Confidence-interval Based Sensing Quality Evaluation for Mobile Sensor Networks

Chaocan Xiang[1], Panlong Yang[1], Bin Yu[2], Chang Tian [1],
Hong He[3], Jibin Guo[1], Yan Xiong[4]
[1] PLA University of Science and Technology, China
[2]PLA Unit 95028, China
[3]Logistic Engineering University, China
[4]University of Science and Technology of China, China
xiang.chaocan@gmail.com, panlongyang@gmail.com, 13407117375@163.com,
tianchang163@163.com, hehong116@sina.com, xcc123dbs@163.com,
yxiong@ustc.edu.cn

ABSTRACT

Sensing quality evaluation is fundamentally important for mobile sensor network. However, due to the inherent sensing uncertainty in mobile sensor networks and the unavailability of the ground truth, achieving effective and accurate evaluation on sensing quality is extremely challenging. In this paper, we propose a confidence-interval based sensing quality evaluation method, leveraging the Fisher information and the asymptotic normality property of maximum likelihood estimation. The simulation results demonstrate our method can evaluate the sensing quality more reasonably and accurately than the status quo method. Further, our evaluation asymptotically approaches to the ground truth with the stepwise movements of sensors.

1. INTRODUCTION

Recent years, mobile sensor networks have been widely investigated, especially in pollution source localization in aquatic environments, so as to adapt to the dynamic diffusion of pollution sources and reduce the sensing cost[1, 2, 3, 4]. Most of current methods focus on how to improve the sensing quality of mobile sensor network, *e.g.* how to schedule the sensor movements to improve the localization accuracy[2]. However, few studies concentrate on evaluating the sensing quality of mobile sensor network. It is noted that, in this paper, the sensing quality indicates the source localization accuracy in the typical application of pollution source localization.

Sensing quality evaluation is greatly important for mobile sensor network. For example, it could be an effective feedback for making efficient moving policies, so as to improve the sensing quality iteratively[5]. Nevertheless, it is challenging to evaluate the sensing quality of sensor network for two reasons. First, as the ground truth values for reference are unknown in prior, sensing errors turn to be unmeasurable. Even worse, as the sensors' measurements are bearing uncertain noise, it is greatly difficult to evaluate the sensing quality. Recently, Wang *et al.* [5] propose a profiling metric, intuitively using the sum of reciprocals of the Cramer-Rao Bound (CRB) [6] to evaluate the sensing quality of mobile sensor network. Nevertheless, this metric fails to evaluate the sensing quality accurately, as its evaluation values are not in the same scale as the ground truth.

To address these problems, in this paper, we propose a confidence-interval based sensing quality evaluation method for mobile sensor network. We leverage the asymptotic normality property of Maximum Likelihood Estimation (MLE)[7] and the Fisher information [8] to compute the confidence interval radius of the pollution source, which evaluates the sensing quality accurately. Also, we conduct simulations to demonstrate that, ours can evaluate the sensing quality more accurately and reasonably than the profiling metric[5]. Further more, our evaluation is asymptotically approaching to the ground truth with the sensor movements.

MSCC'14, August 11, 2014, Philadelphia, PA, USA.

http://dx.doi.org/10.1145/2633675.2633679.

2. SYSTEM MODEL AND PROBLEM FORMULATION

2.1 Pollution Diffusion Model and Sensor Measurement Model

A pollution source diffuses with a constant rate in a 2-dimension static aquatic environment. Let $c(x, y, t)$ denote the pollution concentration at the location (x, y) with the diffusion time t. Hence, according to the Fick's law[9], the pollution concentration $c(x, y, t)$ satisfies the following partial equation:

$$\frac{\partial c(x,y,t)}{\partial t} = \lambda \cdot \frac{\partial^2 c(x,y,t)}{\partial x^2} + \lambda \cdot \frac{\partial^2 c(x,y,t)}{\partial y^2} \tag{1}$$

where λ is a constant diffusion coefficient, related to the species of the solvent and diffuser.

Let A and (x_0, y_0) denote the total pollution substance and the location of the pollution source respectively. According to Equ.1, we have:

$$c(x,y,t) = \frac{A}{4\pi\lambda t} \exp(-\frac{d^2(x,y)}{4\lambda t}) \tag{2}$$

where $d(x, y)$ denotes the distance between the measurement location (x, y) and the source location (x_0, y_0). This diffusion model in Equ.2 has been widely used in numbers of approaches[9, 2, 3, 4] and validated by the experiments in [2].

The sensor measurements are subjected to noise, due to their limited sensing capability and the environment noise.

For simplicity, we assume the measurement noise of each sensor follows an $i.i.d.$ Gaussian distribution. Specifically, let $z(x,y,t)$ denote the sensor measurement at the location (x,y) with the diffusion time t. Thus, we have:

$$z(x,y,t) \sim \mathcal{N}(c(x,y,t)+u, \sigma^2) \tag{3}$$

where u denotes the bias of the noise experienced by the sensor, and σ^2 denotes its variance. These two parameters can be easily available. For example, they can be specified by the sensor manufacturer or measured before the real deployment. The above measurement model has been widely used for most of chemical sensors[10, 3, 1, 2].

2.2 Sensing Quality Evaluation Problem of Sensor Network

Let N denote the number of sensors in mobile sensor network. Besides the pollution concentration measurement, the sensor can measure its location, $e.g.$ GPS[1, 5, 4]. Let Z denote the measurement set of mobile sensor network. Thus, $Z = \{(x_k, y_k, z_k), k = 1, \ldots N\}$, where z_k and (x_k, y_k) denote the concentration measurement as well as its location of the k-th sensor respectively.

As the pollution source appears unpredictably and abruptly, the parameters of the pollution source are known, including the total pollution substance A, the diffusion time t and the position (x_0, y_0). Let Θ denote the unknown parameters, $i.e.$ $\Theta = \{A, t, x_0, y_0\}$. In most of studies[2, 10, 3, 9], the head uses the MLE method to estimate the unknown parameters of the pollution source, based on the measurement set Z. According to Equ.2 and 3, we derive the likelihood function of the sensor measurements as:

$$L(Z/\Theta) = \ln\big(\prod_{k=1}^{N} f_k(z/\Theta)\big) \tag{4}$$

$$f_k(z/\Theta) = \frac{1}{\sqrt{2\pi}\sigma} \exp \frac{(z - \frac{A}{4\pi\lambda t} \cdot h_k)^2}{-2\sigma^2} \tag{5}$$

where $f_k(z/\Theta)$ denotes the probability density function of the k-th sensor's measurement, and $h_k = \exp(\frac{d^2(x_k,y_k)}{-4\lambda t})$.

Using the MLE method, we can easily get the parameter estimation $\hat{\Theta}$ of the pollution source, $i.e.$ $\hat{\Theta} = \{\hat{A}, \hat{t}, \hat{x_0}, \hat{y_0}\}$. However, what is the sensing quality of mobile sensor network for localizing the pollution source? In other words, what are the localization errors? In this paper, we focus on studying how to evaluate the sensing quality of sensor network without the ground truth localization of the pollution source.

Sensing Quality Evaluation Problem of Sensor Network: Given a pollution source with the unknown parameters Θ, only based on the measurement set Z of the sensor network, find a metric of sensing quality evaluation, minimizing the distance between this metric and the actual localization error ϵ of the pollution source, $i.e.$ $\epsilon = \| (\hat{x_0}, \hat{y_0}) - (x_0, y_0) \|_2$.

3. CONFIDENCE-INTERVAL BASED SENSING QUALITY EVALUATION METHOD

To solve the sensing quality evaluation problem, we propose a confidence-interval based sensing quality evaluation method for mobile sensor network. Specifically, we leverage the confidence interval radius as a metric, evaluating the sensing quality of mobile sensor network. Moreover, we use the Fisher information and the asymptotic normality property of MLE to compute the confidence interval radius. In the following, we first derive the fisher information of the sensor measurements. And then, based upon this fisher information, we use the asymptotic normality property of MLE to derive the confidence interval radius. Finally, based on these derivation results, we propose the confidence-interval based sensing quality evaluation algorithm.

3.1 Derivation of Fisher Information

The fisher information is a way of evaluating the amount of information that the sensor measurements carry about the parameters of pollution source, on which the probability of these measurements are dependent[8]. Let $I(\Theta)$ denote the fisher information. As Θ includes four unknown parameters, $I(\Theta)$ is a 4×4 matrix, represented by $(J(i,j))|_{4\times4}$. According to the definition of Fisher information[8], we have:

$$J(i,j) = -E(\frac{\partial^2 L(Z,\Theta)}{\partial\theta_i \partial\theta_j}) \quad i,j = 1,2,3,4 \tag{6}$$

where $\theta_i(\theta_j)$ denotes the unknown parameter of the pollution source, $i.e.$ $\theta_i, \theta_j \in \Theta$.

We firstly derive the Fisher information for the measurement of a single sensor. Let $I_k(\Theta)$ denote the Fisher information for the k-th sensor,$(k = 1, 2 \ldots N)$. According to Equ.4 and 5, the likelihood function of the k-th sensor measurement is:

$$L_k(Z/\Theta) = \ln(f_k(z/\Theta)) = \ln(\frac{1}{\sqrt{2\pi}\sigma}) - \frac{(z - \frac{Ah_k}{4\pi\lambda t})^2}{2\sigma^2} \tag{7}$$

In order to facilitate derivation, we substitute the variable β for A, where $\beta = \frac{A}{4\pi\lambda t}$. Then, the unknown parameter vector Θ includes β, t, x_0 and y_0, denoted by $\theta_1, \theta_2, \theta_3$ and θ_4 respectively. As the pollution source location only involves the parameters x_0 and y_0, this substituting process does not affect the derivation results.

$I_k(\Theta)$ is also a 4×4 matrix, represented by $(J_k(i,j))|_{4\times4}$. Substituting Equ.7 into Equ.6, we have:

$$J_k(i,j) = \frac{1}{\sigma^2}(\frac{\partial(\beta h_k)}{\partial\theta_i}\frac{\partial(\beta h_k)}{\partial\theta_j} + E(\beta h_k - z)\frac{\partial^2(\beta h_k)}{\partial\theta_i\partial\theta_j}) \tag{8}$$

As $E(z) = \beta h_k$, $E(\beta h_k - z) = 0$. Then, we have:

$$J_k(i,j) = \frac{1}{\sigma^2} \cdot \frac{\partial(\beta h_k)}{\partial\theta_i} \cdot \frac{\partial(\beta h_k)}{\partial\theta_j} \tag{9}$$

According to Equ.5, h_k involves three unknown parameters, $i.e.$ t, x_0 and y_0, while β is independent of them. Thus, we have:

$$\frac{\partial(\beta h_k)}{\partial\theta_i} = \begin{cases} \beta \cdot \frac{\partial h_k}{\partial\theta_i} & \text{when} \quad \theta_i \neq \beta \\ h_k & \text{when} \quad \theta_i = \beta \end{cases} \tag{10}$$

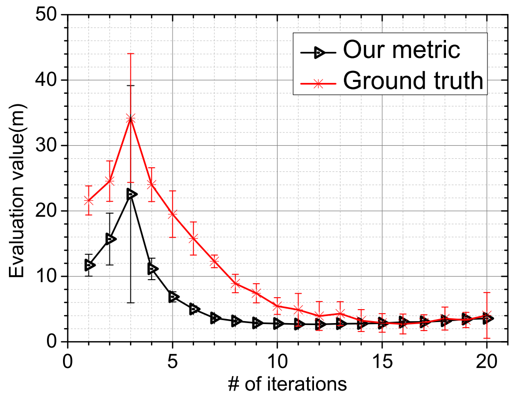
According to Equ.9 and 10, we derive $J_k(i,j)$ as: $(k = 1, 2 \ldots N)$

$$J_k(i,j) = \frac{1}{\sigma^2} \cdot \vartheta_i^k \cdot \vartheta_j^k \qquad i,j = 1,2,3,4 \tag{11}$$

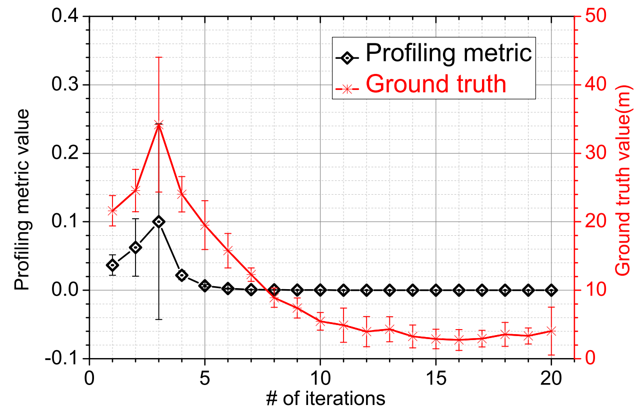
$$\vartheta_i^k = \begin{cases} \frac{A}{4\pi\lambda t} \cdot \frac{\partial h_k}{\partial\theta_i} & i = 2,3,4 \\ h_k & i = 1 \end{cases} \tag{12}$$

According to the additive property of Fisher information[8], we have $I(\Theta) = \sum_{k=1}^{N} I_k(\Theta)$. Thus, according to Equ.11 and 12, we derive $J(i,j)$ as: $(i,j = 1,2,3,4)$

$$J(i,j) = \sum_{k=1}^{N} \frac{1}{\sigma^2} \cdot \vartheta_i^k \cdot \vartheta_j^k \quad i,j = 1,2,3,4 \tag{13}$$



((a)) Our metric verse Ground truth



((b)) Profiling metric verse Ground truth

Figure 1: Comparison between our metric and the profiling metric, under the benchmark of the ground truth.

3.2 Derivation of Confidence Interval Radius

Maximum likelihood estimation has an asymptotic normality property, *i.e.* the distribution of MLE tends to the Gaussian distribution when the scale of the measurements is large[7]. Specifically, the distribution of the estimation error is:

$$\sqrt{N}(\hat{\Theta} - \Theta) \xrightarrow{d} \mathcal{N}(0, I^{-1}(\hat{\Theta})) \tag{14}$$

where $I^{-1}(\Theta)$ is the inverse of Fisher information $I(\Theta)$, which is computed by Equ.13 in the previous section.

According to Equ.13 and 14, the confidence interval of the parameter θ_i ($\forall \theta_i \in \Theta$) with the confidence probability p is derived as:

$$\theta_i \in \left[\hat{\theta}_i \pm \frac{\rho_p}{\sqrt[4]{N}}\sqrt{I^{-1}_{i,i}(\hat{\Theta})}\right], i = 1,2,3,4 \tag{15}$$

where ρ_p denotes the standard score of the confidence probability p. $\hat{\theta}_i$ denotes the estimation value for the unknown parameter θ_i. $I^{-1}_{i,i}(\Theta)$ denotes the element in the i^{th} row and i^{th} column of the matrix $I^{-1}(\Theta)$.

As the position of the pollution source involves two parameters, *i.e.* x_0 and y_0, according to Equ.15, the confidence interval of the source's position is:

$$\{(x,y)\big||x - \hat{x}_0| \le l_x, |y - \hat{y}_0| \le l_y\} \tag{16}$$

where $l_x = \frac{\rho_p}{\sqrt[4]{N}}\sqrt{I^{-1}_{3,3}(\hat{\Theta})}$, and $l_y = \frac{\rho_p}{\sqrt[4]{N}}\sqrt{I^{-1}_{4,4}(\hat{\Theta})}$. $\hat{x}_0$ and $\hat{y}_0$ denote the estimation value of x_0 and y_0 respectively.

According to Equ.16, the confidence interval of the source's position is a rectangle centered at the estimation position of the source. For ease of sensing quality evaluation, we employ a disc of this confidence interval defined by Equ.17, as the confidence interval of the source approximately.

$$\{(x,y)\big|(x - \hat{x}_0)^2 + (y - \hat{y}_0)^2 \le \gamma^2\} \tag{17}$$

where γ is the confidence interval radius as:

$$\gamma = \frac{\rho_p}{\sqrt[4]{N}}\sqrt{I^{-1}_{3,3}(\hat{\Theta}) + I^{-1}_{4,4}(\hat{\Theta})} \tag{18}$$

where ρ_p is a constant determined by the confidence probability. $I^{-1}_{3,3}(\hat{\Theta})$ and $I^{-1}_{4,4}(\hat{\Theta})$ are the CRB of x_0 and y_0 respectively.

3.3 Confidence-interval Based Sensing Quality Evaluation Algorithm

We use the confidence interval radius γ as a metric, evaluating the localization error of the pollution source for the following reasons. The actual location of the pollution source is within the confidence interval centered at its estimation location with a probability. If the probability is very large and close to 1, the actual location is nearly with this confidence interval. As a result, the confidence interval radius can accurately evaluate the error between the actual location and the estimation one of the pollution source. This also conforms to the following intuition. According to Equ.18, the confidence interval radius is determined by the CRB of x_0 and y_0, which approximates the estimation error of x_0 and y_0 respectively[6]. Thus, it is reasonable to use the confidence interval radius as the evaluation metric. We will validate the accuracy of this metric by experiments in the following section.

In summary, based on the above derivations and analysis, we propose the confidence-interval based sensing quality evaluation algorithm (as Algorithm 1).

Algorithm 1: Confidence-interval Based Sensing Quality Evaluation Algorithm

---

Input:
Measurement set of sensors: Z;
Output:
Evaluation value of sensing quality: γ;

1: Given the measurement set of sensors Z, use the MLE method to compute the parameter estimations $\hat{\Theta}$ of the pollution source, according to Equ.4 and 5.
2: Given the parameter estimations $\hat{\Theta}$, calculate the Fisher information matrix $I(\hat{\Theta})$, according to Equ.12 and 13.
3: Given the Fisher information matrix $I(\hat{\Theta})$, calculate the confidence interval radius γ, according to Equ.18.
4: **return** the evaluation value of sensing quality γ;

---

4. EXPERIMENTAL RESULTS

4.1 Simulation Methodology and Settings

Like most of related works[1, 2, 3], we use the following iterative process of mobile sensor network to localize the pollution source in the aquatic environment. In each iteration,

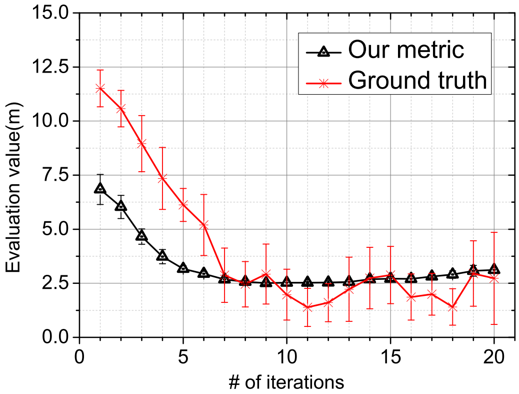


Figure 2: Comparison between our metric and the ground truth when the beginning time of the sensor movements is 1500s.

firstly, all the sensors make measurements, based on which the head estimates the parameters of the pollution source, using the MLE method. Note that, in this simulation, we exploit the Nelder-Mead's algorithm[11] to solve the nonlinear optimization of the likelihood function. And then, the sensors move towards to the estimation location of the pollution source for a constant distance. We evaluate the sensing quality of mobile sensor network in each iteration. We compare our evaluation metric with the profiling metric in [5], which is the most accurate evaluation metric. The compared benchmark is the ground truth of the localization error, which is the error between the actual location of the pollution source and the estimation one.

The simulation settings are as follows. The parameters of the pollution source are that, $A = 0.7 \times 10^6 cm^3, \lambda = 0.5m^2/s$. They are consistent with the real field experiment reports[12]. The standard deviation of the measurement noise for each sensor $\sigma = 1cm^3/m^2$. 20 sensors are randomly deployed in a $150m \times 150m$ square region, and the pollution source is centered at this region, *i.e.* $x_0 = y_0 = 0$. The moving speed of each sensor $\nu = 2.5m/min$, according to the speed of the robotic fish[13]. The sensors start moving after the pollution source diffuses for 1000s. The number of iterations is set to 20. The simulation programs are written in Matlab. All the simulations are executed for 100 times and we get average values.

4.2 Performance Evaluation

We make simulations to analyze the accuracy of our evaluation metric (*i.e.* the confidence interval radius γ), compared with the profiling metric. As shown in Fig.1(a), our metric values are in the same scale as the ground truth. The error between the evaluation value of our metric and the ground truth is small. The average error is 4.76m. Further more, the errors gradually decrease with the iterations. Surprisingly, after the 15-th iteration, our metric values nearly are the same as the ground truth. In addition, the evaluation values of our metric are always below the ground truth. The reason is that, our metric is derived from the Cramer-Rao Bound, which is the lower bound of the estimation errors. In contrast, as shown in Fig.1(b), the profiling metric values only approximately reflect the changing tendency of the ground truth. However, this metric cannot evaluate the sensing quality accurately, as their values are not in the same scale as the ground truth. Thus, our metric evaluates the sensing quality more accurately than the profiling metric. Also, we change the start time of the sensor movements, and analyze the accuracy of our metric again. As shown in Fig.2, similarly, our metric still evaluates the sensing quality accurately. The largest error between our metric value and the ground truth is 4.6m, the average one is 1.5m.

5. CONCLUSION

In this paper, leveraging the asymptotic normality property of MLE and the Fisher information, we propose a confidence-interval based sensing quality evaluation method for mobile sensor network. Simulation results show that, our method can evaluate the sensing quality of mobile sensor network more accurately and reasonably than the current best method. Further, our metric values are gradually close to the ground truth in pace with the sensor movements.

6. ACKNOWLEDGMENTS

This research is partially supported by NSF China under Grants No. 61003277, 61232018, 61272487, 61170216, 61172062, 61228202, 61273210, NSF CNS-0832120, NSF CNS-1035894, NSF EECS-1247944, NSF ECCS-1343306, China 973 Program under Grants No. 2009CB320400, 2010CB328100, 2010CB334707, 2011CB302705.

7. REFERENCES

[1] Aleksandar Jeremicąă and Arye Nehorai. Landmine detection and localization using chemical sensor array processing. *IEEE Transactions on Signal Processing*, 48(5):1295–1305, 2000.

[2] Yu Wang, Rui Tan, Guoliang Xing, and Xiaobo Tan. Accuracy-aware aquatic diffusion process profiling using mobile sensor networks. In *IPSN*, pages 1–12, 2012.

[3] Boaz Porat and Arye Nehorai. Localizing vapor-emitting sources by moving sensors. *IEEE Transactions on Signal Processing*, 44(4):1018 –1021, 1996.

[4] Saravanan Vijayakumaran, Yoav Levinbook, and Tan F. Wong. Maximum likelihood localization of a diffusive point source using binary observations. *IEEE Transactions on Signal Processing*, 55(2):665–676, 2007.

[5] Yu Wang, Rui Tan, Guoliang Xing, Jianxun Wang, and Xiaobo Tan. Profiling aquatic diffusion process using robotic sensor networks. *IEEE Transactions on Mobile Computing*, PP(99), 2013.

[6] Petre Stoica and Arye Nehorai. Music, maximum likelihood, and cramer-rao bound. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 37(5):720–741, 1989.

[7] Peter J. Bickel and Bo Li. Mathematical statistics. *Test*, 15, 1977.

[8] Jorma J Rissanen. Fisher information and stochastic complexity. *IEEE Transactions on Information Theory*, 42(1):40–47, 1996.

[9] Tong Zhao and Arye Nehorai. Detecting and estimating biochemical dispersion of a moving source in a semi infinite medium. *IEEE Transactions on Signal Processing*, 54(6):2213–2225, 2006.

[10] Arye Nehorai, Boaz Porat, and Eytan Paldi. Detection and localization of vapor-emitting sources. *IEEE Transaction on Signal Processing*, 43(1):243–253, 1995.

[11] J. A. Nelder and R. Mead. A simplex method for function minimization. *Computer Journal*, 7(4):308–313, 1965.

[12] Alan J. Elliott. Shear diffusion and the spread of oil in the surface layers of the north sea. *Ocean Dynamics*, 39(3):113–137, 1986.

[13] Xiaobo Tan. Autonomous robotic fish as mobile sensor platforms: Challenges and potential solutions. *Marine Technology Society Journal*, 45(4):31–40, 2011.