Non-Line-of-Sight Error Mitigation for Range Estimation in Dynamic Environments

Qinghua Wang\textsuperscript{1} and Ilangko Balasingham\textsuperscript{1,2,3}

\textsuperscript{1}Department of Electronics and Telecommunications, Norwegian University of Science and Technology, Trondheim N-7491, Norway
\textsuperscript{2}The Interventional Centre, Oslo University Hospital, Oslo N-0027, Norway
\textsuperscript{3}Institute of Clinical Medicine, University of Oslo, Oslo N-0316, Norway

Abstract—Localization is an important component in many applications such as the promising Ultra-Wideband (UWB) wireless sensor network for medical treatment. For the majority of localization technologies, it is essential to measure the ranges between a target and several reference nodes before the target can be localized. Existing range estimation techniques rely on the measurements of time-of-arrival (TOA) and received-signal-strength (RSS) which suffer from environmental change. Dynamic environment such as human mobility can cause non-line-of-sight (NLOS) measurements which will significantly degrade the accuracy of range estimation. Therefore, range estimation methods are desired to be robust to NLOS measurements. In this paper, it is proposed to use hypothesis tests to identify whether there are NLOS measurements mixed in with the measurement dataset. For those NLOS corrupted measurement datasets, a new range estimation method based on a Log-normal model is found to be capable of reducing the range estimation error. Another advantage of this new range estimation method is that NLOS measurements are not required to be excluded from its analysis. However, simulation results show that the range estimation accuracy can be further improved if NLOS measurements are excluded.

Index Terms—Localization, NLOS error, range estimation.

I. INTRODUCTION

Localization is an important service which is demanded by many applications. For example, automatic driving and navigation are based on fine-grained localization. Location information of mobile phone users can be utilized by authorities in case of emergency. In an Ultra-Wideband (UWB) medical sensor network, a miniaturized capsule endoscope needs to be accurately localized in order to have its sensed information to be meaningful [1]. In the past years, many localization strategies have been proposed. A widely used localization strategy is GPS, which provides reliable and cost-effective localization for outdoor applications. GPS uses “triangulation” to calculate a target’s location. To “triangulate”, a GPS receiver measures ranges (distances) to satellites using the travel time of radio signals. To localize a target on earth, range measurements to at least three satellites are required. Similar localization strategies can be used for indoor and even in-body environments. However, in an indoor or in-body environment, radio-based range measurements suffer from obstacles and multi-path effects. A received non-line-of-sight (NLOS) radio signal can greatly mislead the range estimation because it may come from a path which is much longer than the line-of-sight (LOS) range. Therefore, the mitigation of NLOS errors are vital for accurate indoor and in-body range estimation, and requires careful study.

In a dynamic environment such as the indoor and in-body environments, NLOS range measurements can be caused by different sources, including static features in the environment, moving obstacles, or a moving sender/receiver, etc. One example of this scenario is a capsule endoscope swallowed by a patient to inspect the digestive tract [1]. In this example, the capsule endoscope functions as an in-body sensor. It regularly transmits the in-body images wirelessly to those on-body receivers attached to the body. In order to have the received in-body images meaningful, it is vital important to know the accurate location of the capsule endoscope. The location of the capsule endoscope can be estimated through “triangulation” or “multiangulation” based on the ranges measured between the capsule endoscope (i.e. the transmitter) and the receivers on the body. However, the signals received at the on-body receivers can be NLOS due to the tissue movement (e.g. blood) or organ movement (e.g. heart beat). Given the seriousness of any medical error, it is vital important to identify those NLOS corrupted range measurements and mitigate their effects on medical localization. Another example is to measure the range between a radio sender and a radio receiver lying in the two diagonally opposite corners in a lobby hall or in an office environment. Due to the complicated human movement, a LOS signal can be obstructed by different people appearing on different locations and at different time. Obviously, any range estimate based on these NLOS corrupted measurements could lead to a large and unacceptable error.

In the above depicted scenarios, it seems to be very difficult to identify those NLOS measurements and mitigate their effects on range measurement or localization. Fortunately, there are some features we can use to characterize NLOS measurements. One feature is that NLOS measurements are usually larger than LOS measurements, and this is determined by the fact that an NLOS path is always longer than the LOS...
path. Another feature is that NLOS measurements experience a larger variance than that experienced by LOS measurements, and this is determined by the fact that there are much more uncertainty for NLOS measurements than for LOS measurements. In this paper, these two features are used to identify whether a measurement dataset has been corrupted by NLOS measurements or not.

If a measurement dataset is found to be corrupted by NLOS measurements, a good idea is to simply discard this dataset as many papers have suggested [2], [3]. However, this simple approach is not always appropriate. If the range measurement is done in a highly dynamic environment, it can be difficult to acquire a non-NLOS-corrupted measurement dataset. In addition, real-time applications will demand constant range estimation even a non-NLOS-corrupted measurement dataset is not available. Therefore, it is necessary to design some estimators which are not sensitive to NLOS errors while still efficient when there is no or few NLOS errors.

In this paper, a new range estimation method is proposed to provide range estimates based on NLOS corrupted measurement datasets. The assumption is that, besides the possible existence of NLOS measurements, there are an adequate number of LOS measurements in the dataset. Simulation results show that this new range estimation method can achieve a much higher estimation accuracy than what is provided by the sample mean when NLOS errors are present. Although this new range estimation method does not require the identification and the subsequent exclusion of NLOS measurements from the dataset, simulations show that the performance of this new range estimation method can be further improved if NLOS measurements are excluded from analysis.

The paper is organized as follows. Section II presents the hypothesis tests which can be used to identify those LOS or NLOS corrupted measurement datasets. Section III proposes a new range estimation method which estimates the true range from a set of NLOS measurements. In Section IV, simulations are designed to investigate the performance of the proposed range estimation method. Finally, Section V concludes this paper.

II. IDENTIFYING NLOS CORRUPTED MEASUREMENTS

Let \( X = \{x_1, x_2, ..., x_N\} \) be the dataset of \( N \) sample range measurements. There could be NLOS corrupted range measurements mixed in with the \( N \) samples, but there is no a priori information about which sample is NLOS corrupted. For LOS measurements, it has been shown that the pdf can be well characterized by a Gaussian model which is analogous to the situation of corruption due to receiver noise [4], [5]. This leads to the pdf being \( f_{\text{los}}(x) \sim N(\mu, \sigma_{\text{los}}^2) \), where \( \mu \) represents the true range and is always assumed to be a deterministic unknown quantity. Because the LOS measurements have a Gaussian distribution, the sample mean \( \bar{x} = \frac{1}{N} \sum_{n=1}^{N} x_n \) can be used as the unbiased estimate of the true (i.e. LOS) range in the case that all the measurements in the dataset \( X \) are LOS. However, the sample mean is not any more an unbiased estimate when the dataset is NLOS corrupted. Therefore, it is important to identify whether a dataset is corrupted by NLOS measurements or not.

Usually, it is not difficult to identify whether there are NLOS measurements mixed in with the dataset if the measurements are supervised. However, it can be difficult to identify an NLOS corrupted dataset when the measurements are not supervised. In this case, we can refer to a binary hypothesis test [4], [5] in judging whether the dataset has been corrupted by NLOS measurements. Because we are considering a dynamic range measurement environment, the NLOS error itself can be viewed as a random variable. This random variable should only be supported by positive real numbers because an NLOS range is always larger than the LOS range. Therefore, the mean of NLOS corrupted measurements is expected to be larger than the mean of LOS measurements. Meanwhile, the distribution of a random variable has a non-zero variance as opposed to the situation of a deterministic value. Considering that there are also additive receiver noise in NLOS measurements, the distribution of NLOS measurements will have a larger variance than that of the distribution of LOS measurements.

Consequently, the distribution of the mixed LOS and NLOS measurements will expect a larger variance than that of the distribution of a set of pure LOS measurements.

Let \( r \) be the true range and \( \sigma_{\text{los}}^2 \) be the variance of the distribution of LOS measurements. Let \( \mu \) be the population mean and \( \sigma^2 \) be the population variance of the sample. According to the above analysis, a hypothesis test can be built:

\[
H_0 : \mu = r, \sigma^2 = \sigma_{\text{los}}^2 \tag{1}
\]
\[
H_1 : \mu > r, \sigma^2 > \sigma_{\text{los}}^2 \tag{2}
\]

where the null hypothesis \( H_0 \) represents the situation when all measurements are LOS and the sample comes from a Gaussian distribution, and \( H_1 \) represents the situation when there are NLOS measurements mixed in with the dataset and the sample does not come from a Gaussian distribution.

Because the values of \( \mu \) and \( \sigma \) are correlated, the hypothesis can be either tested by comparing \( \mu \) with \( r \) or comparing \( \sigma \) with \( \sigma_{\text{los}} \).

If the true range \( r \) is known or can be estimated, a t-test can be applied:

\[
t = \frac{\bar{x} - r}{s/\sqrt{N}} \tag{3}
\]

where \( N \) is the sample size, \( \bar{x} \) is the sample mean and \( s \) is the sample standard variance. If \( H_0 \) is true, the variable \( t \) has a Student’s t-distribution with \( N-1 \) degrees of freedom. Given a significance level \( \alpha \), \( H_0 \) should be rejected if \( t > t_{\alpha, N-1} \).

Instead, if \( \sigma_{\text{los}} \) is known, a chi-squared test can be applied:

\[
\chi^2 = \frac{(N-1)s^2}{\sigma_{\text{los}}^2} \tag{4}
\]

The variance experienced by NLOS measurements is due to the uncertainty of error sources and the uncertainty of receiver noise. The variance experienced by LOS measurements is only due to the uncertainty of receiver noise.
where $N$ is the sample size and $s^2$ is the sample variance. If $H_0$ is true, $\chi^2$ is a value of the chi-squared distribution with $N - 1$ degrees of freedom. Given a significance level $\alpha$, $H_0$ should be rejected if $\chi^2 > \chi^2_0$.

If there is no precise knowledge about $\delta$ and $\sigma_{\text{LOS}}$, the hypothesis can be tested by comparing the sample mean $\bar{x}$ or the sample variance $s^2$ with an experience value. In this case, $H_0$ is rejected if the sample mean or the sample variance is unusually large.

III. ESTIMATING THE TRUE RANGE FROM A SET OF NLOS-CORRUPTED MEASUREMENTS

Once again, let $X = \{x_1, x_2, ..., x_N\}$ be the dataset of $N$ sample range measurements. It has been pointed out in Section II that a Gaussian model is sufficient at characterizing the pdf of LOS measurements. However, the pdf of $X$ cannot be modeled by a Gaussian model if there are NLOS measurements mixed in with $X$.

Because an NLOS range is normally greater than the true range (except for the case when there is an additive negative noise), it could be expected that the higher the proportion of NLOS corrupted distribution. Therefore, we propose to use a Log-normal model for NLOS measurements. As long as a high proportion of LOS measurements are generated from a Gaussian distribution, the results are similar but not presented here. For another assumption that the distribution of NLOS errors is very dependent on range measurements and we do not have a knowledge of NLOS errors.

The Log-normal model is shown as follows:

$$f(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi x\sigma}} e^{-\frac{(\ln x - \mu)^2}{2\sigma^2}}, \quad x \geq 0$$

where $\mu$ and $\sigma$ are the mean and standard deviation of the variable’s natural logarithm (by definition, the variable’s logarithm follows a Gaussian distribution).

As long as there is a high proportion of LOS measurements in the sample dataset $X$, it should be expected that the highest peak of the pdf of $X$ appears when $x = r$ ($r$ is the true range). This can be explained by the addition of a Gaussian distribution which represents the distribution of LOS measurements, and a uniform distribution or another Gaussian distribution which represents the distribution of NLOS measurements; The peak of the Gaussian distribution representing LOS measurements is still a peak after the addition of the two distributions. In addition, the peak presented by LOS measurements should be the point of global maximum if the proportion of LOS measurements is large enough. Counter examples may happen when the proportion of NLOS corrupted measurements are too large. In that case, the peak (or multiple peaks) representing the distribution of NLOS measurements will overwhelm the peak representing LOS measurements.

When the Log-normal model is used to characterize the mixed distribution of LOS and NLOS measurements, the peak of the pdf (i.e. the point of global maximum) appears at the mode of the distribution. The mode can be found by setting $(\ln f(x))^\prime = 0$. According to the above analysis, we have the true range $r$ to be the mode and

$$r = e^{\mu - \sigma^2}.$$  \hspace{1cm} (6)

Since the expression of $r$ is a function of $\mu$ and $\sigma$, the problem of finding $r$ is transformed to the problem of estimating the parameters $\mu$ and $\sigma$ that specifies the Log-normal model given the observed dataset $X$.

Because we do not have any prior knowledge about the distribution of $\mu$ and $\sigma$, an Maximum Likelihood Estimation (MLE) method can be used here. Because the range measurements are observed independently, the likelihood function which specifies the likelihood of observing the dataset $X$ is:

$$\Lambda(X; \mu, \sigma) = \prod_{n=1}^{N} f(x_n; \mu, \sigma)$$

$$= \prod_{n=1}^{N} \frac{1}{\sqrt{2\pi x_n\sigma}} e^{-\frac{(\ln x_n - \mu)^2}{2\sigma^2}}, \quad x_n \geq 0. \quad (7)$$

According to MLE, the parameter values should be chosen such that the likelihood function is maximized. Therefore, we have

$$[\hat{\mu}, \hat{\sigma}] = \arg \max \Lambda(X; \mu, \sigma).$$

In order to maximize $\Lambda$, the derivatives of the logarithm of $\Lambda$ with respect to $\mu$ and $\sigma$ are computed and set to zero. We get the following MLE estimates:

$$\hat{\mu} = \frac{\sum_{n=1}^{N} \ln x_n}{N} \quad \text{and} \quad \hat{\sigma} = \sqrt{\frac{1}{N} \sum_{n=1}^{N} (\ln x_n - \hat{\mu})^2}. \quad (8)$$

Replacing $\mu$ and $\sigma$ in (6) with their MLE estimates $\hat{\mu}$ and $\hat{\sigma}$, we have an estimation of the true range $r$ in the case that there are NLOS corrupted measurements mixed in with $X$.

IV. SIMULATION RESULTS

In this section, simulation experiments are made to validate the presented range estimation method. We consider a dataset of 100 range measurements of which an unknown proportion of measurements are NLOS corrupted. The LOS measurements are generated from a Gaussian distribution $N(r, \sigma^2)$, where $r$ is the true range and $\sigma$ is the standard variance representing the receiver noise. The simulation of NLOS measurements is a little bit complicated. For each single NLOS measurement, it is generated from a different Gaussian distribution $N(r + \delta, \sigma^2)$, where $\delta$ represents the NLOS error. Because the environment of range measurement is very dynamic, $\delta$ is different for each NLOS measurement.

In our experiments, we assume that $\delta$ spreads between $0$ and $100\sigma$. This is a reasonable range for NLOS errors which can be in the order of hundreds of meters [6], [7]. Because the distribution of NLOS errors is very dependent on range measurement scenarios and we do not have a knowledge on that, we assume that $\delta$ has a uniform distribution on its supporting range. This uniform NLOS error distribution has been considered in [7]. For another assumption that $\delta$ has a Gaussian distribution, the results are similar but not presented here due to space limitation.
One factor that may affect the accuracy of range estimation is the proportion of LOS measurements in the NLOS corrupted dataset. Figure 1 shows the results of range estimation when the proportion of LOS measurements change from 10% to 90%. Figure 2 shows the results of range estimation when the proportion of LOS measurements change from 91% to 100%. Besides that the range estimation method based on the Log-normal model presented in Section III is used to estimate the range, the sample mean of the range measurements is also used to provide an alternative range estimate. In our experiments, we have the true range $r = 100$ meters and the standard variance $\sigma = 1$ meter. It can be seen from the figures that both the Log-normal model based method and the sample mean provide range estimates of which the estimation error decreases as the proportion of LOS measurements increases. More concretely, the sample mean estimate has a linear relationship with the proportion of LOS measurements. The estimate based on the Log-normal model approaches the true range even faster than the sample mean estimate as the proportion of LOS measurements increases. In Figure 1, the estimation error of the Log-normal model based method provides a range estimation accuracy which is about 20-50% higher than the estimate provided by the sample mean. In Figure 2, the Log-normal model based range estimation provides a range estimation accuracy which is about 55% higher than the estimate provided by the sample mean except for a special case when all measurements are LOS. When all measurements are LOS, both the Log-normal based estimate and the sample mean estimate are unbiased.

Although the Log-normal based range estimation does not require NLOS measurements being excluded from its analysis, it is known from the above simulation results that the exclusion of NLOS measurements will lead to a high proportion of LOS measurements and thus high range estimation accuracy. Therefore, we simulate two scenarios where heuristic methods are used to exclude NLOS measurements. In one scenario, the range measurements are sorted in ascending order and only the first half of the sorted measurements are used to estimate the true range. The results are shown in Figure 3. In the other scenario, the proportion of LOS/NLOS measurements is assumed to be a priori and only a proportion of the largest measurements equal to that of NLOS measurements are excluded from analysis. The results are shown in Figure 4. Both heuristic methods are based on the observation that NLOS measurements are usually larger than LOS measurements. However, NLOS measurements are not perfectly excluded in both scenarios as there exist NLOS measurements not excluded and LOS measurements excluded by mistake. It is shown in Figure 3 that both the Log-normal based estimation and the sample mean method have their estimation accuracies improved when the proportion of excluded measurements is less than the proportion of NLOS measurements. When the proportion of the excluded measurements is higher than the proportion of NLOS measurements, the range estimates are negatively biased. However, the biased estimates will be bounded by the variance representing the receiver noise which is usually much less harmful than NLOS errors. When the proportion of the excluded measurements is equal to the proportion of NLOS measurements, near-optimal accuracies are achieved by both range estimation methods as it is shown in Figure 4. Once again, the estimates are negatively biased and the biases are bounded by the variance representing the receiver noise. It is also shown that the biases vanish as the proportion of LOS measurements increases. This is explained by the fact that relatively less LOS measurements are excluded by mistake and relatively less NLOS measurements are retained by mistake when the proportion of LOS measurements is high.

To conclude, both the Log-normal based range estimation and the sample mean method can have their estimation accuracies improved when NLOS measurements are excluded. When the proportion of NLOS measurements is known, simply excluding a proportion, equalling to that of NLOS measurements, of the largest measurements from analysis is an easy and efficient way of roughly eliminating the influence of NLOS measurements. When the knowledge about the proportion of

---

2Let $p$ be the proportion of LOS measurements. Let $\delta$ be the expected NLOS error. The estimate according to the sample mean has its expectation being $r + (1-p)\delta$. 

---

**Fig. 1.** Range estimation results when the proportion of LOS measurements change from 10% to 90%. (a) Estimating range using the Log-normal model; (b) Estimating range using the sample mean. In both (a) and (b), the solid line shows the mean of 100 simulations and the dotted lines show the 95% confidence interval.

**Fig. 2.** Range estimation results when the proportion of LOS measurements change from 91% to 100%. (a) Estimating range using the Log-normal model; (b) Estimating range using the sample mean. In both (a) and (b), the solid line shows the mean of 100 simulations and the dotted lines show the 95% confidence interval.
has been shown in Figure 3. It is believed that of NLOS measurements. The reason is that retaining NLOS measurements by mistake is much more harmful than excluding LOS measurements by mistake as it is always safe to exclude NLOS measurements from analysis is an efficient way of roughly eliminating the influence of NLOS measurements. NLOS measurements is not precise, it is always safe to exclude a larger proportion of the largest measurements than what is believed to be that of NLOS measurements. The reason is that retaining NLOS measurements by mistake is much more harmful than excluding LOS measurements by mistake as it has been shown in Figure 3.

V. CONCLUSIONS

Localization is an important component required by many applications such as the promising UWB-based medical sensor network. In order to localize a target, ranges between the target node and reference nodes must be measured. GPS is an example of accurate localization where satellites are used as reference nodes. The problem with this range-based localization technology is that multi-path radio transmission and obstacles can cause non-line-of-sight (NLOS) errors in the measured ranges, especially when the range measurement is taken inside a room or even inside a body (as many medical applications do). Based on an NLOS corrupted range measurement dataset, the estimated range could deviate a lot from the true range. Therefore, it is vital important to identify whether a range measurement dataset has been corrupted by NLOS errors or not. Meanwhile, it may not always be appropriate to discard those NLOS corrupted range measurement datasets (e.g. if there are real time requirements for the range estimation or if it is very difficult to collect non-corrupted range measurement datasets). Therefore, it is also desired that accurate range estimation can take advantage of NLOS-corrupted range measurement datasets. In this paper, hypothesis tests are designed to test whether a given range measurement dataset is corrupted by NLOS errors or not. For those NLOS corrupted range measurement datasets, they are suggested to be modeled by Log-normal distributions. In addition, the mode of the fitted Log-normal distribution is found to coincide with the true range. As a result, the problem of finding the true range is transformed to the problem of estimating parameters of the fitted Log-normal distribution. Simulation results show that the proposed range estimation method can reduce the estimation error by 20-55% compared to the estimates provided by the sample mean. Although NLOS measurements are not required to be excluded from range estimation and it is actually the advantage of the proposed range estimation method, further simulations show that range estimation accuracy can be further improved if NLOS measurements are excluded. It is also found in this paper that simply excluding a proportion of the largest measurements from analysis is an efficient way of roughly eliminating the influence of NLOS measurements.

REFERENCES


