Performance Characterization of Random Proximity Sensor Networks

Agostino Capponi
Department of Computer Science
California Institute of Technology
Pasadena, CA 91125 USA
acapponi@cs.caltech.edu

Lance Kaplan
U.S. Army Research Laboratory
Adelphi, MD 20783 USA
lkaplan@ieee.org

Concetta Pilotto
Department of Computer Science
California Institute of Technology
Pasadena, CA 91125
pilotto@cs.caltech.edu

***Abstract*— In this paper, we characterize the localization performance and connectivity of sensors networks consisting of binary proximity sensors using a random sensor management strategy. The sensors are deployed uniformly at random over an area, and to limit the energy dissipation, each sensor node switches between an active and idle state according to random mechanisms regulated by a birth-and-death stochastic process. We first develop an upper bound for the minimum transmitting range which guarantees connectivity of the active nodes in the network with a desired probability. Then, we derive an analytical formula for predicting the mean-squared localization error of the active nodes when assuming a centroid localization scheme. Simulations are used to verify the theoretical claims for various localization schemes that operate only over connected active nodes.**

I. Introduction

The emerging world of sensor networks has imposed a different approach to the tracking problem [1]. A large number of tiny, inexpensive sensors are spread over a monitored area. Each sensor node operates untethered, has a microprocessor and a limited amount of memory for signal processing. Each sensor has severe constraints on the battery power, and can only communicate wirelessly with a small number of sensors within its radio communication range. Efficient localization can only be possible by an appropriate sensor collaboration aimed at increasing the estimation quality while controlling the power/bandwidth consumption and robustness against failures. Such an approach is known in the literature as *Collaborative Signal and Information Processing* [2]. In this paper we restrict our attention to *binary proximity sensors*, *i.e.*, sensors which can only provide one bit of information regarding the presence of a target in their neighborhood. Individually they are unable to localize the target, but collectively they can communicate their binary detection decision over the network and localize the target by appropriately fusing the received information. To conserve the battery, a power management policy is employed where each sensor switches between two different states, *idle* or *active*, using a a random mechanism based on birth-and-death processes. Traditional protocols such as on-demand [3] and rendezvous wake-up mechanisms require sensors to synchronize their clocks so that a common wake-up schedule can be decided. This problem becomes particularly difficult in multi-hop sensor networks where nodes do not directly communicate with each other. Random wake-up mechanisms such as the one proposed in this paper do not require any synchronization between sensors both for the sleeping and awaking stage. Each sensor switches state independently of any other in the network and the switching times are determined randomly. While such protocols are simplistic and do not exploit features of the network configuration, they are cheap, easy to implement, scalable and independent of the particular network topology.

The main purpose of the paper is to predict the probability of connectivity and localization performance of the sensor network given design parameters for the transmission range and random energy management protocol. These expressions can be used in the design of sensor networks to meet certain specifications, *e.g.*, root mean squared (rms) localization error performance. The rest of the paper is organized as follows. Section II presents the random energy management strategy used for deciding transitions between the states of the sensors. Section III provides an upper bound for the minimum transmitting range required to achieve connectivity of the network with a desired probability. Section IV develops an analytical formula for the mean-squared error of the centroid localization algorithm. Results for the simulated network model are provided in Section V. Finally, Section VI concludes the paper.

II. Random Energy Management Protocol

Each sensor can either be in an active or idle state, and each sensor makes the active-to-idle or idle-to-active transitions according to two independent exponential distributions over time characterized by mean $\frac{1}{\mu}$ and $\frac{1}{\lambda}$, respectively. The parameters $\lambda, \mu \geq 0$ will be referred to as the *sensor activation* and *sensor deactivation* rate, respectively. Let's denote by N the number of deployed sensors and define $X(t)$ to be the number of active sensors in the network at time t. Then $\{X(t), t \geq 0\}$ defines a *birth-and-death* process taking on integer values in $[0, N]$ with parameters λ_n, μ_n defined as

$$\lambda_n = \begin{cases} (N-n)\lambda & 0 \leq n \leq N \\ 0 & \text{otherwise} \end{cases}, \qquad \mu_n = \begin{cases} n\mu & 0 \leq n \leq N \\ 0 & \text{otherwise} \end{cases}. \tag{1}$$

The parameters $\{\lambda_n\}_{n\in N}$ are referred to as *network activation* rates. If n sensors are active, then each of the remaining $N-n$ idle sensors awake at a rate λ, we have that $\lambda_n = (N-n)\lambda$. The parameters $\{\mu_n\}_{n\in N}$ are referred to as *network deactivation* rates. If n sensors are active, then each of them will transition to the idle state at a rate μ, and it follows that $\mu_n = n\mu$.

We focus our attention on the long-term behavior of the network and denote by $\eta_n = \lim_{t\to\infty} \mathbb{P}(X(t) = n)$ the stationary probability that n sensors in the network are active. Using the theory of birth-and-death processes [4] we obtain

$$\eta_n = \frac{\frac{\lambda^n}{\mu^n}\binom{N}{n}}{1+\sum_{n=1}^{N}\frac{\lambda^n}{\mu^n}\binom{N}{n}} = \frac{\frac{\lambda^n}{\mu^n}\binom{N}{n}}{(\frac{\lambda+\mu}{\mu})^N}, \quad 0 \le n \le N. \tag{2}$$

Hence, for $t\to\infty$, the average number of sensors which are active, denoted by $E[X] := \lim_{t\to\infty} E[X(t)]$, is

$$E[X] = \sum_{n=0}^{N} n\eta_n = \frac{\sum_{n=1}^{N} n\frac{\lambda^n}{\mu^n}\binom{N}{n}}{1+\sum_{n=1}^{N}\frac{\lambda^n}{\mu^n}\binom{N}{n}} = N\frac{\lambda}{\lambda+\mu}. \tag{3}$$

III. Critical Transmitting Range

A. Preliminaries

Let I_1, I_2 be two intervals on the real line. A *random geometric graph* $G[I_1, I_2]$ is a graph where each node is uniquely identified by its x, y coordinates which are selected according to some probability distribution on $I_1 \times I_2$ and an edge between two nodes occurs if and only if their distance is smaller than a certain threshold r [5]. For our application, the distance is Euclidean. A *connected component* C of the graph is a maximal set of nodes such that a path exists for any pair of nodes in C. A graph G is said to be *connected* if it consists of only one connected component, or equivalently, for any pair of nodes u, v in G, there exists a path from u to v in G. For a given $\epsilon > 0$, we denote by $\kappa_{1-\epsilon}(G[I_1, I_2])$ the minimum value of r above which $G[I_1, I_2]$ is connected with probability larger than $1-\epsilon$. The *degree* of a node v in the graph is defined to be the number of nodes which are connected to v by an edge. The degree of the graph G is defined to be the minimum degree of any of its nodes. If a node has zero degree we say that the node is *isolated*. We denote by $\delta_{1-\epsilon}(G[I_1, I_2])$ the minimum value of r above which $G[I_1, I_2]$ has strictly positive degree with probability larger than $1-\epsilon$. Clearly, the event that the degree of G is strictly positive is a necessary condition for the connectivity of the network, thus $\delta_{1-\epsilon}(G[I_1, I_2]) \le \kappa_{1-\epsilon}(G[I_1, I_2])$.

B. Derivation of the minimum transmitting range

We assume that the N sensors are deployed uniformly at random in a square of unit area. The network may be modeled by a dynamic geometric random graph $G[I, I]$, $I = [0, 1]$, where sensors are the nodes and the edge between nodes x and y exists if and only if 1) the distance between x and y is smaller than a certain threshold r and 2) both x and y are currently active sensors. In constrast to the theory of geometric random graphs where the occurrence of edges are known once the positions of the nodes are known, in our case the occurrence of edges between nodes is random and determined by the activation and deactivation rates of the sensors. Our objective is to compute the minimum transmitting range which guarantees that the network be connected with a given desired probability. We are interested in minimizing the transmitting range because the energy consumed by a node for communication is directly dependent on its transmitting range [6]. We first give a bound for $\delta_{1-\epsilon}(G[I, I])$ when using the (toroidal) Euclidean distance and then show experimentally in Section V that this is very close to $\kappa_{1-\epsilon}(G[I, I])$ for sensor networks consisting of more than thirty sensors.

Lemma 1: Let $\epsilon > 0$ and G a graph on N nodes. Then $\delta_{1-\epsilon}(G[I, I]) \le \sqrt{-\frac{\log(\epsilon/E[X])}{\pi E[X](N-1)/N}}$.

Proof: We denote by $A = \pi r^2$ the communication area of a sensor whose transmission range is r, and by P_{isol} the probability that a given node in the graph is isolated. Then using the union bound it holds that

$$P_{isol} \le \sum_{n=1}^{N} \eta_n \; n(1-A)^{n-1}, \tag{4}$$

since in a graph consisting of n nodes, we have that a given node is isolated if all the remaining $n-1$ nodes fall outside its communication area. Inserting (2) in (4) leads to

$$P_{isol} \le \sum_{n=0}^{N-1}(n+1)\frac{(\frac{\lambda}{\mu})^{n+1}\binom{N}{n+1}}{(\frac{\lambda+\mu}{\mu})^N}(1-A)^n, \tag{5}$$

which can be rewritten using the binomial identity $\binom{N}{k+1} = \frac{N}{k+1}\binom{N-1}{k}$ and some algebraic manipulations as

$$P_{isol} \le \frac{\lambda N}{\lambda+\mu}\left(1-\frac{A\lambda}{\lambda+\mu}\right)^{N-1}. \tag{6}$$

Letting $q = \frac{A\lambda}{\lambda+\mu}$, then (6) can be rewritten as

$$\frac{\lambda N}{\lambda+\mu}(1-q)^{N-1} = \frac{\lambda N}{\lambda+\mu}\exp(-(N-1)\log(1-q)). \tag{7}$$

Similar to the approach used in [7] for random graphs, the Taylor series expansion for the second term on the right hand side of (7) can be rewritten as

$$\exp(-(N-1)q)\exp\left(-(N-1)q^2\left(\frac{1}{2}+\frac{q}{3}+\dots\right)\right). \tag{8}$$

Choosing $A = -\frac{\log(\epsilon/E[X])}{E[X](N-1)/N}$, we obtain that the term in (8) is smaller than $\frac{\epsilon}{E[X]}$, and therefore, $P_{isol} < \epsilon$ from due to (6). Since $A = \pi r^2$, then

$$r = \sqrt{-\frac{\log(\epsilon/E[X])}{\pi E[X]\ (N-1)/N}} \tag{9}$$

guarantees that the probability of non-isolation is greater than $1-\epsilon$ and $\delta_{1-\epsilon}(G[I, I]) \le r$. □

In the proof of Lemma 1, $A = \pi r^2$ for the (non-toroidal) Euclidean distance except near the edges, and for all cases $A >= \frac{\pi r^2}{4}$ with equality being satisfied when the sensor is located in a corner. Therefore, taking the effect of the edges into account, we will have that

$$\delta_{1-\epsilon}(G[I,I]) = c\sqrt{\frac{A}{\pi}}, \tag{10}$$

for some $1 \leq c \leq 2$ when using the standard Euclidean distance.

IV. Mathematical model of Localization

This section derives the theoretical expression for the expected mean squared error (MSE) of the sensor network for the case that the target is located in the center of the surveillance region of radius R and the centroid estimator is used for localization. Furthermore, we assume that there is zero latency in message transmission; thus, if the transmitting range is chosen as in (10) and ϵ is small, all sensors will have the same set of decisions at all instants with very high probability. The location of the i-th sensor can be described in Cartesian coordinates (x_i, y_i) where, without loss of generality, the target is located at the origin. Alternatively, the location of the sensor can be represented via polar coordinates relative to the target (r_i, θ_i), where r_i and θ_i are the range and bearing to the target, respectively. For binary sensors, the sensor model is the probability of detection $P_d(r_i)$, where r_i denotes the position of the i-th sensor relative to the target. The centroid estimator is a baseline technique that localizes a target as the centroid of the active sensors that detect the target, *i.e.*,

$$(\hat{x}(\boldsymbol{d}), \hat{y}(\boldsymbol{d}))^T = \frac{\sum_{i=1}^{N_a} d_i (x_i, y_i)^T}{\sum_{i=1}^{N_a} d_i}, \tag{11}$$

where N_a is the number of active nodes and $\boldsymbol{d}$ is the vector of detection results where the i-th element d_i corresponds to whether or not the i-th sensor detected the target. In other words, $d_i = 1$ if the i-th active sensor detected the target, and $d_i = 0$ otherwise.

The expression for the MSE of the estimator is averaged over all possible $2^{N_a} - 1$ realizations of sensor network detections, *i.e.*,

$$\text{MSE} = \sum_{N_a=1}^{N} \eta_{N_a} \sum_{j=1}^{2^{N_a}-1} P_j(\boldsymbol{r}) \parallel (\hat{x}(\boldsymbol{d}_j), \hat{y}(\boldsymbol{d}_j)^T) \parallel^2, \tag{12}$$

where $\|\cdot\|$ is the l_2-norm, $\boldsymbol{d}_j$ is the j-th realization of detections, $P_j(\boldsymbol{r})$ is the probability of $\boldsymbol{d}_j$, and η_i was previously defined in (2). Note that localization via proximity sensors is undefined for the case that no nodes detect the target. Thus, the probability of the j-th realization of detections is

$$P_j(\boldsymbol{r}) = \frac{1}{P_{Net}} \prod_{k=1}^{N_a} (P_d(r_k))^{d_{k,j}} (1 - P_d(r_k))^{1-d_{k,j}}, \tag{13}$$

were $\boldsymbol{r} = (r_1, r_2, \ldots, r_n)$ is the vector of all relative sensor/target distances, $d_{k,j}$ is the detection result associated to the k-th sensor for the j-th realization, and $P_{Net} = 1 - \prod_{k=1}^{N_a}(1 - P_d(r_k))$ is the probability that at least one detection took place.

The expected value of the inner argument of (12) over all possible node configurations for realization $\boldsymbol{d}_j$ is

$$B = \frac{E[P_j(\boldsymbol{r}) \sum_{h=1}^{N_a} \sum_{k=1}^{N_a} d_{h,j} d_{k,j} r_h r_k \cos(\theta_h - \theta_k)]}{(\sum_{h=1}^{N_a} d_{h,j})^2}. \tag{14}$$

Since the nodes are independently distributed uniformly over a circular region of radius R, r_i and θ_i for $i = 1, \ldots, N_a$ are statistically independent. Furthermore, the range to the target has a density $f(r) = \frac{2r}{R^2}$ for $0 \leq r \leq R$, and the bearing has density $g(\theta) = \frac{1}{2\pi}$ for $0 \leq \theta < 2\pi$. Because $E[\cos(\theta_i - \theta_j)] = \delta_{ij}$ and $\boldsymbol{d}_j$ is independent of network configuration, (14) simplifies to

$$B = \frac{\sum_{k=1}^{N_a} d_{k,j} E[P_j(\boldsymbol{r}) r_j^2]}{(\sum_{k=1}^{N_a} d_{k,j})^2}. \tag{15}$$

Notice that r_j, $j = 1, \ldots, N_a$, are independent and identically distributed. Under the assumption that P_{Net} is one (which becomes reasonably good as the number of sensors N_a increases), (15) further simplifies to

$$B = \frac{1}{N_d} T_2 T_0^{N_d - 1} (1 - T_0)^{N_a - N_d}, \tag{16}$$

where

$$N_d = \sum_{i=1}^{N_a} d_i, \tag{17}$$

$$T_2 = \frac{2}{R^2} \int_0^R r^3 P_d(r) dr, \tag{18}$$

$$T_0 = \frac{2}{R^2} \int_0^R r P_d(r) dr. \tag{19}$$

Since there are $\binom{N_a}{N_d}$ realizations d_j corresponding to N_d detections, we obtain

$$E[\text{MSE}] = \sum_{N_a=1}^{N} \eta_{N_a} \frac{T_2}{T_0} \sum_{N_d=1}^{N_a} \binom{N_a}{N_d} \frac{1}{N_d} T_0^{N_d} (1 - T_0)^{N_a - N_d}. \tag{20}$$

V. Simulations

In the simulations, we use the following sensor model

$$P_d(r) = \exp\left(\log(P_{FA}) \left(1 + \frac{SNR_1}{r^2}\right)^{-1}\right), \tag{21}$$

where P_{FA} is the false alarm probability, and SNR_1 is the normalized signal-to-noise ratio at a distance of one meter. The model given by equation (21) represents the performance of the energy detector when interrogating Swerling I or II targets for one coherent interval [8]. For all simulations, $P_{FA} = 0.001$, $SNR_1 = 67.6$db, $\lambda = 2$, $\mu = 1$, and the deployment area is 1 km^2. The purpose of the experiments is to validate the bound for connectivity given by (10) and the ability of the formula for the mean-squared error given by (20) to predict the performance of centroid estimator and the maximum likelihood estimator (MLE) algorithm proposed in [9].

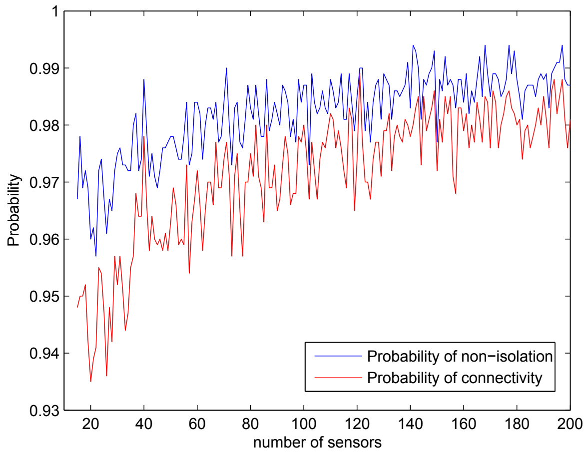

Fig. 1. Probability of non-isolation and connectivity for sensor networks ranging from ten to one-hundred sensors

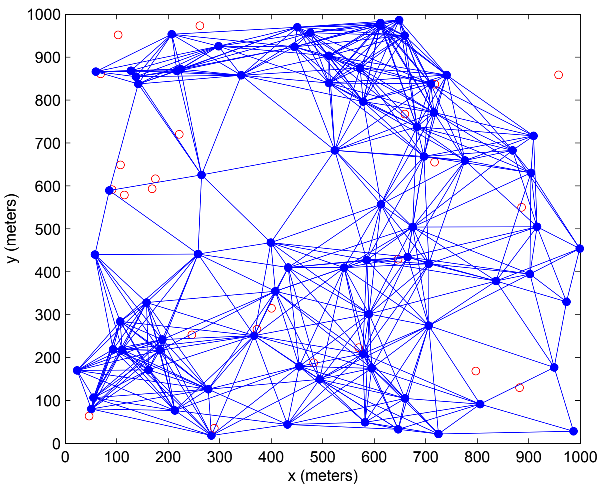

Fig. 2. Network consisting of one-hundred sensors; the sensors are indicated by circles; the blue circles are the active sensors and the red circles with empty interior are the idle sensors. An edge occurs between two active sensors if their distance is smaller than 269 meters.

We ran one-thousand Monte-Carlo simulations fixing $\epsilon = 0.03$ and empirically calculated that the smallest constant c in (10) that guarantees the network is connected with probability larger than 97% is $c \approx 1.5166$. Figure 1 reports the estimated probability of connectivity and non-isolation as a function of the number of sensors. It appears from Figure 1 that the transmitting range given by equation (10) is overly pessimistic as the number of sensors increases and a smaller range may be chosen to guarantee a 97% probability. A snapshot of the emulated sensor network consisting of one-hundred sensors and critical transmitting range equal to $1.5166\sqrt{\frac{A}{\pi}} = 269$ meters is reported in Figure 2.

To evaluate the localization performance, we generated one-thousand configurations of nodes uniformly distributed in the squared area, each of them switching state according to random energy management protocol. Figure 3 shows the square root of the average mean-squared error, *i.e.*, rms, as a function of the number of sensors in the network. To achieve a connectivity probability of greater than 97%, the transmission range is set to 269 meters. For each configuration, we computed the rms error when the target is located at the center of the area. We then averaged the error over the one-thousand configurations. When the active portion of the network happened to be disconnected, we computed the estimate by two methods: 1) using all active nodes and 2) using the largest connected component of active nodes. For all runs, the number of connected components in the graph was at most two, which occurred with probability smaller than 5%.

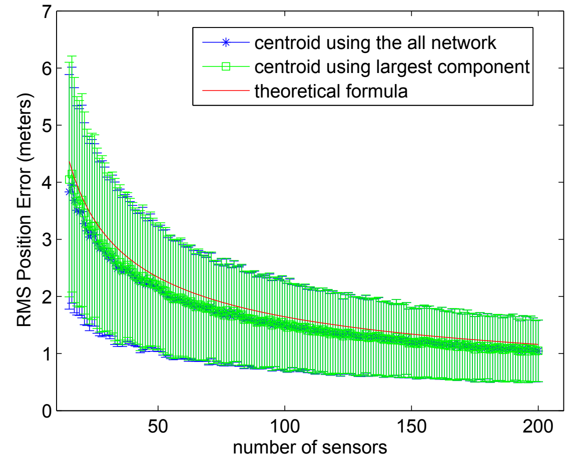

Fig. 3. Comparison between the rms error obtained from simulation and the theoretical formula. The sensor configuration ranges from N=10 to N=200 sensors.

The error bars represent the spread about the mean values. It appears from the graph that the error is predominantly inversely proportional to the number of sensors in the network. Furthermore, the error curve obtained when all active detectors in the network contribute to the estimate resembles very closely to the one obtained when only the detections coming from the active sensors in the largest connected component contribute to the estimate. Figure 3 also compares the rms error obtained from the numerical simulations with the theoretical expression (20) evaluated using the same parameters and choosing R in such a way that the area of the circle equals the area of the square deployment region, *i.e.*, $R = \frac{1}{\sqrt{\pi}} = 564.2$ meters. It appears clearly from the graph that (20) agrees very nicely with the simulated results.

Next, we compare by means of one-thousand Monte-Carlo runs the rms errors of the centroid and MLE. For execution of the MLE, the likelihood that the target is located at (x, y)

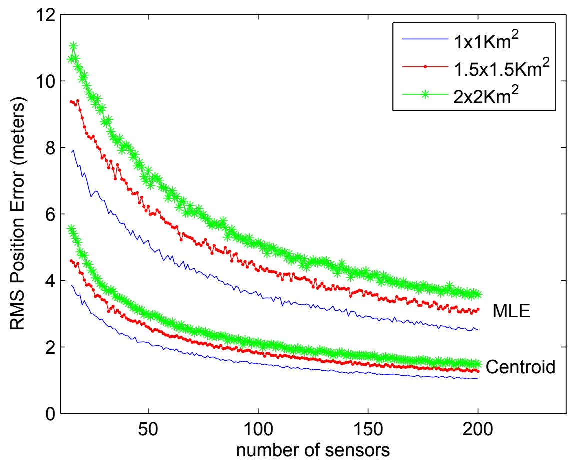

Fig. 4. Comparison between the rms error of the centroid and MLE. The sensor configuration ranges from N=10 to N=200 sensors.

is defined is

$$P(\boldsymbol{d}, x, y) = \prod_{i=1}^{N_a} P_d(r_i(x,y))^{d_i}(1 - P_d(r_i(x,y)))^{1-d_i} \quad (22)$$

where $r_i(x,y) = \sqrt{(x-x_i)^2 + (y-y_i)^2}$. We use the same parameters as in the previous experiment and again assume that the target is located at the center of the area. We consider areas of size 1 km^2, 1.5×1.5 km^2 and 2×2 km^2. For this experiment, we computed the estimate using the detection of all active sensors in the network. As shown earlier, the rms obtained using only the subset of sensors in the largest connected component would have been approximately the same. The results of the experiments are reported in Figure 4. Clearly, the centroid demonstrate smaller error when the target is located at the center of the area.

A natural question is what happens as the target move closer to the edge of the area. In the next set of experiments, we generated one-hundred random configurations, each consisting of one-hundred sensors uniformly deployed within an area of 1 km^2. We evaluated the rms error for both the centroid and MLE. The graphs in Figure 5 indicate the contours of the error surface. The centroid exhibits better performance than the MLE near the center of the surveillance region. However, the centroid quickly degrades as the target moves closer to the edge of the surveillance region. On the other hand, the MLE maintains performance near the edge, where it clearly outperforms the centroid estimator.

VI. Conclusions and Future Work

The simulations presented in this paper clearly demonstrate the efficacy of (20) in predicting the rms error of the centroid estimator algorithm. Such an expression can be used along with (10) to design a sensor network with a simple random management protocol where each active sensor can communicate with any other active sensor in the network with very large

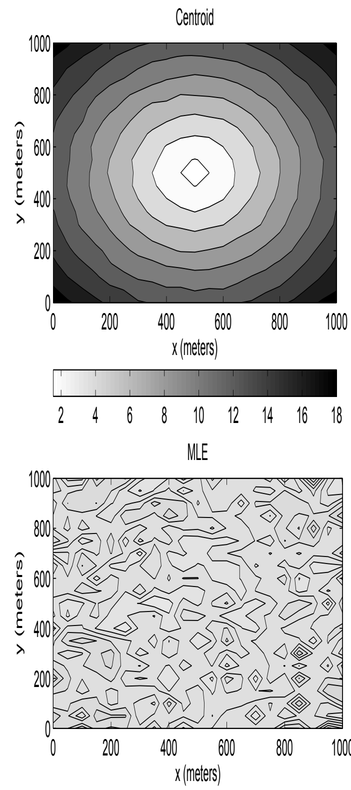

Fig. 5. Error surface for the rms error of the centroid and MLE. The number of sensors is N=100.

probability and achieve a desired level of estimation error. In a future work, we plan to determine how to determine λ and μ using a proper battery model in order to optimize the lifetime of the network while achieving a specified level of localization performance. Furthermore, we hope to determine an expression to predict the performance of the MLE method.

References

[1] Z. Feng, S. Jaewon, and R. James, "Information-driven dynamic sensor collaboration for target tracking," *IEEE Signal Processing Mag.*, vol. 2, no. 19, pp. 61–72, 2002.

[2] F. Zhao, J. Liu, J. Liu, L. Guibas, and J. Reich, "Collaborative signal and information processing: An information directed approach," *Proc. IEEE*, vol. 8, no. 91, pp. 1199–1209, 2003.

[3] E. Shih, S.-H. Cho, N. Ickes, R. Min, A. Sinha, A. Wang, and A. Chandrakasan, "Physical layer driven algorithm and protocol design for energy-efficient wireless sensor networks," in *Proceedings of the 7th Annual ACM/IEEE International Conference on Mobile Computing and Networking, MobiCom,Italy*. ACM/IEEE, 2001, pp. 272–287.

[4] S. Ross, *Stochastic Processes*, 2nd ed. Wiley Series in Probability and Mathematical Statistics, 1996.

[5] M. Penrose, *Random geometric graphs*. Oxford University Press, 2003.

[6] P. Gupta and P. Kumar, "The capacity of wireless networks," *IEEE Trans. Inform. Theory*, vol. 45, no. 2, pp. 338–404, 2000.

[7] D. West, *Introduction to Graph Theory*. Prentice Hall of India, 1999.

[8] S. Blackman and R. Popoli, *Design and Analysis of Modern Tracking Systems*. Artech House, 1999.

[9] A. Rodriguez, M. Lazaro, and L. Tong, "Target local estimation in sensor networks using range information," in *IEEE Sensor Array and Multichannel Signal Processing workshop*, 2004.