CONDITIONS FOR GUARANTEED CONVERGENCE IN SENSOR AND SOURCE LOCALIZATION

Barış Fidan, Soura Dasgupta and Brian D. O. Anderson

ABSTRACT

This paper considers localization of a source or a sensor from distance measurements. We argue that linear algorithms proposed for this purpose are susceptible to poor noise performance. Instead given a set of sensors/anchors of known positions and measured distances of the source/sensor to be localized from them we propose a potentially nonconvex weighted cost function whose global minimum estimates the location of the source/sensor one seeks. The contribution of this paper is to provide nontrivial ellipsoidal and polytopic regions surrounding these sensors/anchors of known positions, such that if the object to be localized is in this region localization occurs by globally convergent gradient descent. This has implication to the deployment of sensors/anchors to achieve a desired level of geographical coverage.

Key words: Localization, Sensors, Global Convergence, Optimization, Gradient Descent

1. INTRODUCTION

Over the last few years, the problem of source/sensor localization has assumed increasing significance (see [1] for application scope). Specifically *source localization* refers to a set of sensors estimating the precise location of a source using information related to their relative position to the source. In *sensor localization* a sensor estimates its own position using similar information relative to several nodes of known positions called *anchors*. This information can be distance, bearing, power level (indirectly related to distance) and time difference of arrival (TDOA). *In this paper, we focus on distances only.* In abstract terms our goal is the following. Given known 2 or 3- dimensional vectors $x_1, \cdots, x_N$ ($N > 2$ and $N > 3$ in 2 and 3 dimensions respectively) and an unknown vector y^*, estimate the value of y^*, from the measured distances $d_i = \|y^* - x_i\|$. In the source localization problem, y^* represents the position of the unknown source, and the x_i the positions of the sensors seeking to estimate its location. In the sensor localization problem, the x_i are the positions of the anchors, and y^* the position of the sensor estimating its own position.

Such distances can be estimated through various means. For example if a source emits a signal, the signal intensity and the characteristics of the medium provides a distance estimate. In this case with A the source signal strength, and η the power loss coefficient, the signal strength at a distance d from the source is given by

$$s = A/d^\eta. \tag{1.1}$$

Thus, A, s and η provide d. Alternatively, a sensor may transmit signals of its own, and estimate the distance by measuring the time it takes for this signal reflected off the target to return.

In 2-dimensions, localization from distance measurements generically requires that distances of y^* from at least three non-collinearly situated x_i be available. To be precise, with just two distances, the position can be determined to within a binary ambiguity. Occasionally, *a priori* information may be available which will resolve that ambiguity. Otherwise, a third distance is needed. In three dimensions, one generically needs at least four non-coplanar x_i.

As explained in section 2 if the d_i uniquely specify y^*, y^* can be estimated using linear algorithms, [2],[3]. In practice, as also explained in section 2, such a linear algorithm with certain geometries may deliver highly inaccurate estimates with noisy measurements of the distances, even when the noise is small. On the other hand, several papers adopt a nonlinear estimation approach, [4]-[8]. Many of these directly work with (1.1), with known A and η. Thus rather than assuming that the d_i are directly available, they work with the received signal strength at several sensors and formulate a minimization problem to obtain y^*. Such minimization problems are inevitably non-convex and manifested with locally attractive false optima. Among these, [5]-[6] conduct searches *in two dimensions*, that as noted in [7] are sensitive to spurious stationary points, while [7] provides search alternatives involving the so called Projection on Convex Sets (POCS) approach also in two dimensions, as does [8] though in a framework that is more general. Barring [7], none attempts to characterize conditions under which convergence obtains. The POCS approach of [7], however, has the unique solution of y^* in the noise free case iff y^* is in the convex hull of x_i (i.e $\mathrm{Co}(x_1, \cdots, x_N)$).

In this paper, we work directly with the d_i and with certain weights $\lambda_i > 0$, seeking to obtain y^* as the y minimizing

$$J(y) = 0.5 \sum_{i=1}^{N} \lambda_i \left(\|x_i - y\|^2 - d_i^2 \right)^2, \tag{1.2}$$

We show in section 2 that this is in general a non-convex problem. The goal of this paper is to provide conditions on x_i and y^* under which the gradient descent minimization of (1.2) is globally convergent. Such conditions have important practical import as they provide a benchmark of how to deploy sensors (anchors) to cover a geographical region to localize sources (sensors) potentially located throughout the region. We show that these regions extend beyond $\mathrm{Co}(x_1, \cdots, x_N)$.

Section 2 provides more background and examples showing the nonconvexity of (1.2). In section 3, given $\lambda_i > 0$ and x_i, we characterize a nontrivial *ellipsoidal* set members of which are

Fidan and Anderson are with Australian National University and National ICT Australia Ltd, Locked Bag 8001, Canberra ACT 2601 Australia, Baris.Fidan and Brian.Anderson@nicta.com.au. Dasgupta is with the Dept. of ECE, University of Iowa, Iowa City, IA 52242, USA, dasgupta@engineering.uiowa.edu. Supported by NSF grants ECS-0225530 and ECS-0622017, and by National ICT Australia, which is funded by the Australian Government's Department of Communications, Information Tehcnology and the Arts and the Australian Research Council through the *Backing Australia's Ability* initiative and the ICT Centre of Excellence Program.

guaranteed to be localized through the gradient descent minimization of (1.2). In section 4 given x_i, we quantify a nontrivial polytopic set for which there exist $\lambda_i > 0$ such that members of this set can be similarly localized. While these analyses are deterministic, section 5 demonstrates the performance of gradient descent minimization under noisy distance measurements. It is after all an important aim of this paper to present an algorithm initially derived to cope with noiseless measurements, but able to cope with noisy ones in a graceful manner. Section 6 is the conclusion.

2. BACKGROUND AND PRELIMINARIES

In this section we discuss linear algorithms, key past results and examples demonstrating the convexity of (1.2).

2.1. Linear algorithms

We discuss now the practical ramifications of linear localization algorithms. Consider three non-collinear x_i in 2-dimesions and and equations

$$\|x_i - y^*\|^2 = d_i^2, \quad \text{for } i \in \{1,2,3\}. \tag{2.1}$$

Subtracting the first equation from the remaining two one obtains,

$$2\begin{bmatrix} (x_1-x_2)^T \\ (x_1-x_3)^T \end{bmatrix} y^* = \begin{bmatrix} \|x_1\|^2 - \|x_2\|^2 + d_2^2 - d_1^2 \\ \|x_1\|^2 - \|x_3\|^2 + d_3^2 - d_1^2 \end{bmatrix}. \tag{2.2}$$

For non-collinear x_i, $det([(x_1-x_2),(x_1-x_3)]^T) \neq 0$, i.e. y^* can be solved uniquely. But, the solution is invariant if for any α the d_i^2 are replaced by $d_i^2 + \alpha$, suggesting and verified by example in [9], that such linear algorithms may have poor noise performance.

2.2. Nonlinear algorithms

In [5],[6], the model of (1.1) is used to obtain y^* by optimizing (through an iterative serach procedure as opposed to analytical calculation) a cost function related to (1.1). As noted in [7] convergence properties are impaired by spurious optima, conditions for whose absence have not been obtained.

A reformulation is used in [7] to apply the POCS approach. This results in a problem whose unique solution (given noiseless measurements of the relevant distances) equals the desired y^* iff y^* is in the convex hull of the x_i, even if the x_i are non-collinear and regardless of their number. Related POCS algorithms are in [8]. Simulations provided in [8] show that their algorithms may converge even if y^* is not in the convex hull of the x_i, without providing conditions under which such convergence can occur.

2.3. Preliminaries of (1.2)

Our standing assumption below ensures that $J(y) = 0$ iff $y = y^*$.

Assumption 2.1 *In two dimensions the x_i, $i \in \{1, \cdots, N\}$, are non-collinear. In three dimensions they are non-coplanar.*

We will seek to find conditions under which

$$\partial J(y)/\partial y = \sum_{i=1}^{N} \lambda_i(\|y - x_i\|^2 - d_i^2)(y - x_i) = 0 \text{ iff } y = y^*. \tag{2.3}$$

Under (2.3), for every $\|y[0]\| \leq C$, there exists μ^* dependent on C such that for all $\mu \leq \mu^*$, the gradient descent algorithm

$$y[k+1] = y[k] - \mu\, \partial J(y)/\partial y|_{y=y[k]} \tag{2.4}$$

is globally convergent. Even if (2.3) fails, however, practical convergence will still occur if false stationary points are locally unstable. Thus consider the 2-dimensional case where $\lambda_i = 1$, $x_1 = [-1,0]^T$, $x_2 = [0,-1]^T$, $x_3 = 0$, and $y^* = [-1,-1]^T$ depicted in fig. 1(a). In this case $d_1^2 = d_2^2 = 1$ and $d_3^2 = 2$. Some routine algebra shows that $\partial J/\partial y = 0$ has two solutions: $y = y^*$ and the spurious value $y = 0$. In (2.4), with $y = [y_1, y_2]^T$, $y_1[k+1] - y_1[k]$ equals

$$-\mu\left[2y_1^2[k] + (y_1[k] + y_2[k])^2 + 3y_1[k](y_1^2[k] + y_2^2[k])\right].$$

If $y_1[k] < 0$, with $y[k]$ close to the origin, $y_1[k+1] < y_1[k]$, exhibiting the local instability of $y = 0$. Thus in practical terms, the local instability of this stationary point will make it unattainable, as the slightest noise will drive the solution off it.

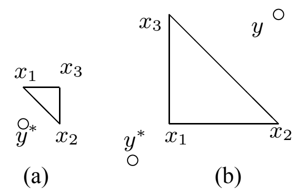

Fig. 1. (a) False unstable stationary point at x_3. (b) False stable stationary point at y

On the other hand, there are examples were the spurious stationary points are locally stable. This occurs, [10] if the Hessian

$$\frac{\partial}{\partial y}\left[\sum_{i=1}^{N} \lambda_i(\|y - x_i\|^2 - d_i^2)(y - x_i)\right] \tag{2.5}$$

is positive definite at such a stationary point. To this end choose $\lambda_i = 1$, $x_1 = [1,1]^T$, $x_2 = [1,3]^T$, $x_3 = [3,1]^T$, and the true $y^* = 0$ depicted in fig. 1(b). In this case $d_3^2 = d_2^2 = 10$ and $d_1^2 = 2$. It turns out $y = [3,3]^T$ is a spurious stationary point. Evaluated at this point one can verify that (2.5) is positive definite. It is thus important to consider conditions under which (2.3) holds.

3. GUARANTEED CONVERGENCE FOR GIVEN WEIGHTS

We provide conditions under which for fixed $\lambda_i > 0$ (2.3) holds. Call the N-vector $u_N = [1, \cdots 1]^T$, $\mathcal{X} = [[x_1, \cdots, x_N]^T, u_N]^T$, and $\bar{y} = [y^{*T}, 1]^T$. We prove an initial result.

Lemma 3.1 *Consider x_i in 2 or 3-dimensions with assumption 2.1 holding. Then for every y^* there exist scalar β_i obeying*

$$\sum_{i=1}^{N} \beta_i = 1 \tag{3.1}$$

for which

$$\sum_{i=1}^{N} \beta_i x_i = y^*. \tag{3.2}$$

Proof: (3.1),(3.2) is equivalent to $\mathcal{X}\beta = \bar{y}$, where $\beta = [\beta_1, \cdots, \beta_N]^T$. Thus, we need only to show that rank$[\mathcal{X}] = \dim(\bar{y})$. If the contrary were true then there exists a nonzero $\theta = [\bar{\theta}^T, \theta_{n+1}]^T \in \Re^{n+1}$ (where $\theta_{n+1} \in \Re$ and $n = \dim(x_i)$) for which $\theta^T\mathcal{X} = 0$, i.e. $\bar{\theta}^T x_i = \theta_{n+1}$ for all $i \in \{1, \cdots, N\}$. Noting that $\theta \neq 0$, this implies that in 2-dimensions the x_i are collinear and in 3-dimensions co-planar, violating assumption 2.1. ■

If $\beta_i \geq 0$ for all i, then $y^* \in \text{Co}\{x_1, \cdots, x_N\}$. Further, for $N > 3$ in 2-dimensions and $N > 4$ in 3-dimensions, the β_i that obey (3.1), (3.2) are nonunique. We now develop a condition involving β_i, and λ_i, to ensure (2.3). Define $\tilde{x}_i = x_i - y^*$ and $\tilde{y} = y - y^*$. Then (2.3) holds iff

$$\sum_{i=1}^{N} \lambda_i(\|\tilde{y} - \tilde{x}_i\|^2 - \|\tilde{x}_i\|^2)(\tilde{y} - \tilde{x}_i) = 0 \Leftrightarrow \tilde{y} = 0. \quad (3.3)$$

Further because of (3.1), (3.2),

$$\sum_{i=1}^{N} \beta_i \tilde{x}_i = 0. \quad (3.4)$$

With $e_i = \tilde{y} - \tilde{x}_i$, $E = [e_1, \cdots, e_N]$, $\Lambda = \text{ diag } \{\lambda_1, \cdots, \lambda_N\}$, $\partial J / \partial y$ becomes:

$$\sum_{i=1}^{N} \lambda_i(\|\tilde{y} - \tilde{x}_i\|^2 - \|\tilde{x}_i\|^2)(\tilde{y} - \tilde{x}_i) = 2E\Lambda E^T \tilde{y} - E\Lambda u_N \tilde{y}^T \tilde{y}.$$

Because of (3.4), and (3.1) $E\beta = \tilde{y}$. Thus (2.3) holds iff

$$2E\Lambda E^T E\beta = E\Lambda u_N \beta^T E E^T \beta \Leftrightarrow E\beta = 0$$

i.e. with $P = 2\Lambda - \Lambda u_N \beta^T$,

$$2EPE^T E\beta = 0 \Leftrightarrow E\beta = 0, \quad (3.5)$$

Then we provide a sufficient condition on P that ensures (2.3).

Lemma 3.2 *Consider y^*, $x_1, \cdots x_N$ with assumption 2.1 in place. Suppose $\beta = [\beta_1, \cdots, \beta_N]^T$, obeys (3.1) and (3.2), $\lambda_i > 0$, and P is as above. Then (2.3) holds if $P + P^T > 0$.*

Proof: We need to show that (3.5) is true, if $P + P^T > 0$. Now, $EPE^T E\beta = 0$ implies that $\beta^T E^T EPE^T E\beta = 0$. As $P + P^T > 0$, one has $EE^T \beta = 0$, i.e $\beta^T EE^T \beta = \tilde{y}^T \tilde{y} = 0$. ■

This is only a sufficient condition. Even if $P + P^T$ is not positive definite, (2.3) will hold unless some $\tilde{y} \neq 0$ is in the null space of EPE^T which itself depends on $\tilde{y}$. Yet we show below that it does quantify a nontrivial domain where (2.3) holds. First an intermediate lemma.

Lemma 3.3 *Consider two $N \times 1$ vectors a, b. Then $4I - ab^T - ba^T > 0$ iff $\sqrt{(b^T b)(a^T a)} < 4 - a^T b$.*

Proof: The characteristic polynomial of $ab^T + ba^T$ is $s^{N-2}(s^2 - 2(b^T a)s + (a^T b)^2 - (a^T a)(b^T b)$. Thus the smallest eigenvalue of $4I - ab^T - ba^T$ is $4 - b^T a - \sqrt{(a^T a)(b^T b)}$, whence the result. ■

Lemmas 3.3 and 3.2, (3.1) and the fact below prove Theorem 3.1.

$$P + P^T > 0 \Leftrightarrow 4I - \Lambda^{1/2} u_N \beta^T \Lambda^{-1/2} - \left(\Lambda^{1/2} u_N \beta^T \Lambda^{-1/2}\right)^T > 0.$$

Theorem 3.1 *Consider in 2 or 3-dimensions, x_i, obeying assumption 2.1, $\lambda_i > 0$, Λ as above, any y^* and β obeying (3.1) and (3.2). Then (2.3) holds if:*

$$(\beta^T \Lambda^{-1} \beta)(u_N^T \Lambda u_N) = \left(\sum_{i=1}^{N} \beta_i^2 / \lambda_i\right)\left(\sum_{i=1}^{N} \lambda_i\right) < 9. \quad (3.6)$$

This is in fact a necessary and sufficient condition for $P + P^T$ to be positive definite, though from the argument above only a sufficient condition for (2.3).

For unity weights, i.e. $\lambda_i = 1$, and $N < 9$, $\text{Co}(x_1, \cdots, x_N)$ is a *proper subset* of this set. Indeed if $y^* \in \text{Co}(x_1, \cdots, x_N)$ then in (3.1) and (3.2) $\beta_i \geq 0$, and $\sqrt{\beta^T \beta} \leq \beta^T u_N = 1$. Thus as $\Lambda = I$, $(\beta^T \Lambda^{-1} \beta)(u_N^T \Lambda u_N) \leq 8$. Recalling that in 2 and 3-dimensions it suffices to have $N = 3$ and $N = 4$, respectively, for given x_i satisfying assumption 2.1, the set characterized by Theorem 3.1 can be chosen to be significantly larger than their convex hull. This means that in the sensor (resp. source) localization problem, just a few anchors (resp. sensors) will achieve substantial geographical coverage. We now show that the set of y^* satisfying (3.1),(3.2) and (3.6) is an *ellipsoid*.

Theorem 3.2 *For every $\lambda_i > 0$, and x_i, obeying assumption 2.1, the set of y^* for which scalar β_i satisfying (3.1),(3.2) and (3.6) exist, is a nonempty ellipsoid, determined entirely by x_i and λ_i.*

Proof: Note, (3.6) holds for some choice of λ_i iff it holds with any $\delta > 0$, for $\delta\lambda_i$. Thus without loss of generality choose λ_i so that $u_N^T \Lambda u_N = 1$. Then $\beta_i = \lambda_i$, satisfies (3.1) and (3.6). Thus, the set is nonempty. Recall (3.1) and (3.2) combine to give $\mathcal{X}\beta = \bar{y}$. Assumption 2.1 together with an argument similar to that in the proof of Lemma 3.1 ensures that rank of $\mathcal{X}$ is the same as the dimension of $\bar{y}$. Then β satisfying (3.1) and (3.2) is completely characterized by $\beta = G[(F\bar{y})^T, z^T]^T$ where F and G are nonsingular matrices determined entirely by x_i and z is any suitably dimensioned vector. Thus with A a positive definite symmetric matrix, determined by Λ, and the x_i, (3.6) becomes

$$\min_z \left([\bar{y}^T F^T, z^T] A [\bar{y}^T F^T T, z^T]^T\right) < 9. \quad (3.7)$$

Partition A as

$$A = \begin{bmatrix} A_{11} & A_{21}^T \\ A_{21} & A_{22} \end{bmatrix}.$$

Observe $B = [A_{11} - A_{21}^T A_{22}^{-1} A_{21}] > 0$. It is readily seen that (3.7) becomes $\bar{y}^T F^T B F \bar{y} < 9$. Thus the set of $\bar{y} = [y^{*T}, 1]^T$ is an ellipsoid. Then as long as the the set of y^* is non empty it is an elliposid determined entirely by x_i and λ_i.

■

4. CHOOSING THE WEIGHTS

In this section we quantify the set of y^* for which there exist λ_i so that (2.3) holds. Indeed suppose y^* is much closer to x_1 than the other x_i. Intuition suggests that one should emphasize d_1 more than the other d_i, by choosing λ_1 to be relatively larger. The results of this section should be viewed in this context. We first present the following Theorem that follows from (3.6) and the Cauchy-Schwarz inequality.

Theorem 4.1 *Under the hypotheses of Theorem 3.1 there exist $\lambda_i > 0$ for which $P + P^T > 0$ iff*

$$\sum_{i=1}^{N} |\beta_i| < 3. \quad (4.1)$$

Further under (4.1)

$$\lambda_i = |\beta_i| \quad (4.2)$$

always guarantee $P + P^T > 0$.

This set is polytopic in β and because of (3.2) also in y^*. Further as under (3.1), $\beta_i \geq 0$ for all i, guarantees (4.1), Co$\{x_1, \cdots, x_N\}$ is a proper subset of thus polytope. As with Theorem 3.2 this polytope can be determined entirely from the x_i. Consider the 2-dimensional example depicted in figure 2 where 1,2,3 represent the x_i locations. Choose $\beta_1 = 0$. Then y^* satisfying (3.1) and (4.1) are in (4,5). Here [2,3] is a subset of (4,5), and the lengths of the segments joining 4 and 2, 3 and 5 and 2 and 3 are all equal. By similarly extending [1,3] and [1,2], one could come up with a hexagon 6,7,4,8,9,5 that defines the desired polytope. It is also

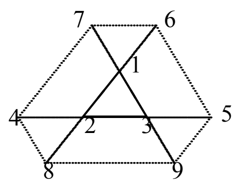

Fig. 2. Illustration of the polytope suggested by Theorem 4.1

interesting to note that if y^* is very close to x_1 then β_1 is much larger than the remaining β_i and the choice in (4.2) forces λ_1 to be much larger than the remaining λ_i. Yet (4.2) is not the only choice one can make. The more one enters the interior of this polygon, the more the choices of λ_i, and indeed the larger the region where a common set of λ_i guarantees (2.3). This in particular has implications to the positioning of the x_i. Thus, with a potentially rough estimate of the position of a source, groups of sensors can collaborate to determine whether they can settle upon a λ_i which ensures (2.3). This provides guidance on how to deploy fewer sensors to achieve greater coverage.

5. SIMULATIONS

We present two examples where (2.4) is applied when: $x_1 = [1,0,0]^T$, $x_2 = [0,2,0]^T$, $x_3 = [-2,-1,0]^T$, $x_4 = [0,0,2]^T$, and $x_5 = [0,0,-1]^T$. We use a $\mu = 0.02$, assume a Gaussian noise perturbs each distance measurement and perform 400-step-long runs with various noise levels. We plot the square of the estimation error (averaged over such 500 runs for each fixed SNR) as a function of the SNR. The results when $y^* = 0$, and $y^* = [-1,1,1]^T$ are in fig. 3 and fig. 4 respectively. Both show very good localization ability despite the fact that noise is present and that in the second case $y^* \notin$ Co$(x_1, \cdots, x_5)$. In both case $P + P^T > 0$.

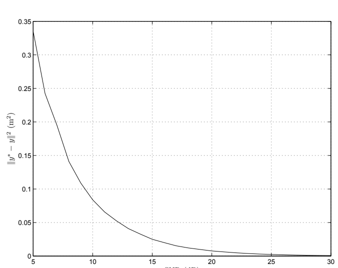

Fig. 3. Simulation example when $y^* = 0$

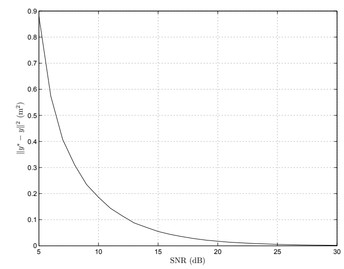

Fig. 4. Simulation example when $y^* = [-1,1,1]^T$

6. CONCLUSION

We have studied conditions under which a localization algorithm involves a globally convergent gradient descent minimization problem. In particular given a set of nodes with known positions (e.g. sensors), this algorithm seeks to localize an object (e.g. a source) whose distances from these nodes are available. Given a set of such sensors and a set of weights we characterize a nontrivial ellipsoidal geographical region surrounding these sensors such that a source lying in this region can be localized through a minimization described above. We also characterize a polytopic region for which there exist weights that permit similar localization.

7. REFERENCES

[1] *IEEE Signal Processing Magazine*, July 2005.

[2] Tarighat, A., and Sayed, A.H., "Network based wireless l.ocalization", *IEEE Signal Processing Magazine*, pp. 24-40, July 2005.

[3] Dandach, S., Fidan, B., Dasgupta, S. and Anderson, B.D.O., "Adaptive source localization with mobile agents", in *Proceedings of CDC*, pp. 2045-2050, San Diego, CA, December 2006.

[4] Patwari, N., Ash, J.N., Kyperountas, S., Hero, A.O., Moses, R.L., and Correal, N.S., "Locating the nodes: cooperative localization in wireless sensor networks", *IEEE Signal Processing Magazine*, pp. 54 - 69, July 2005.

[5] Rabbat, M.G. and Nowak, R.D., "Decentralized source localization and tracking", in *Proceedings of ICASSP*, Montreal, Canada, May 2004.

[6] Rabbat, M.G. and Nowak, R.D., "Distributed optimization in sensor networks", in *Proceedings of 3rd International Symposium on Information Processing in Sensor Networks*, Berkeley, CA, April 2004.

[7] Blatt, D. and Hero, A.O., "Sensor network source localization via projection onto convex sets (POCS)", in *Proceedings of ICASSP*, Philadelphia, P.A., March 2005.

[8] Rydstrom, M., Strom, E. G., and Svensson, A., "Robust sensor network positioning based on projection onto circular and hyperbolic convex sets (POCS)", in *Proceedings of SPAWC*, Cannes, France, July, 2006.

[9] Cao, M., Anderson, B.D.O., and Morse, A.S., "Sensor network localization with imprecise distance measurements", *Systems and Control Letters*, **vol. 55**, pp. 87-93, 2006.

[10] Khalil, H.K., *Nonlinear Systems*, Prentice Hall, 2002.