ROBUST LOCALIZATION IN SENSOR NETWORKS WITH ITERATIVE MAJORIZATION TECHNIQUES

*Sayit Korkmaz and Alle-Jan van der Veen**

TU Delft, Fac. EEMCS, Mekelweg 4, 2628 CD Delft, The Netherlands

ABSTRACT

Self localization in sensor networks with measurements that include outliers is an important problem. E.g., distance measurements based on non-line-of-sight observations can be quite wrong. If not handled properly, such outliers can greatly influence the positioning accuracy. To achieve robustness we consider positioning with Huber estimators. The Huber cost function interpolates between the ℓ_1 and the ℓ_2 norms. The minimization of the Huber cost function can be efficiently obtained via iterative majorization techniques, with the advantageous property of guaranteed convergence to a local minimum.

***Index Terms*—** Iterative majorization, robust location estimation, time-of-arrival, sensor networks, localization, Huber cost function

1. INTRODUCTION

In this paper we are concerned with the self localization problem in sensor networks [1]. Mathematically speaking, the problem is not new. It has long been an issue of great interest to obtain a low dimensional representation from pairwise similarity measurements among objects, which is studied in statistics as multidimensional scaling (MDS) [2]. The node location estimation problem can also be formulated as non-linear programming hence a range of iterative techniques are applicable [3]. An important issue here is the problem of divergence. For a large number of sensors, computational complexity is relevant: the number of unknown parameters in a sensor network with N nodes can be as large as $3N$, and applying nonlinear programming may require the inversion of $3N \times 3N$ matrices.

Iterative majorization techniques emerged as an important way of solving the multidimensional scaling problem. These algorithms have the advantage of simple implementation and guaranteed convergence to a local minimum. E.g., SMACOF [2] is an iterative majorization based solution of the least squares multidimensional scaling problem. A distributed version of SMACOF was developed in [4], henceforth referred to as Costa's algorithm. The main advantage of Costa's algorithm is the fact that it does not require matrix inversion and can be implemented in a parallel and distributed manner.

Self localization in sensor networks with measurements that include outliers apparently has not received much interest. It is however a relevant issue. Specifically for indoor environments we may expect erroneous measurements due to non-line-of-sight propagation. A method based on the EM algorithm is provided [5]. The main drawback of this method is the need to specify outlier statistics—this information may not be available. When outlier statistics are not known a common technique is to use M-Estimators [6]. These estimators modify the least squares cost function in such a way that large residuals are down-weighted. A popular choice would be to use the ℓ_1 norm. This line of research in the context of MDS has been investigated by Heiser in 1987 [7]. Nevertheless there is a singularity problem in the way majorization has been done. To circumvent this problem Heiser proposes to use Huber cost function instead of ℓ_1 norm but provides no solution in [7]. Instead Heiser cites [8] for further discussions. In [8] Heiser did not explicitly consider the sensor network positioning problem but studied a somewhat related problem called *reciprocal location problem*. In Heiser's words "... the problem of locating two p dimensional configuration of points with respect to each other in such a way that the weighted sum of interpoint distances is minimal." Heiser solved the problem by iterative majorization techniques, leading to the LARAMP algorithm.

In this paper we follow up on this work and propose to use this little-known technique for sensor localization in the presence of outliers. Each step in the resulting iterative majorization problem is subsequently solved using Costa's algorithm. This maintains the advantage that it does not require matrix inversion or SVD, and can be implemented in a parallel and distributed manner.

2. PROBLEM FORMULATION

A typical sensor network localization is as follows. Let δ_{ij} denote a measurement of the distance between nodes i and j. The aim is to find the elements of the matrix $\mathbf{X}$ whose ith column $\mathbf{x}_i$ contains the Cartesian coordinates of the ith sensor node. $d_{ij}(\mathbf{X})$ denotes the true distance between node i and node j based on the distance matrix $\mathbf{X}$. We assume that there are $N + M$ sensor nodes in the network. The first N sensors have no or imperfect a priori knowledge about their position. The last M sensors have perfect knowledge of their location. These sensor nodes are anchor nodes. In this formulation the matrix $\mathbf{X}$ takes the following form:

$$\mathbf{X} = [\mathbf{x}_1, ..., \mathbf{x}_N, \mathbf{x}_{N+1}, ..., \mathbf{x}_{N+M}] \, .$$

Related to this, let us define the following index sets:

$$S = \{1, 2, ..., N\}$$
$$S_i = \{i + 1, ..., N + M\} \, .$$

The position estimation problem can be formulated as a minimization problem of a variety of functions. A common choice is the least squares cost function,

$$f^{(1)}(\mathbf{X}) = \sum_{i \in S} \sum_{j \in S_i} w_{ij} (\delta_{ij} - d_{ij}(\mathbf{X}))^2 \, ,$$

where w_{ij} is a weight for the connection between node i and j. If there is no connection between these nodes (the distance between

*This research was supported in part by NWO-STW under the VICI programme (DTC.5893).

i and j is not measured), set $w_{ij} = 0$. The solution to the resulting minimization problem is considered in statistics and a particular solution based on iterative majorization algorithms is the SMACOF algorithm [2].

A second cost function, with a reputation of robustness to outliers, is obtained by replacing the ℓ_2 norm by ℓ_1, i.e.,

$$f^{(2)}(\mathbf{X}) = \sum_{i \in S} \sum_{j \in S_i} w_{ij} |\delta_{ij} - d_{ij}(\mathbf{X})| \,.$$

This line of research is investigated in [7]. As noted, the function is not differentiable at its minimum and is hard to majorize, leading to a degeneracy that makes the problem incomplete and numerically unstable. Heiser [7] mentions this problem and refers to [8] where he proposes to use Huber's cost function.

Huber's cost function interpolates between the ℓ_2 norm minimization and ℓ_1 minimization, i.e.,

$$f^{(3)}(\mathbf{X}) = \sum_{i \in S} \sum_{j \in S_i} w_{ij} \rho(\delta_{ij} - d_{ij}(\mathbf{X})) \,,$$

where the Huber cost function $\rho(\cdot)$ is given as

$$\rho(y) = \begin{cases} y^2 \,, & |y| < k \\ 2k|y| - k^2 \,, & |y| \geq k \,. \end{cases}$$

For small k, this is equivalent to $f^{(2)}(\mathbf{X})$, but it avoids the singularity at 0. Although the problem is not entirely solved, important insights are provided in [8] on how the majorization function can be chosen.

In a different development, Costa et al. [4] consider the case where about some nodes we have some stochastic a priori knowledge on their location: a mean location $\bar{\mathbf{x}}_i$ and a variance $1/r_i$. This leads to a minimization of the following cost function,[1]

$$f^{(4)}(\mathbf{X}) = \sum_{i \in S} \sum_{j \in S_i} w_{ij} (\delta_{ij} - d_{ij}(\mathbf{X}))^2 + \sum_{i \in S} r_i \|\mathbf{x}_i - \bar{\mathbf{x}}_i\|^2 \,.$$

We refer to the solution to this problem as Costa's algorithm.

We aim to solve the following combined problem in this paper:

$$f^{(5)}(\mathbf{X}) = \sum_{i \in S} \sum_{j \in S_i} w_{ij} \rho(\delta_{ij} - d_{ij}(\mathbf{X})) + \sum_{i \in S} r_i \|\mathbf{x}_i - \bar{\mathbf{x}}_i\|^2 \,.$$

3. MAJORIZATION

3.1. Iterative majorization algorithm

In this section we briefly discuss the basic idea behind the iterative majorization technique [2] and how it is applied to the problem of positioning. Let us assume that we want to find the minimum of the function $f(\mathbf{X})$. We want to construct a function $g(\mathbf{X}, \mathbf{Y})$ such that it satisfies

$$\begin{aligned} g(\mathbf{X}, \mathbf{Y}) &\geq f(\mathbf{X}) \\ g(\mathbf{Y}, \mathbf{Y}) &= f(\mathbf{Y}) \,. \end{aligned}$$

Beyond this, it is essential that $g(\mathbf{X}, \mathbf{Y})$ is easier to minimize (for fixed $\mathbf{Y}$). In this formulation $\mathbf{Y}$ denotes a supporting point, i.e., the current best estimate of $\mathbf{X}$. In the algorithm, the successor supporting point is obtained as

$$\mathbf{X}^u = \arg\min_{\mathbf{X}} g(\mathbf{X}, \mathbf{Y}) \,.$$

From here we can state that

$$f(\mathbf{X}^u) \leq g(\mathbf{X}^u, \mathbf{Y}) \leq g(\mathbf{Y}, \mathbf{Y}) = f(\mathbf{Y}) \,,$$

which proves convergence to a local minimum. The resulting algorithm is defined as follows:

1. Set $\mathbf{Y} = \mathbf{Y}_0$ where $\mathbf{Y}_0$ denotes the initial starting point
2. Find update $\mathbf{X}^u = \arg\min_{\mathbf{X}} g(\mathbf{X}, \mathbf{Y})$
3. If $f(\mathbf{Y}) - f(\mathbf{X}^u) < \epsilon$, stop ($\epsilon$ is a small positive constant)
4. Set $\mathbf{Y} = \mathbf{X}^u$ and go to step 2.

The problem is reduced to finding a suitable majorizing function $g(\mathbf{X}, \mathbf{Y})$.

3.2. Majorizing the cost function $f^{(5)}(\mathbf{X})$

Let

$$h_{ij}(\mathbf{X}) = |\delta_{ij} - d_{ij}(\mathbf{X})|$$

and define

$$f_{ij}(\mathbf{X}) = w_{ij} \rho(\delta_{ij} - d_{ij}(\mathbf{X})) \,.$$

In this notation we can write one term in $f^{(3)}(\mathbf{X})$ as

$$f_{ij}(\mathbf{X}) = \begin{cases} w_{ij} h_{ij}^2(\mathbf{X}) \,, & h_{ij}(\mathbf{X}) < k \\ 2k w_{ij} h_{ij}(\mathbf{X}) - w_{ij} k^2 \,, & h_{ij}(\mathbf{X}) \geq k \,. \end{cases}$$

The following majorizing function is proposed (the form is analogous to the one proposed in [8] for a different h_{ij}):

$$g_{ij}(\mathbf{X}, \mathbf{Y}) = \begin{cases} w_{ij} h_{ij}^2(\mathbf{X}) \,, & h_{ij}(\mathbf{Y}) < k \\ \dfrac{k w_{ij}}{h_{ij}(\mathbf{Y})} h_{ij}^2(\mathbf{X}) + k w_{ij} h_{ij}(\mathbf{Y}) - k^2 w_{ij} \,, & h_{ij}(\mathbf{Y}) \geq k \,. \end{cases}$$

It remains to show that the criteria for a majorizing function are satisfied. The proof follows that of [8]. First note that

$$g_{ij}(\mathbf{Y}, \mathbf{Y}) = f_{ij}(\mathbf{Y}) \,.$$

We now need to prove that

$$g_{ij}(\mathbf{X}, \mathbf{Y}) \geq f_{ij}(\mathbf{X}) \,.$$

Depending on $h_{ij}(\mathbf{X})$ and $h_{ij}(\mathbf{Y})$, there are four cases to be considered separately.

In the first case assume that $h_{ij}(\mathbf{X}) < k$ and $h_{ij}(\mathbf{Y}) < k$. In that case by definition we have equality: $g_{ij}(\mathbf{X}, \mathbf{Y}) = f_{ij}(\mathbf{X})$.

In the second case assume that $h_{ij}(\mathbf{X}) \geq k$ and $h_{ij}(\mathbf{Y}) < k$. Then

$$\begin{aligned} & (k - h_{ij}(\mathbf{X}))^2 \geq 0 \\ \Leftrightarrow \quad & k^2 + h_{ij}^2(\mathbf{X}) - 2k h_{ij}(\mathbf{X}) \geq 0 \\ \Leftrightarrow \quad & w_{ij} h_{ij}^2(\mathbf{X}) \geq 2k w_{ij} h_{ij}(\mathbf{X}) - k^2 w_{ij} \,. \end{aligned}$$

It follows that $g_{ij}(\mathbf{X}, \mathbf{Y}) \geq f_{ij}(\mathbf{X})$.

[1] In fact, they consider a straightforward extension of this which includes multiple measurements.

In the third case now assume that $h_{ij}(\mathbf{X}) \geq k$ and $h_{ij}(\mathbf{Y}) \geq k$. Then we need to prove that

$$\frac{kw_{ij}h_{ij}^2(\mathbf{X})}{h_{ij}(\mathbf{Y})} + kw_{ij}h_{ij}(\mathbf{Y}) - k^2w_{ij} \geq 2kw_{ij}h_{ij}(\mathbf{X}) - k^2w_{ij} \,.$$

We can do this in a very simple manner,

$$\begin{aligned}
\Leftrightarrow \quad & \frac{kw_{ij}h_{ij}^2(\mathbf{X})}{h_{ij}(\mathbf{Y})} + kw_{ij}h_{ij}(\mathbf{Y}) \geq 2kw_{ij}h_{ij}(\mathbf{X}) \\
\Leftrightarrow \quad & \frac{h_{ij}^2(\mathbf{X})}{h_{ij}(\mathbf{Y})} + h_{ij}(\mathbf{Y}) \geq 2h_{ij}(\mathbf{X}) \\
\Leftrightarrow \quad & h_{ij}^2(\mathbf{X}) + h_{ij}^2(\mathbf{Y}) \geq 2h_{ij}(\mathbf{X})h_{ij}(\mathbf{Y}) \\
\Leftrightarrow \quad & (h_{ij}(\mathbf{X}) - h_{ij}(\mathbf{Y}))^2 \geq 0 \,.
\end{aligned}$$

The last statement is always true. Finally, in the fourth case, assume that $h_{ij}(\mathbf{X}) < k$ and $h_{ij}(\mathbf{Y}) \geq k$, so that

$$0 \leq h_{ij}(\mathbf{X}) < k \leq h_{ij}(\mathbf{Y}) \,.$$

From here we may write that

$$\begin{aligned}
\Leftrightarrow \quad & \frac{h_{ij}(\mathbf{X})}{k} < 1 \leq \frac{h_{ij}(\mathbf{Y})}{k} \\
\Leftrightarrow \quad & \frac{h_{ij}^2(\mathbf{X})}{k^2} < 1 \leq \frac{h_{ij}(\mathbf{Y})}{k}
\end{aligned}$$

and conclude that

$$\Leftrightarrow \quad h_{ij}^2(\mathbf{X}) \leq kh_{ij}(\mathbf{Y}) \,.$$

If $h_{ij}(\mathbf{Y}) > k$, we can multiply both sides by $(h_{ij}(\mathbf{Y}) - k)$ to obtain

$$\begin{aligned}
\Leftrightarrow \quad & (h_{ij}(\mathbf{Y}) - k)h_{ij}^2(\mathbf{X}) \leq kh_{ij}(\mathbf{Y})(h_{ij}(\mathbf{Y}) - k) \\
\Leftrightarrow \quad & h_{ij}(\mathbf{Y})h_{ij}^2(\mathbf{X}) \leq kh_{ij}^2(\mathbf{X}) + kh_{ij}^2(\mathbf{Y}) - k^2h_{ij}(\mathbf{Y}) \\
\Leftrightarrow \quad & h_{ij}^2(\mathbf{X}) \leq \frac{kh_{ij}^2(\mathbf{X})}{h_{ij}(\mathbf{Y})} + kh_{ij}(\mathbf{Y}) - k^2 \\
\Leftrightarrow \quad & w_{ij}h_{ij}^2(\mathbf{X}) \leq \frac{kw_{ij}h_{ij}^2(\mathbf{X})}{h_{ij}(\mathbf{Y})} + kw_{ij}h_{ij}(\mathbf{Y}) - w_{ij}k^2 \,,
\end{aligned}$$

which shows $g_{ij}(\mathbf{X}, \mathbf{Y}) \geq f_{ij}(\mathbf{X})$. It remains to consider the case where $h_{ij}(\mathbf{Y}) = k$. In that case the final equation becomes an equality. This proves that $g_{ij}(\mathbf{X}, \mathbf{Y})$ satisfies the criteria for a majorizing function.

A majorizing function for $f^{(5)}(\mathbf{X})$ is now immediately obtained as

$$g(\mathbf{X}, \mathbf{Y}) = \sum_{i \in S} \sum_{j \in S_i} g_{ij}(\mathbf{X}, \mathbf{Y}) + \sum_{i \in S} r_i \|\mathbf{x}_i - \bar{\mathbf{x}}_i\|^2 \,.$$

We can optimize $f^{(5)}(\mathbf{X})$ via the iteration in section 3.1. In each step, for fixed $\mathbf{Y}$, the majorized cost function reduces to a nonlinear least squares problem of the same form as $f^{(4)}(\mathbf{X})$.

4. A SUMMARY OF COSTA'S ALGORITHM

At every step of the iterative majorization algorithm we need to solve a nonlinear least squares problem of the form $f^{(4)}(\mathbf{X})$. For this we propose to use Costa's algorithm [4], as it is of low complexity. This algorithm is again based on an iterative majorization. For completeness, we summarize this algorithm for the case at hand.

For $p = 1, 2, \cdots$ until convergence, compute

$$\mathbf{x}_i^{(p+1)} = a_i\Big(r_i\bar{\mathbf{x}}_i + \mathbf{X}^{(p)}\mathbf{b}_i^{(p)}\Big) \,, \quad i = 1, \cdots, N$$

where

$$a_i^{-1} = \sum_{j=1, j\neq i}^{N} w_{ij} + \sum_{j=N+1}^{N+M} 2w_{ij} + r_i$$

and $\mathbf{b}_i^{(p)} = [b_{i1},\ b_{i2},\ \cdots,\ b_{i,N+M}]^T$, the entries of which are given as

$$\begin{aligned}
b_{ij} &= w_{ij}\Big(1 - \delta_{ij}/d_{ij}(\mathbf{X}^{(p)})\Big) \,, \qquad j \neq i, j \leq N \\
b_{ii} &= \sum_{j=1, j\neq i}^{N} w_{ij}\delta_{ij}/d_{ij}(\mathbf{X}^{(p)}) + \sum_{j=N+1}^{N+M} 2w_{ij}\delta_{ij}/d_{ij}(\mathbf{X}^{(p)}) \\
b_{ij} &= 2w_{ij}\Big(1 - \delta_{ij}/d_{ij}(\mathbf{X}^{(p)})\Big) \,, \qquad j > N \,.
\end{aligned}$$

5. SIMULATIONS

The following simulation was performed to provide some insight into the performance of the proposed algorithms.

Fig. 1 shows the true locations of 100 sensors in a sensor network. We assume that all the sensors have connectivity with each other, in other words $w_{ij} \neq 0$. Sensors numbered from 91 to 100 have known coordinates and the rest of sensors have unknown coordinates. All the sensor nodes with known coordinates have y-coordinate 10, in other words they lie on the upper line of the sensor network.

Fig. 2 shows the estimated locations for this sensor network, for a case where 4 nodes are malfunctioning and any pairwise distance measurement that involves these malfunctioning nodes is an outlier measurement which was obtained by adding 10 to the true distance. The noise variance on the pairwise distance measurements is set at 0.1 for all measurements. In the proposed algorithm, the Huber threshold was set to $k = 0.01$. As can be seen from the figure, the malfunctioning nodes do not disturb the location estimates of the rest of the nodes. This is the robustness result we aimed to achieve. (However there is no hope for localizing the outlier nodes.)

For comparison, Fig. 3 shows the estimated locations based on the least squares cost function $f^{(4)}$, using Costa's algorithm. We used the same outlier setup as in the previous section. As can be seen from the figure the outlier nodes significantly influence the estimation accuracy of the whole sensor network except the anchor nodes. This result demonstrates the advantage of using robust algorithms.

6. CONCLUSION

As shown in the simulations, outliers can greatly influence positioning accuracy for algorithms based on least squares cost functions. The Huber cost function provides a robust estimator, which can be calculated using iterative majorization algorithms. Although this gives good performance, the price paid for robustness is increased computational complexity. In our experience, the proposed algorithm may have slow convergence.

7. REFERENCES

[1] N. Patwari, J. N. Ash, S. Kyperountas, A. O. Hero, R. L. Moses, and N. S. Correal, "Locating the nodes," *IEEE Signal Processing Magazine*, pp. 54–69, July 2005.

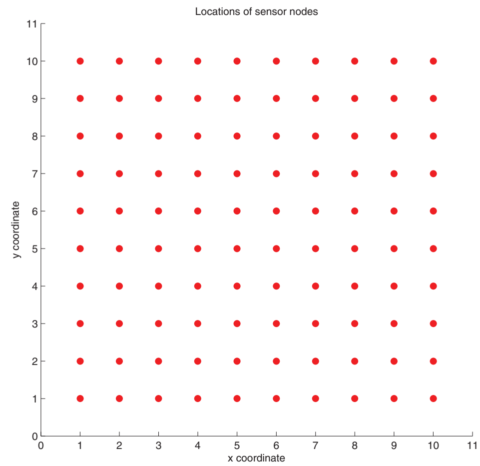

Fig. 1. Exact coordinates of the nodes in the network

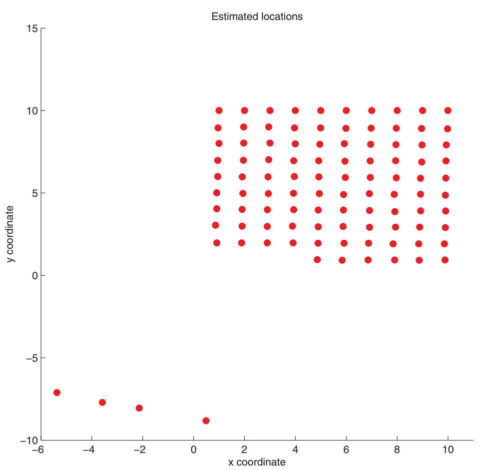

Fig. 2. Position estimation using the proposed algorithm. There are 4 outlier nodes and 10 reference nodes (top row). The outliers do not influence the location estimates of the other nodes.

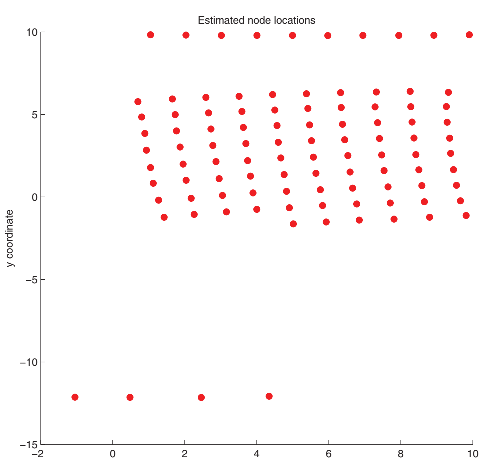

Fig. 3. Position estimation using least squares localization. There are 4 outlier nodes and 10 reference nodes (top row). The outliers greatly influence the location estimates.

[2] I. Borg and P. Groenen, *Modern multidimensional scaling: Theory and applications*, Springer, New York, 1997.

[3] D. P. Bertsekas, *Nonlinear Programming*, Athena Scientific, Belmont, Massachusetts, 1999.

[4] J. A. Costa, N. Patwari, and A. O. Hero, "Distributed weighted multidimensional scaling for node localization in sensor networks," *ACM Transactions on Sensor Networks*, vol. 2, no. 1, pp. 39–64, Feb. 2006.

[5] J. N. Ash and R. L. Moses, "Outlier compensation in sensor network self localization vie the EM algorithm," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing ICASSP*, Philadelphia, US, Mar. 2005, pp. 749–752.

[6] P. J. Rousseeuw and A. M. Leroy, *Robust regression and outlier detection*, Wiley, New York, 1987.

[7] W. J. Heiser, "Multidimensional scaling with least absolute residuals," in *Proceedings of the first conference of the international federation of classification societies (IFCS)*, Aachen, Germany, June 1987, pp. 455–462.

[8] W. J. Heiser, *Notes on the LARAMP algorithm*, Internal report, Dept. of Data Theory, Leiden University, 1987.