Hypothesis Testing and Iterative WLS Minimization for WSN Localization under LOS/NLOS Conditions

Giuseppe Destino, Davide Macagnano and Giuseppe Thadeu Freitas de Abreu
Center for Wireless Communications
University of Oulu, Finland
E-mails: [destino, macagnan, giuseppe]@ee.oulu.fi

***Abstract*—We propose a novel non-parametric solution for accurate distance-based source localization in wireless sensor networks (WSN's). The proposed technique includes a method to detect whether or not ranging is affected by bias due to non-line-of-sight (NLOS) conditions, requiring no *a-priori* knowledge of distance estimate statistics. Instead, we exploit the triangular inequality property of the Euclidean space and employ hypothesis testing (HT) in order to derive confidence levels on the observations and classify each link in the network as LOS or NLOS. These confidence levels are then incorporated in the formulation of an iterative WLS (IWLS) algorithm for WSN localization. The combination of the two contributions proves a powerful WSN localization algorithm, that is robust to noise, bias and erasure (incompleteness) over ranging data.**

I. Introduction

Location awareness has recently arisen as a key feature for several applications envisioned for wireless communications systems, especially wireless sensor networks (WSNs). Environmental monitoring, health-care systems and logistics are just a few application scenarios where the knowledge of the node location may play an important role. The challenge is that although these applications require accurate location estimation, their operation environments often present harsh conditions to be dealt with.

Several existing radio technologies allow devices to perform distance measurements (ranging), based on signal parameters such as received power (strength) and time-of-arrival (ToA). One of such technologies is Ultra-wide Band (UWB) which offers ToA-based ranging with sub-centimeter accuracy [1] under favorable, *i.e.* line-of-sight (LOS), channel conditions. In presence of multipath, *i.e.* non-line-of-sight (NLOS) conditions, however, UWB ToA-based ranging accuracy decreases substantially, because secondary paths may add constructively, resulting in stronger signals than that of the direct path, which causes ranging algorithms to false-detect [2], leading to distance estimates typically affected by positive errors. For similar reasons, ranging techniques based on received signal strength (RSSI) are also afflicted by positive errors, which are often referred to as bias [3].

Distance-based localization algorithms perform poorly in the presence of biased distance estimates resulting from NLOS channel conditions. In this article, the weighted least square (WLS) localization algorithm proposed in [4], [5] is further developed in order to improve its accuracy in presence of bias. In particular, hypothesis testing [6] is utilized to infer the presence of strong bias in the acquired distance measurements.

Our approach differs significantly form that of *e.g.*, [7], [8] and [3, sec. IV], in that our method exploits the fourth property of the Euclidian distance matrix [9] to detect bias utilizing data not from 1, but 3 different point-to-point links simultaneously.

The approach is also both non-parametric and “fuzzy”, in that no a-priori knowledge of the distribution of range estimates is assumed, and in that the result of the hypothesis test is ultimately used to adjust the weight of the corresponding link into the WLS. This extends the WLS concept in which the distance estimate corresponding to a link is weighed taking into account not only the reliability of that estimate alone, but also the reliability of the portion of the graph where it belongs.

Once LOS and NLOS channels are identified the location estimates can be improved via iterations of the WLS, designed to compensate for the bias affecting the NLOS links.

II. Problem Statement

We consider the problem of localizing a network of N devices embedded in a η-dimension Euclidean space from the set of distances amongst them. The network is represented by a graph $G(V,E)$, where the set of vertices $V = \{v_i\}$ and the set of edges $E = \{e_{ij}\}$ are associated to, respectively, the devices and the link between v_i and v_j [10].

A unique realization of such a graph is characterized by a matrix $\mathbf{X} \in \mathbb{R}^{N\times\eta}$, where the i-th row $\mathbf{x}_i \in \mathbb{R}^\eta$ indicates the coordinates of the vertex v_i. Let $\{v_1,\ldots,v_A\}$ and $\{v_{A+1},\ldots,v_N\}$ denote, respectively, the set of anchor (An) and target (Tn) nodes. Anchor and target nodes differ in which the location of An's are known, while those of Tn's are not.

Associated to the graph $G(V,E)$ are the the *Euclidean Distance Matrix* (EDM) and the *connectivity* (or *adjacent*) matrix, respectively denoted by $\mathbf{D}$ and $\mathbf{C}$ and defined by

$$d_{ij} = \sqrt{<\mathbf{x}_i - \mathbf{x}_j; \mathbf{x}_i - \mathbf{x}_j>}, \tag{1}$$

$$c_{ij} = \begin{cases} 1 & \text{if } e_{ij} \in E, \\ 0 & \text{otherwise}, \end{cases} \tag{2}$$

where $<;>$ denotes the inner product between two vectors.

In WSN's, distance estimates are usually obtained from ToA- or RSS-based ranging algorithms, and are affected by both small variable errors (due to noise) and unknown, typically positive, biases (due to NLOS conditions) [2], [11]. Recent research has indicated that noise-related and bias errors on distance estimates are characterized by different statistical parameters [12].

The ranging model considered in this article is therefore

$$\tilde{d}_{ij} = d_{ij} + n_{ij} + \varrho_{ij}, \tag{3}$$

where n_{ij} and ϱ_{ij} indicates the noise and bias, respectively.

978-1-4244-2110-7/08/$25.00 ©2007 IEEE

The distance-based WSN localization problem concerns the estimation of the coordinates $\mathbf{X}$ of the nodes interconnected as indicated by $\mathbf{C}$, given a set of ranging measurements modeled by (3), collected into EDM's $\mathbf{D}$.

Several distance-based localization techniques are discussed in [3], [13]. In general, the performance of such algorithms degrade in the presence of strong bias, *i.e.*, NLOS channels. Techniques to mitigate the effect of bias in distance-based localization include embedded bias mitigation, a-posteriory bias compensation, and LOS/NLOS identification algorithms.

An example of an embedded bias mitigation method can be found in [14], where a distributed minimization technique is employed to optimize a Bayesian log-likelihood cost function built with a-priori knowledge of the ranging model. The problem with this approach is obviously that performance is directly dependent of the accuracy of the ranging model [14].

In contrast, bias compensation can be performed in a non-parametric way. One such method is proposed in [15], where Linear Programming (LP) is employed in the estimate of $\mathbf{X}$, with LOS and NLOS measurements incorporated in the problem formulation and the definition of the feasibility regions. The compensation method of [15] is non-parametric, but assumes that LOS and NLOS links are identified a-priori.

Non-parametric LOS/NLOS identification techniques are therefore required in order to take advantage of flexibility and robustness offered by non-parametric bias compensation techniques such as the one described above.

In [5], a WLS localization algorithm was proposed in which distance estimates are weighed according to confidence scores computed also in a non-parametric fashion. A key advantage of the WLS algorithm of [5] is its ability to cope with a significant amount of erasures onto the EDM's, incomplete connectivity in the network, that is, the likely event that not every node is directly connected to all others. In contrast, the LP algorithm proposed in [15], for instance, requires that *every* anchor node be connected to *all* target nodes.

The performance of the original WLS algorithm degrades in the presence of bias onto distance estimates, as shown in [16]. The problem addressed in this paper is therefore how to improve the aforementioned WLS algorithm so that it can cope with NLOS conditions, while maintaining its robustness against connectivity incompleteness.

III. Preliminaries

In [4], [5], the concept utilized to weight the influence of each link within the least-squares minimization was to compute a corresponding confidence score based on the reliability of the associated distance estimate. One straightforward alternative to extend that idea would be to apply a link-wise NLOS mitigation technique such as those described in [7], [8] and [3, sec. IV], in order to take into account the presence of outliers in the computation of the weights.

The problem is that link-wise NLOS mitigation algorithms rely on the assumption that only *some* of the distance estimates are affected by bias. Instead, we exploit the 4-th property of EDM's [9] and use hypothesis testing [6] onto three different links simultaneously, so as to identify portions of the network that would be most affected by the presence of bias. In other words, NLOS identification is done indirectly and relatively, meaning that the likelihood of NLOS detection is proportional to the amount of bias and dependent on the structure of the triangle formed by the 3 links (see section IV-A).

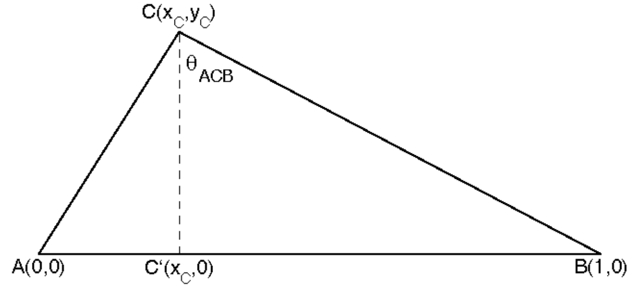


Fig. 1. The triangle ACB has a unitary base AB and a variable location of the vertex C, uniquely determined by AC' and θ_{ACB}.

The idea behind this approach is that even if the error processes affecting the estimation of the 3 sides of a triangle are similar, the error distribution on the vertex-coordinates estimates may be different, depending on the shape of the triangle. In order to illustrate the idea, a pseudo-analytical study of this effect is presented in this section.

Consider a triangle with vertices v_A, v_B and v_C, at the coordinates $(0,0)$, $(0,1)$ and (x,y), respectively, as shown in figure 1, such that base of the triangle is fixed to the unity and the location of the vertex v_C determines the angle θ_ACB, opposite to the base (and vice-versa). Let the sides of the triangle be measured under the effect of noise and bias, as per the ranging model of equation (3), such that the distributions of the distance measurements are $\Gamma(a_\gamma, b_\gamma)$ with $a_\gamma = d+\varrho_d/b_\gamma$, $b_\gamma = \sigma_d^2/(d+\varrho_d)$, the random bias ϱ_d uniformly distributed in $[0.1, 0.3]$m, and the standard deviation $\sigma_d = 0.1$m.

Next, let $\hat{\mathbf{X}}^{(\ell)}$ denote the matrix containing the estimate of the locations of v_A, v_B and v_C, for the ℓ-th set of measurements of the three sides of the triangle, computed using the metric multidimensional scaling (MDS) algorithm [17].

If the error of the estimates of the coordinates of v_A, obtained at the ℓ-th realization is denoted by $\varepsilon_\mathrm{A}^{(\ell)} = \|\mathbf{x}_\mathrm{A} - \mathbf{x}_\mathrm{A}^{(\ell)}\|_2$, where $\|\cdot\|_2$ denotes the norm-2, then the deviation of the estimation error of $\mathbf{x}_\mathrm{A}$, over L realizations is

$$\sigma_\mathrm{A} = \sqrt{\frac{1}{L-1}\cdot\sum_{\ell=1}^{L}(\varepsilon_\mathrm{A}^{(\ell)} - \bar{\varepsilon}_\mathrm{A})^2}, \tag{4}$$

where $\bar{\varepsilon}_\mathrm{A}$ is the mean error.

The deviations σ_B and σ_C of the estimation errors of $\mathbf{x}_\mathrm{B}$ and $\mathbf{x}_\mathrm{C}$ can be computed similarly. For a given model of distance estimates and a given set of parameters, the values of σ_A, σ_B and σ_C, for asymptotically large L depend only on the shape of the triangle, which is determined by θ_ACB and x_C.

In figure 2, the deviations σ_A, σ_B and σ_C for an isosceles triangle, and the total deviation $\sigma_\mathrm{A} + \sigma_\mathrm{B} + \sigma_\mathrm{C}$ for three different triangles are plotted against θ_ACB. The upper part of the figure shows that the shape of the triangle influences on the deviations. In particular, it is found that obtuse triangles, especially isosceles, yield larger coordinate estimation errors.

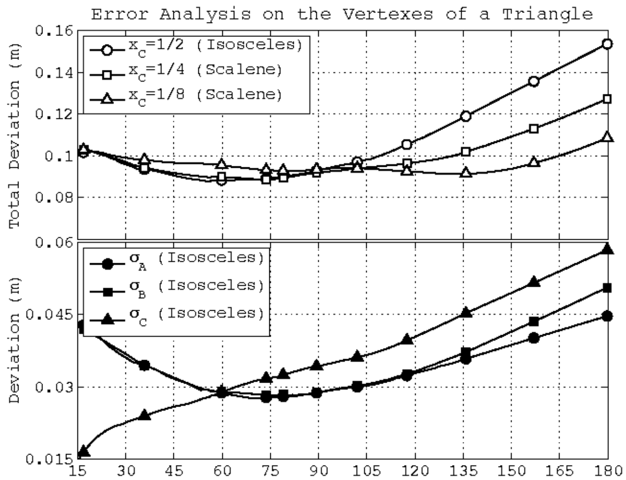


Fig. 2. Estimation of the errors at each node of a triangle ABC, with links affected by 0 bias and $\sigma_d = 0.1$m std of Gamma distributed noise.

Though not elaborated further, similar results are obtained if different ranging models are applied. The lesson that can be learned from this pseudo-analysis is that any of the commonly utilized ranging models, ranging errors have a stronger effect on the triangulation-based coordinate estimation errors when the three vertices to be locates form obtuse triangles.

The bottom part of figure 2 complements to the above by illustrating that the deviations of the estimates of the three vertices are not generally identical. In particular, when the vertices form very wide obtuse triangles, the error on the vertex opposite to the wide base (here represented by v_C), is more subject to error than the other two.

IV. LOS/NLOS Identification Algorithm

It is known that in meshy sensor networks with a sufficiently large average node-degree, *i.e.*, where each node is connected to several others, accurate node localization can be achieved even when only a subset of the available ranging information is used. Evidence to this fact was given in [18] and [5], where it was shown that localization accuracy under full ranging information is nearly the same as that attained under roughly 60% of completeness.

Combining this knowledge with the results of the preceding section, it is fair to expect that in a meshy network, which contains multiple triangles of different shapes, better localization accuracy can be achieved by prioritizing edges that are not associated with the error propagation-prone obtuse triangles. Of course this intuitive notion does not replace the fact that the larger ranging errors do degenerate localization accuracy. It is imperative, therefore, that both notions be taken into account jointly, in balance with one another.

In light of the above, our approach to LOS/NLOS identification is the following. First, employ hypothesis testing in order to capture which portions a graph $G(V,E)$ are more susceptible to error propagation, given both the dispersion of link-wise distance estimation errors *and* the triangular structures formed with adjacent edges. Then, a confidence score is assigned to each edge based on the result of the statistical test, and compared to a threshold in order to infer whether the edge is under NLOS conditions or not.

A. Hypothesis Testing

Our hypothesis test is formulated around the claim [6]:

Claim 1: Triangles are formed by every clique (v_i, v_j, v_k) of a graph $G(V, E, \bar{\mathbf{D}})$, with [4]

$$\bar{\mathbf{D}} = \mathcal{A}(\{\tilde{\mathbf{D}}\}) = [\bar{d}_{ij}^{\,2}]\,, \tag{5a}$$

$$\bar{d}_{ij} \triangleq \frac{1}{Q}\sum_{q=1}^{Q} \tilde{d}_{ij:q}, \tag{5b}$$

where $\{\tilde{\mathbf{D}}\}$ is the set of EDM samples and Q its cardinality.

In order to test this claim, the complementary inequality of the so-called 4-th property of EDM's [9] is used:

$$0 \leqslant d_{ij} - d_{jk} - d_{ik}. \tag{6}$$

The *null* hypothesis (H_0) indicates the veracity of the 4-th property, while the contrary is expressed by the *alternative* (H_1). Mathematically,

$$\mu\left(\{t_{ij,k}\}\right) \overset{H_1}{\underset{H_0}{\gtrless}} 0, \tag{7}$$

in which the formulation of the null hypothesis $H_0 : \mu \leqslant 0$ will be replaced by the equality, in order to imply a conservative decision on the result of the test statistic. The variable $\mu(\{t_{ij,k}\})$ returns the average of

$$t_{ij,k} = \tilde{d}_{ij} - \tilde{d}_{ik} - \tilde{d}_{jk}, \tag{8}$$

over all $P_{ij,k}$ combinations of $\{\tilde{d}_{ij}\}$, $\{\tilde{d}_{jk}\}$ and $\{\tilde{d}_{ik}\}$.

Even if the sets, $\{\tilde{d}_{ij}\}$, $\{\tilde{d}_{jk}\}$ and $\{\tilde{d}_{ik}\}$ contain relatively few elements, the number of combination of these elements becomes sufficiently large that, by force of the Central Limit Theorem, the distribution of the variable $\bar{t}_{ij,k}$ is approximately Gaussian, regardless of the distributions of the distance estimates. Furthermore, for the same reason, the variance of $t_{ij,k}$ can be accurately estimated by its sample variance $S^2_{t_{ij,k}}$.

Under these assumptions, the *test statistic* for the triangle formed by the vertices v_i, v_j and v_k, or *score* of $\bar{t}_{ij,k}$, is

$$Z_{ij,k} = \frac{\bar{t}_{ij,k} - 0}{S_{t_{ij,k}}/\sqrt{P_{ij,k}}}. \tag{9}$$

The *rejection* region (RR) of this hypothesis test is

$$RR : Z_{ij,k} \geqslant z_{\alpha_{ij,k}}, \tag{10}$$

where $z_{\alpha_{ij,k}}$ is the *critical value* of $Z_{ij,k}$ for a significance α

$$\Phi(z_{\alpha_{ij,k}}) = P_{\mathrm{E}}[Z_{ij,k} \geqslant z_{\alpha_{ij,k}}] = \int_{z_{\alpha_{ij,k}}}^{\infty} \phi(Z)\, dZ = \alpha, \tag{11}$$

where $\phi(Z)$ and $\Phi(z)$ are the standard normal distribution and its cumulative, respectively, and where $P_{\mathrm{E}}[\cdot]$ denotes the probability of the event defined in the bracket.

The objective of conventional hypothesis [6] testing is to decide, with a certain confidence $\beta = 1-\alpha$, whether or not the observations contain sufficient evidence of the veracity of the null hypothesis H_0. In contrast, our objective is to evaluate the maximum level of confidence β_{M} provided by the observations that the null hypothesis H_0 is true.

Since the maximum confidence associated with the score is at the boundary of the rejection region, we have

$$\beta_{\mathrm{M}_{ij,k}} = 1 - \Phi(Z_{ij,k}), \tag{12}$$

for the triangle (v_i, v_j, v_k), where $Z_{ij,k} = z_{\alpha_{ij,k}}$.

Due to the formulation of the hypothesis test described above, $\beta_{\mathrm{M}_{ij,k}}$ are typically small if $\bar{d}_{ij} \gg \bar{d}_{ik}$ and/or $\bar{d}_{ij} \gg d_{jk}$, which may result from the following two possibilities.

The first is that (v_{ij}, v_{ik}, v_{jk}) form an obtuse triangle, such that $d_{ij} \gg d_{ik}$ and/or $d_{ij} \gg d_{jk}$. Then, any distance-localization algorithm could potentially be affected by the error propagation induced by obtuse triangles, as discussed in section III and summarized in figure 2.

The second possibility is that $\bar{d}_{ij} \gg d_{ij}$ and/or $\bar{d}_{ik} \ll d_{ik}$ and/or $\bar{d}_{jk} \ll d_{jk}$, which indicates that a strong bias has affected the measurement of d_{ij} and/or d_{ik} and/or d_{jk}. Then, in the context of the WLS optimization-based localization algorithm here considered [4], [5], it is desirable to give less importance to the distance measurements $\tilde{d}_{ij}$ in the computation of the estimate $\hat{\mathbf{X}}$ of the entire set of node locations $\mathbf{X}$, such that $\hat{\mathbf{X}}$ depends less on $\tilde{d}_{ij}$.

B. Triangular Confidence Weight

In light of the later remarks, the confidences $\beta_{\mathrm{M}_{ij,k}}$ obtained with the hypothesis test described above could be used to weight the importance of the edge e_{ij} in the localization of the vertices of the Graph $G(V, E)$. First, however, one must notice that any edge e_{ij} can form triangles with multiple vertices v_k to which v_i and v_j are simultaneously connected to. In order to be conservative in selecting one out such $\beta_{\mathrm{M}_{ij,k}}$'s we define the *triangular confidence weight* of the edge e_{ij} as the minimum $\beta_{\mathrm{M}_{ij,k}}$, for all permissible k, that is

$$h_{\mathrm{T}_{ij}} \triangleq \min_{k} \beta_{\mathrm{M}_{ij,k}} \quad \text{subject to } \{c_{ij}, c_{jk}, c_{ik}\} = 1, \tag{13}$$

with $0 \leqslant h_{\mathrm{T}_{ij}} \leqslant 1$.

Notice that to each edge e_{ij} for which a triangular confidence weight is computed is associated one and only one triangle, identified by $k|\beta_{\mathrm{M}_{ij,k}} = h_{\mathrm{T}_{ij}}$. The angle opposite to e_{ij} in this triangle will be hereafter denoted $\theta_{\mathrm{T}_{ij}}$ and referred to as the *characteristic angle* of e_{ij}.

Of course, triangular confidence weights are not computed for all edges of the Graph. In particular, in addition to the connectivity condition embedded in equation (13), a triangular confidence weight is computed for an edge e_{ij} only if (the triangular confidence weight of all remaining edges is arbitrarily set to the unit):

a) e_{ij} is *not* known to be in LOS,
b) e_{ik} and e_{jk} are *not* known to be in NLOS.

The latter conditions prevent the computation of non-unitary triangular confidence weight for edge e_{ij} that are known to be in LOS conditions, and that an edge knowingly affected by bias influence the confidence on the distance estimate of d_{ij}.

Notice also that the triangular confidence weight is *complementary* – rather than substitute – to link-reliability weights defined in [4], such that the product of both weights is actually used in the modified WLS algorithm described in the sequel.

C. Indirect LOS/NLOS Identification

Due to the geometry of the problem, explored in section III, and the properties of the confidence levels β, triangular confidence weights can also be used in the identification of LOS/NLOS channels based on the argument that the absence/presence of strong bias onto e_{ij} increases the likelihood that $h_{\mathrm{T}_{ij}} \to 1$, and $h_{\mathrm{T}_{ij}} \to 0$, respectively. Since the intended classification is binary (LOS/NLOS), it is fair to divide the decision region in the middle, such that an edge will be considered to be LOS if $h_{\mathrm{T}_{ij}} > 0.5$, or NLOS if $h_{\mathrm{T}_{ij}} \leqslant 0.5$.

Of course, since the correlation of the events (LOS, $h_{\mathrm{T}_{ij}} > 0.5$), or equivalently (NLOS, $h_{\mathrm{T}_{ij}} \leqslant 0.5$) is not unitary, identification errors will occur. In particular, the presence of bias onto the estimates of d_{ij} may not be detected if its value is small and/or if all triangles formed with the edge e_{ij} are acute, leading to the classification of a NLOS channel as a LOS. Hereafter, this type of error will be referred to as the error of *type I.*

Likewise, even if e_{ij} is a LOS edge, there is a likelihood, proportional to the value of the largest θ_{ikj} amongst all the triangles formed with e_{ij}, that the errors on the estimates of d_{ij}, d_{ik} and/or d_{jk} may result in the false identification of e_{ij} as a NLOS edge. This type of error will be referred to as the error of *type II.*

Errors of type I may degrade the accuracy of the WLS localization algorithm, since estimates affected by (possibly) substantial bias are not properly compensated for [16]. In contrast, errors of type II may be either beneficial – in light of the results of section III, when the error is caused by obtuse triangles – or insignificant, when the culprit is the small scale of the bias. Based on this argument, one could conclude that increasing the boundary of the LOS/NLOS decision region to $\zeta > 0.5$ would be a conservative strategy to improve the accuracy of the localization algorithm.

It is know, however, that the localization accuracy of the WLS and SDP algorithms decrease with the incompleteness of the network [5], due to conditions on the uniqueness of Graph realization [19]. Based on this argument, setting $\zeta < 0.5$ so as to allow for less-biased edges – in fact those under the effect of strongest bias – to be compensated simultaneously, may be a better strategy to follow.

All in all, the optimum partition of the LOS/NLOS region depends on the compensation technique to be subsequently used and is a matter that deserves further investigation. For the time being, we have found heuristically that for the WLS algorithm, with completeness in of no less than 60%, the value $\zeta \approx 0.45$ yields the best results.

D. Performance of TCW-based LOS/NLOS Identification

The probability that an edge affected by bias ϱ, with characteristic angle θ is properly classified is

$$P_{\mathrm{C}}(\theta, \varrho) = Pr[h_{\mathrm{T}} > \zeta | \varrho = 0] + Pr[h_{\mathrm{T}} \leqslant \zeta | \varrho \neq 0]. \tag{14}$$

Although hard to calculate directly, $P_{\mathrm{C}}(\theta, \varrho)$ can be computed approximately as follows. Consider L random Graph realizations with corresponding classification of all edges $e_{ij}^{(\ell)}$ and let $u_{ij}^{(\ell)} = 1$ if $e_{ij}^{(\ell)}$ is correctly classified, and $u_{ij}^{(\ell)} = 0$

otherwise. Next, partition the interval $[0, \pi]$ and the range $\Delta\varrho = (\max\{\varrho_{ij}^{(\ell)}\} - \min\{\varrho_{ij}^{(\ell)}\})$ of bias encountered over all $e_{ij}^{(\ell)}$, into B equi-sized bands $\Delta\theta(b) \triangleq [(b-1)\pi/B, b\pi/B]$ and R equi-sized bands $\Delta\varrho(r) \triangleq [(r-1)\Delta\varrho/R, r\Delta\varrho/R]$, respectively, with $b = 1, \cdots, B$ and $r = 1, \cdots, R$.

Let $\Omega_{b,r}$ be the total number of edges $e_{ij}^{(\ell)}$ such that $\varrho_{ij}^{(\ell)} \in \Delta\varrho(r)$ and $\theta_{\mathrm{T}_{ij}}^{(\ell)} \in \Delta\theta(b)$; and define the number of successful LOS/NLOS classifications within the (b, r)-th bin as

$$\omega_{b,r} \triangleq \sum_{ij,\ell} u_{ij}^{(\ell)} | \varrho_{ij}^{(\ell)} \in \Delta\varrho(r) \;\; \text{and} \;\; \theta_{\mathrm{T}_{ij}}^{(\ell)} \in \Delta\theta(b). \qquad (15)$$

Then, the probability $\bar{P}_{\mathrm{C}}(b, r)$ of correctly identifying an edge with bias characteristic angle in the interval $\Delta\theta(b)$ and bias in the range $\Delta\varrho(r)$ can be estimated by

$$\bar{P}_{\mathrm{C}}(\theta_{\mathrm{T}} \in \Delta\theta(b), \varrho \in \Delta\varrho(r)) \approx \frac{\omega_{b,r}}{\Omega_{b,r}}, \qquad (16)$$

with equality holding at $\lim_{L\to\infty}$.

The quantity $\bar{P}_{\mathrm{C}}(\theta_{\mathrm{T}} \in \Delta\theta(b), \varrho \in \Delta\varrho(r))$ is representative of the probability $P_{\mathrm{C}}(\theta, \varrho)$. In particular,

$$P_{\mathrm{C}}(\theta, \varrho) = \lim_{\substack{B\to\infty \\ R\to\infty}} \bar{P}_{\mathrm{C}}(\theta_{\mathrm{T}} \in \Delta\theta(b), \varrho \in \Delta\varrho(r)). \qquad (17)$$

Contour plots of equation (16), computed with 200 random realization of a 4-An-10-Tn network with 100% completeness in a 2-dimensional Euclidean are shown in figure 3.

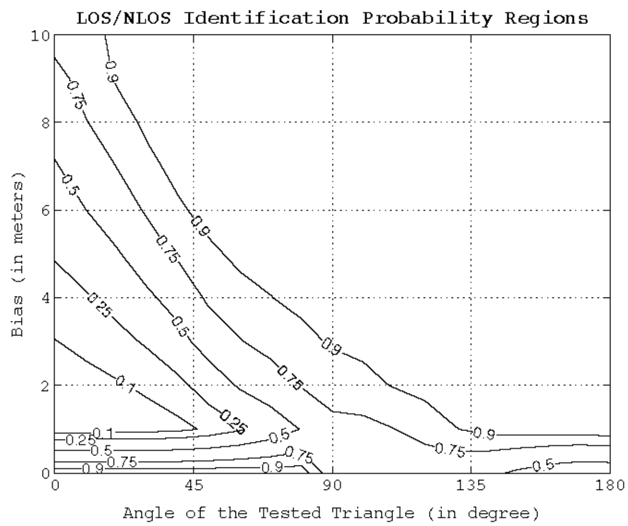


Fig. 3. LOS/NLOS identification probability regions for a network with 4 anchors, 10 targets, $\sigma_d = 1m$, bias $\mathcal{U}$, 25% NLOS links and threshold 0.45.

The tests were conducted with 10 distance measurements per edge, generated with the model described in section II and a LOS/NLOS decision region partitioned at $\zeta = 0.45$. The figure confirms the expectations raised earlier in this subsection that the hypothesis testing-based LOS/NLOS identification technique is most successful when obtuse triangles dominate the decision, and/or when the distance estimates are affected by strong bias. This outcome is particularly good in light of the properties of the WLS algorithm and the pseudo-analytical results of section III, since the bottom-left corner of figure 3 indicate the region where bias is less significant and where error propagation due to triangulation is smaller.

V. Iterative WLS

It has been shown [5] that the WLS localization algorithm unites the computational efficiency of least-squares optimization with the robustness to incomplete and imperfect ranging information provided by the weighted SDP technique [20]. In this subsection, an iterative WLS (IWLS) algorithm is presented, which incorporates the hypothesis testing LOS/NLOS identification approach here proposed in order to improve localization accuracy.

Although not optimal, the rather simple IWLS serves the purpose of illustrating the potential of the combination of the WLS and HT-based identification.

The WLS objective utilized in the ℓ-th iteration of the IWLS algorithm is given by [5]

$$\min_{\hat{\mathbf{X}}^{(\ell)} \in \mathbb{R}^{\eta \times N}} \left\| \mathbf{H}_D^{(\ell)} \circ \mathbf{H}_{\mathrm{T}}^{(\ell)} \circ \left(\mathcal{A}(\{\check{\mathbf{D}}\}^{(\ell)}) - \hat{\mathbf{D}}^{(\ell)} \odot \hat{\mathbf{D}}^{(\ell)} \right) \right\|_{\mathrm{F}}^2, \qquad (18)$$

where $\odot$ denotes the Hadamard product, the first WLS computation is counted as the 0-th iteration; $\mathbf{H}_D^{(\ell)}$ is the original weighing matrix computed as described in [4] using $\{\check{\mathbf{D}}\}^{(\ell)}$; $\mathbf{H}_{\mathrm{T}}^{(\ell)}$ is the matrix of triangular confidence weights; $\hat{\mathbf{D}}^{(\ell)}$ is the EDM constructed from $\hat{\mathbf{X}}^{(\ell)}$ and $\{\check{\mathbf{D}}\}^{(\ell)}$ is the set of virtual EDM's samples defined as

$$\{\check{\mathbf{D}}\}^{(\ell)} \triangleq \begin{cases} \{\tilde{\mathbf{D}}\} & \text{if } \ell = 0, \\ \{\{\tilde{\mathbf{D}}\}, \mathbf{D}_{\mathrm{A}_1}, \cdots, \mathbf{D}_{\mathrm{A}_\ell}\} & \text{if } 1 \leqslant \ell \leqslant w, \\ \{\tilde{\mathbf{D}}_{\ell-w+1}, \cdots, \tilde{\mathbf{D}}_Q, \mathbf{D}_{\mathrm{A}_1}, \cdots, \mathbf{D}_{\mathrm{A}_\ell}\} & \text{if } w < \ell \leqslant w+Q, \\ \{\mathbf{D}_{\mathrm{A}_{\ell-w-Q}}, \cdots, \mathbf{D}_{\mathrm{A}_\ell}\} & \text{if } \ell > w+Q, \end{cases} \qquad (19)$$

where w is a sliding window-size and $\mathbf{D}_{\mathrm{A}_\ell}$ is defined by

$$\mathbf{D}_{\mathrm{A}_{\ell+1}} = \overline{\{\check{\mathbf{D}}\}^{(\ell)}} \odot \mathbf{C}_{\mathrm{LOS}}^{(\ell)} + \hat{\mathbf{D}}^{(\ell)} \odot (\mathbf{I} - \mathbf{C}_{\mathrm{LOS}}^{(\ell)}) \qquad (20)$$

where $\mathbf{I}$ is an all-one matrix and the (i, j)-th element of $\mathbf{C}_{\mathrm{LOS}}^{(\ell)}$ is 1 if e_{ij} has been identified as LOS, or 0 otherwise.

The stop-criterion of the iteration is

$$\max\{S_{ij}^{(\ell)}\} \leqslant \epsilon, \qquad (21)$$

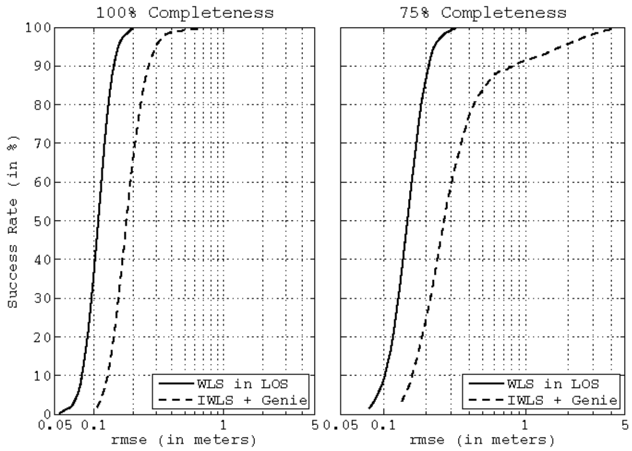
where ϵ is a threshold that indicates the admissible variation onto the virtual distance estimates.

VI. Simulation Results

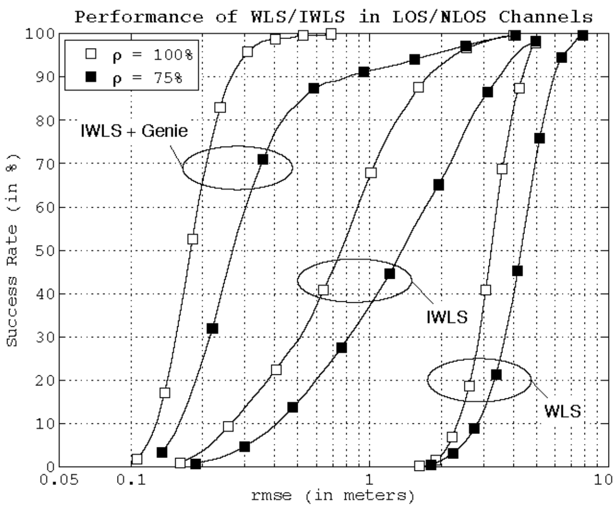
The localization accuracy attained with the IWLS under mixed LOS/NLOS conditions (25% NLOS) identified perfectly (with a Genie) is compared against the WLS under pure LOS conditions in figure 4(a) for a 4-An-10-Tn network with uniformly distributed $\sigma_{d_{ij}}$ and $\varrho_{d_{ij}}$. Plots of the root-mean-square-error ξ of $\hat{\mathbf{X}}$ for 200 realizations with $\epsilon = 0.05$, obtained under 100% and 75% of completeness are shown. It is found that the performance of the two algorithms is similar. For instance, for 100% completeness $\xi \approx 0.1$m and 0.18m with probability 50% for the WLS and the IWLS, respectively.

$$\xi \triangleq \left\| [\mathbf{X}]_{\mathrm{UL}:(N-A)\times\eta} - [\hat{\mathbf{X}}]_{\mathrm{UL}:(N-A)\times\eta} \right\|_{\mathrm{F}} / \sqrt{N-A} \qquad (22)$$

Finally, we compare the IWLS with HT-based LOS/NLOS identification using $\zeta = 0.45$, against the IWLS with a Genie and the WLS algorithms, under mixed LOS/NLOS conditions (25% NLOS), for 100% and 75% completeness.



(a) Effect of iteration of the WLS.



(b) Iteration and LOS/NLOS identification combined.

Fig. 4. Success rate and localization error at a network with 4 anchors, 10 targets, $\sigma_d \in \mathcal{U}(0.1, 1)$m, bias $\rho_d \in \mathcal{U}(1, 10)$m, 25% NLOS.

The results, shown in figure 4(b), demonstrate that the combination of hypothesis testing for LOS/NLOS identification with the iterative extension of the WLS algorithm can effectively mitigate the effect of bias in the distance estimates.

If the curves for the IWLS with a Genie are taken as a reference of the attainable improvement, it is found that the IWLS with HT-identification can reap from 70% to 80% improvement over the WLS algorithm. For example, with 50% error probability and 100% completion, the IWLS algorithm brings the localization error from 3.2m down to 0.75m. In comparison to the possible 0.18m attainable with a Genie, and the 0.1m attainable with the WLS under pure LOS conditions, indicating that the method reaps 81% and 79%, respectively, of the total localization accuracy improvement possible.

References

[1] S. Gezici, Z. Tian, G. B. Giannakis, H. Kobayashi, A. F. Molisch, H. V. Poor, and Z. Sahinoglu, "Localization via ultra-wideband radios," *IEEE Signal Processing Mag.*, vol. 22, no. 4, pp. 70–78, 2005.

[2] L. Joon-Yong and R. Scholtz, "Ranging in a dense multipath environment using an UWB radio link." *IEEE J. Select. Areas Commun.*, vol. 20, pp. 1667–1683, December 2002.

[3] B. F. Guoqiang Mao and B. D. O. Anderson, "Wireless sensor network localization techniques," 2006, accepted.

[4] G. Destino and G. Abreu, "Localization from imperfect and incomplete ranging," in *Proc. IEEE* 17^{th} *Int. Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC'06)*, Sep. 11-14 2006.

[5] ——, "Sensor localization from WLS optmization with closed-form gradient and Hessian," in *Proc. IEEE* 49^{th} *Annual Globecom Conference (GLOBECOM'06)*, Nov. 27 - Dec. 1 2006.

[6] D. Zwillinger and S. Kokoska, *CRC: Standard Probability and Statistics Tables and Formulae*. Chapman & Hall/CRC, 1999.

[7] C. Chong, F. Watanabe, and I. Guvenc, "NLOS identification and mitigation for UWB localization systems," in *IEEE Wireless Communications and Networking Conference (WCNC)*, Mar. 2007, pp. 1571–1576.

[8] S. Gezici, H. Kobayashi, and H. Poor, "Non-parametric non-line-of-sight identification," in *in Proc. IEEE 58th Veh. Tech. Conf. (VTC)*, 2003.

[9] J. Dattorro, *Convex Optimization and Euclidean Distance Geometry*. Meboo Publishing, 2005.

[10] K. H. Rosen, *Handbook of Discrete and Combinatorial Mathematics*. CRC Press, 2000.

[11] M. Wylie and J. Holtzmann, "The non-line-of-sight problem in mobile location estimation," in *5th IEEE International Conference on Universal Personal Communications*, vol. 2, 1996, pp. 827–831.

[12] N. Patwari, R. J. O. Dea, and Y. Wang, "Relative location in wireless networks," *in Proc. IEEE Veh. Tech. Conf. (VTC)*, May 2001.

[13] N. Patwari, J. Ash, S. Kyperountas, R. Moses, and N. Correal, "Locating the nodes: Cooperative localization in wireless sensor networks," *IEEE Signal Processing Magazine*, pp. 54–69, 4 2005.

[14] B. Denis, L. He, and L. Ouvry, "A flexible distributed maximum log-likelihood scheme for UWB indoor positioning," in *IEEE 4th Workshop on Positioning, Navigation and Communication*, 2007.

[15] S. Venkatesh and R. M. Buehrer, "A linear programming approach to NLOS error mitigation in sensor networks," *Proceedings of the Fifth International Conference on Information Processing in Sensor Networks (IPSN)*, pp. pp. 301–308, pril 2006.

[16] G. Destino, D. Macagnano, G. Abreu, B. Denis, and L. Ouvry, "Localization and tracking for LDR-UWB systems," in *Proc. IST Mobile & Wireless Communications Summit*, 2007.

[17] T. F. Cox and M. A. A. Cox, *Multidimensional Scaling*, 2nd ed. Chapman & Hall/CRC, 2000.

[18] A. Y. Alfakih, H. Wolkowicz, and A. Khandani, "Solving euclidean distance matrix completion problems via semidefinite programming," *Journal on Comp. Opt. and Apps.*, vol. 12, no. 1, pp. 13 – 30, 1999.

[19] B. Hendrickson, "Conditions for unique graph realization," *SIAM Journal on Comp.*, vol. 21, no. 1, pp. 65–84, Feb 1992.

[20] A. M.-C. So and Y. Ye, "Theory of semidefinite programming for sensor network localization," in *ACM Proceedings of the* 16^{th} *Annual ACM-SIAM Symposium on Discrete Algorithms*, 2005, pp. 405 – 414.