A Non-Line-of-Sight Error Mitigation Algorithm in Location Estimation

Pi-Chun Chen
pcchen@winlab.rutgers.edu
Wireless Information Network Laboratory (WINLAB)
Department of Electrical and Computer Engineering
Rutgers University

Abstract — The location estimation of mobile telephones is of great current interest. The two sources of range measurement errors in geolocation techniques are measuring error and Non-Line-of-Sight (NLOS) error. NLOS errors, derived from the blocking of direct paths, have been considered as a killer issue in the location estimation. In this paper we develop an algorithm to mitigate the NLOS errors in location estimation when the range measurements corrupted by NLOS errors are not identifiable. This algorithm, referred to as Residual weighting algorithm (Rwgh), assumes that the number of range measurements is greater than the minimum number required. The Rwgh algorithm is compared with two other generic algorithms and tested under various scenarios and NLOS error models, which includes deterministic and random NLOS error models. In both cases, Rwgh has shown the effectiveness of NLOS error mitigation in location estimation.

I. INTRODUCTION

Providing the accurate location information of the Mobile Station (MS) for Emergent 911 call has become mandatory in the States [1], and is of great interest for location related applications. Geolocation techniques, such as Time of Arrival (TOA), Time Difference of Arrival (TDOA), and signal strength measurements are often used to obtain distance information (range measurements) between transmitters and receivers. The MS location estimate can be computed based on a set of range measurements. The assumption in applying these geolocation techniques is that received signals propagate through Line-of-Sight (LOS) paths. Violation of this assumption will introduce NLOS (Non-Line-of-Sight) errors in range measurements and will lead to erroneous location estimate.

A LOS path, also called a direct path, is a straight line path that connects the transmitter and the receiver. With the absence of the LOS path, the transmitted signal could only reach the receiver through reflected, diffracted, or scattered paths. The NLOS error (in unit of distance) is defined as the excessive traveling distance with respect to the direct path.

Accordingly, NLOS errors are always positive, and could range from a small number to 1300m [2][3], depending on the propagation environment. Under certain circumstances, NLOS range measurements, the range measurements corrupted by NLOS errors, could bias the true range substantially and cause an inaccurate location estimate.

To protect location estimates from NLOS error corruption, the NLOS identification technique has been proposed. The idea is to find some distinct properties of NLOS range measurements and develop hypothesis tests to segregate NLOS measurements from LOS measurements. For instance, reference [3] observes that NLOS range measurements have greater variance than the variance of LOS range measurements, and develops a hypothesis test to identify NLOS range measurements based on a consecutive sequence of range measurements. With complete knowledge of range measurement error models, more effective hypothesis tests have been developed in resolving the NLOS measurements. Reference [4] models NLOS and LOS range measurements as Gaussian random variables, and derives the theoretical framework for nonparametric and parametric hypothesis tests.

The major problem associated with the NLOS identification approach is the lack of a NLOS error model and the absence of field measurements that relate NLOS errors to signal strength measurements. Field measurements done in the past concentrated on the delay spread, but not on the NLOS-LOS channel condition. This limits the implementations of NLOS identification technique.

The research reported in this work considers the situation in which NLOS measurements are unrecognizable. We propose a NLOS mitigation technique, referred to as Residual Weighting Algorithm (Rwgh), to alleviate the effect of NLOS error in the location estimate. The main advantage of Rwgh is that no statistical models or prior information on LOS-NLOS channel conditions are needed. The only assumption is that the number of the range measurements has to be greater than the minimum required.

The paper is organized as follows. Section II derives the Rwgh algorithm step by step. Section III compares the performance of Rwgh with other algorithms under various scenarios and NLOS error models. The NLOS error models include both deterministic and random NLOS errors to order to thoroughly apprehend the performance of Rwgh. Simulation results and conclusions are presented in Section IV and V.

II. NLOS MITIGATION

The MS location problem is formulated as an estimation problem. The observation space is a set of range measurements (assume TOA measurements in this paper), and the parameters that need to be estimated are the geographical coordinates of the MS. Without any prior information on the statistics of the range measurements, a LS estimator is used to determine the location of the MS. That is,

$$\hat{x} = \arg \min_x \sum_{i \in S} (r_i - ||x - X_i||)^2.$$  (1)

||x|| denotes the norm operation over a vector.
||x - X_i|| represents the distance between vectors x and X_i,

where $S$ is the BSs index set, $r_i$ is the range measurement from the MS to i-th BS, $i \in S$, $x = [x, y]^T$ MS position in Cartesian coordinates, $\hat{x}$ is the estimate of vector x, $X_i = [X_i, Y_i]^T$ the coordinate vector of the i-th BS position in Cartesian coordinates, $(r_i - ||x - X_i||)$ is called the i-th residual for a particular x.

The range measurement, $r_i$, is contaminated by two types of errors: measuring errors and NLOS errors. The measuring

---

*Since June 1999, the author has joined Telcordia Technologies as a research scientist in Red Bank, New Jersey.
error results from the measuring processes in a noisy channel, and can be improved with better Signal-to-Noise Ratio (SNR) of the received signals or receiver structures. The NLOS error derives from the blocking of the direct paths, and can’t be improved with general approach.

Equation (1) implies that the LS location solution, \( \hat{x} \), is the estimate that minimizes the sum of the residual squares over the data set. For convenience, we define \( R_{res}(x; S) \) as the sum of the residual squares of \( x \) over the range measurements set \( S \).

\[
R_{res}(x; S) = \sum_{i \in S} (r_i - (x - X_i))^2
\]  
(2)

Thus, (1) could be rewritten as

\[
\hat{x} = \arg \min_x R_{res}(x; S).
\]  
(3)

The LS location solution, \( \hat{x} \), is the estimate that minimizes \( R_{res}(x; S) \). That is

\[
R_{res}(\hat{x}; S) = \min_x R_{res}(x; S).
\]  
(4)

Equation (4) states that, in the least square sense, a good estimate is the one with minimum residual.

When there is a NLOS range measurement in a data set, the residual of the estimate is likely to be greater than the residual when there is not a NLOS range measurement. This observation is easy to perceive but difficult to prove. We illustrate this point with a simple example. Consider a set of error-free range measurements. The location estimate will be accurately determined using the LS algorithm and the residual will be zero. Now if we introduce a NLOS error in one of the range measurements, the location estimate will deviate from the optimal one and cause a greater residual. Furthermore, if we consider the other error source, the measuring error, the situation will become more complicated. It is possible to find some counter examples where NLOS errors in a data set could result in a smaller residual than the residual when only measuring errors are present. However, those special cases only occur with low probability. If the range measurement error is NLOS error limited, the observation remains true.

Proceeding along this line, we associate the quality of an estimate \( \hat{x} \) with its \( R_{res}(\hat{x}; S) \) and will show how our proposed Rwgh algorithm could alleviate the effect of NLOS errors.

The assumption that the number of range measurements is greater than the minimum required grants us the freedom of performing various combinations of range measurements. To determine a two dimensional MS location, the minimum number of measurements (TOAs) required is three. For example, if we have \( M \) \((M>3)\) range measurements from each of a different BS, we can group those range measurements in various ways subject to the constraint that the number of range measurements in each group is no less than 3. Take \( M=5 \) for example, there are 16 eligible range measurement combinations. One could choose all 5 range measurements, or select 4 out of 5, or select 3 out of 5. That is,

1. select 5 out of 5 : \( \binom{5}{5} = 1 \) combination
2. select 4 out of 5 : \( \binom{5}{4} = 5 \) combinations
3. select 3 out of 5 : \( \binom{5}{3} = 10 \) combinations

Overall, there are 16 different range measurement combinations. Applying the LS estimator on these combinations (using equation(1)), we can obtain 16 MS location estimates, which are denoted as intermediate location estimates.

Since these groups are commutatively constructed, some of the groups will contain no NLOS measurements or less NLOS measurements than the others. If we could rely more on the estimates derived from those “good” groups, the impact of NLOS measurements can be reduced. As \( R_{res}(x; S) \) is a good indicator in evaluating the quality of an estimate \( x \), we will weight the estimate \( \hat{x} \) based on its \( R_{res}(\hat{x}; S) \). However, the numbers of range measurements in the groups are different. Therefore we define the normalized \( R_{res} \) to remove the dependence on the size of the group.

\[
\tilde{R}_{res}(\hat{x}; S) = \frac{R_{res}(\hat{x}; S)}{\text{Size of } S}
\]  
(5)

As a good estimate tends to have a smaller \( \tilde{R}_{res} \), we weight each estimate according to its \( \tilde{R}_{res} \). In consequence, the final Rwgh location estimate is the linear combination of the intermediate location estimates weighted inversely to their \( \tilde{R}_{res} \).

By doing this, Rwgh automatically suppresses NLOS measurements in the final location estimate, and thereby achieves the goal of NLOS mitigation. In summary, the Rwgh algorithm consists of the following steps.

1. Given \( M \) \((M>3)\) range measurements (from \( M \) different BSs), form

\[
N = \sum_{i=3}^{M} \binom{M}{i}
\]  
(6)

Each combination is represented by a BS index set \( S_k \) \((k = 1, 2, ..., N)\).

2. For each combination, compute the intermediate LS estimate of \( x \),

\[
\hat{x}_k = \arg \min_x R_{res}(x; S_k)
\]  
(7)

and \( \tilde{R}_{res} \)

\[
\tilde{R}_{res}(\hat{x}_k; S_k) = \frac{R_{res}(\hat{x}_k; S_k)}{\text{size of } S_k}, \quad \forall k.
\]  
(8)

3. Find the final estimate of \( x \) as the weighted linear combination of the intermediate estimates from step 2. The weight is inversely proportional to \( \tilde{R}_{res} \) of the estimate. Mathematically,

\[
\hat{x} = \frac{\sum_{k=1}^{N} \tilde{R}_{res}(\hat{x}_k; S_k)^{-1} \hat{x}_k}{\sum_{k=1}^{N} \tilde{R}_{res}(\hat{x}_k; S_k)^{-1}}
\]  
(9)

III. PERFORMANCE EVALUATION

The performance of the Rwgh is evaluated through simulations. Figure 1 shows the simulation structure. Range measurements are generated by adding measuring errors to the true ranges. To generate NLOS measurements, NLOS errors are inserted to range measurements in addition to the measurement errors. The generation of NLOS errors is crucial in evaluating Rwgh. In order to have a general representation of NLOS errors, we adopt two types of NLOS error models, deterministic and random NLOS error models. Both models supplement each other in the understanding of Rwgh performance.

The inputs to the location algorithms are range measurements and the corresponding BSs coordinates. The performance criteria of the algorithms is chosen as the Root Mean Square Error (RMSE),

\[
\text{RMSE} = \sqrt{\sigma_x^2 + \sigma_y^2}
\]  
(10)

where \( \sigma_x^2 \) and \( \sigma_y^2 \) are the error variances of the location estimate along x and y directions in Cartesian coordinates.

We compare Rwgh with two other generic algorithms under three scenarios described in the next two sections.
Additionally, we also introduce three other variations of Rwhg algorithms to explore the potential improvement. These variations are Rmin, Rwhg(1/4), and Rwhg(1/2). The details of these algorithms are described below.

**Whole**: The location estimate is the LS estimate based on all the range measurements available without any selection.

**Good**: The location estimate is the LS estimate based on only the LOS range measurements.

**Rwhg**: As described in Section II.

**Rwhg(1/2)**: The same as Rwhg but rank the intermediate estimates in terms of their residuals in increasing order and only use the first half of the ranked intermediate estimates to form the linear weighted combination.

**Rwhg(1/4)**: The same as Rwhg but rank the intermediate estimates in terms of their residuals in increasing order and only use the first quarter of the ranked intermediate estimates to form the linear weighted combination.

**Rmin**: The same as Rwhg but find the estimate as the intermediate estimate with minimum residual.

### B. Simulation Examples

Three scenarios considered in our simulations are:

1. Case 4/1: One out of four range measurements is contaminated by NLOS error.
2. Case 5/1: One out of five range measurements is contaminated by NLOS error.
3. Case 5/2: Two out of five range measurements are contaminated by NLOS error.

### C. Measurement Error Models

All range measurements are corrupted by measuring errors. In case of a NLOS channel condition, the range measurement is further degraded by a NLOS error.

#### C.1 Measuring Error Model

We model the measurement error as a Gaussian random variable with zero mean and standard deviation of 60m. The credibility of the Gaussian assumption is verified in [5].

#### C.2 NLOS Error Models

- **NLOS Error as a Deterministic Variable**
  NLOS errors depend on the propagation environments and change from time to time. However, at a fixed time instance, the NLOS error could be treated as a constant. The deterministic NLOS error model allows us to parameterize the NLOS error in the simulations. In this group of simulations, we vary NLOS errors from 100m to 1300m, and calculate the RMSE as a function of NLOS errors.

- **NLOS Error as a Random Variable**
  Random NLOS errors can be derived from the delay profiles described by a probability density function of excessive propagation delay with respect to a direct path. NLOS errors are obtained as the excessive delay multiplied by the speed of light. Three frequently used delay profiles suggest models for generating the random NLOS errors. They are an exponential, a uniform, and a delta random variable. In this paper, we only present the results with the exponential model, the results with the other models can be found in [6]. The Exponential Distribution error model [7][8] is,

\[
D(\tau) = \begin{cases} 
\frac{1}{\tau_{rms}} e^{-\frac{\tau}{\tau_{rms}}}, & \tau > 0 \\
0 & \text{otherwise} 
\end{cases}
\]

where \(\tau_{rms}\) is the delay spread, which depends on the propagation environment. Reference [9] suggests that \(\tau_{rms}\) is lognormal.
distributed and could be further characterized by four environmental dependence variables.

\[ \tau_{rms} = T_1 d^\epsilon \xi \]  

where

- \( T_1 \): the median value of \( \tau_{rms} \) at \( d = 1 \) km
- \( d \): the distance between the transmitter and receiver in kilometers
- \( \epsilon \): an exponent that lies between 0.5-1.0
- \( \xi \): a lognormal random variable. Specifically, 10 \( \log \xi \) is a Gaussian random variable having zero mean and a standard deviation, \( \sigma_\xi \), that lies between 2-6 dB.

The four environment types are Bad urban, Urban, Suburban, and Rural. Their parameter settings are listed in Table I.

<table>
<thead>
<tr>
<th>Environment Type</th>
<th>( T_1 (\mu s) )</th>
<th>( \epsilon )</th>
<th>( \sigma_\xi )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bad Urban</td>
<td>1.0</td>
<td>0.5</td>
<td>4 dB</td>
</tr>
<tr>
<td>Urban</td>
<td>0.4</td>
<td>0.5</td>
<td>4 dB</td>
</tr>
<tr>
<td>Suburban</td>
<td>0.3</td>
<td>0.5</td>
<td>4 dB</td>
</tr>
<tr>
<td>Rural</td>
<td>0.1</td>
<td>0.5</td>
<td>4 dB</td>
</tr>
</tbody>
</table>

**TABLE I**

**PARAMETER VALUES FOR FOUR DIFFERENT ENVIRONMENTAL TYPES**

IV. SIMULATION RESULTS

The performance of Rwhg algorithm not only depends on the measuring and NLOS errors, but also on the geographical location of the MS. To remove the MS location dependence in the performance evaluation, all the results presented here are the averaged results over the MS positions that are uniformly distributed in the coverage area.

A. Deterministic NLOS Errors

Figure 3 - Figure 5 show the comparisons of RMSE as a function of NLOS error among six different algorithms for the three scenarios. 60m and 30m of the standard deviation of the measuring error are used.

As expected in all figures, the Whole algorithm results in the poorest RMSE performance among the six algorithms. Rwhg and its three variations (Rmin, Rwhg(1/4), Rwhg(1/2)) have very similar performance. By and large, Rwhg outperforms its variations. The RMSE of the Whole algorithm is linearly proportional to the NLOS error, which is consistent with the observation in [2]. The slope of these lines represents the RMSE error increment caused by 1 meter increase in NLOS error. Case 5/1 has the smallest slope, which implies that the NLOS error has less impact on the location estimate when there are more LOS range measurements in the data set. Case 5/2 has the biggest slope, and is the worst case among the three.

The performance of the Good algorithm sets the optimum performance that all the algorithms could achieve. The RMSE of Rwhg is close to the optimum with a small standard deviation of measuring errors. The RMSE of Rwhg saturates as the NLOS error increases, and is about 50-100m above the optimum in Case 4/1 and Case 5/1. However, in Case 5/2, the RMSE increases as the NLOS error increases but the rate of increase is approximately half of that of the Whole algorithm. This substantiates Rwhg algorithm’s NLOS mitigation effect.

B. Random NLOS Errors

The simulation results of random NLOS errors are presented in 3 figures. Figures 6, 7, and 8 illustrate the results of Case 4/1, Case 5/1, and Case 5/2 with exponential delay profiles respectively. Each figure contains four groups of bar plots. Each group corresponds to one of the four environment types. Within a group, the RMS values derived from the six algorithms are represented by the heights of six bars. The standard deviation of the measuring error is 60m.
deviation of measurement errors in this group of simulations is 60m.

On the whole, the performance ranking from the worst to the best with respect to the environment types are: Bad Urban, Urban, Suburban, and Rural. As for the algorithm used, the ranking order from the worst to the best is: Whole, Rmin, Rwgh(1/4), Rwgh(1/2), Rwgh, and Good. Similar to the results of deterministic NLOS errors, the Rwgh and its variations have very close RMSE values.

Compared with the Whole algorithm, the performance improvement of the Rwgh varies for different scenarios and environment types. Generally speaking, in an urban environment, the RMSE is improved by approximately 1000 m; in a bad urban environment, the improvement even goes up to 3000 - 5000 m. For a suburban environment, the improvement is around 600 - 900 m; as for a rural environment, in which NLOS errors are not severe and do not leave much space for improvement, the RMSE is decreased by 100 - 200m.

V. SUMMARY

The NLOS problem has long been considered a killer issue in location estimation. The proposed Rwgh algorithm provides a way of reducing the NLOS errors conveyed into the location estimate without prior information on error models.

The simulation results provide evidence on the NLOS error mitigation effect of Rwgh algorithm. For example, from the results of the deterministic NLOS error model,

- The performance improvement of Rwgh is substantial when NLOS measurements are not distinguishable.
- Rwgh performs as if only LOS measurements are used when the standard deviation of measurement errors is small.
- The error increment of location estimation error per NLOS error is dramatically reduced by Rwgh.

The results of random NLOS errors show that

- The Rwgh algorithm reduces the RMSE error by hundreds or even thousands of meters.

REFERENCES


measurements are less than three as well. From the simulation results, we found that Rwgh algorithm still greatly outperforms the Whole algorithm (especially in the Bad urban environment). To avoid mistakenly using a location estimate derived from a data set primarily composed of NLOS measurements, we could choose the estimates with their residuals lower than a certain threshold. The choice of the threshold could be derived from the location accuracy requirements or Cramer-Rao Lower Bound (CRLB) of the estimator.