Time-of-Arrival Estimation in Non-Line-Of-Sight Environments

Sinan Gezici, Hisashi Kobayashi and H. Vincent Poor[1]
Department of Electrical Engineering
Princeton University
Princeton, NJ 08544
{sgezici,hisashi,poor}@princeton.edu

Serdar Yüksel and Tamer Başar[2]
Coordinated Science Laboratory
University of Illinois at Urbana-Champaign
1308 West Main Street, Urbana, IL
{yuksel,tbasar}@decision.csl.uiuc.edu

Abstract — **A new estimator is proposed to extract the true distance between a mobile and a base station from a set of time-of-arrival (TOA) data, corrupted by unknown non-line-of-sight (NLOS) errors and Gaussian measurement noise. Characteristics of the estimator are discussed for a class of NLOS errors with variance considerably greater than that of the measurement noise, which is usually the case in practice. A quantization approach is used to estimate the probability density function of the observed TOA measurements and the TOA estimator is computed from this density estimate.**
***Keywords*: Wireless location, NLOS error, parameter estimation, quantization.**

I. Introduction

In the study of mobile positioning in wireless communication systems, accurate location estimation of a mobile station has proven to be an important problem. Multipath, non-line-of-sight (NLOS) propagation and multiple access interference are often the main sources of errors in geolocation, and make mobile positioning challenging. Among these error sources, NLOS is usually considered to be the most crucial one and several practical algorithms to reduce the effects of this error have been proposed in the literature. In [2] a hypothesis testing approach is used to determine NLOS base stations and then the line-if-sight (LOS) reconstruction is achieved using the history of the measurements based on the hypothesis testing results. In [1] statistics of the measurement noise are used to estimate the distances between the mobile and the base stations. However, the assumption made that measurements lower than the true distance are due to the measurement noise only may not hold in general. In [3] an algorithm is used to mitigate the NLOS errors when no prior information is available assuming that there are more than three base stations for geolocation purpose.

In this paper, a time of arrival (TOA) estimator is proposed for a class of NLOS error statistics. It is shown that when the NLOS error has a much larger variance than the Gaussian measurement error, which is usually the case [1], and when its probability density function (pdf) is approximately constant around the origin, then the first derivative of the pdf of the measurements has a peak near the true TOA. This observation is the key to the estimator proposed in this paper.

The remainder of this paper is organized as follows. Section II describes the estimator and discusses its properties. Use of the estimator in a geolocation problem is addressed in Section III, followed by simulation results in Section IV. Finally, Section V presents some concluding remarks.

II. Time of arrival estimation

The estimator we propose is based on the estimation of the pdf of the observed set of TOA measurements, which are the outcomes of a random process with the following characteristics:

$$Y = d + E + N \tag{1}$$

where d is the LOS delay to be estimated, N is the measurement error, which is modelled as a Gaussian random variable with zero mean and known variance σ^2, and E is the NLOS incurred delay taking a value between 0 and some value L, possibly infinite.

Define $Z = d + E$. Note that the pdf of Z will be zero for $Z \leq d$, and the pdf of the observation Y will be the convolution of the pdf of Z and the pdf of the Gaussian noise, N.

Depending on the properties of the pdf of E, we can make the following property.

Property 2.1 *If the density of the NLOS error E is uniform with a variance much larger than the variance of the Gaussian noise, then the probability density of Y has its second derivative equal to zero (saddle point) at almost the true time-of-arrival, d. Furthermore, the assertion remains valid for a non-uniform NLOS error with bounded pdf, that is, $f_E(x) \leq B$ for all x, if the pdf of the NLOS error is almost constant between 0 and α, where $\alpha < L$ satisfies $\int_0^{\alpha}(1-x^2/\sigma^2)e^{-\frac{x^2}{2\sigma^2}}dx \ll \sqrt{2\pi}\sigma^3/B$ and is considerably larger than the standard deviation of the Gaussian measurement noise, σ.*

An argument supporting this property is as follows. Since $Z = d + E$, the pdf of Y at a point t is

$$f_Y(t) = \int_d^{d+L} f_Z(\tau) f_N(t-\tau) d\tau, \tag{2}$$

where f_Z is the pdf of Z and f_N is the density associated with the Gaussian noise. Taking the derivative with respect to t, we have

$$\frac{df_Y(t)}{dt} = \int_d^{d+L} f_Z(\tau) \frac{df_N(t-\tau)}{dt} d\tau. \tag{3}$$

Setting the second derivative to zero to find the point that achieves the maximum increment in the density of Y, we obtain

$$\frac{d^2 f_Y(t)}{dt^2} = \int_d^{d+L} f_Z(\tau) \frac{d^2 f_N(t-\tau)}{dt^2} d\tau = 0. \tag{4}$$

Since

$$\frac{d^2 f_N(t-\tau)}{dt^2} = -\frac{1}{\sqrt{2\pi}\sigma^3}\left(1 - \frac{(t-\tau)^2}{\sigma^2}\right) e^{-\frac{(t-\tau)^2}{2\sigma^2}}, \tag{5}$$

[1]This research is supported in part by the U.S. National Science Foundation under grant CCR-99-79361, and in part by the New Jersey Center for Wireless Telecommunications.

[2]This research is supported in part by the U.S. National Science Foundation under grants CCR-00-85917 and ECS 02-25481.

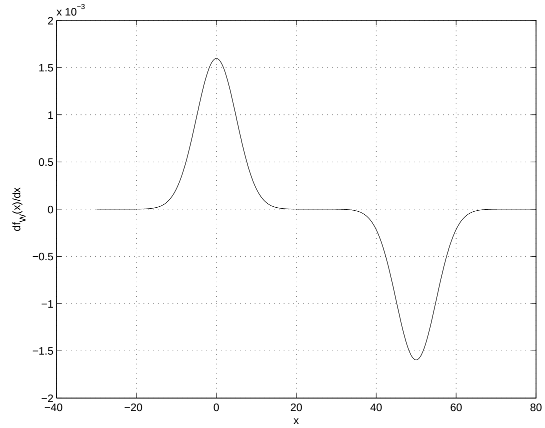

Figure 1: First derivative of the sum of the NLOS error and the Gaussian error for $L = 50$m and $\sigma^2 = 25$m.

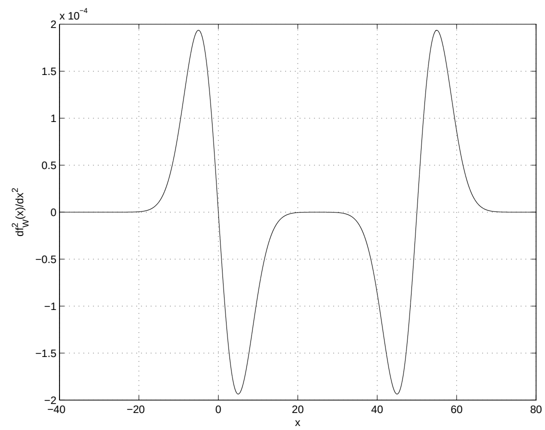

Figure 2: Second derivative of the sum of the NLOS error and the Gaussian error for $L = 50$m and $\sigma^2 = 25$m.

equation (4) leads to

$$\int_d^{d+L} f_Z(\tau)\frac{1}{\sqrt{2\pi}\sigma}\sigma^2 e^{-\frac{(t-\tau)^2}{2\sigma^2}}\,d\tau = \int_d^{d+L} f_Z(\tau)\frac{1}{\sqrt{2\pi}\sigma}(t-\tau)^2 e^{-\frac{(t-\tau)^2}{2\sigma^2}}\,d\tau. \tag{6}$$

If the NLOS error E has a uniform distribution, the associated pdf can be taken out of the integral and cancelled. Furthermore, if the Gaussian noise has almost all of its power between $-L$ and L, then at $t = d$ the left hand side is equal to $\frac{\sigma^2}{2}$ which is equal to the right-hand side.

Thus, for sufficiently large L, the second derivative of the pdf of the observed random variable is approximately zero at true TOA, d. This point is the first (and only, if the NLOS is uniform) positive peak of the first derivative of the observed pdf.

The fact that the second derivative is close zero at d means that the true TOA is close to the saddle point since the second derivative of the pdf of the measurements changes considerably around that point (see Figure 2 where $d = 0$), which can be seen from the non-negligible negativity of the third derivative at $t = d$:

$$\frac{d^3 f_Y(t)}{dt^3} = -\int_d^{d+L} f_Z(\tau)\frac{(\tau-d)}{\sqrt{2\pi}\sigma^5}\left(3 - \frac{(\tau-d)^2}{\sigma^2}\right) e^{-\frac{(\tau-d)^2}{2\sigma^2}}\,d\tau.$$

If the NLOS distribution is not uniform, then the arguments above will not hold in general, especially when the measurement noise variance is comparable to the NLOS error variance. However, in practice, as is argued in [1] and [2], the measurement noise variance is usually much smaller than the NLOS error variance, and the pdf of the NLOS error can be regarded as practically constant for an interval comparable to σ and the above result still holds approximately as shown below.

After a change of variables and using $f_E(x) = f_Z(x + d)$, the expression (6) at $t = d$ can be expressed as

$$\int_0^L f_E(x)\frac{1}{\sqrt{2\pi}\sigma^3}(1 - x^2/\sigma^2)e^{-\frac{x^2}{2\sigma^2}}\,dx. \tag{7}$$

If we consider the integrals from 0 to α and α to L separately, we have approximately

$$f_E(0)\int_0^\alpha \frac{1}{\sqrt{2\pi}\sigma^3}(1 - x^2/\sigma^2)e^{-\frac{x^2}{2\sigma^2}}\,dx + \int_\alpha^L f_E(x)\frac{1}{\sqrt{2\pi}\sigma^3}(1 - x^2/\sigma^2)e^{-\frac{x^2}{2\sigma^2}}\,dx. \tag{8}$$

Due to the conditions in the proposition, the first term in (8) is approximately zero. Also since L is very large compared to σ, $\int_0^L \frac{1}{\sqrt{2\pi}\sigma^3}(1 - x^2/\sigma^2)e^{-\frac{x^2}{2\sigma^2}}\,dx$ is approximately zero. Therefore, $|\int_\alpha^L (1 - x^2/\sigma^2)e^{-\frac{x^2}{2\sigma^2}}\,dx| \ll \sqrt{2\pi}\sigma^3/B$ as well. Hence, the second term is also approximately zero.

Therefore, when the pdf of the NLOS error is almost constant around the origin and has much larger variance than the measurement noise variance, the true TOA will be very close to the point where the second derivative is zero, and might be assumed to be equal to that saddle point without much loss of performance.

As a worst limiting case, where the assumption regarding the relationship between the variances completely fails, the NLOS error density is impulsive around the origin, meaning that there are only LOS observations. In this case the second derivative will be zero at the point $d - \sigma$. Thus, in the worst case for a decreasing density of NLOS, the error in the estimation will be σ. Actually, in this case, not the second derivative but the first derivative being equal to zero characterizes the true TOA. However, it is usually possible to distinguish LOS measurements from the NLOS measurements by some decision criterion [5]. Hence the LOS case can be considered separately.

The estimator is characterized by the pdf of the observed random variable as follows: A pdf estimation algorithm is employed, for which a histogram approach is used, which corresponds to interpreting the observed data as the input to a uniform quantizer with a specified bin size. The number of observations per quantization bin is used to compute the probability mass function (pmf) of the data. The smaller bin edge value corresponding to the quantizer bin having the first local peak in the difference between the current bin and the preceding bin is the estimator output. Note that the choice of the bin size of the quantizer is an important parameter.

The estimator can be formulated as follows. Given m independent identically distributed (iid) measurements $Y_1, ..., Y_m$, we quantize these measurements into N_b bins. Assume that the ith bin extends from $(i-1)l_b$ to il_b where l_b is the bin size. Let $S_1, ..., S_{N_B}$ denote the number of samples in the bins and $D_i = S_i - S_{i-1}$ the difference in the number of samples between adjacent bins. Then, the index of the bin including the true TOA is estimated as

$$\hat{b} = \arg\min_i \{D_i - D_{i-1} > 0 \text{ and } D_{i+1} - D_i \leq 0\}, \quad (9)$$

and $(i-1)l_b$ gives the TOA estimation. Note that it is not the reconstruction value but the bin edge that gives the true TOA estimation, since the bin edge characterizes the increment better.

With this definition of the estimator, we have the following property:

Property 2.2 *Assume that the NLOS error is uniformly distributed with support much larger than the Gaussian measurement error standard deviation. Then, as the number of samples goes to infinity, the error of the estimator is almost surely confined to the interval* [-0.5 l_b, 0.5 l_b] *for small values of* l_b.

Let m denote the number of samples. As $m \longrightarrow \infty$, $S_i/m \longrightarrow p_i$ almost surely, where

$$p_i = \int_{(i-1)l_b}^{il_b} f_Y(y)dy, \quad (10)$$

with Y given by (1). Thus, $(D_i - D_{i-1})/m \longrightarrow p_i - 2p_{i-1} + p_{i-2}$ almost surely.

Let the kth bin contain the true TOA. From the above analysis of the pdf of the measurements, we know that the pdf increases until some point with increasing first derivative and after almost the true TOA it continues to increase with a decreasing first derivative. In other words, the pdf of the measurements is strictly convex before the true TOA and strictly concave after that for some time.

Consider $D_{k+1} - D_k$. For large m, we have

$$\begin{aligned} & (D_{k+1} - D_k)/m \approx p_{k+1} + p_{k-1} - 2p_k \\ = & \int_{kl_b}^{(k+1)l_b} f_Y(y)dy + \int_{(k-2)l_b}^{(k-1)l_b} f_Y(y)dy - 2\int_{(k-1)l_b}^{kl_b} f_Y(y)dy \\ = & \int_{(k-1)l_b}^{kl_b} [f_Y(y+l_b) + f_Y(y-l_b) - 2f_Y(y)]dy. \quad (11) \end{aligned}$$

Since the true TOA is in the kth bin, we can express d as $d = (k-1)l_b + \Delta$ where $0 \leq \Delta \leq l_b$. Defining $W = N + E$ as the sum of the NLOS and measurement errors and using the fact that $f_W(x) = f_Y(x+d)$, the last integral can be written as:

$$\int_{-\Delta}^{l_b - \Delta} [f_W(y+l_b) + f_W(y-l_b) - 2f_W(y)]dy. \quad (12)$$

When the support of the uniform error is much larger than the standard deviation of the Gaussian measurement error and the bin size is sufficiently small, the value of the integral (12) is approximately zero for $\Delta = l_b/2$. From the convexity/concavity of the pdf around the origin, we also have for $\Delta < l_b/2$, $D_{k+1} - D_k < 0$ and for $\Delta > l_b/2$, $D_{k+1} - D_k > 0$.

Since $D_i - D_{i-1} = \int_{-\Delta-(k-i+1)l_b}^{-\Delta-(k-i)l_b} [f_W(y+l) + f_W(y-l) - 2f_W(y)]dy$, the convexity of $f_W(.)$ to the left of the origin and concavity of it to the right of the origin to some extent together with the previous result implies that $D_i - D_{i-1} > 0$ for $i = k, k-1, ...$ and $D_i - D_{i-1} < 0$ for $i = k+2, k+3, ...$ Therefore, the only bins that can satisfy the conditions of the estimator (9) are the kth or the $(k+1)$st bins: For $\Delta < l_b/2$, the kth bin and for $\Delta > l_b/2$, the $(k+1)st$ bin satisfies the conditions of the estimator. Therefore, for $\Delta < l_b/2$, the TOA estimate is $(k-1)l_b$ and for $\Delta > l_b/2$, it is kl_b. Since $d = (k-1)l_b + \Delta$, the error is between $-0.5l_b$ and $0.5l_b$.

Note that for pdf's satisfying the alternative conditions stated in Property 2.1, Property 2.2 is still approximately true.

The histogram can be interpreted as the quantization of the derivative of the pdf of the measurements. In this case, the increments can approximately be represented by $f'_Y((i-1)l_b)\, l_b^2 = p_i - p_{i-1}$ for small bin sizes. Thus the derivative is represented by rectangular bins, and the bin edge of the quantizer can be regarded as the representative sample for the derivative.

The derivative of the density function of the total error with a uniform NLOS noise is depicted in Figure 1. As is seen, the derivative function goes to zero after the peak, and then has a negative peak. This figure can be used to introduce a bound on the bin size for the estimator. The estimator will compare the competing derivatives for the peak and output the peak as the estimated bin. The bins around the upper part of the quantizer will be the candidates for the maximum value, and if those bins do not interfere with the lower portion (negative values) of the derivative curve, the estimation outcome will be at most half the quantizer bin size away from the true TOA (which is 0 in the figure). Assuming symmetry in the upper part of the curve with respective to the peak of the derivative, the bin sizes less than the difference between the peak and the point where the derivative crosses zero, will cause a distortion not greater than $\frac{l_b}{2}$, whereas larger bin sizes may result in interference with the values in the negative portion and cause an error.

The difference between the saddle point (peak of the derivative) of the density of the observation and the maximum point (where the derivative is zero) is on the order of a few standard deviations of the noise. Thus the quantizer bin size should be upper bounded by a few σ's.

III. Mobile Location

Location of a mobile in a wireless communications system can be determined from TOA measurements from a number of base stations (BSs). Traditionally, the location of the mobile is estimated by a least squares (LS) estimator [4], which is the optimal estimator when the mobile is LOS to all the BS's. When the NLOS error is not known, this approach is still employed and the NLOS observations are given less weight in order to mitigate the effect of the NLOS error.

If the locations of the mobile and the BS's are denoted by $\mathbf{p}_m$ and $\mathbf{p}_1, ..., \mathbf{p}_b$ respectively with b being the number of BS's, the location estimate is given by the $\mathbf{p}$ that minimizes

$$\sum_{i=1}^{b} w_i(\|\mathbf{p} - \mathbf{p}_i\| - vY_i)^2, \quad (13)$$

where $\|.\|$ denotes L2-norm, v is the speed of light, Y_i is the TOA measurement from the ith BS and w_i is the weight that reflects the reliability of the measurement from the ith BS, which might be considered to be inversely proportional to the variance of the measurements.

In NLOS situations, even though the NLOS BS is given smaller weight, the error will still be significant unless there

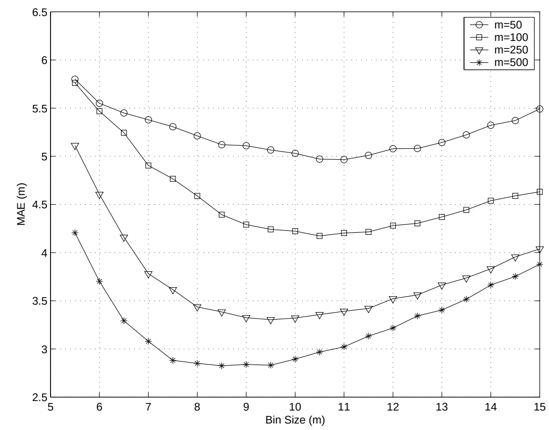

Figure 3: Mean Absolute Error versus bin size for different sample sizes for the uniform NLOS error, $[0, 50\text{m}]$, and Gaussian measurement error with variance 25m^2.

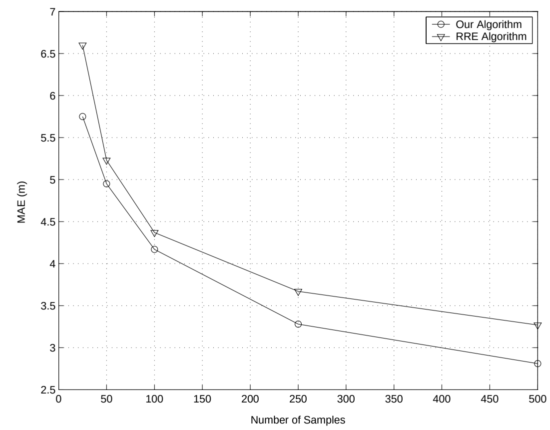

Figure 4: Mean Absolute Error for different algorithms for the uniform NLOS error, $[0, 50\text{m}]$, and Gaussian measurement error with variance 25m^2.

are at least three LOS BS's (for two-dimensional location). Therefore, we suggest employing our TOA estimation technique for the NLOS BS's before using the LS estimator. We assume that the NLOS BS's can be determined from the measurements using a variance test [5].

Consider a series of TOA measurements from a NLOS BS. To apply the TOA estimation technique, consider a window around the specific TOA measurement and take those measurements as input to the estimator. Note that the samples should be sufficiently spaced in time so that they are independent. Then, for NLOS BS's Y_i in equation (13) is replaced by the result of the TOA estimation and the final location is estimated by the LS scheme.

IV. Simulation Results

Consider the model in (1) where the Gaussian measurement error has variance σ^2 and the NLOS error is uniform between 0 and L. Assume that m iid TOA measurements are taken according to this model and the location of the mobile does not change considerably during these measurements. This assumption means that the NLOS error is assumed to be independent from sampling instant to sampling instant, which is due to the dynamic environment and/or the movement of the mobile, but the mobile's location can be considered to be constant for the purposes of TOA estimation. These two seemingly inconsistent assumptions can be reconciled since the NLOS error variations are small-scale variations while the mobile's location is a large-scale property.

For the simulations, the observed TOA data have been converted to distance measurements for convenience. The variance of the Gaussian measurement error is taken to be 25 m^2. In Figure 3, the NLOS error is taken to be uniformly distributed between 0 and 50 m and the mean absolute error (MAE) is calculated for different bin sizes, l_b. For each bin size, 20,000 trials were performed with the true distance uniformly distributed between some range so that the probability that the true distance is at any location in a bin is always uniformly distributed. From the figure, it is seen that there is an optimal bin size for each sample size, m. For small bin sizes, number of samples may not be sufficient to represent the derivative value correctly resulting in higher errors. For large bin sizes, even though the correct bin can be chosen, the estimation value may not be very close to the true distance due to the large size of the bin. Also note that when the sample size increases, the optimal value of bin size decreases.

Figure 4 compares the performance of the estimator to the Robust Range Estimation (RRE) algorithm [1]. The error of the RRE algorithm is higher than the error of our estimator since the assumption used in the RRE algorithm, namely that the measurements lower than the true distance comes from the Gaussian error assumption, may cause some small errors. However, that algorithm does not require specific properties of the NLOS pdf which are assumed for our estimator. But in practical situations the conditions on the pdf are not so tight and are satisfied by most NLOS error distributions.

In Figure 5, the NLOS error is taken to be exponentially distributed with mean 20 m and the mean absolute error (MAE) is calculated for different bin sizes. For each bin size, 20,000 trials were performed with the probability that the true distance is at any location in a bin again uniformly distributed. From the figure, the same observations as the uniform NLOS error case can be made. Also the algorithm clearly works for the exponential NLOS error as well since the variance of the NLOS error is considerably larger than the Gaussian measurement error variance.

Figure 6 compares the two estimators. Again the algorithm proposed in this paper achieves lower error values. However, in this case, the results are closer since the NLOS pdf is exponentially distributed and Proposition 2.1 applies since it has a large variance compared to the Gaussian error.

Now consider the mobile tracking scenario shown in Figure 7, in which there are three BS'a at $(0, 0)$, $(300, 0)$ and $(150, 259.8)$ and the mobile moves from $(150, 80)$ to $(250, 80)$ on the solid line shown with a velocity of 10 m/s. Twenty TOA measurements are taken per second. The first BS is assumed to be NLOS with an exponential NLOS error with mean 20 m, and the other BS's are assumed to be LOS. The Gaussian measurement error variance is 25 m^2 for all measurements. 20 samples around a given sample are considered for the TOA estimation for the NLOS BS and a bin size of 12 m is used for the quantization. Also the weights are taken to be inversely proportional to the variances of the measurements for the LS

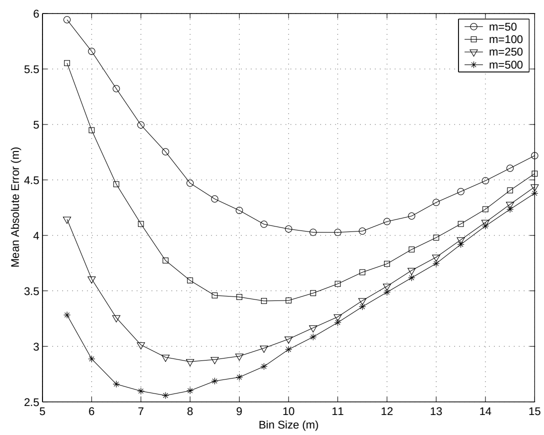

Figure 5: Mean Absolute Error versus bin size for different sample sizes for the exponential NLOS error with mean 20m and Gaussian measurement error with variance $25m^2$.

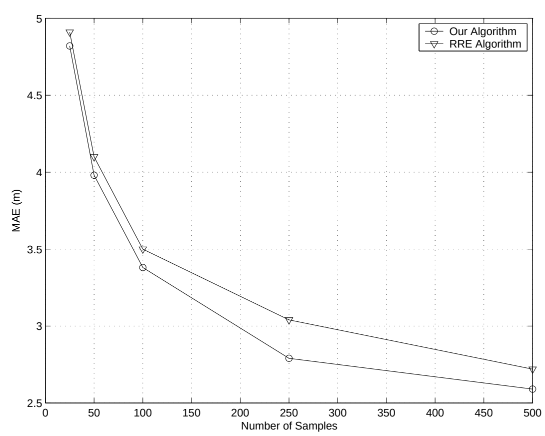

Figure 6: Mean Absolute Error for different algorithms for the exponential NLOS error with mean 20m and Gaussian measurement error with variance $25m^2$.

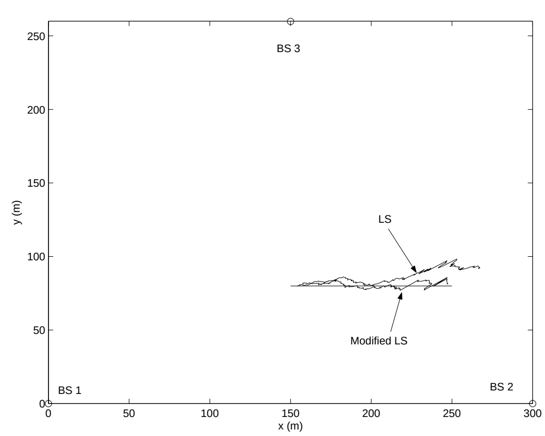

Figure 7: Mobile Tracking Scenario. Mobile moves from left to right with velocity 10m/s and 20 TOA measurements are taken per second. Only BS1 is NLOS.

estimator. From the figure it is seen that employing our TOA technique for the NLOS BS improves the location accuracy. The MAE for the traditional LS algorithm is 10.6 m whereas it is 2.93 m for the modified scheme.

V. Conclusion

In this paper we have introduced a practical algorithm to estimate true TOA of a signal between a mobile and a base station in the presence of NLOS noise with unknown statistics and Gaussian noise with a variance considerably smaller than that of the NLOS error. It has been shown that under reasonable conditions, TOA estimation can be performed by locating the first peak of the first derivative of the pdf of the measurements. A histogram approach can then be employed to approximate the pdf of the measurements and thus obtain this estimate. The simulations are performed to asses the performance of the estimator. Our simulations show the superiority of this approach to existing methods.

References

[1] M. P. Wylie-Green and S. S. Wang, "Robust range estimation in the presence of the non-line-of-sight error," *Proc. IEEE 54th Vehicular Technology Conference, Fall 2001* vol. 1, Atlantic City, NJ, pp. 101-105.

[2] M. P. Wylie and J. Holtzman, "The non-line of sight problem in mobile location estimation," *Proc. 5th IEEE International Conference on Universal Personal Communications, 1996*, Cambridge, MA, pp. 827 -831.

[3] P. C. Chen, "A non-line-of-sight error mitigation algorithm in location estimation," *Proc. IEEE Wireless Communications and Networking Conference, 1999*, vol. 1, New Orleans, LA, pp. 316-320.

[4] J. Caffery, Jr. and G. Stuber, "Subscriber location in CDMA cellular networks," *IEEE Trans. Vehicular Technology*, vol. 47, pp. 406-416, May 1998.

[5] J. Borras, P. Hatrack and N. B. Mandayam, "Decision theoretical framework for NLOS identification," *Proc. IEEE Vehicular Technology Conference, Spring 1998* vol. 2, Ottawa, Canada, pp. 1583-1587.