) dense sensor close to road, (c) loose sensor located at one side, (d) loose sensor located at two side.

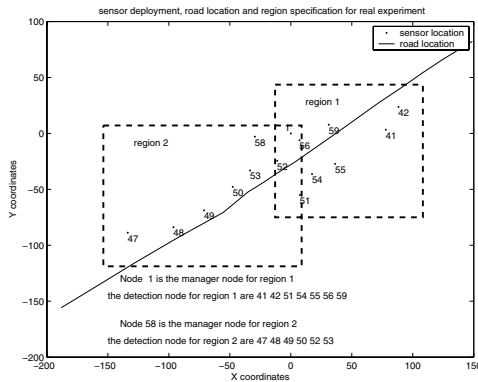
## 4 Experiments and Simulations

### 4.1 Experiments

**Sensor Network System.** The raw signals were recorded by 29 sensor nodes deployed along the road in the sensor field, CA in November 2001, sponsored by the DARPA ITO SensIT project. The data we used to evaluate  $EBL$  algorithms were taken from 15 acoustic sensors recording the signatures of AAV vehicle going from east to west during a time period of  $\sim 2$  minutes. Figure 3 shows the road coordinates and sensor node positions, both supplied by the global positioning system (GPS). The sensor field is divided into two regions as shown in the above figure. Region 1 is composed of node 1, 41, 42, 46, 48, 49, 50, 51. Region 2 is composed of node 52, 53, 54, 55, 56, 58, 59. In region 1, node 1 is chosen as manager node, others are detection node. In region 2, node 58 is chosen as manager node, others are detection node.



**Fig. 2.** CRB for different sensor deployment shown as Fig. 1, (a) CRB for target 1, (b) CRB for target 2.



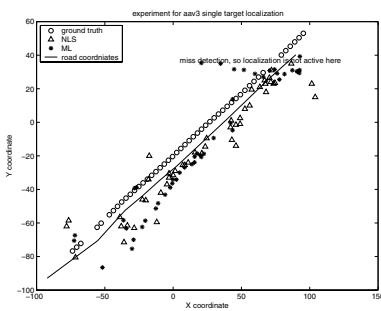
**Fig. 3.** Sensor deployment, road coordinate and region specification for experiments

The region is activated by our tracking algorithm implemented by Kalman filter. When the region is activated, multi-modality node detection and region detection are performed. Once region detection announces the detection of the target, *EBL* localization algorithm is activated and performed to locate the targets using the most recent reported acoustic energy, noise mean and variance from its detection nodes. The sampling frequency is  $f_s = 4960\text{Hz}$ . The energy is computed by averaging the  $T=0.75\text{sec}$  non-overlapping data segment (3720 data points).

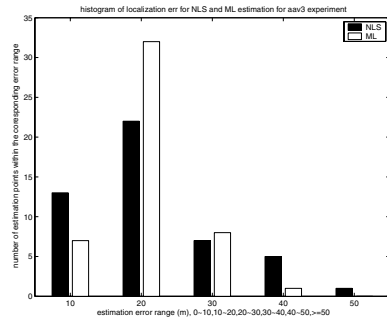
Fig. 4(a) shows the AAV ground truth and the localization results based on the *ML* algorithm with projection solution and *NLS* algorithm. *MR* search is used to estimate the location. The grid size we chose is:  $4*4$ ,  $2*2$ ,  $1*1$ . Note that the missing ground-truth points in this figure are the miss-detection points by

our multi-modality detector and therefore, there is no localization operation at these points.

To evaluate *ML* algorithm and compare it with *NLS* we proposed in [13] for *EBL* problem, we first compute the localization errors defined as the Euclidian distance between the location estimates and the true target locations for the time when region detection is announced. The true target location can be determined since they must be positioned on the target trajectory which can be extracted from GPS log. These localization errors are then grouped into different error range, i.e.,  $0 \sim 10, 10 \sim 20, \dots, 40 \sim 50, \geq 50$ . We call it as error histogram of our localization algorithm. Fig. 4(b) shows this localization error histogram for AAV localization.



(a) AAV ground truth and localization estimation results



(b) Estimation error histogram for AAV experiment data

**Fig. 4.** AAV ground truth, localization estimation results and estimation error histogram based on *ML* algorithm with projection solution and *NLS* algorithm (MR search is used, grid size is  $4 \times 4, 2 \times 2$  and  $1 \times 1$ ). Estimation results look bias from the ground-truth, see discussion for reasoning)

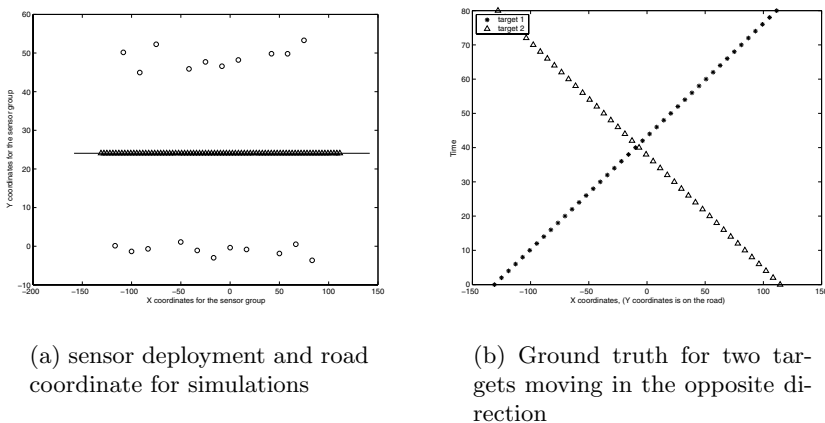
From experiment, we can see that, overall, both *ML* and *NLS* algorithms perform well estimations of target location. It verifies that inverse distance square acoustic energy decay model is suitable in the free space in normal situation. *ML* algorithm with projection solution outperforms to *NLS* algorithm in the sense that it has less estimation error. Besides, there are some points having estimation error bigger than 30 meters using *NLS* estimation. This says that *NLS* is not as stable as *ML* estimation with projection solution. However, *NLS* algorithm needs less bandwidth. This is because *NLS* doesn't use noise variance for its estimation while *ML* algorithm does need it. So, for *NLS* algorithm, we save about 1/4 bandwidth. (For *ML* estimation, detection nodes need to report acoustic energy, noise mean, noise variance and multi-modality binary node detection results in every 0.75 second).

Fig. 4 shows that the localization estimation results look bias from the real ground-truth. This can be caused by inaccurate GPS measurement. It can be seen that ground truth also looks bias from the road while in the experiment, vehicle moved along the road. We can also see that estimation results are closer and less biased to the road than to the ground-truth. In our works on parameter sensitivity analysis, we found that *EBL* algorithms are sensitive to the sensor gain, sharp background noise or sensor faults. Inaccurate measurement or estimation of these parameters can cause the estimation bias.

## 4.2 Simulations

**Simulations for *ML* Estimation with *EM* Solution and *Projection* Solution for Multi-Target Localization.** The simulation parameters were designed according to the experiment data described previously. From the experiment, we know that, for AAV vehicle, the maximum average energy ( $y_{max}$ ) received by the acoustic sensor is around  $y_{max} = 1.324e^{11}/2^{30}$  when the sensors are deployed as Fig. 3. Note that with this sensor deployment, the closest distance from the sensor to the road is about 15 meters. Using energy decay model, we know that AAV acoustic source energy is about  $S = 2.78e^4$ . Noise mean  $\mu$  detected by our *CFAR* detector is from  $0.01y_{max}$  to  $0.04y_{max}$ . Noise variance is in the range of  $(1 \sim 2)\mu$ . So, the maximum SNR received by the acoustic sensor for AAV is around  $14db$  to  $20db$ .

Follow the experiment data, we designed a sensor field with twenty-one sensors scattering along the two sides of the road. The size of sensor field is  $200 \times 80$  meter<sup>2</sup>. Simulations were conducted by moving two-targets in opposite direction in the above sensor field. Fig. 5 shows the sensor deployment for this sensor field and the ground truth of the two targets.



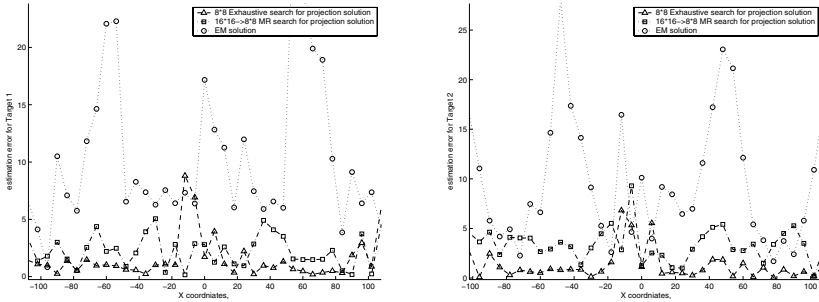
**Fig. 5.** Sensor deployment, road coordinate and ground truth of the two targets for the simulation

Two cases were run in the simulation. For case 1, the source energy of the two targets are similar, i.e.,  $S_1 = S = 2.78e^4$ ,  $S_2 = 1.2S_1$ . For case 2, the source energy of the two targets have significant difference, i.e.,  $S_1 = S$ ,  $S_2 = 3S_1$ . For each simulation case, we assume the acoustic source energy,  $S_1$  and  $S_2$ , keep constant in the simulation time.

Background noise  $\varepsilon_i$  was generated by *matlab* noise generator. It is gaussian distributed with mean uniformly distributed on  $(0.01 \sim 0.04)y_{max}$  and variance uniformly distributed on  $(1 \sim 2)\mu$ .

Using *ML* estimation algorithm with *projection* solution and *EM* solution, we can solve the location of the two targets at every time period. For *projection* solution, grid size we used for the *exhaustive* search is  $8 * 8$ . The grid size of *MR* search is  $16 * 16$ ,  $8 * 8$ . Direct Monte Carlo simulations were performed with 100 trials for each simulation case.

To evaluate the *ML* algorithm with *projection* solution and *EM* solution for energy based multi-target source localization problem, we calculate the localization error which is defined as the distance error between the target ground-truth and the mean of estimation of 100 trials of the Direct Monte Carlo simulation at each target ground-truth. We also calculate the estimation standard deviation (*std*), and compare the *std* with the root of CRB. Fig.6 and Fig.7 show the estimation error and *std* of case 1.



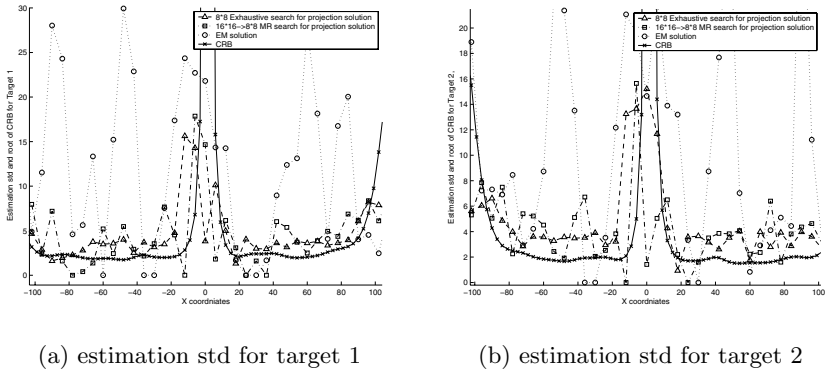
(a) estimation error for target 1

(b) estimation error for target 2

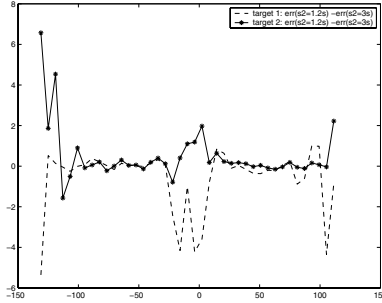
**Fig. 6.** Estimation errors for *projection* solution using *MR* search, *projection* solution using *exhaustive* search and *EM* solution.(a) target 1, (b) target 2 (sensor deployment and ground truth for the two targets are shown in Fig. 5; Noise is uniformly distributed from  $0.01y_{max}$  to  $0.04y_{max}$ ,  $S_1 = y_{max}$ ,  $S_2 = 1.2S_1$ ).

To evaluate the effects of different target energy ratio ( $S_2/S_1$ ) on our algorithm, we calculate the estimation error difference for the two cases at every estimation point. The result was shown in Fig. 8.

From Fig.6 and Fig.7, we know that, for *projection* solution, *exhaustive* search has better performance than *MR* search since *MR* search is just approximate



**Fig. 7.** Estimation standard deviation for *projection* solution using *MR* search, *projection* solution using *MR* search using *exhaustive* search and *EM* solution. (a) target 1, (b) target 2 (simulation condition is the same as Fig.6 )



**Fig. 8.** Estimation error difference when two targets have different  $S_2/S_1$

ML estimation. Yet, the degradation using *MR* search is small for most of time. Performance of *EM* solution is much worse than that of the projection solution. This is because *EM* solution is easy to track into the local minimum. Besides, *EM* solution is sensitive to the initial condition. The advantage of *EM* solution is that it has much less computation complexity. For most of time, it can get results within 6 or 7 iteration and so, it has less computation complexity while exhaustive search and *MR* search for the *projection* solution need to search in the entire effective sensor region. Fig.7 shows that the estimation variance of *ML* estimation with projection solution approaches its CRB. It concludes that *ML* estimation with projection solution is the optimum solution for the *EBL* problem when the prior probability of target location is unknown.

From Fig.8, we know that estimation error decreases for the target with higher acoustic energy. For the target with lower acoustic energy, estimation error increases, especially at the area where the two targets are close to each other or at the boundary of the sensor field. This is because the target with low energy is more ambiguous when the source of other target is much stronger.

The above multi-target localization is based on the assumption that we already detect the active region where the target is inside by our region detect algorithm and we know the number of targets in that active region. The region detection has been developed by the combination of CFAR detector, classification, data fusion and decision fusion. The number of multi-targets can be determined by multi-modality detection we have developed and space-time analysis of the energy sequences that we are developing recently.

## 5 Conclusion and Future Work

Collaborative energy-based acoustic source localization method has been presented. This method is based on the inverse distance square acoustic energy decay model under certain conditions. *ML* algorithm with different solutions is proposed to estimate multi-target source location. *CRB* is derived and be used for sensor deployment analysis. Experiments and simulations were conducted to evaluate the *ML* algorithm with different solutions and to compare the *ML* algorithm with *NLS* algorithm. Results show that energy based *ML* estimation using *projection* solution and *MR* search is robust, accurate, efficient.

Overall, *EBL* algorithms need low communication bandwidth since each sensor only reports energy reading, noise mean and variance (*NLS* doesn't need variance), and detection binary results to the manager node at every time period rather than at every time instant. Besides, it is power efficient. For detection node, it only calculates the average energy in the time period and performs energy-based CFAR detector (simple algorithm). For manager node, it performs simple voting algorithm and decision fusion algorithm. Manager node performs localization algorithm only if the region detection announces the targets. *ML* estimation with projection solution and MR search under the reduced search region saves the computation burden and so, saves the manager node battery further more. Detection node energy computation requires averaging of instantaneous power over a pre-defined time interval. Hence it is less susceptible to parameter perturbations, and so, the algorithm is robust.

From *CRB* analysis, we know that the performance of our localizer is related to the sensor deployment. The performance of the region detection is also related to the sensor deployment. Besides, our initial research on target number prediction based on the space-time analysis also shows the importance of the sensor deployment. In addition, sensor deployment also affects the sensitivity of the parameter perturbation. Therefore, the sensor deployment is very important in our sensor network communication. The optimum sensor deployment based on all these consideration will be conducted as our future work.

## References

1. Li, D. Wong, K.D., Hu, Y. H., Sayeed, A. M.: Detection, classification, and tracking of targets. *IEEE Signal Processing Magazine*, **19**, (2002), 17–29
2. Haykin, S.: *Array Signal Processing*, Englewood-Cliffs, NJ: Prentice-Hall, (1985)

3. Taff, L. G.: Target localization from bearings-only observations, *IEEE Trans. Aerosp. Electron.*, **3**, issue 1, (1997) 2–10
4. Oshman, Y., and Davidson, P.: Optimization of observer trajectories for bearings-only target localization, *IEEE Trans. Aerosp. Electron.*, **35**, issue 3, (1999), 892–902
5. Kaplan, K. M., Le, Q., and Molnar, P.: Maximum likelihood methods for bearings-only target localization, *Proc IEEE ICASSP*, **5**, (2001), 3001–3004
6. Carter G. C.: *Coherence and Time Delay Estimation*, IEEE Press, 1993.
7. Brandstein, M., and Silverman, H.: A localization-error-based method for microphone-array design, *Proc. ICASSP'96*, Atlanta, GA, (1996), 901–904
8. Brandstein, M. S., Adcock, J. E., and Silverman, H. F.: A closed form location estimator for use with room environment microphone arrays, *IEEE Trans. Speech and Audio Processing*, vol. 5 (1997), 45–50
9. Yao, K., Hudson, R. E., Reed, C. W., Chen, D., and Lorenzelli, F.: Blind beamforming on a randomly distributed sensor array system, *IEEE J. Selected areas in communications*, **16** (1998) 1555–1567
10. Reed, C.W., Hudson, R., and Yao, K.: Direct joint source localization and propagation speed estimation. In *Proc. ICASSP'99*, Phoenix, AZ, (1999) 1169–1172
11. Special issue on time-delay estimation, *IEEE Trans. ASSP* **29**, (1981)
12. Smith, J.O., and Abel, J.S.: Closed form least square source location estimation from range difference measurements. *IEEE Trans. ASSP* **35** (1987) 1661–1669
13. Hu, Y.H., and Li, D.: Energy based source localization, *IEEE Trans. ASSP*, submitted, 2002
14. Kinsler, L.E., et al.: *Fundamentals of Acoustics*. NY, NY: John Wiley and Sons, Inc., 1982
15. Rabiner, L. R.: A Tutorial on Hidden Markov models and selected applications in speech recognition, *Proceedings of the IEEE*, vol. **77** **2**, (1989) 257–287
16. Bishop, L.M.: *Neural Network for Pattern Recognition*, Oxford University press, 1995, Chapter 2.