16th European Signal Processing Conference (EUSIPCO 2008), Lausanne, Switzerland, August 25-29, 2008, copyright by EURASIP

SEMI-PARAMETRIC GEOLOCATION ESTIMATION IN NLOS ENVIRONMENTS

Ulrich Hammes, Eric Wolsztynski, Abdelhak M. Zoubir

Signal Processing Group, Institute of Telecommunications, Technische Universität Darmstadt
Merckstrasse 25, 64283, Darmstadt, Germany
phone: + (49) 6151-16 4595, fax: + (49) 6151-16 3778, {hammes, ewol, zoubir}@spg.tu-darmstadt.de
web: www.spg.tu-darmstadt.de

ABSTRACT

The position of a stationary target can be determined using triangulation in combination with time of arrival measurements at several sensors. In urban environments, none-line-of-sight (NLOS) propagation leads to biased time estimation and thus to inaccurate position estimates. Here, a semi-parametric approach is proposed to mitigate the effects of NLOS propagation. The degree of contamination by NLOS components in the observations, which result in asymmetric noise statistics, is determined and incorporated into the estimator. The proposed method is adequate for environments where the NLOS error plays a dominant role and outperforms previous approaches that assume a symmetric noise statistic.

1. INTRODUCTION

Several techniques for geolocation finding based on received signal strength (RSS), angle of arrival (AOA), time difference of arrival (TDOA) and time of arrival (TOA) exist in the literature, e.g. [4]. Here, we deal with TOA measurements, where several TOA estimates are obtained from different sensor positions (Rx) to determine the position of the target (Tx). If there is a direct line of sight (LOS) path between Tx and Rx, accurate position estimates can be obtained using triangulation. However, in urban areas or hilly terrain, no direct LOS pathes may be available due to multiple reflections at buildings and other obstacles, leading to longer propagation pathes which result in an overestimated TOA. This bias in the time estimate leads to a biased position estimate of the target. For these reasons, robust techniques are needed to cope with the NLOS impairments.

Various methods exist to mitigate the effects of NLOS components. In [1] a parametric approach based on detection of NLOS components is presented, whereas in [3, 7] the authors consider non-parametric approaches that rely on kernel density estimates (KDE) of the underlying noise probability density function (pdf). Since the NLOS error is always greater than zero, the overall noise distribution tends to be asymmetric in case of NLOS propagation. However, in [7] the assumption of a symmetric pdf has been made in order to achieve consistency in the symmetric model. Since the symmetric assumption may not be satisfied in reality due to NLOS components, one may do better when the knowledge of a skewed distribution is incorporated into the estimator which is considered here. In particular, we propose a NLOS mitigation method by adapting semi-parametric constructions, suggested in previous work [5], to asymmetric noise densities. The proposed approach can be summarised in the following way:

-An estimator determines the level of skewness of the underlying distribution and based on this level, the data are transformed into an approximately Gaussian sample.

-Non-parametric KDE is performed in the transformed domain and an estimate of the true pdf in the original domain can be obtained via back-transformation.

-This estimate is used to determine the location of the target based on the maximum likelihood principle.

The paper is structured as follows: Section 2 introduces the signal model and Section 3 presents the algorithms based on transformation density estimation. In Section 4 simulation results are shown and Section 5 concludes the paper.

2. SIGNAL MODEL

We consider a single moving sensor in a two dimensional plane where N TOA measurements can be captured between the moving sensor Rx and the stationary target Tx. The ith distance measurement is obtained by multiplying the ith time estimate by the speed of light. The measured distances r_i are of the form

$$r_i = d_i + g_i + \mathrm{NLOS}_i, \qquad i = 1, \ldots, N, \tag{1}$$

where d_i is the true distance, g_i is sensor noise having density f_G which we assume to be Gaussian with variance σ_g^2. $NLOS_i$, independent of g_i, are positive random variables with density f_{NLOS}. In [4] f_{NLOS} was assumed to be Gaussian, whereas in [8] an exponential model was considered.

Here however, we consider the general noise components $g_i + \mathrm{NLOS}_i$ from (1) as a random variable having a density

$$f_0 = (1-\varepsilon) f_G + \varepsilon \mathscr{H}, \tag{2}$$

where $0 \leq \varepsilon \leq 1$ is the degree of contamination by NLOS components and $\mathscr{H} = f_G * f_{\mathrm{NLOS}}$, where $*$ denotes the convolution. We assume i.i.d. samples from f_0 for the perturbations but we do not make any assumption on f_{NLOS} for the construction of our algorithms. The position of the target is denoted as (x, y) and the position of the sensor of the ith distance measurement is given as (x_i, y_i). In the absence of noise, r_i can be expressed as

$$\begin{aligned} r_i^2 &= (x_i - x)^2 + (y_i - y)^2 \\ &= K_i - 2x_i x - 2y_i y + x^2 + y^2, \end{aligned} \tag{3}$$

where $K_i = x_i^2 + y_i^2$ [7]. For any distance measurement r_i, (3) can be written as

$$r_i^2 - K_i = -2x_i x - 2y_i y + R, \tag{4}$$

This work was supported by the Deutsche Forschungsgemeinschaft (DFG).

where $R^2 = x^2 + y^2$. Taking $\theta = [x, y, R]^\top$ (4) can be written in matrix form:

$$\mathbf{y} = \mathbf{S}\theta, \tag{5}$$

where

$$\mathbf{y} = \begin{bmatrix} r_1^2 - K_1 \\ r_2^2 - K_2 \\ \vdots \\ r_N^2 - K_N \end{bmatrix}$$

$$\mathbf{S} = \begin{bmatrix} -2x_1 & -2y_1 & 1 \\ -2x_2 & -2y_2 & 1 \\ & \vdots & \\ -2x_N & -2y_N & 1 \end{bmatrix}.$$

In case of measurement noise and noise due to NLOS propagations, the terms r_i^2 from (1) in the vector $\mathbf{y}$ become

$$r_i^2 = d_i^2 + n_i(d_i), \quad n_i(d_i) = 2d_i(g_i + \mathrm{NLOS}_i) + (g_i + \mathrm{NLOS}_i)^2 \tag{6}$$

The additive term $n_i(d_i)$ corresponds to random nonlinearities unexplained by the model. To model the pdf f of $n_i(d_i)$ parametrically would require knowledge of f_G, f_{NLOS} and of the true parameter d_i. Here, we prefer to leave f unspecified in the construction of the estimator (hence taking a semi-parametric approach).
In the presence of noise, (5) can be expressed as

$$\mathbf{y} = \mathbf{S}\theta + \mathbf{n} \tag{7}$$

where $\mathbf{n}$ is assumed to be a vector of i.i.d. random variables containing the elements $n_i(d_i)$. The maximum likelihood estimator (MLE) for θ is then given by

$$\hat{\theta}_{MLE} = \arg\min_{\theta} \sum_{i=1}^{N} -\log f\left(y_i - \sum_{k=1}^{3} \mathbf{S}_{ik}\theta_k\right). \tag{8}$$

Solving the derivative of (8) for zero, we obtain

$$\sum_{i=1}^{N} \mathbf{S}_{(ik)}\varphi\left(y_i - \sum_{k'=1}^{3} \mathbf{S}_{(ik')}\theta_{k'}\right) = 0, \quad k = 1, \ldots, 3, \tag{9}$$

where $\varphi = -f'/f$ is the location score function, where f' is $df(u)/du$. If f is Gaussian, (9) becomes least-squares estimation. However, in practice we can neither assume that f is Gaussian nor that it is known, and least-squares approaches may degrade severely due to NLOS components. Moreover, the minimax approach considered in [6] is robust but suboptimal for a particular noise distribution. Instead, we consider semi-parametric estimators which adapt to the underlying noise statistics.
Since the parameters of interest are not independent of each other, i.e. $R^2 = x^2 + y^2$, improvements on the accuracy of the parameter estimates may be achieved by incorporating this relationship into the estimator, as in [2]. This step is left out in what follows.

3. ADAPTIVE ESTIMATION

3.1 General Concept

The general approach we consider to estimate θ consists in approximating the MLE as performed in (9), using the residuals $\hat{\mathbf{n}} = \mathbf{y} - \mathbf{S}\hat{\theta}_0$. They are obtained from a preliminary estimate $\hat{\theta}_0$ (e.g. least-squares $\hat{\theta}_{LS}$), and used to estimate f and its derivative non-parametrically to obtain the estimate $\hat{\varphi}$. This estimate is used to solve Equation (9). Several methods for estimating φ exist in the literature. In [12] a method based on KDE that uses local smoothing was considered. It is conceptually and computationally heavy but achieves good performance with respect to linear or minimax [6] approaches. In [5], conceptually simpler semi-parametric approaches based on transforming the residuals by a parametric function are considered and can achieve performance similar to the non-parametric one from [12]. In general terms, the pdf in the original domain, f_U, is obtained by back-transformation, i.e.,

$$f_U(u) = f_W(t(u,\lambda)) \left| \frac{d\, t(u,\lambda)}{d\, u} \right|, \tag{10}$$

where $t(u,\lambda)$ is the transformation function and f_W the pdf of the transformed data. The shape parameter λ of the transformation function is chosen such that the data in the transformed domain is approximately Gaussian. Then KDE with global smoothing is sufficient to obtain an appropriate estimate of the transformed pdf f_W and an estimate of the true pdf can be obtained via (10). Likewise, an estimate of f' is obtained from differentiating (10) with respect to u.
Even though the above mentioned approaches are designed for symmetric noise environments, in [7] it is shown that the non-parametric approach using local smoothing in [12] leads to significant improvements with respect to conventional techniques in non-symmetric noise. In [7] the geolocation problem is considered, and the underlying distribution in (1) is asymmetric if a sufficient number of NLOS components are taken into account. If the a priori knowledge of asymmetry is incorporated into the estimator, better performance can be expected. Note that in some cases, e.g. when no NLOS components are contained in the observations, f may remain symmetric. Hence, an estimator is needed that takes into account the degree of asymmetry of the underlying data automatically.
Once the score function is estimated, a solution step can be constructed in two different ways to solve for (7) [6]. The first approach (A) is based on iterative Newton-Raphson steps, as in [12], using *modified residuals*, the second approach (B) uses iterative weighted least-squares (*modified weights*). In B, the weights w_i are determined as the absolute value of the ratio between the estimated score function at the residual $\hat{n}_i$ and the residual itself, i.e., $w_i = |\hat{\varphi}(\hat{n}_i)/\hat{n}_i|$. The flowchart in Figure (1) illustrates the steps for each of the two approaches. From an initial estimate $\hat{\theta}_0$ the residuals are determined and an estimate of the score function is obtained. This estimate is used to update an estimate of the parameter of interest θ using either approach A or approach B. The algorithm stops when the previous and actual estimates are close enough together. In [7] it has been shown that the weighting leads to improvements on the parameter estimate.

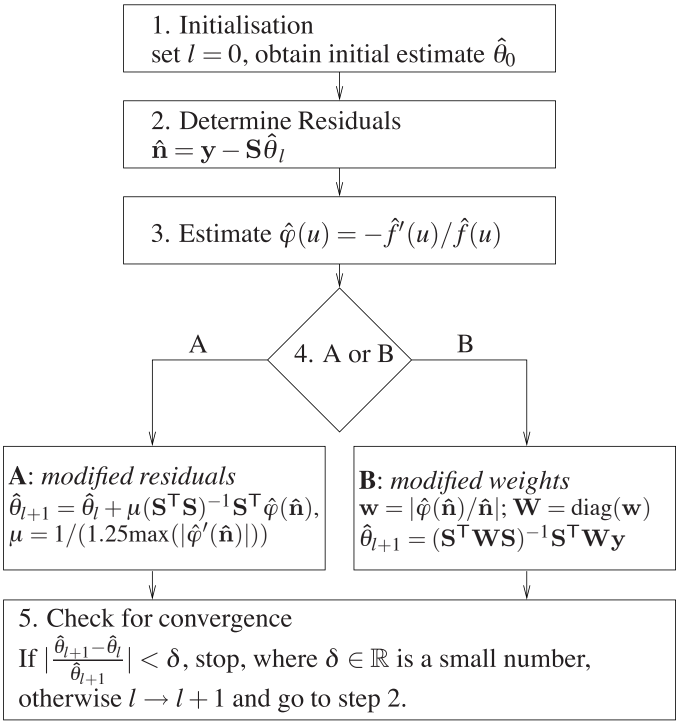

Figure 1: *Adaptive algorithm*

3.2 Semi-parametric method for non-symmetric noise

The approach we consider here is based on a modification of a method suggested in [5] where the symmetry assumption was met, leading to an anti-symmetric score function estimate. In this section, we relax the symmetry constraint and propose a different construction of the score function used in Step 3 of the adaptive algorithm (Figure 1).
Since the residuals are supposed to be asymmetric, we transform them into $t(\hat{\mathbf{n}},\lambda)$ using a transformation given in [11] that suits asymmetric data such that the data after the transformation are approximately Gaussian. The function and a selection scheme for its parameter are described in the following subsection. Since we have almost normal data after transformation, we symmetrise the transformed residuals into $\hat{\mathbf{n}}_{WS} = [-t(\hat{\mathbf{n}},\lambda) \quad ,+t(\hat{\mathbf{n}},\lambda)]$ in the transformed domain. This allows us to double the sample size which improves small sample performance. The KDE of $\hat{\mathbf{n}}_{WS}$ is computed in order to obtain $\hat{f}_W$. An estimate of the original pdf is obtained by back-transformation using (10) where f_W is replaced by its estimate. The global bandwidth used for KDE is constructed using the plug-in rule $\hat{h} = 1.06\hat{\sigma}_W N^{-1/5}$ [9], where $\hat{\sigma}_W = 1.483\mathrm{mad}[t(\hat{\mathbf{n}},\hat{\lambda})]$ and $\mathrm{mad}[\cdot]$ is the median absolute deviation. Furthermore, as in [5], if f is to be assumed unimodal, a constraint of a unimodal density can be met for $\hat{f}$ in order to improve small sample performance. This can be achieved by increasing the global bandwidth iteratively by $\eta = 1.05$ until a unimodal pdf is obtained [9]. Note that it may be unnecessary and even inadequate to make this assumption in some environments. In what follows, we do not consider the unimodal constraint.

3.3 Transformation function and shape control

The shape of the transformation function $t(u,\lambda)$ given by

$$t(u,\lambda) = \begin{cases} [(|u|+1)^\lambda - 1]/\lambda, & \lambda \neq 0,\ u \geq 0 \\ \log(u+1), & \lambda = 0,\ u \geq 0 \\ -[(-u+1)^{2-\lambda} - 1]/(2-\lambda), & \lambda \neq 2,\ u < 0 \\ -\log(-u+1), & \lambda = 2,\ u < 0, \end{cases} \qquad (11)$$

depends on the parameter λ that must be selected [11]. Unlike [5] where the modulus transformation is used to correct for symmetric heavy-tailed distributions, $t(u,\lambda)$ is appropriate for reducing kurtosis and skewness of a given data set. For $\lambda < 1$, $t(u,\lambda)$ is concave, for $\lambda > 1$ it is convex and linear for $\lambda = 1$. This means that, if $\lambda < 1$, the function decreases the right tail of a distribution and increases the left tail whereas for $\lambda > 1$ the right tail is increased and the left tail is decreased. Figure 2 illustrates the shape of $t(u,\lambda)$ for different values of λ. Since one can assume that the NLOS error is always positive, which corresponds to a right skewed noise statistic, we constrain the parameter λ to be always smaller than one so that positive outlying values are transformed closer to the core of the data. This facilitates the use of KDE with a global bandwidth. An MLE for the parameter λ under the assumption that the data after transformation is Gaussian is

$$\begin{aligned} \hat{\lambda} \quad = \quad & \arg\max_{\lambda} \left\{ -\frac{N}{2}\log\hat{\sigma}_W^2(\lambda) \right. \\ & \left. + (\lambda - 1)\sum_{i=1}^{N} \mathrm{sign}(\hat{n}_i)\log(|\hat{n}_i|+1) \right\}, \end{aligned} \qquad (12)$$

where $\hat{\sigma}_W^2(\lambda)$ is the sample variance of the transformed data [11].

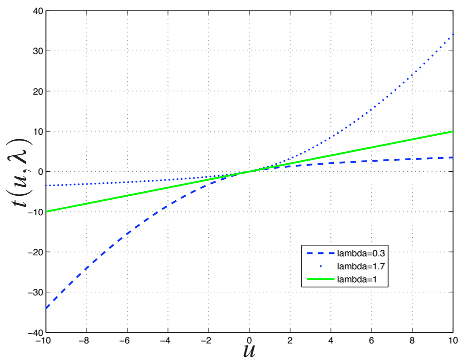

Figure 2: *Transformation function* [11] *for different values of* λ.

4. SIMULATIONS

We consider a single sensor Rx that is moving around a fixed target Tx collecting TOA estimates at different positions based on the signal model which has been introduced in Section 2. The initial position (x_1,y_1) of Rx is (10km,10km) and the position of the target (x,y) is (5km,5km). Rx is moving on a trajectory around Tx and measures $N = 50$ TOA estimates. Simulations are performed over $MC = 10000$ Monte Carlo runs.
We compare the performance of eight different estimators:

the least-squares (LS), an iterative re-weighted least squares (IRWLS), the non-parametric estimator suggested in [12] which uses symmetrised residuals (non-param. sym.) and the same one without symmetrisation (non-param.). These estimators proceed as indicated by path A in the algorithm in Figure 1. Furthermore, we consider using the non-parametric construction of [12] with path B in Figure 1, i.e. by using the score function estimate to compute the weights of a weighted least squares approach. Estimation of these weights is performed both with symmetrised residuals (IRWLS non-param. sym.) and without symmetrisation (IRWLS non-param.). These methods are compared with the proposed transformation-based estimator using path A (Semi-param.), i.e. modifying the residuals, and path B (IRWLS-semi-param.), i.e. computing weights from this score estimate for weighted least squares.

Since we fix a certain number of NLOS components, some observations are distorted by thermal noise only and others are affected by both thermal noise and the NLOS components. The positions of the NLOS components in the observations are set randomly (uniformly). As a performance metric, we use the mean error distance (MED)

$$\text{MED} = \frac{1}{MC}\sum_{j=1}^{MC}\sqrt{(x-\hat{x}^{(j)})^2+(y-\hat{y}^{(j)})^2} \qquad (13)$$

where $\hat{x}^{(j)}$, $\hat{y}^{(j)}$ are the jth geolocation estimates. Figure 3 illustrates the impact of the number of NLOS components on the MED for noise generated as in (2), with Gaussian measurement noise with standard deviation $\sigma_g = 150m$ and NLOS components that are exponentially distributed. Here the standard deviation of the NLOS components is set to $\sigma_{\text{NLOS}} = 409m$; this is a typical order of magnitude in many applicative examples, see e.g. [10]. The case where only LOS components are observed (LOS-only) is used as a benchmark for comparison of the different estimators. We observe that the least-squares estimators LS and IRWLS suffer from a performance loss when the number of NLOS components increases. Both non-parametric methods using symmetrisation, i.e. 'non-param sym.' and 'IRWLS non-param sym.', show to be more robust, as was shown in [7]. We can observe that non-symmetrisation of the residuals leads to improvement of the non-parametric approach (non-param), using path A. This method also compares with the transformation-based semi-parametric method using path A (semi-param). Finally, Figure 3 illustrates that the weighted least-squares using this semi-parametric construction for computing the weights (IRWLS semi-param.) leads to the best estimation in this comparison, and significantly outperforms the other methods when more than 22 NLOS components are present. However, the proposed methods using both A and B suffer from a small performance loss in environments where a small amount of NLOS components are observed. The fact that 'IRWLS non-param.' fails compared to all other methods, and in particular with 'IRWLS semi-param.', may be explained by the effect of a small sample size on the estimation of the weights. Although not shown here, a significant increase in variance of the estimated weights obtained from 'IRWLS non-param.' can be observed, and is notably higher than that observed from 'IRWLS semi-param'. Only $N = 50$ samples are used to perform KDE for 'IRWLS non-param.', whereas the semi-parametric counterpart 'IRWLS semi-param.' uses $N = 100$ samples since symmetrising the residuals in the transformed domain. The transformation function also yields a KDE with smooth tails which facilitates estimation of the weights.

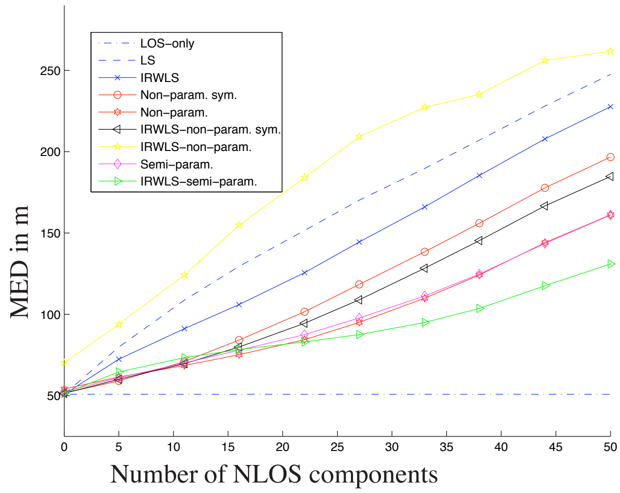

Figure 3: *MED versus number of NLOS components.* f_G *is a zero-mean Gaussian with* $\sigma_g = 150m$ *and* f_{NLOS} *is an exponential pdf with* $\sigma_{\text{NLOS}} = 409m$.

Figure 4 illustrates the cumulative distribution function of the geolocation estimates for the different algorithms when 38 NLOS components are present in each observation sample. We observe that the proposed construction using *modified weights* (IRWLS semi-param), i.e. path B, stands as more accurate than the other methods in terms of MED. The non-symmetrising 'non-param.' and the 'semi-param.' schemes seem to behave almost identically, and, although they are outperformed by 'IRWLS semi-param', they also improve the techniques already available in the literature (i.e. the least-squares procedures, 'non-param. sym.' and 'IRWLS non-param. sym.').

Let us now consider a noise model that consists of two Gaussian distributions, i.e. in model (2) the sensor noise remains Gaussian f_G with standard deviation $\sigma_g = 150m$, and the NLOS components are generated according to a shifted Gaussian density f_{NLOS} with mean $\mu = 763m$ and $\sigma_{\text{NLOS}} = 453m$. The order of magnitude of these settings is standard in many measurements, see e.g. [10]. Simulation results in terms of MED vs number of NLOS components are shown in Figure 5. Here again, we can observe distinct improvement obtained from the path-B-scheme 'IRWLS semi-param.', and then also path-A-methods 'semi-param.' and (non-symmetrising) 'non-param.', over all other competing methods. Also, the weighted least-squares approach 'IRWLS non-param.' (i.e. the path-B-method without symmetrising the residuals in the non-parametric construction) leads again to higher MED in the example. These results suggest that the proposed methods 'semi-param.' and 'IRWLS semi-param.', and also the suggestion of not symmetrising the non-parametric path-A-type construction, provide geolocation that is both more accurate and more robust than previous methods. Scheme 'IRWLS semi-param.' always seems to yield the best perfomance in our simulations.

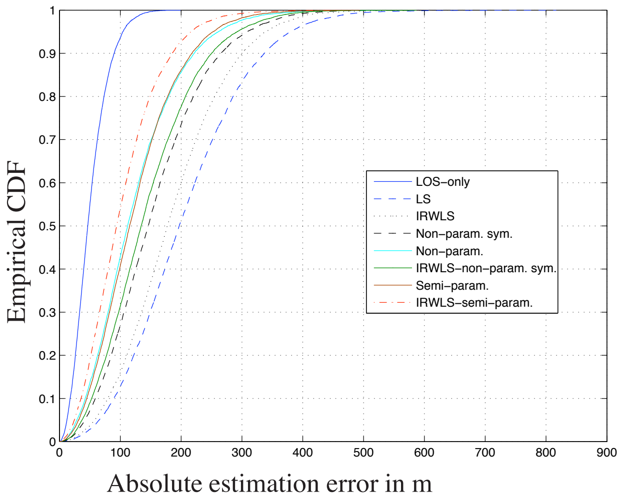

Figure 4: *Cumulative distribution functions of the location estimates for 38 NLOS components.* f_G *is a zero-mean Gaussian with* $\sigma_g = 150m$ *and* f_{NLOS} *is an exponential pdf with* $\sigma_{\mathrm{NLOS}} = 409m$.

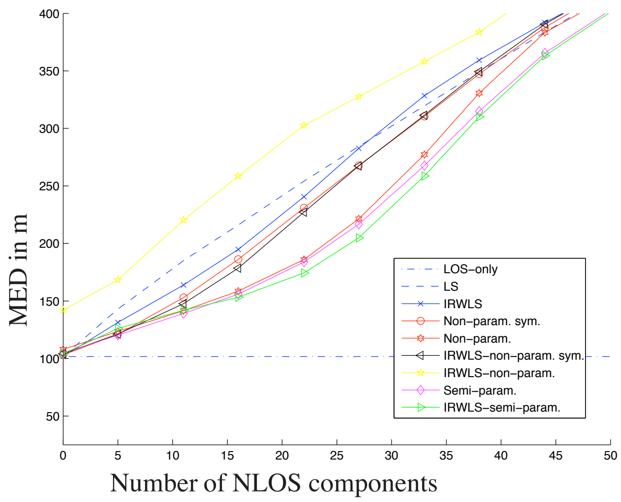

Figure 5: *MED versus number of NLOS components.* f_G *is a zero-mean Gaussian pdf with* $\sigma_g = 150m$ *and* f_{NLOS} *is a shifted Gaussian with* $\mu = 763m$ *and* $\sigma_{\mathrm{NLOS}} = 453m$.

5. CONCLUSION

Transformation-based semi-parametric estimators have been considered for mitigation of the effects of NLOS propagation in order to perform accurate geolocation of a target. Since the underlying noise pdf is affected by NLOS errors, it tends to be asymmetric. The proposed estimators proceed to estimate the underlying nuisance pdf using some density transformation that suits asymmetric samples. In particular, the shape of the transformation is controlled and adapts to the level of asymmetry of the data. Using this density estimate, two approaches have been considered, performing either by reproducing the maximum likelihood principle from the density estimate directly, or by computing the weights of a weighted least-squares approach.

It is shown that the proposed methods outperform their competitors in environments where the NLOS error significantly affects the observations. Improvement on a previous non-parametric construction which involves KDE with local smoothing was also suggested here and seems to compete well, although not as robustly as the new semi-parametric weighted least squares scheme. Summarising the results, it seems that robustness of the latter becomes more and more obvious w.r.t. its competitors as the level of skewness of the underlying noise distribution (i.e. the more NLOS errors are observed) increases.

More realistic noise parameters and correlation structures of the overall noise, e.g. using Markov chains, are currently under investigation. The impact of λ on the MED is under study. In case there is a significant impact on the MED, another estimator for λ may be selected in order to improve performance.

6. ACKNOWLEDGEMENTS

The authors would like to thank Dr. Chin Heng Lim from Temasek Labs, Singapore, for helpful discussions.

REFERENCES

[1] Y. Chan, W. Tsui, H. So, and P. Ching. Time-of-arrival based localization under nlos conditions. *IEEE Transactions on Vehicular Technology*, 55(1):17–24, 2006.

[2] Y.T. Chan and K.C. Ho. A simple and efficient estimator for hyperbolic location. *IEEE Transactions on Signal Processing*, 42(8):1905–1915, 1994.

[3] S. Gezici, H. Kobayashi, and H.V. Poor. Nonparametric non-line-of-sight identification. In H. Kobayashi, editor, *IEEE 58th Vehicular Technology Conference*, volume 4, pages 2544–2548 Vol.4, 2003.

[4] F. Gustafsson and F. Gunnarsson. Mobile positioning using wireless networks: possibilities and fundamental limitations based on available wireless network measurements. *Signal Processing Magazine, IEEE*, 22(4):41–53, 2005.

[5] U. Hammes, E. Wolsztynski, and A. M. Zoubir. Transformation-based robust semi-parametric estimation. *IEEE Signal Processing Letters*, 2008. To appear.

[6] P. Huber. *Robust Statistics*. Wiley, 1981.

[7] C.-H. Lim, A. M. Zoubir, C.-M. S. See, and B.-P. Ng. A robust statistical approach to non-line-of-sight mitigation. In *Proceedings of the 16th IEEE Workshop on Statistical Signal Processing (SSP'07)*, 2007.

[8] M. McGuire, K.N. Plataniotis, and A.N. Venetsanopoulos. Data fusion of power and time measurements for mobile terminal location. *IEEE Transactions on Mobile Computing*, 4(2):142–153, 2005.

[9] B. Silverman. *Density Estimation for Statistics and Data Analysis*. Chapman & Hall, 1986.

[10] M. P. Wylie and J. Holtzman. The non-line of sight problem in mobile location estimation. In *5th IEEE International Conference on Universal Personal Communications,*, volume 2, pages 827–831, 1996.

[11] I.-K. Yeo and R. A. Johnson. A new family of power transformations to improve normality or symmetry. *Biometrika*, 87:954–959, 2000.

[12] A. M. Zoubir and R. Brcich. Multiuser detection in heavy tailed noise. In *Digital Signal Processing: A Review Journal*, volume 12(2-3), pages 262–73, April-July 2002.