Parameter Estimation in Wireless Sensor Networks with Normally Distributed Sensor Gains

Zhenxing Luo

Abstract—Wireless sensor networks (WSN) have attracted significant attention recently. The distributed estimation problem is an important research topic in WSNs. In the distributed estimation problem, the fusion center estimates an unknown parameter based on information gathered from sensors. Usually, it is assumed that sensors have identical gains. However, this may not be true due to manufacture errors or environmental influence. In this paper, we assume sensor gains follow normal distribution and present a maximum likelihood estimation (MLE) approach for distributed estimation in WSNs with normally distributed sensor gains. Moreover, the Cramer-Rao lower bound (CRLB) corresponding to this MLE approach is also derived. Simulation results showed that the root mean square (RMS) estimation errors given by this MLE approach were close to the CRLB if the variance of the sensor gains is small. If the variance of the sensor gains was large, the RMS estimation errors were not close to the CRLB.

Keywords—Distributed estimation, maximum likelihood estimation, Gaussian distribution, wireless sensor networks.

I. INTRODUCTION

Due to a vast number of applications, wireless sensor networks (WSNs) have gained significant attention [1]-[13]. Usually, a WSN consists of many sensors which will send gathered information to a fusion center [14]. After acquiring information from sensors, the fusion center can perform many tasks such as tracking, detection and distributed estimation [14]. Distributed estimation is of particular interest because it serves as a cornerstone for many other tasks.

In the distributed estimation problem, the fusion center tries to estimate unknown parameters based on information from sensors. To estimate one unknown parameter, a maximum likelihood estimation (MLE) approach was presented in [15][16]. Then, the MLE approach was extended to consider imperfect communication channels in [17]. However, in [15]-[17], the sensor gains were assumed to be identical while in many applications, sensor gains are not identical due to manufacture errors or environmental influences.

In this paper, we assume that sensor gains follow a Gaussian distribution. This assumption is valid when the sensor gains are determined by a sum of many factors. If this is true, the sensor gains will follow a Gaussian distribution [18]. Then, the distribution information of sensor gains is incorporated into the MLE framework to address heterogeneous sensor gains.

The main contribution of this paper is a MLE approach for distributed estimation in WSNs with normally distributed sensor gains. Moreover, the Cramer-Rao lower bound (CRLB) corresponding to this MLE approach is also derived. Simulation results showed that when the variance of the Gaussian distribution is small, root square mean (RMS) estimation errors were close to the CRLB. When the variance of the Gaussian distribution was large, RMS estimation errors deviated from the CRLB.

This paper is organized in the following way. Section II presents the MLE approach for distributed estimation in WSNs with normally distributed sensor gains. Section III presents the CRLB, followed by simulation setup in Section IV. Section V provides results and analysis. Finally, Section VI delivers concluding remarks.

II. DISTRIBUTED ESTIMATION METHOD IN WIRELESS SENSOR NETWORKS WITH NORMALLY DISTRIBUTED SENSOR GAINS

A WSN consists of a fusion center and many sensors (Figure 1). After sensors measure the parameter \( \theta \), sensors will quantize the measurement according to a set of pre-determined thresholds \( \gamma \). Sending quantized data instead of analogy data to the fusion center can save communication bandwidth and sensor energy [14]. The fusion center can use a MLE approach to estimate the unknown parameter \( \theta \) based on information from sensors. Now, this MLE approach is discussed in details.

![Figure 1: Distributed estimation system diagram](image)

Following the setup in [11][14]-[16], we use \( N \) sensors to estimate the unknown parameter \( \theta \). However, sensor gains are not identical and follow the normal distribution. The gain of the sensor \( i \) is \( G_i \), and \( G_i \) follows Normal distribution with mean \( u \) and variance \( \sigma_i^2 \):

\[
f(G_i) = \frac{1}{\sqrt{2\pi}\sigma_i} e^{-\frac{(G_i-u)^2}{2\sigma_i^2}}, \quad G_i \in (-\infty, +\infty).
\]

The \( i \)th sensor receives the signal from \( \theta \) and this signal can be denoted as \( a_i \), which is defined as

\[
a_i = G_i \theta.
\]

Because of the presence of noises, the signal arrives at the \( i \)th sensor is \( s_i \), which can be defined as
Parameter Estimation in Wireless Sensor Networks with Normally Distributed Sensor Gains

The noise \( w_i \) in (3) is a Gaussian noise with zero mean and variance \( \sigma_i^2 \).

The probability density function (PDF) of the sum of two Normally distributed random variables also follows Normal distribution [18]. Therefore, we can have

\[
 f(s) = \frac{1}{\sqrt{2\pi} \sigma_1 \sigma_2} e^{-\frac{(s-m)^2}{2\sigma_1^2 + \sigma_2^2}},
\]

After a sensor acquires the measurement, the sensor quantizes the measurement into a decision \( m \) according to a set of pre-determined thresholds \( \gamma_i \),

\[
 \gamma_i = [\gamma_{i0}, \gamma_{i1}, \ldots, \gamma_{iL}].
\]

The quantization process can be expressed by

\[
 m = \begin{cases} 
 0 & -\infty < s < \gamma_{i1} \\
 1 & \gamma_{i1} < s < \gamma_{i2} \\
 \vdots & \vdots \\
 L-2 & \gamma_{i(L-2)} < s < \gamma_{i(L-1)} \\
 L-1 & \gamma_{i(L-1)} < s < \infty 
\end{cases}
\]

For a specific \( \theta \), the probability that \( m_i \) is equal to \( 1 \) is

\[
 p_d(\gamma_i, \theta) = R(\gamma_i) - R(\gamma_{i(L-1)}) = \int_{\gamma_{i1}}^{\gamma_{i2}} f(s) ds.
\]

Then, sensors will send the decision vector \( \mathbf{M} = [m_1, m_2, \ldots, m_{N-1}, m_N] \) to the fusion center, and the fusion center estimates \( \theta \) by finding the \( \theta \) value to maximize

\[
 \ln p(\mathbf{M} | \theta) = \sum_{i=1}^{L} \sum_{j=0}^{L} \delta(m_i - l) \ln p_d(\gamma_i, \theta)
\]

where

\[
 \delta(x) = \begin{cases} 
 1, & x = 0 \\
 0, & x \neq 0 
\end{cases}
\]

The maximum likelihood estimator can be expressed as

\[
 \hat{\theta} = \max_{\theta} \ln p(\mathbf{M} | \theta).
\]

If an unbiased estimate of \( \theta \) exists, the CRLB can be calculated by

\[
 \mathbf{J} = -E \left[ \nabla_{\theta} \nabla_{\theta}^T \ln p(\mathbf{M} | \theta) \right].
\]

III. PERFORMANCE EVALUATION-CRAMER-RAO LOWER BOUND

If the estimation result of (12) is unbiased, the \( \mathbf{J} \) matrix can be derived by an approach similar to that in [6][11][14], which is

\[
 \frac{\partial^2 \ln p(\mathbf{M} | \theta)}{\partial \theta^2} = \sum_{i=1}^{L} \sum_{j=0}^{L} \left[ \frac{\delta(m_i - l)}{p_d(\gamma_i, \theta)} \left( \frac{\partial p_d(\gamma_i, \theta)}{\partial \theta} \right) \right]^{2}.
\]

We can use \( E[\delta(m_i - l)] = p_d(\gamma_i, \theta) \) to simplify (14). Then, (14) can be expressed as

\[
 E \left[ \frac{\partial^2 p(\mathbf{M} | \theta)}{\partial \theta^2} \right] = \sum_{i=1}^{L} \sum_{j=0}^{L} \frac{1}{p_d(\gamma_i, \theta)} \left( \frac{\partial^2 p_d(\gamma_i, \theta)}{\partial \theta^2} \right)^2.
\]
As for the effect of the $\sigma_1$ value on RMS estimation errors and the CRLB, we can see that when $\sigma_1$ value was small, RMS estimation errors and the CRLB were low, and the RMS estimation errors were close to the CRLB (Figure 1). When $\sigma_1$ value was large, RMS estimation errors and CRLB were also large, and the RMS estimation errors were not close to the CRLB. The reason is that if the variance of sensor gains is too large, which means sensor gains vary dramatically from one to another, the MLE approach cannot accommodate so different sensor gains and estimation performance will suffer.

As for the effect of $\gamma$ on RMS estimation errors and CRLB, we can see that when $\gamma$ was from 10 to 13, the RMS estimation errors were close to the CRLB. When $\gamma$ was greater than 13, RMS errors deviated from the CRLB (Figure 3). This highlights the importance of the $\gamma$ value. The reason is that if large $\gamma$ value is used, not enough sensors will send 1s to the fusion center, the estimation performance will suffer. If low $\gamma$ value is used, too many sensors will send 1s to the fusion center, estimation performance will also suffer.

**Table 1: NEES values for different $\sigma_1$ values ( $\sigma_2 = 0.2$, $u = 0.1$, $\theta = 5$, $\gamma = 5$ and 100 runs)**

<table>
<thead>
<tr>
<th>$\sigma_1$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>NEES</td>
<td>0.9654</td>
<td>0.9879</td>
<td>0.9994</td>
<td>1.4028</td>
<td>2.7526</td>
</tr>
</tbody>
</table>

**Figure 2: RMS estimation errors compared to the CRLB ( $\sigma_1 = 0.2$, $u = 0.1$, $\theta = 5$, $\gamma = 5$, 100 runs and different $\sigma_1$ value)**

**Figure 3: RMS estimation errors compared to the CRLB ( $\sigma_1 = 1$, $\sigma_2 = 0.2$, $u = 0.1$, $\theta = 5$, 100 runs and different $\gamma$ values)**

**VI. CONCLUSION**

In this paper, a MLE approach for distributed estimation in WSNs with normally distributed sensor gains was presented. This approach can alleviate performance degradation cause by heterogeneous sensor gains, which follow the Gaussian distribution. Simulation results showed that the RMS errors given by this MLE approach were close to the CRLB if the variance of the sensor gain was small. In many applications, the sensor gains are influenced by many variables and the overall effect of these variables is the sum of these variables. If the sensor gains will follow the Gaussian distribution, our MLE approach can be used to alleviate performance degradation.

**REFERENCES**