Distributed Range Difference Based Target Localization in Sensor Network

Chartchai Meesookho and Shrikanth Narayanan
Department of Electrical Engineering
Viterbi School of Engineering
University of Southern California
Email: cmeesook@usc.edu, shri@sipi.usc.edu

***Abstract*— Target localization is a key application in the sensor network context. Of the various conventional methods can be applied, and have been proposed, the Range Difference (RD) Based method is attractive due to improved accuracy and ease of implementation it affords. While the basic concepts of the RD based method can be adopted to the case of sensor networks, the data acquisition and aggregation procedures need to be formulated and characterized subject to the energy constraint. The challenge is to design an efficient algorithm that is economical and still accurate. In this paper, based on range difference localization method, we propose a distributed algorithm which allows the time delay estimation to be carried out at each participating sensor. The acquired data is fused using a sequential least squares scheme which enables the appropriate sensor selection based on the current estimate. The results, evaluated using realistic simulation models, illustrate that the distributed localization produces smaller error and consumes less energy than the centralized method. The advantage of distributed localization in terms of the accuracy becomes more conspicuous when the number of participating sensors is small while the energy saving increases when the number of participating sensors is large. The proposed method accuracy is also more robust to decreasing target signal energy and the instantaneous error from the sequence of estimates can be approximated and used to reconcile the cost and the system performance.**

I. Introduction

Target localization is one of the key motivating applications for implementing sensor networks. A large number of sensors enable the redundancy of the observations and close proximity to the target, and thus, improving the chances for improved target localization and tracking performance. Some example applications include localizing military vehicles in a battlefield and tracking wild animals in their natural habitat. Recently, conventional target localization methods have been applied to the case of sensor networks. The Range Difference (RD) based method is particularly attractive in this context [1], [2] since it offers better ease of implementation than the Maximum Likelihood (ML) estimator [3], [4], is more accurate than energy based localization [5], and does not require the prior knowledge of the signal generated by the target. While the basic concepts of the RD based method can be adopted to the sensor networks problem, the data aggregation procedure needs to be developed and characterized. In traditional systems such as radars and microphone arrays, time series data collected from each sensor, the fundamental information needed in the process, is assumed to be available at the central processing unit without the concern for the cost incurred in gathering such information. However, due to the characteristics of sensor networks, which are typically battery-powered and wireless, the energy expenditure for time series data exchange between sensors should be taken into account. The challenge is to design an efficient algorithm that is economical and still accurate. In [1], localization was implemented on a sensor array testbed but the communication cost was not considered. The cluster-based architecture for acoustic target tracking was studied in [2]. Nonetheless, the system performance subject to the communication protocol within the cluster was not addressed. We believe that the impact of the designed algorithm on the system efficiency is highly dependent on what specific method is implemented. In this paper, based on range difference localization, we propose a distributed algorithm which allows time delay estimation to be carried out at each participating sensor so that the amount of energy incurred for time series data transmission can be decreased. The acquired data, which are the range differences, are fused using a sequential least squares scheme. The sequential nature allows for efficient sensor selection based on the current estimate at each time step, thus, enabling accuracy to be improved. The results, evaluated using realistic models and conditions, illustrate that the distributed localization produces smaller error and consumes less energy than the centralized method. Notably, the advantage of distributed localization in terms of the accuracy becomes more significant when the number of participating sensors is small while the energy saving increases when the number of participating sensors is large. The proposed method is also robust in that its accuracy is less affected by a target signal with low energy (lower SNR) and the instantaneous error from the sequence of estimates can be approximated and used to reconcile the cost and the system performance.

II. Clustering for target localization

Since a centralized global processing of information or measurements gathered from all sensors does not seem to be attractive, or may be feasible, especially in a large and dense sensor field due to the demands of high communication cost, an appropriate solution is to divide the sensors into a number of smaller groups to operate on the tasks where each group has a local processing unit. The fusion and compression of locally processed data can save the energy used for the transmission of raw data to the base station or the end user. These groups of sensors are commonly called clusters and one sensor is selected to be a cluster head playing the role of a local processing unit. Generic energy-efficient clustering protocols for decentralized processing in sensor network can be found in [6]. A target localization problem using clustering technique may require a somewhat more specific formulation. Since the ultimate goal is to identify a particular point which is the most likely target location, the critical information should be available in the corresponding target area. Hence, only the most informative cluster may be needed and that motivates the design of an efficient cluster formation at the particular time and place. Cluster forming protocols for the purpose of target localization and tracking have been presented in [7] where the Dynamic Space-Time Clustering (DSTC) algorithm was proposed to work as a cluster forming protocol based on Closest Point of Approach (CPA). Information Driven Sensor Querying (IDSQ) introduced by Zhao et al [8] allows the maximum information gain for the dynamic clustering using an information utility measure. Dynamic clustering for acoustic target tracking was presented in [9]. A sparsely placed high-capability sensor scenario (expected to play cluster head roles) is assumed and a cluster is formed when the acoustic signal strength detected by the cluster head exceeds a predetermined threshold. The cluster's priority is to integrate the measurements collected at each cluster member to represent the earning knowledge from each cluster. It is fairly application specific in order to describe the mechanism of the information management associated with the cluster formation. Generally, all members communicate with the cluster head either by direct or multi-hop communication. The signal processing functions are carried out at the cluster head before transferring compressed data to the base station or end user. We will call this a centralized processing scheme. A specific application such as target localization using conventional methods, however, requires a circumspect design in order to obtain an accurate and cost-effective system. The main reason is that the observation at each sensor is commonly time series data. To individually transmit such data from all cluster members to be processed at cluster head entails a large amount of overall communication cost particularly when the cluster is designed to be large to reach the localization accuracy requirement. An alternative is to apply distributed processing by some means depending upon the characteristics of the utilized methods for a particular application. We exploit a well-known scheme, the range difference based localization, for distributed processing in sensor networks and demonstrate that the system performance can be improved when compared with the centralized method.

III. Range Difference Based Least Square Localization

Range Differences (RD) can be derived from Time Difference of Arrival (TDOA) estimation through the relationship between distance and traveling speed of the signal over a medium. Time delay estimation technique [10] is the fundamental tool used to determine TDOAs. We will assume the existence of an optimal time delay estimator producing estimated TDOA perturbed by additive noise to model uncertainty. There have been a number of RD-based approaches proposed in the past [11], [12]. We focus on a closed-form least square method proposed in [11] since it was reported to be more efficient than the other schemes and was shown to approach the Cramer Rao Bound (CRB) in high Signal to Noise Ratio (SNR) environments. Let N sensors be assigned to participate in the localization process located at coordinates $\{(x_1, y_1), \ldots (x_N, y_N)\}$. Assuming the target is located at $\mathbf{z}_s = (x_s, y_s)$, the differences of the distance between sensors i and j where $i, j = 1, \ldots, N$ and the source denoted by d_{ij} can be obtained by the basic relation: $d_{ij} = D_i - D_j$ where $D_i = \sqrt{(x_s - x_i)^2 + (y_s - y_i)^2}$. RDs with respect to one arbitrary reference sensor are typically used. Without the loss of the generality, we select (x_1, y_1) to be the location of reference sensor. The time series data collected from the other sensors together with the received signal at the reference sensor can produce the TDOA estimates and RDs can be derived from TDOAs using the knowledge of signal traveling speed. In the real application, however, the actual RDs are not available since there are some errors from TDOA estimation. Consequently, We have $\hat{d}_{i1} = d_{i1} + n_{i1}, \quad i = 1 \ldots N$. The TDOA estimate obtained by generalized cross correlation with Gaussian data is asymptotically normally distributed in high SNR environment [13]. Therefore, the RD estimate is also Gaussian and we assume $n_{i1} \sim N(0, \sigma_{i1}^2)$. The Localization problem can be formulated as a linear least squares problem, $\mathbf{A}\boldsymbol{\theta} = \mathbf{b}$, where

$$\begin{aligned}
\mathbf{A} &= \begin{bmatrix} x_2 & y_2 & \hat{d}_{12} \\ \vdots & \vdots & \vdots \\ x_N & y_N & \hat{d}_{N1} \end{bmatrix} \\
\boldsymbol{\theta} &= \begin{bmatrix} x_s \\ y_s \\ R_s \end{bmatrix}, \quad \mathbf{b} = \frac{1}{2}\begin{bmatrix} R_1^2 - \hat{d}_{12}^2 \\ \vdots \\ R_N^2 - \hat{d}_{N1}^2 \end{bmatrix} \\
R_i &= \sqrt{(x_i - x_1)^2 + (y_i - y_1)^2}
\end{aligned}$$

These linear least square equations can be solved by a batch approach and the solution, $\hat{\boldsymbol{\theta}} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{b}$. However, we can update $\hat{\boldsymbol{\theta}}$ without having to resolve the linear equations by a sequential least squares procedure [14] which can be described by letting $\mathbf{A}[n] = [\mathbf{A}[n-1]\mathbf{a}^T[n]]^T$ and $\mathbf{b}[n] = [b[1] \quad b[2] \quad \ldots \quad b[n]]^T$. The sequential least square estimator becomes

$$\hat{\boldsymbol{\theta}}[n] = \hat{\boldsymbol{\theta}}[n-1] + \mathbf{K}[n]\big(b[n] - \mathbf{a}^T[n]\hat{\boldsymbol{\theta}}[n-1]\big) \quad (1)$$

$$\mathbf{K}[n] = \frac{\boldsymbol{\Sigma}[n-1]\mathbf{a}[n]}{1 + \mathbf{a}^T[n]\boldsymbol{\Sigma}[n-1]\mathbf{a}[n]}$$

$$\boldsymbol{\Sigma}[n] = \big(\mathbf{I} - \mathbf{K}[n]\mathbf{a}^T[n]\big)\boldsymbol{\Sigma}[n-1] \quad (2)$$

and index n corresponds to the n^{th} sensor.

IV. Data Model

A static acoustic target generating a WSS Gaussian random observation process, $s(t)$, is assumed where the intensity attenuates at the rate that is inversely proportional to the distance from the target. Perturbed by additive Gaussian measurement noise $w_i(t)$, the received signal at i^{th} sensor is given by $x_i(t) = \frac{s(t-\tau_i)}{D_i} + w_i(t)$. The energy can be calculated by averaging over a time window $T = \frac{M}{f_s}$ where M is the number of samples and f_s is the sampling frequency as $y_i[k] = \frac{1}{M}\sum_{j=(k-1)M+1}^{kM} x_i^2[j]$. Assuming $s(t)$ and $w(t)$ are independent, we get

$$\mathbf{E}\{y_i[k]\} = \frac{\mathbf{E}\{s^2(t)\}}{D_i^2} + \mathbf{E}\{w^2(t)\} \tag{3}$$

$$\text{var}\{y_i[k]\} = \frac{2(\frac{\mathbf{E}\{s^2(t)\}}{D_i^2} + \mathbf{E}\{w^2(t)\})^2}{M} \tag{4}$$

Let $s(t) \sim N(0, \sigma_s^2)$ and the noise at each sensor has the same distribution so that $w_i(t) \sim N(0, \sigma_w^2)$. The signal PSD ($G_s(f)$), the noise PSD ($G_w(f)$), and the coherence are assumed to be flat over a bandwidth Δf Hz centered at frequency f_0. The SNR at each sensor, $\frac{G_{s,i}(f)}{G_{w,i}(f)} = \frac{\sigma_s^2}{D_i^2\sigma_w^2}$. According to [10], CRB of the TDE estimate is the following

$$\sigma_{ij}^2 \geq \left\{ \frac{8T\pi^2}{3}\left[\frac{C_{ij}}{1-C_{ij}}\right]\left[(f_0 + \frac{\Delta f}{2})^3 - (f_0 - \frac{\Delta f}{2})^3\right]\right\}^{-1}$$

$$\text{where} \quad C_{ij} = \frac{1}{(1+\frac{G_{s,i}(f)}{G_{w,i}(f)}^{-1})(1+\frac{G_{s,j}(f)}{G_{w,j}(f)}^{-1})}$$

It is simple to derive that the variance of the estimate can be in the form, $\sigma_{i1}^2 = \sigma_1^2 + \alpha D_i^2$, where

$$\begin{aligned} \sigma_1^2 &= \frac{3D_1^2}{8T\pi^2((f_0+\frac{\Delta f}{2})^3 - (f_0-\frac{\Delta f}{2})^3)\text{SNR}_0} \\ \alpha &= \frac{3(1+\frac{D_1^2}{\text{SNR}_0})}{8T\pi^2((f_0+\frac{\Delta f}{2})^3 - (f_0-\frac{\Delta f}{2})^3)\text{SNR}_0} \end{aligned} \tag{5}$$

and SNR_0 denotes $\frac{\sigma_s^2}{\sigma_w^2}$. Please note that σ_{i1}^2 is the variance of TDE between i^{th} sensor and the reference sensor as assumed in the previous section. Such variance is proportional to D_i^2 where the constants, σ_1^2 and α are functions of D_1^2. Therefore, with a fixed reference sensor, TDE with respect to the farther sensors from the target is less accurate.

V. Distributed Localization

From the description of the range difference based localization method, we can note that there are two key steps which are TDOA estimation and target localization obtained by solving least square equations. In a Centralized scheme, both steps take place at the cluster head. The cluster head should be a reference sensor and TDOAs with respect to the cluster members can be obtained through time delay estimation. The distributed localization concepts can be adopted by enabling some processes to occur at each participating sensor, not just at the cluster head. If time series data collected at the cluster head is transmitted to the participating sensors, time delay estimation can be operated there. Broadcasting the data from one reference sensor to many participating sensors is expected to require less total communication overhead than in the opposite direction. Solving least square equations encompasses two mechanisms depending on whether batch or sequential procedure is applied. Batch estimator requires all measurements available at the same time whereas sequential estimator needs only the estimate obtained from the $(n-1)^{th}$ sensor and a TDOA corresponding to the n^{th} sensor. The latter, however, demands less computational complexity as it does not have to deal with matrix inversion which might be burdensome when the matrix is large due to a large number of participating sensors. Another advantage is that the current estimate can be used as the prior information to properly select the next participating sensor. According to the data model that the variance of time delay estimation is proportional to the square distance between the sensor and the target, the preferred sensors can be simply selected by considering the nearest sensors to the current estimate. Consequently, combining the ideas of distributed processing for time delay estimation and sequential least square localization is expected to improve the localization performance in terms of both communication cost savings and accuracy. By using the notations defined in the previous section, we propose the following algorithm:

1) The sensor which receives the highest average signal energy in a certain time window is selected to be an initial sensor. Please note that the term "initial sensor" is used to call the sensor that starts the process instead of "cluster head".
2) The initial sensor broadcasts collected time series data to at most k nearest neighbors within the maximum radio range where k is the initial expected number of participating sensors. There might be a possibility that less than k sensors can be reached depending on the coverage of the radio range and the density of the sensor field.
3) Each neighbor operates time delay estimation using time series data collected at the sensor and the one sent from the initial sensor to estimate TDOAs.
4) The initial estimate is obtained by using batch estimator based on TDOAs computed by the three nearest neighbors. The neighbor might be requested to broadcast time series data received from the initial sensor if there are less than three sensors that have already received it.
5) $k-4$ nearest sensors to the initial estimate achieved from the batch estimator are expected to participate in the sequential least square method. It becomes $k-i$ nearest sensors for the following estimate where i is a number of sensors that are already included in the localization process.
6) The route of sequential estimator is constructed from

the sensors described in the previous step by convex hull insertion algorithm for Traveling Salesman Problem (TSP) route [15].

7) The estimate is updated in the sequential fashion by using (1). Only $\hat{\boldsymbol{\theta}}$ and $\boldsymbol{\Sigma}$ are needed to be sent across the sensors.
8) Every time the estimate is updated, the route is also updated based on the new estimate and the decreasing number of remaining participating sensors.
9) If the next sensor in the path has not received time series data from the cluster head, the current sensor will broadcast the data to the next sensor including at most $k-j$ nearest sensors located in the radio range where j is the number of sensors which already received the data.

VI. Experimental Results

For the experimental simulation, we assume a 100x100 square meter sensor field with 2,500 sensors randomly and uniformly placed. A static acoustic target location is assumed to have a uniform distribution within the sensor field and generates a 3500-4500 Hz signal. The data observation time for time delay and average energy estimate is 1 second where sampling rate is 9000 Hz. σ_s^2 and σ_w^2 are 100 and 1, thus $\mathrm{SNR}_0 = 100$. The numbers of participating sensors considered are 10,15,20 and 25. The radio range is 5 meters. The estimation error is evaluated by the Mean Square Error (MSE), $(\hat{x}_s - x_s)^2 + (\hat{y}_s - y_x)^2$. 100 Monte Carlo trials were conducted to average the effect of target location and sensor topology. For each topology, 100 runnings are used to average the noise. We assume that the average energy of the received signal at each sensor has a normal distribution with the mean and variance defined in (3) and (4). The sensor which receives the highest energy is selected to be the initial sensor for the distributed method and the cluster head for the centralized method.

Figure 1 illustrates the MSE for both centralized and distributed localization with different numbers of participating sensors. It is obvious that the error caused by the distributed method is smaller than the centralized method. The reason is that the distributed algorithm uses the immediate estimate as the prior information to select the sequence of the participating sensors that include the nearest sensors to the estimated target location. Since the closer sensors to target provide the more accurate time delay estimation as described in the previous section, the localization accuracy is improved when compared with the centralized method which preselects the participating sensors before achieving the result. The difference in the error caused by the two algorithms becomes smaller when the number of participating sensors increases since a large number of sensors make the coverage of region activated by both algorithms more similar. We also study the performance of both schemes when σ_s^2 is varied which corresponds to the variation of the energy of the signal (SNR) created by the target. Values of $\sigma_s^2 = 10, 40, 70,$ and 100 are used and the number of participating sensors is 15. The results in Figure 2 illustrates that the centralized scheme suffers more than distributed scheme when σ_s^2 decreases. This is also affected by the different participating sensors selected by both schemes. When σ_s^2 is small, SNR_0 is small and it makes α in (5) become large. σ_{i1}^2, thus, is proportional to D_i^2 with a higher rate. The distance between sensors and target becomes more influential on the localization performance. The sensors selected by the centralized scheme which are likely to be farther from the target than those selected by distributed scheme, therefore, produce larger error.

We use the first order radio model described in [6] for the communication overhead evaluation. To exchange k-bit data between a distance d, the radio expends $E_{Tx} = E_{elec} * k + \epsilon_{amp} * k * d^2$ and $E_{Rx} = E_{elec} * k$. E_{Tx} and E_{Rx} represent the energy dissipated by the transmitting and receiving data, respectively. $E_{elec} = 50$nJ/bit is the energy used to run the transmitter or receiver circuitry and $\epsilon_{amp} = 100$pJ/bit/m^2 is for the transmission amplifier. We assume that each data sample and the data representing each element in $\hat{\boldsymbol{\theta}}$ and $\boldsymbol{\Sigma}$ requires 8 bits to be encoded. Thus, for time series data exchange, $k =$ (Sampling Rate) x (Observation Time) x (bits/sample)= $9000 * 1 * 8 = 72000$ bits. For the communication required in sequential least square, $k =$ (Summation of number of elements in $\hat{\boldsymbol{\theta}}$ and $\boldsymbol{\Sigma}$) x (bits/sample)= $(3 + 9) * 8 = 96$ bits. As E_{Tx} and $E_{Rx} \propto k$, we can note that the energy spent to exchange the time series data is much more than the amount used for sequential estimate update. Figure 3 illustrates the energy dissipated by both schemes. It is can be noticed that the distributed method enables significant amount of energy savings compared to the centralized method, particularly when the number of sensors is large. This can be simply explained by considering, firstly, the communication between reference sensor and L participating sensors within the radio range. The centralized scheme requires L transmissions in order to obtain the time delay estimates with respect to all participating sensors while only one transmission is enough for distributed algorithm. The extra cost for distributed algorithm is what is used to sequentially aggregate TDOAs extracted from time delay estimates at each sensor. Such communication cost is relatively small as pointed out above. Secondly, for the sensors outside the radio range of the reference sensor, in the centralized scheme, multi-hop communication is required and the data transmission from such sensors still has to finally reach the reference sensor. On the other hand, in the distributed scheme, such sensors can obtain the data from those who already received the data located within the first hop from the reference sensor. That requires shorter total communication distance and lower energy than what is needed in the centralized scheme. Hence, the advantage of distributed method becomes more conspicuous when multiple hops are needed or the number of participating sensors is increases as illustrated Figure 3. We also study the convergence issue for the distributed method by considering the localization error at each iteration comparing it with the distance difference between the consecutive estimates. Figure 4 shows that both amounts are highly correlated. The advantage of this scenario

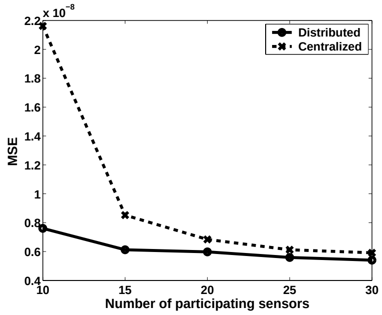

Fig. 1. MSE vs. number of sensors: Distributed method produces smaller error than centralized method.

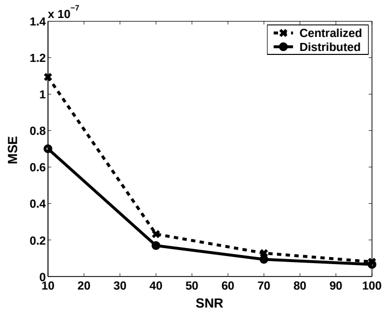

Fig. 2. The accuracy of the distributed method is less affected by a low energy target signal than the centralized method.

is that we can approximately evaluate the current error from the sequence of estimates. This is important if we want to save unnecessary cost when the accuracy has already reached acceptable or required levels.

VII. Conclusion and Future Work

We proposed a distributed algorithm, based on range difference localization method, which allows time delay estimation to be carried out at each participating sensors. TDOAs computed from time delay estimates are fused using a sequential least squares scheme which enables the appropriate sensor selection based on the current estimate. The results illustrate that the distributed localization produces smaller error and consumes less energy than centralized method. The advantage of distributed processing becomes more conspicuous for error considerations when the number of participating sensors is small and the better energy saving is obtained when increasing number of participating sensors. The proposed method is also more robust to decreasing target signal energy and the instantaneous error from the sequence of estimates can be approximated and used to reconcile the cost and the system performance. In the future we aim to study the effect of time synchronization errors on time delay estimation, and thus, the localization performance.

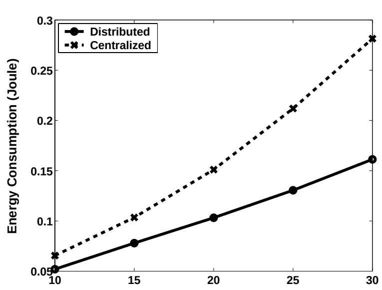

Fig. 3. Energy consumed by centralized method is larger than that consumed in the distributed method.

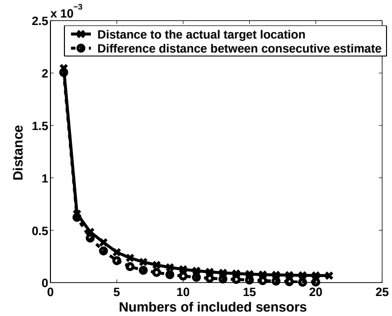

Fig. 4. Distance error and distance between consecutive estimates are highly correlated.

References

[1] J. C. Chen, L. Yip, J. Elson, H. Wang, D. Maniezzo, R. E. Hudson, K. Yao, and D. Estrin, "Coherent acoustic array processing and localization on wireless sensor networks," in *Proceedings of the IEEE*, vol. 91, Aug 2003.

[2] Q. Wang, W. Chen, R. Zheng, K. Lee, and L. Sha, "Acoustic target tracking using tiny wireless sensor devices," in *IPSN 2003*, 2003.

[3] J. C. Chen, R. E. Hudson, and K. Yao, "Maximum-likelihood source localization and unknown sensor location estimation for wideband signals in the near-field," *IEEE Transactions on Signal Processing*, 2002.

[4] R. J. Kozick and B. M. Sadler, "Source localization with distributed sensor arrays and partial spatial coherence," *IEEE Transactions on Signal Processing*, 2004.

[5] X. Sheng and Y. Hu, "Energy based acoustic source localization," in *IPSN*, 2003.

[6] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy-efficient communication protocol for wireless microsensor networks," in *Proc. 33rd Hawaii International Conference on System Sciences*, 2000.

[7] S. Phoba, N. Jacobson, and R. Brooks, "Sensor network based localization and target tracking through hybridization in the operational domains of beamforming and dynamic space-time clustering," in *GLOBECOM 2003*, 2003.

[8] M. Chu, H. Haussecker, and F. Zhao, "Scalable information-driven sensor querying and routing for ad hoc heterogeneous sensor networks," *International Journal of High Performance Computing Applications*, 2002.

[9] W. Chen, J. C. Hou, and L. Sha, "Dynamic clustering for acoustic target tracking in wireless sensor networks," *IEEE Trans. on Mobile Computing*, vol. 3, no. 3, Jul-Sep 2004.

[10] C. H. Knapp and G. C. Carter, "Time delay estimation," in *ICASSP'76*, 1976.

[11] Y. Huang, J. Benesty, and G. W. Elko, "Real-time passive source localization: A practical linear-correction least-squares approach," *IEEE Trans. Speech and Audio Processing*, vol. 9, no. 8, Nov 2001.

[12] Y. T. Chan and K. C. Ho, "A simple and efficient estimator for hyperbolic location," *IEEE Transactions on Signal Processing*, 1994.

[13] G. C. Carter, "Time delay estimation for passive sonar signal processing," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 1981.

[14] S. M. Kay, *Fundamentals of Statistical Signal Processing*. Prentice Hall, 1993.

[15] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to Algorithms*. McGraw-Hill, 2001.