An EM Technique for Multiple Transmitter Localization

Jill K. Nelson
Department of Electrical and Computer Engineering
George Mason University
Fairfax, VA 22030
Email: jnelson@gmu.edu

Maya R. Gupta
Department of Electrical Engineering
University of Washington
Seattle, WA 98195
Email: gupta@ee.washington.edu

***Abstract*— We propose an expectation-maximization (EM) technique for locating multiple transmitters based on power levels observed by a set of arbitrarily-placed receivers. Multiple transmitter localization is of interest for uncoordinated cognitive radio systems, which must identify and transmit over unused radio spectrum without cooperation from conventional transmitters. We employ the EM algorithm to reduce the dimensionality of the maximum-likelihood estimation problem. Because the EM algorithm finds only a locally optimal solution, we explore the use of clustering to generate "smart" initial estimates of the transmitter locations. Simulation results show that, as the number of sensors increases, the proposed EM technique achieves gains of up to an order of magnitude over constricted particle swarm optimization, a popular global optimization technique.**

I. Introduction

As wireless applications advance and expand, available wireless spectrum has become an increasingly scarce resource. As a result, a significant focus of recent research has been the development of efficient methods for identifying and exploiting open bandwidth. Among the most promising methods is opportunistic spectrum access, in which users identify and communicate over unused frequency bands without a conventional license for the spectrum. Because systems that perform opportunistic spectrum access must "think" and adapt to changing conditions, they are also called cognitive radio systems [1].

In conventional (legacy) communication systems, spectrum is licensed to particular users who then have exclusive rights to the allocated bandwidth. Measurement studies have shown, however, that much of the licensed spectrum is underutilized much of the time [1]. The goal of a cognitive radio system is to identify and use the licensed but unused spectrum, thereby increasing efficiency with which bandwidth is employed.

Cognitive radio systems may be designed to cooperate with legacy communication systems, possibly requesting permission to transmit in a certain frequency or time slot, or receiving information about available spectrum. Such systems are known as coordinated cognitive radio systems [2]. In this paper, we focus on uncoordinated cognitive radio systems, which must determine what spectrum is available for transmission without cooperation from any legacy systems operating in the band and geographic region.

A typical goal in the design of uncoordinated cognitive radio systems is that the cognitive radio systems operate with minimal disruption to legacy systems. In order to achieve this goal, cognitive radio nodes must employ opportunistic spectrum access, i.e. they must identify spectral holes in which they can transmit without causing interference to legacy receivers. In this paper, we assume that cognitive nodes identify spectral holes based on received power measurements from the legacy transmitters. In this case, a spectral hole has three aspects: frequency, space, and time. (Signal parameters may be treated as an additional aspect when waveforms are observed [3].) We assume that the cognitive radio system has no prior knowledge of location or activity patterns of legacy transmitters in the region.

A common technique for declaring a spectral hole at the location of a cognitive radio node is based on setting a detection threshold for the power observed by the node. For example, consider two cognitive nodes A and B, each sensing the power in band F_0 in their respective locations. If the power observed at node A and at node B both fall below a threshold value P_{max}, then nodes A and B may communicate in band F_0. The choice of the threshold P_{max} will be influenced by a number of factors, including the level of interference tolerated by legacy receivers, the desired maximum probability of disruption of legacy systems, and the nature of signal propagation in the region of interest.

One of the challenges of spectral hole identification via power measurements from a single transmitter is the effects of fading and shadowing in the propagation environment. Errors due to such effects can be mitigated by sharing measurements among several cognitive radio nodes [4]. For example, a set of S cognitive nodes may each sense the power in band F_0 and determine whether it falls above or below the threshold P_{max}. The detection results can then be shared among the nodes via a control channel (only a binary indicator need be communicated). The likelihood of disrupting a legacy system can be reduced by applying the rule that if any of the S nodes senses power above the threshold, none of the nodes may transmit. If the nodes are near each other geographically, the collection of several measurements will minimize the effects of fading and shadowing to produce a more reliable detection statistic. If the nodes are distributed over a wider area, however, such a rule may result in overly conservative decisions, since a single node near a legacy transmitter will prevent communication by any of the nodes, some of which may be a safe distance from the legacy system's coverage area.

II. Finding Spectral Holes via Transmitter Localization

Recent work has argued that spectral hole identification can be significantly improved over simple detection-based methods [5], [6], which limit cognitive radio nodes to very conservative transmission rules, as described above. One approach to better exploit the spectrum is to use cooperative sensing to estimate the approximate locations of transmitters given the power levels observed by the radio nodes. Once transmitter locations have been estimated, properties of legacy systems, a maximum probability of interference, and government regulations can be employed to determine the radius around each transmitter in which opportunistic communication must be avoided. Additionally, knowledge of transmitter locations allows mobile cognitive radio nodes to identify a spectral hole at a location without taking a measurement at that location. Hence, a cognitive node can determine whether or not it can transmit from a particular location before it reaches its destination. A natural extension would be to consider scenarios in which the primary transmitters are also in motion and to incorporate simple tracking in the transmitter localization algorithm.

In this work, we describe an iterative technique for estimating the location of legacy transmitters based on measurements of received power at the cognitive radio nodes. Significant research has been conducted to develop efficient and effective methods for localization in wireless sensor networks. Note, however, that transmitter localization in a cognitive radio framework poses a more challenging problem due to the lack of cooperation between transmitters and receivers. When sensor localization is performed within a sensor network, measurements may include signals received from the sensor(s) to be localized, as well as signals received at the sensor(s) to be localized. In addition, the transmissions of various sensors may be scheduled such that only a single sensor is transmitting at any time. In contrast, the nodes of an uncoordinated cognitive radio system must estimate the location of legacy transmitters without any cooperation from or communication with the transmitters. When only one transmitter is present, the transmitter location can be determined from three received power measurements via trilateration. However, when there are multiple transmitters contributing unknown proportions of the received power, the non-cooperative localization problem does not admit a straightforward solution.

We propose to estimate the transmitter locations that maximize the likelihood of the observed power measurements, which we show can be efficiently found by an expectation-maximization (EM) algorithm. The remainder of the paper is organized as follows. Section III describes the cognitive radio and legacy transmitter systems considered in our analysis and simulation. The EM technique for localization of multiple transmitters is developed in Section IV, and a method for generating "smart" initial conditions for the EM algorithm is discussed in Section V. Section VI provides simulation results for the EM technique and a global-optimization based localization algorithm; conclusions are presented in Section VII.

III. System Model

Consider M legacy transmitters and N cognitive radio nodes (sensors) located within a square region of unit area. We assume that the locations of the M transmitters are unknown and are denoted by $\theta = [\theta_1 \quad \theta_2 \quad \ldots \quad \theta_M]^T$, where θ_i denotes the two-dimensional location of the ith transmitter. The locations of the N sensors are assumed to be known but arbitrary. The cognitive nodes may be affixed, for example, to vehicles or to individuals, and hence their locations will be defined by the activities of the "carriers" rather than by the best geometry for obtaining reliable transmitter location estimates.

Let $d_j(\theta_i)$ denote the two-dimensional Euclidean distance from the ith transmitter to the jth receiver. We consider a line-of-sight path loss model in which power is assumed to decay at a rate proportional to the square of distance traveled. Note that line-of-sight propagation is a reasonable assumption for television transmission (one of the first bands under consideration for opportunistic spectrum access), since television transmitters are typically quite tall to maximize coverage and avoid clutter [7]. In addition, all transmitters are assumed to transmit with known power P_0. In practice, estimation of transmit powers would be required. Estimates could be based on a combination of measurement data and a priori information about the nature of potential transmitters in the band of interest.

With a line-of-sight model, the power received at the jth sensor from the ith transmitter is given by

$$P_{ij}^R = \frac{P_0}{(\rho d_j(\theta_i))^2}, \tag{1}$$

where ρ is a constant that reflects the carrier frequency and antenna properties. We consider a simple path-loss model to allow clear presentation of the EM-based localization technique. While it is straightforward to extend the concept of maximum likelihood estimation to more sophisticated path-loss models that account for multipath fading and shadowing, deriving a solution via EM techniques is a more challenging problem and is currently under investigation. Certainly, the accuracy of the transmitter location estimates will be a function of the precision with which the signal propagation is modeled, and hence all available a priori information should be used to choose the appropriate path-loss model for any particular senario. In cases in which the cognitive radio system has knowledge of the terrain and/or obstacles in the region of interest, path-loss models such as Longley-Rice, EPM-73, or TIREM [7] may be used to improve localization accuracy.

Consider M transmitters that transmit in a particular frequency band of interest. The observed power is corrupted by additive measurement noise, and hence the power measured in the band of interest by the jth sensor is

$$r_j = w_j + \sum_{i=1}^{M} P_{ij}^R = \sum_{i=1}^{M} h_{ij}, \tag{2}$$

where $h_{ij} = P_{ij}^R + w_{ij}$ denotes the power from the ith transmitter to the jth sensor, and w_j is a zero-mean Gaussian random variable with variance σ^2. To facilitate the mathematical development of the EM localization algorithm, we divide each measurement noise component into M parts: $w_{1j}, w_{2j}, \ldots, w_{Mj}$, which are independent and identically distributed Gaussian random variables with mean zero and variance σ^2/M.

IV. EM Localization Technique

The goal is to determine the maximum likelihood (ML) estimate $\hat{\theta}$ of the locations θ of the M transmitters based on the observed power measurements at each sensor. That is,

$$\hat{\theta} = \arg\max_{\theta} P(r|\theta), \tag{3}$$

where $r = [r_1\ r_2\ \ldots\ r_N]^T$. This problem is one of estimating parameters given superimposed signals; Feder and Weinstein [8] considered EM algorithms for this class of problems, and the transmitter localization problem is an application-specific case of their general development.

Given the transmitter locations θ, the measured sensor power r_j is a Gaussian random variable with mean $\sum_i P_{ij}^R$ and variance σ^2. Let $P_i^R = [P_{i1}^R\ P_{i2}^R\ \ldots\ P_{iN}^R]^T$, such that r is a Gaussian random vector with mean $\mu_r = \sum_i P_i^R$ and covariance matrix $\sigma^2 I_N$, where I_N denotes the $N \times N$ identity matrix. The likelihood is then

$$P(r|\theta) = (2\pi\sigma^2)^{-N/2} \exp\left\{ -\frac{1}{2\sigma^2} \left(r - \sum_{i=1}^{M} P_i^R \right)^T \left(r - \sum_{i=1}^{M} P_i^R \right) \right\}. \tag{4}$$

Taking the log of (4) and ignoring terms independent of θ results in the log-likelihood function

$$L(r|\theta) = -\left(r - \sum_{i=1}^{M} P_i^R \right)^T \left(r - \sum_{i=1}^{M} P_i^R \right). \tag{5}$$

The log-likelihood is a complicated function of the transmitter locations θ and does not admit a straightforward analytic solution, as it often has multiple local maxima. Global optimization methods can maximize (5) directly, but this is difficult, as there are $2M$ unknown parameters, all of which are coupled via an inner product.

An EM algorithm can be used to efficiently reach a local maximum of the log-likelihood function. To apply EM, we must identify "hidden" variables whose likelihood function is simpler to maximize than that of the observed data. The EM algorithm alternates between two steps [9]:

(a) Compute the expectation of the likelihood function of the hidden variables given the observed data and an estimate of the unknown parameters.

(b) Maximize the expectation of the likelihood of the hidden variables over the unknown parameters.

For the multiple transmitter localization problem, we propose to use as hidden variables the set of $M \times N$ unknown measured powers from each transmitter to each receiver $h = [h_1\ h_2 \ldots h_M]$, where $h_i = [h_{i1}\ h_{i2} \ldots h_{iN}]$. The EM algorithm then takes the following form:

1) Generate an initial estimate $\hat{\theta}^0$ for the transmitter locations. Set $k = 1$.
2) Compute $E[L(h)|r, \hat{\theta}^{k-1}]$.
3) Compute $\hat{\theta}^k = \arg\max_\theta E[L(h)|r, \hat{\theta}^{k-1}]$.
4) If $\hat{\theta}^k$ has converged, stop. If not, set $k = k + 1$, and return to step 2.

Next, we detail how to calculate the quantities in the above steps. Let $P^R = [P_1^R\ P_2^R\ \ldots\ P_M^R]^T$. Then akin to (5), the log-likelihood function of h can be written as

$$L(h) = -h^T K^{-1} h + 2{P^R}^T K^{-1} h - {P^R}^T P^R,$$

where $K = \sigma^2 I_{MN}$. The expectation of $L(h)$ given r and $\hat{\theta}_{k-1}$ is

$$\begin{aligned} &E[L(h)|r, \hat{\theta}^{k-1}] \\ &\quad = \ -E[h^T K^{-1} h] + 2E[{P^R}^T K^{-1} h] - E[{P^R}^T P^R]. \end{aligned}$$

Dropping terms independent of θ and denoting $E[h|r, \hat{\theta}^{k-1}]$ as $\hat{h}^k$, we have

$$\begin{aligned} E[L(h)|r, \hat{\theta}^{k-1}] &\propto 2{P^R}^T K^{-1} \hat{h}^k - {P^R}^T P^R \\ &\propto -(\hat{h}^k - P^R)^T K^{-1} (\hat{h}^k - P^R) \\ &= -\frac{\sigma^2}{M} \sum_{i=1}^{M} \left(\hat{h}_i^k - P_i^R \right)^T \left(\hat{h}_i^k - P_i^R \right), \end{aligned} \tag{6}$$

where the second line is obtained by adding the (θ-independent) term $(\hat{h}^k)^T K^{-1} \hat{h}^k$ to complete the square.

To compute $\hat{h}^k$, observe that h and r are jointly Gaussian random vectors and that r can be written in terms of h as $r = Ah$, where A is the row-wise concatenation of M $N \times N$ identity matrices. We use the orthogonality condition [10] to write

$$\begin{aligned} \hat{h}^k &= P^R + KA^T (AKA^T)^{-1} (r - AP^R) \\ &= P^R + \frac{\sigma^2}{NM} \left(r - \sum_{i=1}^{M} \hat{h}_i^k \right), \end{aligned} \tag{7}$$

where $\hat{h}_i^k$ denotes $E[h_i|\hat{\theta}^{k-1}]$. Equation (7) can be substituted into (6) to complete step 2 of the EM algorithm.

Note that, as shown in the last line of (6), $E[L(h)|r, \hat{\theta}^{k-1}]$ is proportional to a sum over M terms, each of which is dependent on the location of only one transmitter. Hence, the maximization in step 3 of the EM algorithm can be performed independently for each transmitter:

$$\hat{\theta}_i^k = \arg\max_{\theta_i} \left(\hat{h}_i^k - P_i^R \right)^T \left(\hat{h}_i^k - P_i^R \right), \quad i = 1, \ldots, M. \tag{8}$$

Direct maximization of the likelihood function of r requires optimizing over $2M$ coupled variables, but the above EM algorithm decouples the transmitter locations, thereby transforming the problem into M 2-dimensional maximizations.

V. Smart Initial Conditions

While the EM algorithm is guaranteed to converge to a local maximum of the likelihood function (5), the function generally has multiple local maxima, and hence the proposed technique may not converge to the global maximum likelihood solution [9]. In order to increase the chances of locating the global maximum of the likelihood function, we employ multiple instantiations of the EM algorithm, each with a different set of initial conditions. In addition, we propose the use of "smart" initial conditions to improve the performance of the EM technique.

Since the EM algorithm converges to a nearby local maximum of the function of interest, the selection of the initial location estimates $\hat{\theta}^0$ plays a critical role in the success of the algorithm in reaching the global maximum of the function. In an effort to generate initial conditions that lie near the global maximum, we have proposed to generate initial transmitter location estimates based on a spatial clustering of the cognitive sensor nodes [6]. The N nodes are grouped into M clusters based on geographic proximity. The power observed by each node in a cluster is assumed to be received from a single (closest) transmitter, and an initial transmitter location estimate is generated for each cluster. If a cluster includes exactly three sensors, the location of the cluster's "effective single transmitter" is determined via trilateration. For clusters with fewer than three sensors, one of the non-unique solutions is chosen; for clusters with more than three sensors, the transmitter location is chosen to minimize the sum of squared differences between measured powers and the power predicted based on (1).

The spatial clustering of the N sensors into M groups is performed using the k-means clustering algorithm [9]. K-means is an iterative algorithm that operates by alternating between computing the centroid of clusters of data points and reassigning data points to clusters based on the centroid to which they are nearest. Like the EM algorithm, the k-means clustering algorithm requires initial conditions, and hence it can provide different spatial clusterings given different random initializations of the algorithm. The ability to produce a variety of clusterings is beneficial for our application, since different clusterings result in different smart initial estimates $\hat{\theta}^0$, which can in turn be used as initial estimates for different instantiations of the EM algorithm. To ensure a variety of clusterings, the k-means algorithm is run for only two iterations for each random start, which generally does not enable the k-means to converge, but is enough iterations to make the sensors in each cluster spatially adjacent.

VI. Simulation Results

We compare the performance of the proposed EM localization technique to particle swarm optimization (PSO), a popular global optimization algorithm [11], [12] via simulation. As a cost function for PSO, we employ the sum of the squared differences between the actual observed power at each sensor and the predicted received power based on estimated transmitter locations [6]. The EM and the PSO localization algorithms are simulated using both random and smart initial conditions, as described above. We used the variant of PSO termed *constriction* PSO.

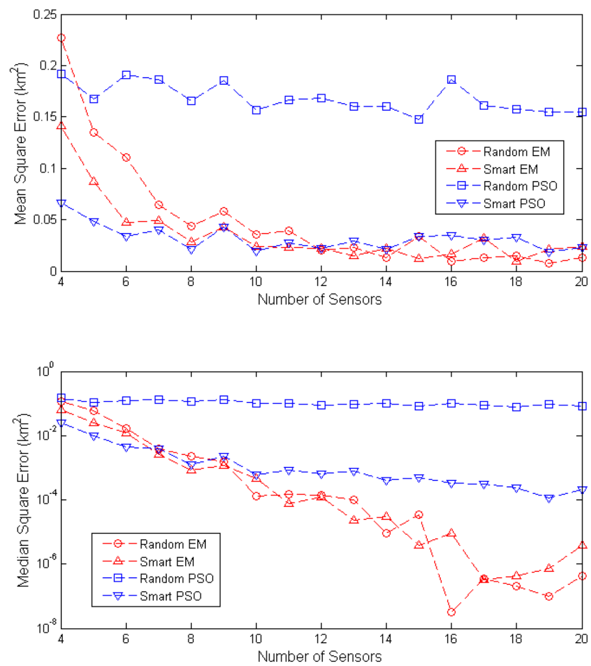

Fig. 1. Simulation results for the estimation of $M = 2$ transmitter locations.

Each localization algorithm was simulated for $M = 2$, 3, and 4 transmitters and for $N = 2M$ to 20 cognitive radio sensors. Two hundred realizations were simulated for each (M, N) pair; for each simulation, the locations of the transmitters and sensors were drawn uniformly from a 1 km $\times$ 1 km square region. For both the EM and PSO approaches, $(2M + 1)^2$ k-means clusterings were performed to generate smart initial conditions, but only the unique clusters were retained. For each realization, the number of random initial conditions was set equal to the number of unique smart initial conditions. To limit the maximum number of computations allowed to perform localization, the EM algorithm was executed for a maximum of 25 iterations, which generally allowed convergence. PSO was allowed $1000M$ guesses at the transmitter locations, which is a common restriction in global optimization experimentation [6], [12].

The performance of the EM and PSO localization schemes is presented in figures 1, 2, and 3 for $M = 2$, 3, and 3 transmitters, respectively, and for random and smart initial conditions. Performance is measured in both the mean and median error in squared distance of the estimated transmitter

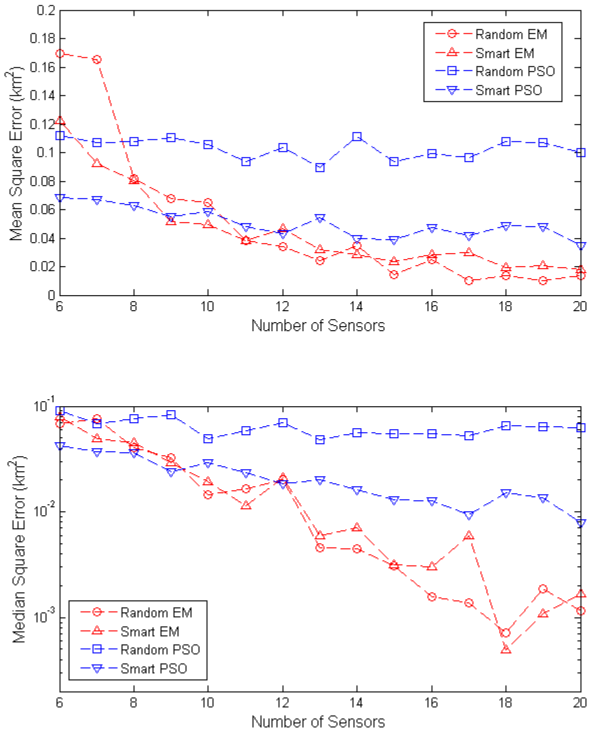

Fig. 2. Simulation results for the estimation of $M = 3$ transmitter locations.

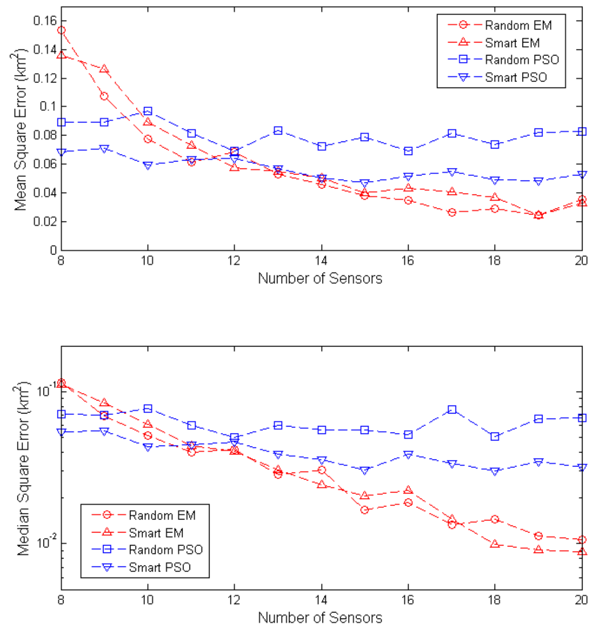

Fig. 3. Simulation results for the estimation of $M = 4$ transmitter locations.

locations relative to the true locations. The figures show that the EM algorithm achieves better performance than PSO-based localization in both mean and median, especially as the number of sensors increases. The EM algorithm shows a larger performance gain in median squared error than in mean squared error, which likely indicates that the EM algorithm significantly outperforms PSO for most iterations but that EM occasionally becomes trapped in particularly poor local maxima that are avoided by the global nature of PSO.

Figures 1, 2, and 3 also reveal that, while smart initial conditions yield significant performance improvements for PSO-based localization, the use of smart initial conditions generally has very little effect on the performance of the EM algorithm. Our analysis shows that the k-means clustering often generates initial estimates that are similar in location. Thus, different smart initial conditions often lead EM to converge to the same local maxima and hence result in no performance improvement. In contrast, PSO may generate different results for the same initial conditions when different initial particle velocities are employed. Smart initial conditions generate a performance improvement for EM only when the number of sensors is very small, a condition in which PSO appears to be the preferable localization technique.

We also employ simulation to study the performance of the EM and PSO localization techniques as a function of measurement noise. For these simulations, we assume $M = 2$ transmitters. Figure 4 presents the mean and median error in squared distance for additive noise powers of $\sigma^2 = 10^{-1}$, 10^{-2}, and 10^{-3}. (While the noise powers are small in an absolute sense, they are generally quite large relative to the power received at a sensor some distance from the legacy transmitters.) Smart initial conditions are used only for the PSO-based scheme, since they result in no performance improvement for EM-based localization.

As expected, the accuracy of the transmitter location estimates generated both localization techniques decreases somewhat as noise power increases. The EM algorithm maintains a significant performance improvement over PSO localization, however, and the performance of the EM technique degrades only slightly as the noise power increases from 10^{-3} to 10^{-1}. It is interesting to note that when noise is present, there is little difference between mean and median measures of squared error. Such results indicate that the estimation error is dominated by the inaccuracy introduced by noise rather than the risk of converging to local maxima.

VII. Conclusions and Open Questions

We have presented an EM-based technique for estimating the locations of multiple transmitters based on measurements of signal strength. The algorithm is applicable to cognitive radio environments in which no cooperation can be assumed between transmitters and sensors. The proposed EM approach reduces a $2M$-dimensional optimization task to M 2-dimensional optimizations. We show via simulation that the EM localization technique performs well and achieves significant gains over PSO-based localization, particularly as

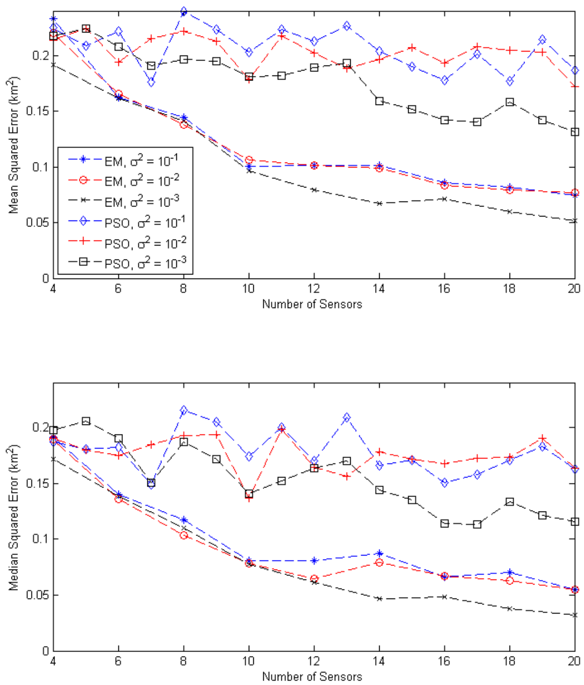

Fig. 4. Simulation results for the estimation of $M = 2$ transmitter locations in the presence of measurement noise.

the complexity of the problem grows. It should be noted that, while we measure performance in terms of the accuracy of transmitter location estimates, the effect of transmitter localization accuracy on the overall performance of the cognitive radio system is of greater interest. This relationship is clearly dependent upon transmission policy, acceptable interference levels, etc., and hence must be studied in such a context. The work we have presented forms a foundation from which to research the important issues of how to perform the estimation without knowledge of the number of transmitters and how to use estimates of transmitter locations to best improve the efficiency of cognitive radio systems.

References

[1] S. Haykin, "Cognitive radio: Brain-empowered wireless communications," *IEEE Journal on Selected Areas in Comm.*, vol. 23, no. 2, pp. 201–220, Feb. 2005.

[2] P. Kolodzy, "Application of cognitive radio technology across the wireless stack," *IEICE Trans. on Comm.*, vol. E88-B, no. 11, pp. 4158–4162, Nov. 2005.

[3] J. Poston, W. Horne, M. Taylor, and F. Zhu, "Ontology-based reasoning for context-aware radios: Insights and findings from prototype development," in *IEEE DYSPAN Symp.*, Nov. 2005, pp. 634–637.

[4] G. Ganesan and Y. Li, "Cooperative spectrum sensing in cognitive radio networks," in *IEEE DYSPAN Symp.*, Nov. 2005, pp. 137–143.

[5] D. Cabric, S. Mishra, D. Willkomm, R. Brodersen, and A. Wolisz, "A cognitive radio approach for usage of virtual unlicensed spectrum," in *Proc. of the 14th IST Mobile and Wireless Comm. Summit*, June 2005.

[6] J. K. Nelson, M. Hazen, and M. R. Gupta, "Global optimization for multiple transmitter localization," in *Proc. of Military Communications Conf. (MILCOM)*, Oct. 2006.

[7] T. S. Rappaport, *Wireless Communications: Principles and Practice*, 2nd ed. Upper Saddle River, NJ: Prentice Hall, 2001.

[8] M. Feder and E. Weinstein, "Parameter estimation of superimposed signals using the EM algorithm," *IEEE Trans. on Acoustics, Speech, and Signal Processing*, vol. 36, no. 4, pp. 477–489, 1988.

[9] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*. New York: Springer-Verlag, 2001.

[10] H. Stark and J. Woods, *Probability, Random Processes, and Estimation Theory for Engineers*, 2nd ed. Englewood Cliffs, NJ: Prentice Hall, 1994.

[11] R. Eberhart and J. Kennedy, "A new optimizer using particle swarm theory," in *Sixth Intl. Symp. on Micro Machine and Human Science*. IEEE, Oct. 1995, pp. 39–43.

[12] J. F. Schutte and A. A. Groenwold, "A study of global optimization using particle swarm," *Journal of Global Optimization*, vol. 31, pp. 93–108, 2005.