A Density-Based Least Squares Formulation of the Sensor Selection Problem for Received Signal Strength Indicator Localization

Ahmad A. Ababneh

Department of Electrical Engineering, Jordan University of Science and Technology, Irbid, Jordan
e-mail: aaababnah@just.edu.jo

Received: July 16, 2017 *Accepted: August 17, 2017*

Abstract— The overall localization error performance in the received signal strength indicator (RSSI) localization depends on several factors that include the number of deployed sensors as well as their deployment positions within the deployment region. In this paper, we study the problem of selecting deployment positions from a set of candidate positions such that localization errors at some specific points of interest are satisfied in a least square error sense under the constraint of a specific number of sensors. To solve this problem, we propose a simple convex formulation of the sensor placement problem as a least squares problem. Instead of expressing error requirements, which can be difficult to evaluate in practice, we use the local sensor deployment density around a point of interest as a cost function to provide the simplified formulation. A numerical evaluation of the proposed approach shows its validity and usefulness.

Keywords— Least squares, Localization, Non-linear estimation, Sensor networks.

I. INTRODUCTION

Localization, whether of a sensor node or a phenomenon of interest (e.g., target), is one of the fundamental underlying processes of a sensor network [1], [2]. The importance of localization stems from the fact that it enables mapping between geographical information and temporal data. This enables the end user to take a targeted proper action (e.g., response to a gas leak) or for sensors to operate correctly (e.g., geographical-based routing).

The focus of this paper is the 'target' localization problem where the goal is to estimate the location of a non-cooperative target of interest based on the measurements of a number of deployed sensors. The estimation error performance depends on several factors (e.g., target energy signature, type and quality of measurements). One of the difficulties associated with the localization problem is the non-linear relationship between measurements (e.g., received signal strength (RSSI), time difference of arrival (TDOA)) and target's true location which makes localization a challenging non-linear estimation problem [3], [4]. This in turn implies the importance of the spatial distribution of sensors with respect to the target's location. In this paper, we focus on sensor positions as an important factor in determining the overall estimation error performance of the sensor network. Thus, it is important to design a suitable strategy for sensor deployment in localization networks.

The sensor placement problem for target localization has received considerable attention over the years [5]-[12]. In the majority of these works, a cost function that is based on some variant of the Fisher information matrix (FIM) or equivalently the Cramer-Rao-Bound (CRB) was used.

The authors in [5] derive optimal positions to minimize estimation error at a single point of interest using only 4 sensors in a 3-D setting. Localization in networks employing time-difference-of-arrival (TDOA) was studied in [6], [7]; and deployment schemes using regular uniform Platonic solids were proposed. The authors in [8]-[10] show that non-uniform deployment schemes can be used to minimize localization errors. One drawback of the above

Corresponding author's e-mail: aaababnah@just.edu.jo

mentioned works is the focus on the minimization of error at only a single point which is not very practical in reality, where sensors monitor a region instead of a point. One work that attempts to address this issue is that of [11], in which sensors are deployed in order to minimize errors at either a single point or multiple points within the region. The authors propose an iterative sensor redeployment algorithm dubbed the RELOCATE algorithm. One shortcoming of RELOCATE is that it restricts sensor movement (i.e., deployment) to the perimeter of the region instead of its interior. Another work that deals with sensor relocation is the one in [12], where an algorithm was proposed to relocate sensors in the ROI to satisfy error requirements. One difference of [12] with our current work is the sequential nature of [20] and the freedom to place sensor. The proposed approach herein is a batch-like approach where sensor placement is restricted to specific locations.

The sensor deployment/placement problem has been also viewed as a sensor selection problem. The sensor selection problem is an NP-hard problem even for modest dimensions. Therefore, several efforts aimed at producing algorithms for effective deployment. For example, in [13] the authors propose an optimal control formulation of sensor placement in detection networks. In [14], the authors propose relaxed convex formulation for sensor placement to be used in linear estimation problems which is not suitable for the localization problem. The work in [15], deals with sensor placement for non-linear measurement models; and several optimization solvers were proposed. The authors use the localization problem as a case study for their approach. However, the localization problem studied does not take the sensing range of sensors into consideration; and it assumes sensors are to be deployed on the perimeter of the ROI but not within the ROI.

In this work, we propose a novel sensor selection approach for localization with emphasis on RSSI localization. In essence, estimation error performance at a point is to be expressed as a function of the number of sensors within a radius around the point (i.e., sensor density). Thus, we can express estimation requirements simply as density requirements. Using density information, we propose to formulate the sensor selection problem for target localization as a least squares problem, which can be readily solved. Given the number and distribution of candidate position points and points of interest, we can assign different levels of importance to different points to adhere with the end user requirements.

II. PROBLEM FORMULATION

Without loss of generality, we consider a 2-D region of interest (ROI) of dimensions $b \times b$. It is assumed that number N points of interest are specified within the ROI that need not be uniformly spaced. Estimation error requirements are assigned to each point, where the requirement of the n^{th} point is denoted as $E_{req}(n)$ and requirements are grouped in vector $\boldsymbol{E}_{req}$. We assume that M homogeneous sensors are to be deployed with a uniform sensing radius R_s. Furthermore, there is a number of K candidate positions where sensors can be placed. We find that $M < K$ in order for the problem to become non-trivial.

There are several types of sensor measurements that can be used for localization. In this paper, and without loss of generality, we focus on the received signal strength indicator (RSSI) measurements. We assume that sensors provide (RSSI) measurements which are easier to attain than other types of measurements (e.g., TDOA) [2]. Let $d(m,t)$ denote the distance between the location of the m^{th} sensors and the target of interest located at the point (p_t). If the target has a uniform energy signature, then the RSSI measurement is given as [16], [17]:

$$z(m,t) = \frac{P_0}{d(t,m)^{\alpha}} + w_m \quad (1)$$

where P_0 denotes the target's initial energy and; α is the decay factor ($1 < \alpha < 2$). The noise $\{w_m, m = 1, \ldots, M\}$ is i.i.d with a Gaussian distribution with zero mean and variance σ^2.

We assume that sensors send their raw analog measurements to a central processor (CP) which performs a location estimation algorithm. We note that error performance depends on the estimation algorithm. However, in order to focus on the problem at hand, we assume that the CP employs the clairvoyant estimator (i.e., maximum likelihood (ML)) to produce the estimate $\hat{p}_t$.

The error performance of the best estimator can be quantified using the Cramer-Rao Bound (CRB) given as (assuming that the target is located at the n^{th} point of interest) [18], [19]:

$$E[(p_t - \hat{p}_t)(p_t - \hat{p}_t)^T] \geq J_n^{-1} \quad (2)$$

The matrix J_n, associated with n^{th} point, is called the Fisher information matrix (FIM); and is given as:

$$J_n = \frac{1}{4\sigma^2} \sum_{i \in I_n} J_{n.i} \quad (3)$$

where I_n denotes the set of sensor indexes that are within the sensing radius R_s of the target's location.

The submatrix $J_{n,i}$ quantifies the amount of information that the i^{th} sensor can provide about the target at the n^{th} point; and is given as [16]:

$$J_{n,i} = \frac{P_0 \alpha^2}{d_{n,i}^4} (p_n - p_i)(p_n - p_i)^T \quad (4)$$

We note that error performance depends on several factors (e.g., decay factor, noise statistics). Additionally, we note that a very important factor is the position of sensors with respect to the points of interest. Thus, in order to achieve a certain error performance it is important to place sensors at specific locations with respect to the points of interest, especially under sensor budget constraints.

We now state the problem that we will study in this paper. Given K candidate sensor positions and M sensors to be deployed ($M < K$), how should we select the deployment positions such that the estimation errors are satisfied in a least square sense, that is:

$$\begin{cases} arg\ min_S \sum (E_{req}(i) - E_s(i))^2 \\ s.t. |s| = M \\ s(k) = \{0,1\},\ 1 \leq k \leq K \end{cases} \quad (5)$$

where the $K \times 1$ vector $\mathbf{s}$ is a selection vector the entries of which are either 0 or 1. If a sensor is present at the k^{th} point then the k^{th} entry $\mathbf{s}(k) = 1$; otherwise, it is set to 0. The cardinality of the vector $\mathbf{s}$ is denoted as $|\mathbf{s}|$; and $E_{\mathbf{s}}(i)$ denotes the achieved estimation error at the i^{th} point corresponding to the selection vector **s.**

The above problem formulation is known as the sensor selection problem which is an NP-hard problem. Note that when $M < K$, the number of possible combinations is on the order of $\binom{K}{M}$ which, even for small values of M and K, makes the problem computationally

intractable. For example, if $K = 15$ and $M = 6$, then there are 5005 possible sensor configurations to choose from; this is computationally intensive to perform. Moreover, the sensor selection problem is non-convex. This results from the non-convex binary nature of the selection vector $\mathbf{s}$. Thus, suitable methods need to be designed to solve such a problem.
In the next section, we present our proposed approach for solving this problem.

III. PROPOSED SOLUTION

One widely used method to solve the sensor selection problem is that of constraint relaxation in which the entries $\mathbf{s}(k)$ of the vector $\mathbf{s}$ are allowed to have values that fall in the interval [0,1]; rounding entries after solution is obtained to either 0 or 1.
In this paper, we propose a formulation of the sensor selection problem in which the cost function is not the exact error as calculated using the CRB bound, but rather a surrogate function which is the sensor density. We note that using sensor density provides a sub-optimal solution to the above problem; however, it is obvious that it helps reduce the complexity of evaluating exact errors and performing matrix inversion at each iteration. Using this new function, the cost function of the sensor selection problem becomes easier to evaluate; and it results in a least square optimization problem which can be easily solved.
We first derive the relationship between localization error and sensor density. We then propose our formulation of the modified sensor selection problem.

A. Localization Error in Terms of Sensor Density

We will first describe localization error in terms of the 'average' of the FIM described earlier. We note that

$$E[(p_t - \hat{p}_t)^T (p_t - \hat{p}_t)] \geq Tr(E(J_n^{-1})) \tag{6}$$

where, Tr() denotes the trace operator. We first calculate the average of entries of the $\boldsymbol{J}_n$ matrix and start with the first entry $\boldsymbol{J}_n(1,1)$ whose average can be written as:

$$E[J_n(1,1)] = \frac{P_0 \alpha^2 N}{4\sigma^2} \frac{1}{N} \sum_{i \in I_n} J_i(1,1) \tag{7}$$

We note that

$$E[J_i(1,1)] = \frac{1}{N} \sum_{i \in I_n} J_i(1,1) \tag{8}$$

Assuming that sensors are uniformly distributed within the sensing zone around a point, the density function of sensors is given as:

$$f_{x_m, y_m}(x_m, y_m) = \frac{1}{\pi R_s^2}, \ \ if \sqrt{x_m^2 + y_m^2} \ \leq R_s$$

Then after some operations (see [20] for more details), the estimation error $E_{res}(n)$ can be written as:

$$E_{res}(n) = \frac{8\sigma^2}{P_0 \alpha^2 \pi (R_s - 1)} \frac{1}{\lambda} \tag{9}$$

where λ denotes sensor density which is given as

$$\lambda = \frac{N}{\pi R_s^2}$$

We note in (9) that the larger the sensor density (i.e., more sensors are available), the lower is the estimation error as expected. Fig. 1 shows a plot of the relation between $E_{res}(n)$ and λ. Equation (9) shows that estimation error is a function of λ. We will use this fact to reformulate the sensor selection problem as in the next section.

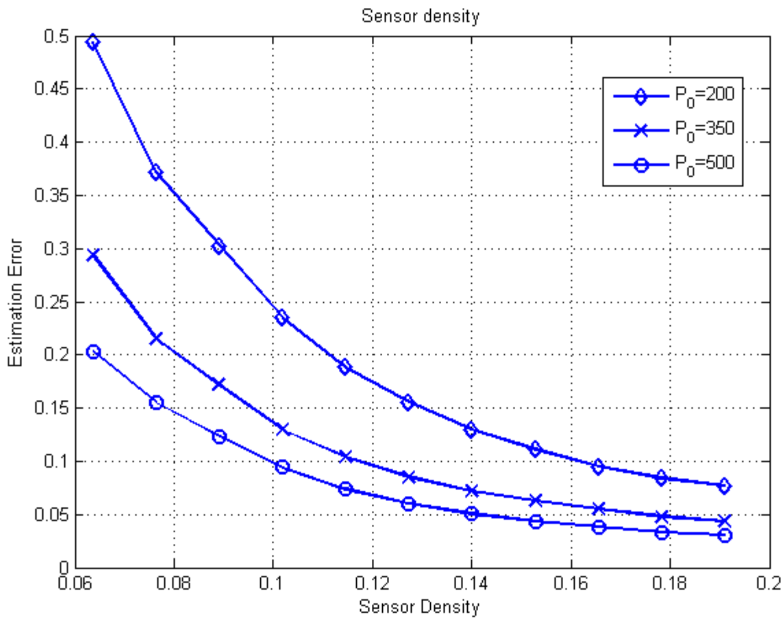

Fig. 1. Localization error vs. sensor density ($R_s = 5, \sigma^2 = 1, \alpha = 1.5$)

B. Least Square Density (LS-Density) Formulation

Based on the discussion above, error requirement can be expressed in terms of the sensor density that can satisfy it. Let the $K \times 1$ vector $\boldsymbol{\lambda}_{req}$ be the corresponding vector of $\mathbf{E}_{req}$, then we propose to formulate the sensor placement/selection problem as follows:

$$\begin{cases} min_s \ \| \gamma As - \lambda_{req} \|_2 \\ \quad s.t. |s| = M \\ s(k) = \{0,1\}, \ 1 \le k \le K \end{cases} \tag{10}$$

where $\| \|_2$ is the second norm operator; and γ is a constant that will be explained later. The matrix $\boldsymbol{A}$ above serves a role similar to an adjacency matrix used in graph theory. The entry $\boldsymbol{A}(i,j)$, in the i^{th} row and j^{th} column, is a function of the i^{th} point of interest and the j^{th} candidate sensor position. There are different methods to assign $\boldsymbol{A}(i,j)$, depending on the importance of a candidate sensor position. For simplicity, we assume that all candidate positions are equally important, thus the entry $\boldsymbol{A}(i,j)$ is given as

$$A(i,j) = \begin{cases} 1, & if \ d(i,j) \le R_s \\ 0, & else \end{cases} \tag{11}$$

In this assignment, the i^{th} row of $\boldsymbol{As}$ quantifies the number of sensors covering (i.e. within the sensing range) the i^{th} point of interest. In order to convert this information to density, we scale $\boldsymbol{As}$ with the factor $\gamma = 1/\pi R_s^2$ which is the area of the sensing zone.

We note that the problem in (5) is transformed into a linear least squares constrained optimization problem which can be readily solved using various programming methods (e.g., Matlab). After solving (10), the largest *M* entries of $\mathbf{s}$ are selected to place the *M* sensors; and the remaining locations are left empty.

In the next section, we examine the performance of the proposed formulation.

IV. Simulation Results and Discussion

In the following experiments, we will investigate the performance of the solution to the sensor selection problem using three metrics. The first performance metric is the average of the squared estimation error difference which is defined as

$$Average\ Error = \frac{1}{N_{c3}} \sum_{i \in I_{c3}} (E_{req}\,(i) - E_s(i))^2$$

where the error difference is taken over the points that are covered with at least 3 sensors. This is intuitive since localization can only be performed in 2-D setups using at least 3 sensors. Estimation error is assumed to be undefined for points with less than 3 sensors.

The second performance metric that we will use is the `coverage hole' ratio (i.e., the percentage of points of interest that are not covered with at least 3 sensors). The combination of both metrics is important since it is not enough to reduce error at a small fraction of points of interest while a larger number of points is not covered and is not incorporated in the average error definition above. Thus, it would be instructive to look at the product of both the average error and the coverage hole ratio to have a clear quantification of the performance of any of the sensor selection methods.

This product term is the third metric that we will be using. The smaller is the product value, the better is the overall performance of the selection process. A large value indicates that either coverage is not adequate; or localization error is high or both.

In addition to our proposed LS-Density solution, we use two other methods for the sake of comparison. The first method is an 'optimal' brute force solution that employs enumeration of `all' possible combinations, compares the average error for all possibilities and chooses the deployment positions with the minimum average error. This approach is suitable for the case of relatively small values of M and K. The second method is a random selection method in which M sensors are selected at random from the K candidate positions.

Experiment 1: Performance vs. number of sensors ($\boldsymbol{M}$):

In the first experiment, we study the performance of different methods. The number of sensors M to be selected out of $K = 12$ positions varies. A uniform error requirement of $E_{req} = 0.1$ is specified at a number of 200 random points of interest within the ROI. This implies that the required sensor density is of $\lambda = 0.12$. The experiment parameters are listed in Table 1 below; and results are depicted in Fig. 2 and Fig. 3 below.

We note that both the error and coverage performance of the proposed LS-Density solution method outperforms that of the random selection approach. Moreover, as *M* is increased, the performance of the LS-Density method approaches that of the optimal enumeration method. This is natural as the number of possible combinations decreases.

TABLE 1
EXPERIMENT PARAMETERS

b	15
K	12
P_0	500
R_s	5
σ^2	1

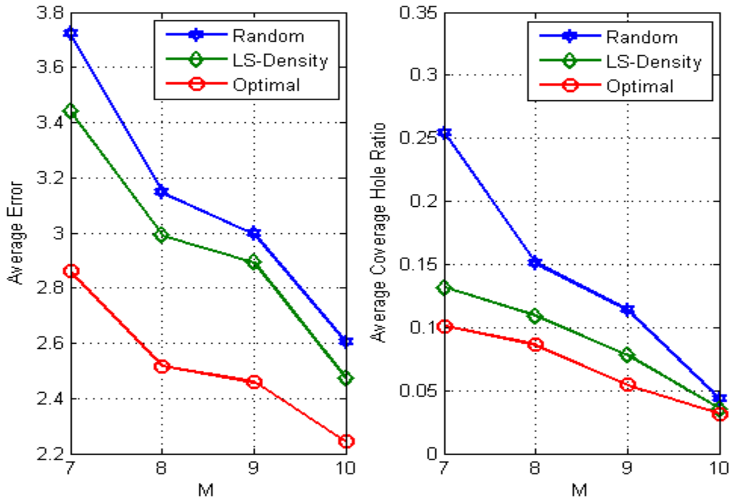

Fig. 2. Average error and coverage hole for different M

However, to have a better quantification of performance, we examine the error-coverage product as shown in Fig. 3. We note that when $M = 7$, the LS-Density offers an overall reduction of almost 50% when compared to the random deployment method. It is also instructive to note that the random approach provides a product almost 300% more than that of the optimal brute search method, whereas the proposed approach yields only 60% more than the optimal. As the number of sensors to be selected is increased (from $M = 7$ to $M = 10$), products become closer to each other due to the reduction in the number of possible combinations.

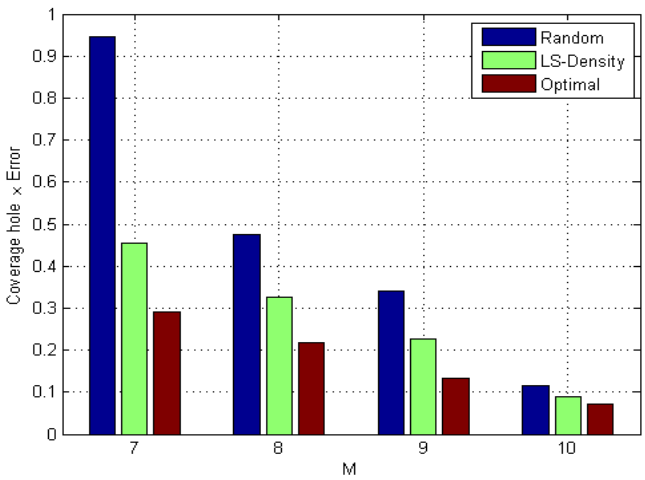

Fig. 3. Average error and coverage hole product for different M

Experiment 2: Performance vs. sensing radius (R_s):

In this experiment, we use the similar parameters to that in the previous experiment but we fix both the number of sensors to $M = 8$ and the estimation error requirement to $E_{req} = 0.1$, and vary the sensing radius (R_s). Fig. 4 and Fig. 5 show the error, coverage and product performance of different approaches.

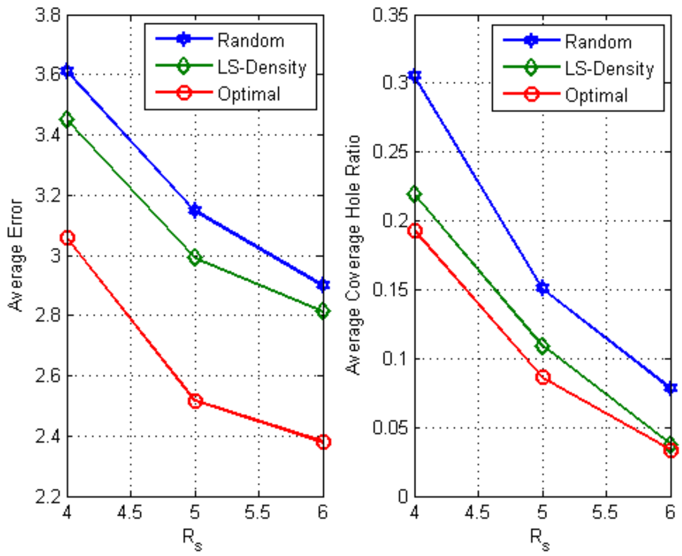

Fig. 4. Average error and coverage hole for different R_s

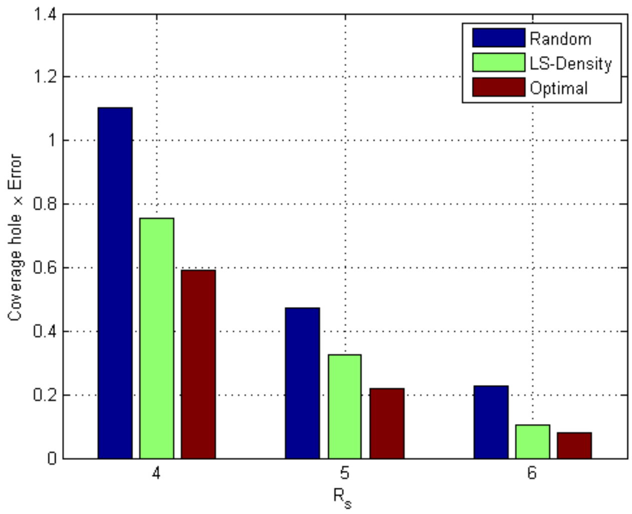

Fig. 5. Average error and coverage hole product for different R_s

It is evident that the coverage performance of the LS-Density is very comparable to the optimal solution; and that it provides a lower error performance than that of the random selection method. In particular, the error-coverage hole product for the LS-Density is only around 25% higher than that of the optimal approach for $R_s = 4$. This is in contrast to the random selection method which is around 85% larger than the optimal for $R_s = 4$. We also note that as R_s is increased, the product performance for all methods becomes closer to each other. This is a result of the fact that more points are covered with more and more sensors.

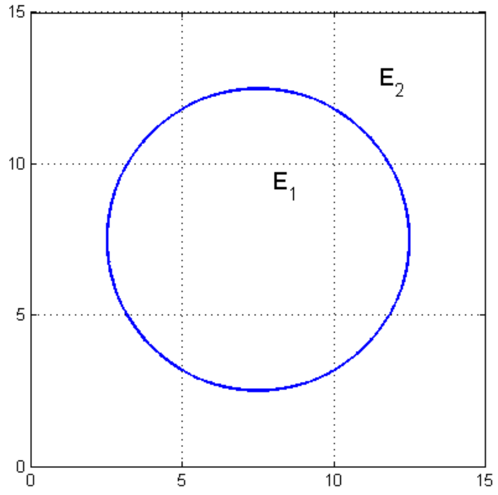

Fig. 6. ROI with different requirements

Experiment 3: Performance vs. different requirements:
In this experiment, we consider ROI with different performance requirements. As shown in Fig. 6 below, points of interest that fall within the central area are assigned a requirement of E_1; and those that are outside the area are assigned a value of $E_2 = 0.15$. We use the same parameters as in the previous experiments with $M = 8$ sensors and $K = 12$ with R_s=5. Simulation results are summarized in Fig. 7 and Fig. 8.

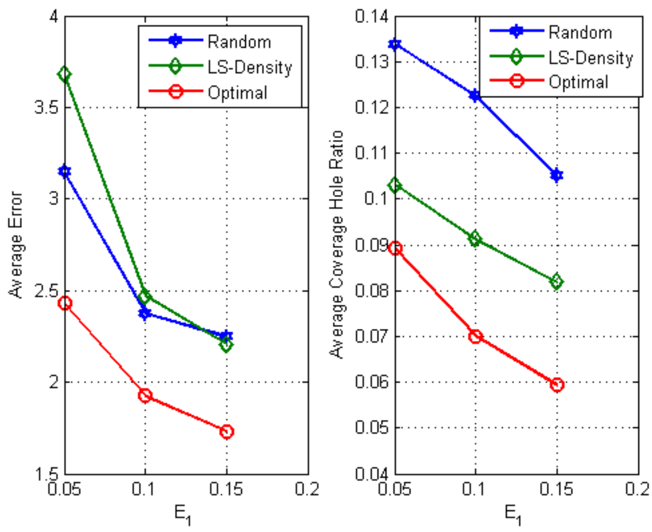

Fig. 7. Average error and coverage hole for different E_1

We note that in this case of non-uniform requirements, random deployment might fair better than the proposed LS-Density with respect to error performance. However, the LS-Density approach compensates through a small coverage hole ratio to cover more points. This is evident in the error-coverage hole product in Fig. 8.

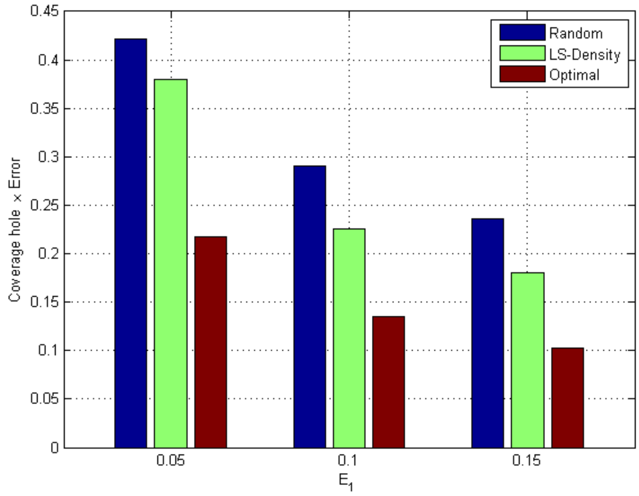

Fig. 8. Average error and coverage hole product for different E_1

V. Conclusion

In this paper, we studied the sensor selection problem for RSSI localization. By expressing estimation error requirements in terms of sensor density information, we were able to transform the non-convex selection problem to another simpler convex least squares formulation, which can be easily solved using a range of computational software. Numerical evaluation of the proposed least squares density based formulation shows its effectiveness in terms of both localization error and coverage performance.

References

[1] H. Karl and A. Willig, *Protocols and Architectures for Wireless Sensor Networks*, Wiley, 2007.

[2] G. Mao, B. Fidan, G. Mao and B. Fidan, *Localization Algorithms and Strategies for Wireless Sensor Networks,* Hershey, PA: Information Science Reference-Imprint of: IGI Publishing, 2009.

[3] X. Sheng and Y. Hu, "Maximum likelihood multiple-source localization using acoustic energy measurements with wireless sensor networks," *IEEE Transactions on Signal Processing*, vol. 53, no. 1, pp. 44-53, 2005.

[4] G. Mao, B. Fidan and B. Anderson, "Wireless sensor network localization techniques," *Computer Networks*, vol. 51, no. 10, pp. 2529-2553, 2007.

[5] T. Hegazy and G. Vachtsevanos, "Sensor placement for isotropic source localization," *Proceedings of the 2nd International Conference on Information Processing in Sensor Networks*, pp. 432-441, 2003.

[6] B. Yang and J. Scheuing, "A theoretical analysis of 2D sensor arrays for TDOA based localization," *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing*, vol. 4, pp. 901-904, 2006.

[7] B. Yang, "Different sensor placement strategies for TDOA based localization," *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing*, vol. 2, pp. 1093-1096, 2007.

[8] A. Bishop, B. Fidan, B. Anderson, P. Pathirana and K. Dogancay, "Optimality analysis of sensor-target geometries in passive localization: part 2-time-of-arrival based localization," *Proceedings of the 3rd International Conference on Intelligent Sensors, Sensor Networks and Information Processing*, pp. 13-18, 2007.

[9] A. Bishop and P. Jensfelt, "An optimality analysis of sensor-target geometries for signal strength based localization," *Proceedings of IEEE 5th International Conference on Intelligent Sensors, Sensor Networks and Information Processing*, pp. 127-132, 2009.

[10] A. Bishop, B. Anderson, B. Fidan, P. Pathirana and G. Mao, "Bearing-only localization using geometrically constrained optimization," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 45, no. 1, pp. 308-320, 2009.

[11] D. Jourdan and N. Roy, "Optimal sensor placement for agent localization," *ACM Transactions on Sensors and Network*, vol. 4, no. 3, pp. 1-4, 2008.

[12] A. Ababneh and F. Zghoul, "A distributed sensor relocation algorithm for target localization problems," *Adhoc & Sensor Wireless Networks*, vol. 35, pp. 301-318, 2017.

[13] A. Ababnah and B. Natarajan, "Optimal control-based strategy for sensor deployment," *IEEE Transactions on Systems, Man, and Cybernetics-part A: Systems and Humans*, vol. 41, no. 1, pp. 97-104, 2011.

[14] S. Joshi and S. Boyd, "Sensor selection via convex optimization," *IEEE Transactions on Signal Processing*, vol. 57, no. 2, pp. 451-462, 2009.

[15] S. Chepuri and G. Leus, "Sparsity-promoting sensor selection for non-linear measurement models," *IEEE Transactions on Signal Processing*, vol. 63, no. 3, pp. 684-698, 2015.

[16] E. Elnahrawy, X. Li and R. Martin, "The limits of localization using signal strength: a comparative study," in *Proceedings of IEEE First Annual IEEE Communications Society Conference on Sensor and AdHoc Communications and Networks,* pp. 406-414, 2004.

[17] R. Niu and P. Varshney, "Target location estimation in sensor networks with quantized data," *IEEE Transactions on Signal Processing*, vol. 54, no. 12, pp. 4519-4528, 2006.

[18] H. Poor, *An Introduction to Signal Detection and Estimation,* Springer-Verlag, 1988.

[19] B. Levy, *Principles of Signal Detection and Parameter Estimation,* Springer, 2008.

[20] A. Ababneh, "Target localization accuracy improvement via sensor mobility," *Parallel, Emergent and Distributed Systems*, pp. 1-21, 2007.