Fundamental Bounds on Position Estimation Using Proximity Reports

Feng Yin and Yuxin Zhao and Fredrik Gunnarsson
Ericsson Research, Linköping, Sweden
Email: firstname.lastname@ericsson.com

Abstract—There is a big trend nowadays toward indoor proximity report based positioning. A binary valued proximity report can be obtained opportunistically through event-triggering, leading to significantly reduced signaling overhead for wireless communications. In this paper, we aim to derive two types of fundamental lower bound, namely the Cramér-Rao bound and the Barankin bound, on the mean-square-error of any proximity report based position estimator. Using the maximum-likelihood estimator as a representative example, we show that the Barankin bound is potentially much tighter than the Cramér-Rao bound and conclude that the Barankin bound ought be better suited for benchmarking any proximity report based position estimator.

Keywords—Barankin bound, Cramér-Rao bound, mean-square-error, position estimation, proximity report.

I. INTRODUCTION

The rapid development of wireless network infrastructures as well as beaconing techniques motivate new-fashioned, low-cost positioning systems. In this paper, we consider event-triggering received-signal-strength (RSS)-proximity report (later called proximity report for brevity) based positioning. Herein, a proximity report is obtained by thresholding a received-signal-strength-indication (RSSI) measurement. Unlike in the conventional paradigm, where a user equipment (UE) sends the measured RSSI values regularly to the core network, a proximity report is triggered only when the UE’s status changes, for instance when the UE is crossing a border to another service region. Such proximity reporting scheme is beneficial in various ways. Among others, the signaling between the UE and the core network can be significantly reduced by sending less frequently binary value instead of 6-8 bits RSSI value.

The contributions of this paper are as follows. We derive two performance lower bounds, namely the Cramér-Rao bound (CRB) and the Barankin bound (BB), on the position estimation mean-square-error (MSE) in generic form for non-cooperative position estimation using proximity reports. We derive closed form MSE expression of the maximum-likelihood estimator (MLE). We show in the simulations with real setup that the MSE of the MLE is much closer to the BB than to the CRB.

The remainder of this paper is organized as follows: Section II discusses on how our contributions are related to prior work. Section III introduces the signal model. Section IV revisits the proximity report based MLE and derives closed form expression of its MSE. Section V derives the CRB and the BB on the MSE of position estimation using proximity reports. Simulation results are given in Section VI. Finally, Section VII concludes this paper.

Throughout this paper, matrices are presented with uppercase letters and vectors with boldface lowercase letters. The operator $[\cdot]^T$ stands for vector/matrix transpose and $[\cdot]^{-1}$ stands for the inverse of a non-singular square matrix. The operator $\text{tr}(\cdot)$ denotes the trace of a square matrix. The operator $\| \cdot \|$ stands for the Euclidean norm of a vector. The operator $E_p(x)(\cdot)$ stands for statistical expectation with respect to a probability density function $p(x)$. The operator $\ln(\cdot)$ stands for the natural logarithm and $\log(\cdot)$ stands for the logarithm to base 10. Further, $\partial/\partial\theta$ denotes the gradient operator and $\mathcal{N}(\mu, \sigma^2)$ denotes a Gaussian distribution with mean $\mu$ and variance $\sigma^2$.

II. RELATION TO PRIOR WORK

Proximity information can be obtained for instance from proximity sensors such as infrared (IR) sensor or from radio system. The earliest proximity based positioning systems were built with IR sensors, e.g., [1]. The idea was then borrowed to radio system using mainly RSS as proximity indicator [2]. Essentially, a proximity measurement is a quantized RSS with two quantization levels [3], [4]. In contrast to a position-related range measurement, such as time-of-arrival or RSS, a proximity measurement only coarsely tells whether or not two devices are within certain communication range hence contains much less information about the actual position of a device to be inferred. In the literature, the positioning algorithms using proximity measurements are called coarse grained algorithms or range-free algorithms. Since the publication of [2], a plethora of proximity based positioning algorithms have been proposed, including the centroid algorithm [2], the approximate point in triangle (APIT) algorithm [5], the maximum-likelihood estimation algorithm [3], the ecolocation algorithm [6], the iterative learning based algorithm [7], and the particle filtering algorithm [8], to mention a few. Most of them were done for large-scale cooperative sensor network localization because of the communication constraints, among others the bandwidth and capacity limitations.

Proximity based positioning is setting new trends nowadays, due to the rapid development of the beaconing techniques on the one hand and its low-complex and low-cost nature for the internet-of-things (IoT) on the other hand. Increasing number of new fashioned proximity based algorithms are expected. To evaluate the goodness of a position estimator, we derive two fundamental lower bounds, namely the Cramér-Rao bound and the Barankin bound, on the estimation MSE in generic form, without assuming unbiasedness. We have the feeling that unbiased, proximity based position estimator generally does not exist for small-scale network, but the bias can be somewhat mitigated by adding dithering noise before quantization [9].
For large-scale network, some position estimator might be closely unbiased due to the shrinked positioning ambiguity regions, see [2, Fig. 1] as an example. In the literature, there exist readily some prior work on the CRB, see for instance [3], [10], but for the covariance matrix of unbiased, proximity based position estimator. The idea to use BB to evaluate a position estimator is not new, existing work include [11], [12], [13], [14]. But as far as we know, this paper is the first to derive a large-error bound for proximity based positioning and further demonstrate that the Barankin bound might be much tighter than the CRB bound on the MSE in practice.

III. SIGNAL MODEL

Consider an indoor localization scenario where in total $N$ reference network nodes are deployed to locate a stationary user equipment (UE). For simplicity, the three-dimensional (3-D) positions of the reference network nodes, $p_i = [x_i, y_i, z_i]^T$, $i = 1, 2, \ldots, N$, are assumed to be precisely known. The 3-D UE position is denoted by $p = [x, y, z]^T$. We assume $\theta = [\theta_1, \ldots, \theta_d]^T$ ($1 \leq d \leq 3$), comprising unknown, to be determined component(s) amid $p$. In the sequel, $\theta$ is assumed to be deterministic. The underlying true parameter is denoted as $\theta_0$. The UE measures received-signal-strength, $r_i$, from the $i$th reference network node and further convert it into a proximity report as follows:

$$c_i \triangleq \begin{cases} 0, & r_i \leq P_{th} \\ 1, & r_i > P_{th} \end{cases},$$

where $P_{th}$ is a trained RSS threshold. One method to obtain a reasonable RSS threshold $P_{th}$ was recently introduced in [15]. To ease our narration in the sequel, we assume that within any sampling time period, $[t, t+1)$, the UE receives a vector of RSS reports from all reference network nodes, $r = [r_1, r_2, \ldots, r_N]^T$. The corresponding vector of proximity reports is defined as $c = [c_1, c_2, \ldots, c_N]^T$. But note that in practice, the UE can hear only a subset of all reference network nodes. In a network centric setup, the vector of proximity reports, $c$, is transmitted to the core network for system configuration as well as for positioning.

In this paper, we adopt a simple propagation model, namely the underlying RSS related to the $i$th reference network node well follows a log-distance path loss model, i.e.,

$$r_i = A_i + 10B_i \log_10 \left( \frac{d_i(p)}{d_0} \right) + e_i, \quad i = 1, 2, \ldots, N,$$

where $A_i$ is a reference path loss measured at the preselected distance $d_0$, $B_i$ is a path loss exponent, the Euclidean distance between the UE’s position, $p$, and the $i$th reference network node’s position, $p_i$, is $d_i(p) \triangleq ||p_i - p||$, and the measurement and modeling error term $e_i$ is assumed to be Gaussian distributed with zero mean and variance $\sigma_e^2$. The propagation model parameters $A_i$, $B_i$, and $\sigma_e^2$, for $i = 1, 2, \ldots, N$, are assumed to be known a priori from offline calibration. The error terms measured for different reference network nodes are assumed to be mutually independent.

IV. MAXIMUM-LIKELIHOOD POSITION ESTIMATOR

We start by defining a position estimator as $\hat{\theta}(c)$, which is a function of the proximity reports, $c$. Among other performance metrics, a direct measure of the estimation error—the mean-square-error (MSE), has been very often applied in comparing different estimators. In the remainder of this section, we take the maximum likelihood estimator (MLE) as an example and derive its MSE in closed form.

Based on our signal model introduced and assumptions made in Section III, the MLE, $\hat{\theta}_M(c)$, is solved by maximizing the likelihood function for the unknown position parameters, $\theta$, given a vector of proximity reports, $c$. More precisely,

$$\hat{\theta}_M(c) = \arg \max_\theta \left\{ \Pr[c; p] = \prod_{i=1}^{N} \Pr[c_i; p] \right\}. \quad (2)$$

Herein, the probability of receiving $c_i$ at the $i$th reference network node is given by

$$\Pr[c_i; p] = \begin{cases} G \left( \frac{P_{th} - \mu \theta_i(p)}{\sigma_i} \right), & c_i = 0 \\ 1 - G \left( \frac{P_{th} - \mu \theta_i(p)}{\sigma_i} \right), & c_i = 1 \end{cases}, \quad (3)$$

where $G(\cdot)$ denotes Gaussian cumulative distribution function (CDF). For notational brevity in subsequent sections, we define $G_i(p) \triangleq G \left( \frac{P_{th} - \mu \theta_i(p)}{\sigma_i} \right)$. In order to find the MLE, we often need to resort to numerical methods, such as Newton-type methods and grid search.

The MSE of the MLE introduced in (2) can be computed as follows:

$$\text{MSE}(\hat{\theta}_M(c)) = \mathbb{E}_{Pr(c|p)} \left[ \left( \hat{\theta}_M(c) - \theta \right) \left( \hat{\theta}_M(c) - \theta \right)^T \right] = b(\theta)b^T(\theta) + \Sigma(\theta),$$

where the bias, $b(\theta)$, and the covariance matrix, $\Sigma(\theta)$, of the MLE are defined respectively by

$$b(\theta) \triangleq \mathbb{E}_{Pr(c|p)} \left[ \hat{\theta}_M(c) - \theta \right],$$

$$\Sigma(\theta) \triangleq \mathbb{E}_{Pr(c|p)} \left[ \left( \hat{\theta}_M(c) - b(\theta) - \theta \right) \left( \hat{\theta}_M(c) - b(\theta) - \theta \right)^T \right].$$

Due to the binary nature of a proximity report, the expectation of the MLE can be expressed as follows:

$$\mathbb{E}_{Pr(c|p)} \left[ \hat{\theta}_M(c) \right] = \sum_{l=0}^{2^N-1} \hat{\theta}_M(c^l) \Pr(c^l; p), \quad (4)$$

where $c^l$ is the $l$th possible vector of proximity reports, containing $N$ binary bits representation of $l$. The maximum-likelihood estimate, $\hat{\theta}_M(c^l)$, can be treated as a constant that is irrespective of the true underlying parameter $\theta_0$. We further assume that the grids are fine enough to guarantee that global maximum of (2) can be reached for any $c^l$, $l = 0, 1, \ldots, 2^N - 1$. Provided that storage permits, we could precompute and store these $2^N$ ML estimates in a look-up table for use in the online positioning phase. The bias and the covariance matrix of the MLE can then be re-written as follows:

$$b(\theta) = \left( \sum_{l=0}^{2^N-1} \hat{\theta}_M(c^l) \Pr(c^l; p) \right) - \theta,$$

$$\Sigma(\theta) = \sum_{l=0}^{2^N-1} \left( \hat{\theta}_M(c^l) - b(\theta) - \theta \right) \left( \hat{\theta}_M(c^l) - b(\theta) - \theta \right)^T \Pr(c^l; p).$$
Thanks to the binary nature of a proximity report, we are able to derive the bias and covariance matrix of the MLE in closed form. However, the computation complexity scales as $O(2^N)$, which forbids the availability for large $N$. This might be also the case when evaluating bias and covariance matrix of other estimators.

After we have designed a parameter estimator, very often we also want to verify whether it is satisfactory or extra effort for tuning is worthwhile. To that end, we can compare the MSE of the designed parameter estimator with some performance lower bound. It is known from the literature, see for instance [16], that estimation lower bounds are placed on the MSE of a parameter estimator by appropriate choice of $\eta$ namely,

$$\text{MSE}(\hat{\theta}(c)) \geq \mathbf{b}(\theta)\mathbf{b}^T(\theta) + T\Gamma^{-1}T^T,$$

(5)

where the information matrix, $\Gamma$, and the translation matrix, $T$, are defined relative to $\eta$ respectively as follows:

$$\Gamma \triangleq \mathbb{E}_{\mathbf{Pr}(c; \mathbf{p})}[\eta\eta^T],$$

(6)

$$T \triangleq \mathbb{E}_{\mathbf{Pr}(c; \mathbf{p})}[ \left( \mathbf{\hat{\theta}(c)} - \mathbf{b}(\theta) - \mathbf{\tilde{\theta}} \right)\eta^T].$$

(7)

In the next section, we will give two different lower bounds distinguished by the selection of $\eta$.

V. FUNDAMENTAL LOWER BOUNDS ON THE MSE

In this section, we derive two different lower bounds on the proximity report based position estimation MSE. They are the Cramér-Rao bound (CRB) and the Barankin bound (BB). It is known that CRB falls into the category of small error bounds, while BB falls into the category of large error bounds. Small/Large-error bound is intended for small/large estimation error. Herein, we assume that all regularity conditions for computing a specific bound are fulfilled.

A. Cramér-Rao Bound (CRB)

When we let

$$\eta = \frac{\partial \mathbf{Pr}(c; \mathbf{p})}{\partial \mathbf{\tilde{\theta}}},$$

the translation matrix, $T$, defined in (7) can be easily proven to be a $d \times d$ matrix of the form

$$T = \left( \frac{\partial \mathbf{\rho}^T(\theta)}{\partial \theta} \right)^T |_{\theta = \theta_0},$$

(8)

where $\rho(\theta) \triangleq \mathbf{b}(\theta) + \mathbf{\theta}$, and the information matrix, $\Gamma$, defined in (6) becomes to the well known Fisher information matrix (FIM) of dimension $d \times d$, namely,

$$\Gamma = \mathbb{E}_{\mathbf{Pr}(c; \mathbf{p})} \left[ \frac{\partial \mathbf{Pr}(c; \mathbf{p})}{\partial \mathbf{\tilde{\theta}}} \left( \frac{\partial \mathbf{Pr}(c; \mathbf{p})}{\partial \mathbf{\tilde{\theta}}} \right)^T \right] |_{\theta = \theta_0}. $$

(9)

The FIM can be expressed, due to the independence assumption on the measurements, as

$$\Gamma = \sum_{i=1}^{N} \mathbb{E}_{\mathbf{Pr}(c_i; \mathbf{p})} \left[ \frac{\partial \mathbf{Pr}(c_i; \mathbf{p})}{\partial \mathbf{\tilde{\theta}}} \left( \frac{\partial \mathbf{Pr}(c_i; \mathbf{p})}{\partial \mathbf{\tilde{\theta}}} \right)^T \right] |_{\theta = \theta_0},$$

$$\triangleq \sum_{i=1}^{N} \begin{bmatrix} f_{i,\theta,\theta_1} & \cdots & f_{i,\theta,\theta_d} \\ \vdots & \ddots & \vdots \\ f_{i,\theta,\theta_1} & \cdots & f_{i,\theta,\theta_d} \end{bmatrix}.$$

(10)

It is easy to derive for $m, n \in \{1, \ldots, d\}$ that

$$f_{i,\theta,\theta_m} = \sum_{c_i \in \{0, 1, \ldots, \}} \frac{\partial \mathbf{Pr}(c_i; \mathbf{p})}{\partial \mathbf{\tilde{\theta}_m}} \frac{\partial \mathbf{Pr}(c_i; \mathbf{p})}{\partial \mathbf{\tilde{\theta}_{n}}} |_{\theta_{m}=\theta_{m_0}, \theta_n=\theta_{n_0}},$$

$$= \frac{50B_i^2}{\pi\sigma_i^2} \frac{\ln^2(\mathbf{G}_i(\mathbf{p})(1-\mathbf{G}_i(\mathbf{p})))}{||\mathbf{p}-\mathbf{p}_l||^4},$$

where $\mathbf{p}_{l,m}$ and $\mathbf{p}_{l,n}$ are respectively the corresponding position components of the $i$th reference network node.

Finally, we note that CRB is equivalent to the first order Bhattacharyya bound. Higher order Bhattacharyya bound is believed to generate somewhat tighter bound than CRB, but at higher cost of computational efforts [17]. Despite of the improved tightness, Bhattacharyya bound is still deemed as a small error bound in the literature [16].

B. Barankin Bound (BB)

Computing a Barankin bound requires a set of test points $\theta_{m}$, where $m = 1, 2, \ldots, M$, to be preselected from $d$-dimensional subspace. For clarity, we further define $\mathbf{p}_{l,m,n}$ to be a 3-D vector which includes $\theta_{m}$ and possibly other $3 - d$ fixed position element(s) $\mathbf{p}\setminus\theta_0, \text{ if } d < 3$. This set of test points decides the tightness of the Barankin bound derived below.

When we set

$$\eta = \begin{bmatrix} \mathbf{Pr}(c; \mathbf{p}_1) - 1, \ldots, \mathbf{Pr}(c; \mathbf{p}_M) - 1 \end{bmatrix}^T,$$

(10)

the translation matrix (of dimension $d \times M$) can be derived based on the results in [16] to be

$$T = \begin{bmatrix} \rho(\theta_1) - \rho(\theta_0), \ldots, \rho(\theta_M) - \rho(\theta_0) \end{bmatrix}^T,$$

and the information matrix $\Gamma$ becomes to the Barankin information matrix (of dimension $M \times M$) whose $(j,k)$th entry is computed as

$$\Gamma_{jk} = \mathbb{E}_{\mathbf{Pr}(c; \mathbf{p})} \left[ \frac{\mathbf{Pr}(c; \mathbf{p}_j)}{\mathbf{Pr}(c; \mathbf{p})} \frac{\mathbf{Pr}(c; \mathbf{p}_k)}{\mathbf{Pr}(c; \mathbf{p})} \right] - 1, \forall j, k \in \{1, \ldots, M\}.$$

Due to the independence assumption on the proximity reports, we can re-write the above expression after some tedious manipulations as

$$\Gamma_{jk} = \frac{N}{\sum_{c_i \in \{0, 1, \ldots, \}} \mathbf{Pr}(c_i; \mathbf{p}_j)\mathbf{Pr}(c_i; \mathbf{p}_k)} - 1. $$

(11)

In order to give a tight lower bound, the test points need to be carefully selected. The optimal set of test points can be found as

$$\theta_{opt}^1, \ldots, \theta_{opt}^M = \arg\max_{\theta_1, \ldots, \theta_M} T\Gamma^{-1}T^T.$$

Unfortunately, this optimization problem is cumbersome for our case. A much easier way of providing a set of sub-optimal test points is to minimize the diagonal elements of the Barankin matrix, $\Gamma_{jj}, \ j = 1, 2, \ldots, M$. Concretely,

$$\theta_{j, opt} = \arg\min_{\theta_j} \Gamma_{jj}, \ \forall j = 1, 2, \ldots, M.$$


According to (11), we have
\[ I_{jj} = \prod_{i=1}^{N} \left( \sum_{c_i \in \{0,1\}} \frac{Pr^2(c_i; \hat{p}_j)}{Pr(c_i; p)} \right) - 1, \]
where it is easy to show
\[ \sum_{c_i \in \{0,1\}} \frac{Pr^2(c_i; \hat{p}_j)}{Pr(c_i; p)} = (\alpha + \beta)G_j^2(\hat{p}_j) - 2\beta G_j(\hat{p}_j) + \beta \]
with \( \alpha \triangleq 1/G_j(p) \) and \( \beta \triangleq 1/(1 - G_i(p)) \) being constants given the actual UE position, \( p \).

We adopt the following algorithm to find the test points. The test points selected in the above way should well represent the location ambiguities that a position estimation algorithm more easily gets stuck at.

Algorithm 1: Test Points Selection for BB Computation

Step-1: Generate a set of \( K \) (with \( K \gg M \)) candidate test points, \( \theta_i \), uniformly and rather densely in the service area.

Step-2: For each candidate test point, evaluate \( I_{kk} \), \( k = 1, 2, \ldots, K \), according to (12) and store them.

Step-3: Find among \( I_{kk} \), \( k = 1, 2, \ldots, K \), all the local minima \( I_{jj} \) and record their indices \( j^* \).

Step-4: Set a selection threshold, \( \varepsilon \), search for all \( j \in \{1, \ldots, j^* - 1, j^* + 1, \ldots, K\} \) that satisfy \( I_{jj} - I_{jj^*} < \varepsilon \) and use \( \theta_{jj} \) as the final test points. Repeat this for all \( j^* \).

Computation of the BB does not require the first order derivatives of \( G_j(p) \) as in the CRB computation, but requires a set of test points to be carefully selected. The computation cost is much higher for the BB computation than for the CRB computation. To ease the test points selection procedure, the Chapman-Robbins bound [18], which is a special case of the above derived Barankin bound using only one test point, can be used; that is
\[ \theta_{1}^{\text{exp}} = \arg\max_{\theta_1} \frac{1}{N} \prod_{i=1}^{N} \sum_{c_i \in \{0,1\}} \frac{Pr^2(c_i; \hat{p}_j)}{Pr(c_i; p)} - 1. \]

The complexity of the above optimization problem has been reduced significantly. However, it was shown in [18, Theorem 1] that the Barankin bound is an increasing function of the number of test points, the resulting Chapman-Robbins bound hence might not be as tight as the Barankin Bound using \( M \gg 1 \) test points.

VI. SIMULATION RESULTS

In this section, we aim to compare the MSE of the MLE derived in Section IV with the two lower bounds derived in Section V. To be realistic, we consider a live Bluetooth low-energy (BLE) network deployed at ERICSSON Research in Linkoping, Sweden. We select in total 25 different UE positions in the service area covered by \( N = 10 \) immobile BLE beacons (serving as reference network nodes). The 3-D positions of the BLE beacons are known \textit{a priori}. The underlying propagation model is approximated by the canonical linear log-distance path loss model with the model parameters calibrated offline. The RSS threshold is set by \( P_{th} = -82 \text{ dBm} \).

TABLE I. BIAS AND MSE OF THE MLE VERSUS TWO LOWER BOUNDS ON THE MSE.

<table>
<thead>
<tr>
<th>( \theta )</th>
<th>MLE (bias, MSE)</th>
<th>CRB</th>
<th>BB</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \theta_2 )</td>
<td>[0.6] - 17</td>
<td>46 - 7</td>
<td>75 - 15</td>
</tr>
<tr>
<td>( \theta_4 )</td>
<td>-2</td>
<td>31</td>
<td>-2</td>
</tr>
<tr>
<td>( \theta_{19} )</td>
<td>-7</td>
<td>15</td>
<td>-1</td>
</tr>
<tr>
<td>( \theta_{24} )</td>
<td>1.5</td>
<td>46</td>
<td>1</td>
</tr>
</tbody>
</table>

![Fig. 1. Service area covered by 10 BLE beacons (marked by blue stars) and 25 different UE positions (marked by red circles).](image)

We have derived two fundamental lower bounds on the mean-square-error of proximity based position estimator. Due to the large variance in the received-signal-strength measurements as well as the thresholding effect, the positioning am-
biguity and eventually the estimation error can be very large. In this case, Barankin bound will be much tighter than the Cramér-Rao bound. Using the Barankin bound as benchmark will avoid extra effort for tuning a specific estimator towards an over-optimistic bound. These have been exemplified by comparing the theoretical mean-square-error of a specific position estimator, namely the maximum-likelihood estimator, with the two derived lower bounds in the simulations. But our statements should apply generally for any proximity based position estimator.

ACKNOWLEDGMENT

This work is funded by the European Union FP7 Marie Curie training programme on Tracking in Complex Sensor Systems (TRAX) with grant number 607400. We acknowledge the support from SenionLab, who provided the BLE beacons as well as associated positioned RSS measurement data.

REFERENCES


