Posterior Cramér-Rao Lower Bound for RSS/TOA-based Indoor Localization Systems

J. Prieto*, S. Mazuelas**, A. Bahillo*, P. Fernández*, R. M. Lorenzo* and E. J. Abril*
*Department of Signal Theory and Communications and Telematic Engineering, University of Valladolid.
Paseo de Belén 15. 47011. Valladolid. SPAIN. Email: javier.prieto@uva.es
**Laboratory for Information and Decision Systems (LIDS), Massachusetts Institute of Technology,
77 Massachusetts Avenue, Cambridge, MA. 02139. USA. Email: mazuelas@mit.edu

***Abstract*—The great success of location-based applications that operate with GPS, has emphasized the necessity of technologies that make up for the lack of position information in harsh environments, as dense urban or indoor areas. Many solutions have been proposed, however, none of them has any widespread acceptance. In this case, it is mandatory to have a mechanism to compare the different proposed systems in terms of performance. In this paper, we derive the Cramér-Rao lower bound (CRLB) on the minimum mean square error for distance estimation in a nonlinear/non-Gaussian localization system. For this purpose, we obtain the density function of the received signal strength (RSS) and time-of-arrival (TOA) measurements of a signal that reaches the MU coming from an access point. The proposed model considers both the random error and the bias caused by multipath and non-line-of-sight (NLOS) propagation conditions present in these harsh environments. Since the relationship among distances in time is likewise considered, the problem is addressed within a Bayesian context. The final CRLB is compared to the mean square error obtained by conventional Kalman filtering techniques.**

***Index Terms*—Cramér-Rao, indoor localization, RSS, TOA.**

I. Introduction

Nowadays, there are more and more mobile devices on the market. So much so, that, for the first time ever, smartphones are outselling personal computers. In this context, localization based services (LBS) are becoming more and more important [1]. These LBS are based on the position provided by GPS or by the localization of an access point (AP) with which the mobile device is communicating. However, a wide range of LBS requires a similar accuracy in harsh environment, as dense urban or indoor areas. For this type of scenarios, several solutions have been proposed, however, none of them seems to dominate over the others [2]. Therefore, it is necessary to utilize an efficient method to compare them in terms of performance, that is, in terms of both accuracy (how close the estimate is to the actual position) and precision (the variability of the estimates due to repetition).

For these harsh environments, conventional estimation techniques, such as maximum likelihood (ML), maximum a posteriori (MAP) or minimum mean square error (MMSE) obtain poor performances [3]. This is due to the presence of nonlinearities or non-Gaussian errors in the models caused by multipath and non-line-of-sight (NLOS) propagation conditions [4]–[6]. In these cases, a suboptimal approach has to be used, such as extended Kalman filters (EKFs), particle filters (PFs), the expectation maximization algorithm (EM), or a mixture of them. These techniques are commonly assessed by means of the mean square error, since it reports information about both the variance or random error (i.e. the precision) and the bias or systematic error (i.e. the accuracy) [7]. To this aim, simulation results are compared to theoretical performance bounds.

The Cramér-Rao (CRLB) bound establishes a lower bound on the mean square error of any unbiased estimator [7]. This bound has been used as a theoretical benchmark for the comparison of implemented suboptimal algorithms and as a measure of the effects of introduced approximations [8]. Therefore, the CRLB is an important design tool used as a predictor of the best achievable performance before implementing a system. However, the computation of this bound is not always an easy task, especially in the nonlinear filtering case [8]. The reliability of this bound depends on how well it reflects all the available information.

In this paper, we obtain the CRLB for a wireless indoor localization system. This system is based on the nonlinear filtering of the received signal strength (RSS) and the time-of-arrival (TOA) of the signals transmitted between several APs, with known positions, and a mobile user (MU) whose position has to be estimated. To this aim, Section II presents the models utilized for the nonlinear filtering problem, and derives a density function for the RSS and TOA measurements that includes both random and systematic error. Section III describes the recursive method for the CRLB computation first proposed in [9], whereas Section IV shows the CRLB obtained by simulation and compares it to the result of applying an EKF to actual RSS and TOA measurements. Finally, Section V analyzes the conclusions drawn from the presented work.

Notations: we denote $N(\mathbf{x};\mu,\mathbf{P})$ the Gaussian density fuction of a random vector $\mathbf{X}$, where μ is the mean vector, and $\mathbf{P}$ is the covariance matrix.

II. Indoor Localization System

This section presents the localization task from RSS and TOA measurements as a nonlinear and non-Gaussian filtering

This research is partially supported by the Directorate General of Telecommunications of the Regional Ministry of Public Works and the Regional Ministry of Education from Castilla y León (Spain), by the Spanish national project LEMUR (TIN2009-14114-C04-03) and by the European Social Fund.

problem. First, we describe the dynamic model that establishes the relationship among distances in time. Afterwards, we detail RSS and TOA models for range estimation, and derived the corresponding densities in order to reflect the systematic error as well as the random error.

We consider a two-dimensional scenario where an MU moving freely has to be located. The MU obtains a set of M RSS and/or M TOA measurements in discrete time instants $\{t_k, k \in \mathbb{N}\}$. These measurements come from the signals transmitted by several APs with know positions, called anchors, to the MU. Then, the position estimation is carried out in two steps: in the first one, the system estimates the distance $d(t)$ to all the access points; in the second step, the position is obtained by trilateration [6]. This paper is focused on the first step, that is, this paper addresses the problem of estimating the distances $\{d[k], k \in \mathbb{N}\}$ to each anchor from the sequence of measurements $\{\mathbf{z}[k], k \in \mathbb{N}\}$.

A. Dynamic model

It is clear that the distance between an MU and an AP in a given time instant is not independent of the distance between them in the immediately previous instant. This correlation over time leads to model the evolution in time of the distance as an analytic function given by its nth order Taylor series expansion. Calling $\mathbf{y}[k]$ the state vector formed by the distance and some of its first derivatives in a time instant t_k, the dynamic model can be approximated by [10]

$$\mathbf{y}[k+1] = \mathbf{F}_k \mathbf{y}[k] + \mathbf{n}_d[k], \tag{1}$$

where $\mathbf{F}_k$ is the transition matrix given by the $(n-m)$th order Taylor expansion for each mth derivative of the distance. The error term $\mathbf{n}_d[k]$ represents the error in the approximation and is commonly modeled as a zero-mean Gaussian random variable with a covariance matrix $\mathbf{Q}_k$ [8], [10].

B. Measurements model

RSS and TOA information are the most conventional metrics utilized to locate an MU, due to the ease of obtaining their values from the transmitted signals [3]–[6]. Hence, these metrics will be the information sources together with the dynamic model utilized to compute the CRLB.

1) Received signal strength, RSS: The distance between the MU and an AP can be inferred from the RSS values since this distance is one of the factors that most affects the strength level. The attenuation caused by the distance between two nodes is known as *path-loss* and is proportional to this distance raised to a certain exponent, called *path-loss exponent* [2]–[5]. However, the RSS values are likewise affected by a wide range of unpredictable factors. In logarithmic units we have,

$$z_s[k] = \alpha_s - 10\beta_s \log_{10}(d[k]) + n_s[k], \tag{2}$$

being $z_s[k]$ the RSS value, β_s the path-loss exponent, and α_s is a constant that depends on several factor like slow and fast fading, gains in the transmitter and receiver antennas, and the transmitted power [3]–[5]. The error term $n_s[k]$, is zero-mean Gaussian in the cases where α_s and β_s perfectly fit their actual values (i.e. in this case the error only includes random error). The value of α_s can be measured or provided by the manufacturer, however, in realistic harsh environments, β_s will not fit its actual value, and n_s will have non-zero mean (the error includes a bias) proportional to the logarithm of the distance [4].

To evaluate the CRLB, we need the density function from which the RSS measurements were generated in the time instant t_k, i.e. we need to know $p(z_s|d[k])$. Given a bias b_s and a distance $d[k]$ we have

$$p(z_s|d[k], b_s) = \begin{cases} \frac{N(z_s;\zeta_s,\sigma_s)}{erf(\sqrt{2})} & if \quad |z_s - \zeta_s| \leq 2\sigma_s \\ \\ 0 & otherwise \end{cases} \tag{3}$$

where $\zeta_s = f_s(d[k]) + b_s$, and $f_s(d[k])$ is given by (2). In this way, the error introduced in the measurements is modeled as a truncated zero-mean Gaussian random variable. We truncate the Gaussian distribution to reflect the fact that the measuring system cannot report RSS values for all $\mathbb{R}$, i.e. the measuring system has a limited range.

By assuming that for common indoor distances $\log_{10}(d[k]) \simeq 1$ and modeling β_s as a Gaussian random variable with mean the actual value and standard deviation σ_β, the bias caused by β_s misestimation can be modeled as,

$$p(b_s) = \begin{cases} \frac{N(b_s;0,\sigma_{sh})}{erf(\sqrt{2})} & if \quad |b_s| \leq 2\sigma_{sh} \\ \\ 0 & otherwise \end{cases} \tag{4}$$

where $\sigma_{sh} = 10\sigma_\beta$.

Therefore, the density function for RSS measurements is obtained by marginalization,

$$p(z_s|d[k]) = \int_{-\infty}^{\infty} p(z_s|d[k], b_s) p(b_s) db_s \tag{5}$$

The result of (5) is a piecewise function where we can distinguish two cases depending on whether σ_s is greater or lower than σ_{sh}. Figure 1 shows the density function for both cases.

2) Time of arrival, TOA: The distance between an AP and the MU can be estimated through a linear transformation of the time that the signal takes to travel from the first to the second node, since the speed of electromagnetic waves in the air can be assumed to be constant and known. However, to avoid nodes synchronization, techniques based on round-trip-times (RTTs) result more attractive [6]. In this case, the relationship between the distance $d[k]$, and the TOA $z_\tau[k]$, has an intercept, that is,

$$z_\tau[k] = \alpha_\tau + \beta_\tau d[k] + n_\tau[k], \tag{6}$$

where α_τ and β_τ are constants that can be estimated in a previous stage to the localization process [6], [11]. The error term n_τ will be zero-mean Gaussian in the case where there is a line-of sight (LOS) between the MU and the AP (i.e. this terms only includes random error), however, n_τ will follow

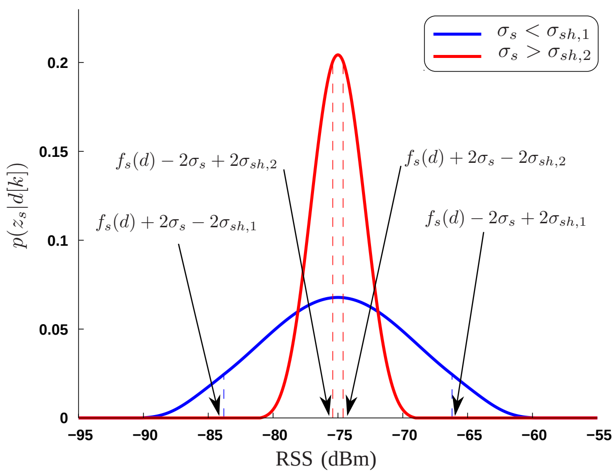

Fig. 1. The central piece of the RSS density varies depending on wheter σ_s is greater or lower than σ_{sh}.

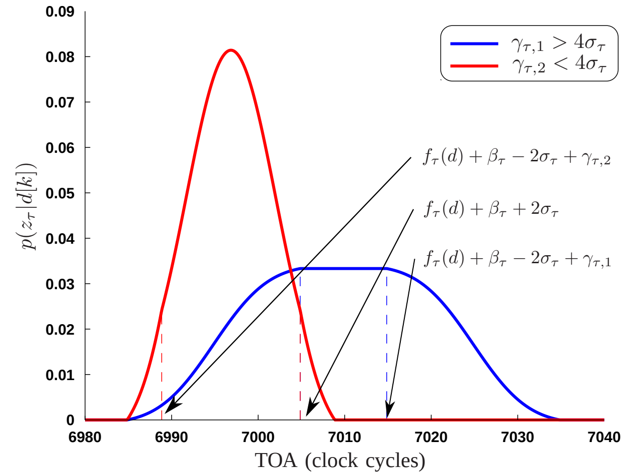

Fig. 2. The central piece of the TOA density varies depending on wheter σ_τ is greater or smaller than $\gamma_\tau/4$.

a positive distribution (positive bias) in the cases of NLOS propagation [4], [5].

Following a parallel procedure to the RSS case, given a bias b_τ and a distance $d[k]$ in the time instant t_k, we have,

$$p(z_\tau|d[k], b_\tau) = \begin{cases} \frac{N(z_\tau;\zeta_\tau,\sigma_\tau)}{erf(\sqrt{2})} & if \quad |z_\tau - \zeta_\tau| \leq 2\sigma_\tau \\ 0 & otherwise \end{cases} \tag{7}$$

where $\zeta_\tau = f_\tau(d[k]) + b_\tau$, and $f_\tau(d[k])$ is given by (6).

In this paper, the NLOS bias is modeled as a positive uniform random variable, however, any other positive distribution can be utilized and the results can be obtained analogously. For the uniform bias,

$$p(b_\tau) = \begin{cases} \frac{1}{\gamma_\tau} & if \quad 0 \leq b_s \leq \gamma_\tau \\ 0 & otherwise \end{cases} \tag{8}$$

being γ_τ the maximum bias caused by NLOS propagation.

As well as in the RSS case, the density function for the TOA measurements is obtained by marginalization,

$$p(z_\tau|d[k]) = \int_{-\infty}^{\infty} p(z_\tau|d[k], b_\tau)p(b_\tau)db_\tau, \tag{9}$$

resulting again in a piecewise function with two different cases depending on whether σ_τ is greater or smaller than $\gamma_\tau/4$. In Fig. 2 we can observe both cases and the most noticeable effect of the bias when the latter is greater than $4\sigma_\tau$.

III. Cramér-Rao lower bound

The CRLB provides a lower bound on the minimum achievable mean square error for any unbiased estimator [7]. However, in the addressed localization problem, within the Bayesian context, there is no true parameter. That is, for time-variant systems, what is estimated is a density function. Van Trees proposed a posterior CRLB (PCRLB) for the Bayesian case [12], since this bound is obtained from posterior distributions [13]. In this case, for each time instant t_k, the PCRLB is given by,

$$\mathbb{E}\{(g(\mathbf{Z}[k]) - \mathbf{y}[k])(g(\mathbf{Z}[k]) - \mathbf{y}[k])^T\} \succeq \mathbf{J}_k^{-1},$$

where $\mathbf{Z}[k]$ denotes all the available measurements up to time t_k, i.e. the set $\{\mathbf{z}[i], i = 1..., k\}$. Moreover, $g(\mathbf{Z}[k])$ is an unbiased estimator of $\mathbf{y}[k]$ and $\mathbf{J}_k$ is the Fisher information matrix (FIM) obtained as,

$$\mathbf{J}_k = -\mathbb{E}\{\nabla_{\mathbf{y}[k]}[\nabla_{\mathbf{y}[k]} \log p(\mathbf{z}[k]|\mathbf{y}[k])]^T\}.$$

Tichasvský et al. proposed in [9] a method for recursive computation of this FIM,

$$\mathbf{J}_{k+1} = \mathbf{D}_k^{22} - \mathbf{D}_k^{21}(\mathbf{J}_k + \mathbf{D}_k^{11})^{-1}\mathbf{D}_k^{12}, \ (k > 0) \tag{10}$$

where for the considered linear-Gaussian dynamic model (1),

$$\begin{aligned} \mathbf{D}_k^{11} &= \mathbf{F}_k^T \mathbf{Q}_k^{-1} \mathbf{F}_k \\ \mathbf{D}_k^{12} &= -\mathbf{F}_k^T \mathbf{Q}_k^{-1} \\ \mathbf{D}_k^{21} &= [\mathbf{D}_k^{12}]^T \\ \mathbf{D}_k^{22} &= \mathbf{Q}_k^{-1} + \mathbf{D}_{k,b}^{22} \end{aligned}$$

where

$$\mathbf{D}_{k,b}^{22} = -\mathbb{E}\{\nabla_{\mathbf{y}[k+1]}[\nabla_{\mathbf{y}[k+1]} \log p(\mathbf{z}[k+1]|\mathbf{y}[k+1])]^T\}.$$

Since $\mathbf{D}_{k,b}^{22}$ is obtained from (5) and (9), this expression has no closed-formed solution and its result has to be obtained by Monte Carlo integration.

Moreover, to start the recursion, the initial FIM, $\mathbf{J}_0$, is obtained by considering the initial density $p(\mathbf{y}_0) = N(\mathbf{y}_0; \mu_0, \mathbf{P}_0)$ and, therefore, $\mathbf{J}_0 = \mathbf{P}_0^{-1}$ [8].

IV. Results

For this section, a state vector formed by the distance and its first two derivatives is considered. Moreover, prior information about these derivatives is incorporated by modeling them as zero-mean Gaussian random variables with standard deviation

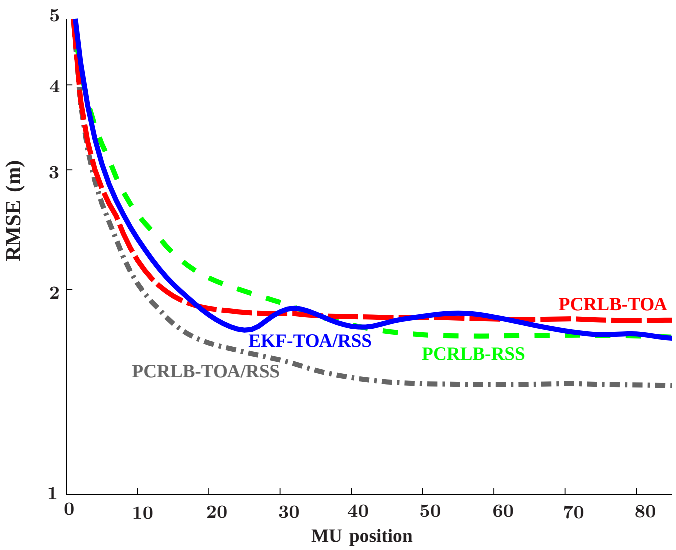

Fig. 3. Since the EKF is a suboptimal approach, this method does not reach the PCRLB obtained for TOA/RSS data fusion.

$\sigma_{d'} = 0.5$ m/s and $\sigma_{d''} = 0.75$ m/s^2, respectively. In the dynamic model, a standard deviation of $\sigma_{d^{(3)}} = 1$ m/s^3 for the third derivative of the distance is considered.

We simulate a random trajectory followed by an MU. In this trajectory, the minimum and maximum distances between the MU and an AP with fixed position are 10 m and 30 m, respectively. The MU receives a set of $M = 10$ RSS and $M = 10$ TOA measurements with respect to the AP every second during 85 seconds. To generate those measurements, expressions (5) and (9) are utilized. We select, in the RSS case, $\sigma_s = 1.6$ dBm and $\sigma_{sh} = 3$ dBm for the truncated Gaussian distributions regarding the random error and the bias, respectively. In the TOA case, $\sigma_\tau = 3.5$ clock cycles and $\gamma_\tau = 3$ clock cycles, are selected for the truncated Gaussian and uniform distributions, corresponding to the random error and the bias, respectively. These are error values obtained in previous works for an IEEE 802.11b/g network [4], [6].

Figure 3 shows the square root of the accumulated PCRLB on the mean square error in range estimation, for each position of the simulated trajectory. We select the accumulated PCRLB in each position since the performance of localization systems is usually assessed by means of the mean square error obtained in all the positions of the trajectory. We show the PCRLB for the cases where only RSS data are available, the case where the MU receives only TOA measurements, and the case where the MU receives both TOA and RSS data. The latter reflects the improvement achieved in the PCRLB by TOA/RSS data fusion. Moreover, we likewise show the root mean square error (RMSE) obtained in range estimation by means of an EKF for the fusion of RSS and TOA measurements. For the EKF, the bias is modeled as a Gaussian $\mathcal{N}(0, \sigma_{sh})$ in the RSS case, and as a Gaussian $\mathcal{N}(\gamma_\tau/2, \gamma_\tau/4)$ in the TOA case.

As reflected in Fig. 3, despite the usage of an EKF to address the nonlinearity problem, the densities given by (5) and (9) differ from the Gaussian densities used by the EKF. Therefore, this suboptimal selection leads to the error that appears in the EKF result compared to the PCRLB.

V. Conclusions

This paper has analyzed the CRLB for the range estimation stage carried out before localizing a mobile user. To this aim, we have derived the density function for the RSS and TOA measurements received by the MU with respect to several anchors. This density takes into account both the random error in the measurements, and the bias caused by multipath and NLOS propagation. The range estimation problem is addressed within the Bayesian framework, by considering the relationship among distances in time. In this case, the CRLB has to be obtained from posterior distributions. This PCRLB can be utilized to compared different algorithms.

In this paper, the PCRLB is obtained from the information provided by RSS and TOA measurements and is compared to the mean square error obtained by an EKF that fuses TOA and RSS data. The goodness of this bound is reflected by: 1) on the one hand, the improvement achieved by the means of the RSS and TOA data fusion; 2) on the other hand, the error introduced by the EKF compared to the PCRLB, since this filter is not the optimal solution when the measurements are generated from the developed density functions.

References

[1] A. Küpper, *Location-Based Services: Fundamentals and Operation*. Chichester, UK: John Wiley & Sons, 2005.

[2] F. Gustafsson and F. Gunnarsson, “Mobile Positioning Using Wireless Networks,” *IEEE Signal Processing Magazine*, vol. 22, no. 4, pp. 41 – 53, July 2005.

[3] N. Patwari, A. O. Hero III, M. Perkins, N. S. Correa, and R. J. O'Dea, “Relative location estimation in wireless sensor networks,” *IEEE Trans. on Signal Processing*, 2003.

[4] S. Mazuelas, A. Bahillo, R. M. Lorenzo, P. Fernández, F. A. Lago, E. Garcia, J. Blas, and E. J. Abril, “Robust Indoor Positioning Provided by Real-Time RSSI Values in Unmodified WLAN Networks,” *IEEE Journal of Selected Topics in Signal Processing*, vol. 3, no. 5, pp. 821 – 831, October 2009.

[5] Y. Qi, “Analysis of Wireless Geolocation in a Non-Line-of-Sight Environment,” *IEEE Transactions on Wireless Communications*, vol. 5, no. 3, pp. 672 – 681, March 2006.

[6] A. Bahillo, S. Mazuelas, R. M. Lorenzo, P. Fernández, J. Prieto, R. Durán, and E. Abril, “Accurate and Integrated Location System for Indoor Environments Based on IEEE 802.11 Round-Trip Time Measurements,” *EURASIP Journal of Wireless Communications and Networking*, vol. 2010, pp. 1 – 13, February 2010.

[7] L. L. Scharf, *Statistical Signal Processing: Detection, Estimation, and Time Series Analysis*. New York:Addison-Wesley Publishing Co., April 1990.

[8] B. Ristic, S. Arulampalam, and N. Gordon, *Beyond the Kalman Filter. Particle Filters for Tracking Applications*. Artech House Publishers, 2004.

[9] P. Tichavský, C. H. Muravchik, and A. Nehorai, “Posterior cramér-rao bounds for discrete-time nonlinear filtering,” *IEEE Transactions on Signal Processing*, vol. 46, no. 5, pp. 1386 – 1396, May 1998.

[10] Y. Bar-Shalom, X. R. Li, and T. Kirubarajan, *Estimation with Applications to Tracking and Navigation: Theory Algorithms and Software*. John Wiley & Sons, Inc., July 2001.

[11] S. A. Golden and S. S. Bateman, “Sensor measurements for wi-fi location with emphasis on time-of-arrival ranging,” *IEEE Trans. on Mobile Computing*, vol. 6, p. 14, 2007.

[12] H. L. Van Trees, *Detection, estimation, and modulation theory: Part I*. New York: Wiley, 1968.

[13] N. Bergman, “Recursive Bayesian Estimation, Navigation and Tracking Applications,” Ph.D. dissertation, Dept. Elect. Eng., Linköping University, Linköping, Sweeden, 1999.