Guaranteed Robust Distributed Estimation in a Network of sensors
Jean-Benoist Léger, Michel Kieffer

To cite this version:

HAL Id: hal-00549168
https://hal.archives-ouvertes.fr/hal-00549168
Submitted on 21 Dec 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
GUARANTEED ROBUST DISTRIBUTED ESTIMATION IN A NETWORK OF SENSORS

J. B. Leger and M. Kieffer

L2S – CNRS - SUPELEC - Univ Paris-Sud
3 rue Joliot-curie, 91192 Gif-sur-Yvette, France
jean-benoist.leger@ens-cachan.fr, kieffer@lss.supelec.fr

ABSTRACT

This paper proposes a guaranteed robust bounded-error distributed estimation algorithm. It may be employed to perform parameter estimation from data collected in a network of wireless sensors. The algorithm is robust to an arbitrary number of outliers. Using interval analysis, one is able, provided that the network is connected, to evaluate at each sensor, an outer approximation of the set of all parameter values which are consistent with a given number of measurements, and with noise bounds. An application to a robust distributed source localization problem is considered.

Index Terms— Bounded-error estimation, distributed estimation, interval analysis, network of sensors, outliers, robust estimation.

1. INTRODUCTION

A network of wireless sensors (NWS) is a set of autonomous devices, with limited computing capability and autonomy, exchanging information via a wireless channel. Various types of sensors may be considered, e.g., for measuring pressure, temperature, sound, vibration, motion... Many applications (environment monitoring, home automation, traffic control) may take advantage of NWS, see, e.g., [1, 2].

Challenging problems arise when considering parameter or state estimation using measurements provided by a NWS. Two main types of estimation techniques may be considered. In centralized approaches, all measurements obtained by the sensors are transmitted to a central processing unit (CPU), see, e.g., [3]. Many data have then to be send to a given point of the network. Moreover, this solution is not robust to a failure of the CPU, since the estimate is only available at that point of the network. Alternative distributed estimation techniques for constant [4] and time varying parameters [5, 6] have been provided. In this case, each sensor is responsible for the processing of its measurement and of data provided by neighboring sensors. An increased robustness to failure of the CPU is thus obtained.

Nevertheless, distributed solutions may not be very robust against erroneous measurements provided by some defective sensors. Albeit robust estimators have been proposed in a centralized context, using bounded-error estimation [7] or linear programming [8], the extension of these techniques to a distributed context is far from being trivial.

This paper considers distributed bounded-error estimation in a NWS [9]. Measurement noise is assumed to be bounded with known bounds, and one aims at evaluating the set of all values of the parameter vector which is consistent with the measurements and the bounds on the measurement noise. It provides a guaranteed and robust estimator in a distributed context using interval analysis [10]. By guaranteed, it is meant that no parameter value consistent with a fixed number of measurements is missed.

Section 2 recalls the robust bounded-error parameter estimation problem in a NWS. The centralized approach is described in Section 3 to provide a reference to the distributed approach presented in Section 4. Implementation issues are considered in Section 5 before providing simulation results in Section 6.

2. PROBLEM FORMULATION

Consider a network of \( N \) sensors spread in an environment. The aim is to provide the estimation of an unknown parameter vector \( \mathbf{p}^* \in \mathbb{P} \) using measurements \( \mathbf{y}_i, i \in [1, N] \) provided by each sensor of the network. The measurement \( \mathbf{y}_i \) is assumed to be linked to \( \mathbf{p}^* \) via the measurement model

\[
\mathbf{y}_i = \mathbf{f}_i(\mathbf{p}^*) + \mathbf{e}_i \quad i \in [1, N]
\]

where \( \mathbf{e}_i \) is assumed to remain in some known interval vector (box) \([\mathbf{e}, \mathbf{e}]\). Introducing \([\mathbf{y}] = \mathbf{y}_1 - [\mathbf{e}, \mathbf{e}]\), one has

\[
\mathbf{f}_i(\mathbf{p}^*) \in [\mathbf{y}].
\]

Bounded-error parameter estimation [11, 12] aims at characterizing the set \( S_0 \) of all parameter values which are consistent with the measurements, the measurement model, and the bounds on the measurement noise, i.e.,

\[
S_0 = \{ \mathbf{p} \in \mathbb{P} | \forall i \in [1, N], \mathbf{f}_i(\mathbf{p}) \in [\mathbf{y}_i] \}.
\]

When they are some outliers, e.g., in the case of a defective sensors, for some measurements, the noise may not remain in its bounds, \( S_0 \) may be then empty. In such situations, one may define a set estimator for \( \mathbf{p} \) robust to \( q \) outliers as follows

\[
\mathcal{S}_q = \{ \mathbf{p} \in \mathbb{P} | \lambda(\mathbf{p}) \geq N - q \}
\]

where

\[
\lambda(\mathbf{p}) = \sum_{i=1}^{N} I_{[y_i]}(\mathbf{f}_i(\mathbf{p}))
\]

and

\[
I_A(x) = \begin{cases} 
1 & \text{if } x \in A \\
0 & \text{else}
\end{cases}
\]

Guaranteed inner and outer approximations \( \overline{S}_q \) and \( \underline{S}_q \) of \( S_q \) may be obtained for any value of \( q \) using interval analysis [10], provided...
that all measurements are collected at a CPU, to which all measurements yi and models fi have been transmitted. By guaranteed, it is meant that one is able to numerically prove that in \( \mathbb{P} \setminus S_q \), all values of the parameter vector are inconsistent with at least \( q + 1 \) measurements, and in \( S_q \), all values of the parameter vector are consistent with at least \( N - q \) measurements.

The aim of this paper is to propose a distributed robust bounded-error estimator, i.e., to provide an estimation algorithm which is able to evaluate at each sensor \( i \) of the network an outer approximation \( S_{q,i} \) of \( S_q \) using only a subset of the measurements available in the network. The aim is to be robust to a failure of the CPU, to compute at each sensor of the network partial estimates with a only subset of the measurements, and if possible, to reduce the amount of data exchanged within the network. In what follows, the network is assumed to be entirely connected, i.e., any sensor of the NWS is able to exchange information with any other sensor, in one or several hops.

3. ROBUST CENTRALIZED APPROACH

In this approach, all sensors send their measurements and measurement functions to a central processing unit. The robust bounded-error approach presented in [7] is briefly recalled here to serve as reference of the distributed approach detailed in Section 4.

The robust estimator is based on the notion of inclusion function, introduced by interval analysis [13, 10]. Consider a function \( f : \Omega \rightarrow \mathbb{R}^n \), an inclusion function \( f \) for \( f \) has to be such that

\[
\forall [x] \subset \Omega \quad f([x]) \supset f([x]).
\]

The natural inclusion function is an inclusion function obtained by replacing all occurrences of the variable \( x \) in the formal expression of \( f(x) \) by the interval counterpart \([x]\). It allows to compute an outer-approximation of the range of \( f \) over any interval \([x] \subset \Omega \). For more details, see [13, 10].

Assuming that an inclusion function \( \lambda \) for \( \lambda \) in (5) is available, one may use the SIVIA algorithm [14] to evaluate an inner approximation \( S_q \) and an outer approximation \( S_q \) of \( S_q \) consisting of unions of non-overlapping boxes of \( \mathbb{P} \), \( S_n \), and \( S_q \) are initialized as \( \emptyset \).

Starting with a working LIFO list \( \mathcal{W} \) of boxes to be processed, initially containing the box \([p_0] = \mathbb{P}\), SIVIA extracts a box \([p]\) from \( \mathcal{W} \) and applies the following tests.

- If \( \lambda([p]) \in [N - q, 0, N, -q] \), then all parameters in \([p]\) are consistent with at least \( N - q \) measurements or more and \([p]\) is stored in \( S_q \) and \( S_n \).
- If \( \lambda([p]) \in [0, N - q] \), then all parameters in \([p]\) are not consistent with \( q + 1 \) measurement or more, and \([p]\) is dropped.
- If the size of \([p]\) is larger than some parameter \( \varepsilon \), \([p]\) is bisected into two subboxes \([p']\) and \([p'']\), which are stored in \( \mathcal{W} \).
- If the size of \([p]\) is smaller than \( \varepsilon \), \([p]\) is stored into \( S_q \).

One of the interesting feature of this approach is that it is not necessary to specify a priori the sensors which are defective. Only the number \( q \) of erroneous data the estimator has to be robust to has to be specified. The approach considered by GOMNE [7] consists in starting with \( q = 0 \) and increasing \( q \) until a non-empty solution set \( S_q \) is obtained. Note that with this approach, the solution set \( S_q \) is only guaranteed to contain the true parameter value \( p \) if the number of outliers is actually less than \( q \). In what follows, the NWS is assumed to be entirely connected, i.e., each sensor is able to exchange information with any other sensor of the network, in one or several hops.

4. IDEALIZED ROBUST DISTRIBUTED APPROACH

In this context, each sensor has to process its own measurement and information transmitted by neighboring sensors. One aims at characterizing \( S_q \) in a guaranteed way, as in the centralized approach. Consider the subset of measurement indexes \( J \subset [1, N] \), and define the set

\[
S_q \cap \bigcup_{I \subset J, \text{card}(I) = \text{card}(J) - q} \cap \bigcup_{i \in I} \mathbb{P}_i,
\]

of all parameters consistent with \( \text{card}(J) - q \) or more measurements provided by sensors with index in \( J \), with where \( \mathbb{P}_i = \{ p \in \mathbb{P} | f_i(p) \in [y_i] \} \), the set of parameters consistent with the measurement provided the sensor \( i \) and \( \text{card}(J) \), the cardinal number of the set \( A \). One may easily verify that \( S_q = S_q^{1[N]} \) and

\[
\forall J_1 \subset J_2 \subset [1, N], \quad S_q^{J_1} \supset S_q^{J_2} \supset S_q \tag{9}
\]

Assume that a sensor has evaluated \( S_q^{J_1} \) and that \( S_q^{J_2} \) has been provided by one of its neighbors. According to (9), to obtain a better outer-approximation of \( S_q \), the sensor has to compute \( S_q^{J_1 \cup J_2} \). If \( J_1 \cap J_2 \neq \emptyset \), there is no simple relation between \( S_q^{J_1} \cap S_q^{J_2} \) and \( S_q^{J_1 \cup J_2} \). Now, if \( J_1 \cap J_2 = \emptyset \), one may easily prove that \( S_q^{J_1} \cap S_q^{J_2} \supset S_q^{J_2} \), but both sets are not equal in general. In fact, to compute \( S_q^{J_1 \cup J_2} \), all \( S_q^{J_1} \), \( \cdots \), \( S_q^{J_1} \), \( \cdots \), \( S_q^{J_2} \) are needed, see Appendix A.1. Thus, each sensor has to transmit \( S_q^{J_1} \), \( \cdots \), \( S_q^{J_2} \) in place of only \( S_q^{J_1} \).

For any subset of indexes \( J \subset [1, N] \), consider the set \( \Gamma_q^{J} = \{ S_q^{J_1}, S_q^{J_1 \cup J_2} \} \), see Figure 1. In what follows, sensors send and receive such sets, and try to compute \( \Gamma_q^{J} \) to obtain \( S_q^{1[N]} = S_q \).

The number of tolerated outliers \( q \) is assumed to be fixed a priori.

**Fig. 1.** Representation some set \( \Gamma_q^{J} = \{ S_q^{J_1}, S_q^{J_1 \cup J_2} \} \)

Initially, each sensor \( i \) processes its own measurement, to get \( \mathbb{P}_i \) and \( \Gamma_i^{(1)} = (\mathbb{P}_i, \mathbb{P}_i, \cdots, \mathbb{P}_i) \). Then, it broadcasts this first estimate to its neighboring sensors and receives similar structures. After a first round of communication, the \( i \)-th sensor is able to improve its estimates as follows. If \( J_1 \cap J_2 = \emptyset \) (combination constraint), then \( \Gamma_i^{(1)} \) and \( \Gamma_i^{(2)} \) can be used to get \( \Gamma_i^{(1)} \) by computing each \( S_q^{J_1 \cup J_2} \) as

\[
\forall q' \in [0, q], \quad S_q^{J_1 \cup J_2} = \bigcup_{q_1 + q_2 = q'} S_q^{J_1} \cap S_q^{J_2}, \tag{10}
\]

see Appendix A.1.

For the next round of communication, each sensor broadcasts the best \( \Gamma_q \) (with the largest \( \text{card}(J) \)). Once all sensors have exchanged improved estimates, new improvements are possible. The two phases (estimation and communication) may be performed until convergence, i.e., until all sensors have obtained \( S_q^{1[N]} = S_q \), which occurs in finite time the proof is not provided here due to lack of space. Computations may also be stopped at any time, each sensor of the network having an outer-approximation of \( S_q \), which improves when more data are exchanged.
In Section 4, sets such as $\mathcal{S}_q^J$ are considered, in order to be able to determine a guaranteed outer approximation of the solution $\mathcal{S}_q^I$. Such outer-approximation may consists of any simple geometric shape, such as ellipsoids, polytopes, or unions of non-overlapping boxes or subpavings [10], which are considered here.

A single subpaving of $\mathcal{P}$ can be used to represent $\Gamma_q^J$. A subpaving may be easily described by a binary tree. Each leaf of the tree has to be labeled with $\ell$ to indicate that the corresponding box is a subset of $\mathcal{S}_\ell^J$. This subpaving implementation of $\Gamma$ as labeled binary trees, allows computation for each sensor to becomes only unions and intersections of subpavings.

Each node stores intermediate results with $\Gamma_q^J = \left( \mathcal{S}_0^J, \mathcal{S}_1^J, \ldots, \mathcal{S}_q^J \right)$ in place of $\Gamma_q^J$ and transmitted to neighboring nodes. Efficient routing protocols, such as Optimized Link State Routing Protocol [15], may be used to satisfy the recombination constraint more easily. The node of the network use multipoint distribution relays (MPR) for transmission. For a given node, only a subset of its neighbors relay its message. The selection of MPRs is adaptive and done in real time. For more details, see [15]. We impose here a dynamic hierarchical structure where a sensor selects his MPRs and sends his sets $\Gamma$ only to its MPRs.

6. EXPERIMENTAL PART

A simple single source localization in a 2D-environment with a NWS is considered, see Figure 2. A network of 9 regularly-spaced nodes is considered. Each sensor measures the power it receives from the source. All measurement errors are bounded: for a received power $y_i$ by the $i$-th sensor, the noise-free measurement is assumed to belong to the interval $[y_i - w, y_i + w]$ with $w = 1.7$. Two outliers are introduced by hand, concerning Sensors 4 and 6. The location of the source $\mathbf{p}^* = (\theta_1, \theta_2)$ has then to be estimated.

The following measurement model is considered for the $i$-th sensor

$$y_{m,i} = \frac{P_0}{d((\theta_1, \theta_2), (\theta_{1i}, \theta_{2i}))^\eta}$$

where $(\theta_{1i}, \theta_{2i})$ is the location of the sensor $i$, and where $d(P_1, P_2)$ is the distance between $P_1$ and $P_2$. Moreover, $P_0 = 1$ and $\eta = 2$ are assumed to be known.

For a number of outliers $q \in \{0, 1, 2\}$, the centralized robust bounded-error estimator for $(\theta_1, \theta_2)$ provides the results represented in Figure 2. In distributed approach, Figure 3 describes the estimates obtained by the 4-th sensor. The sets of $\Gamma_q^J$ are outer approximation of sets of $\Gamma_q^I$.

Increasing the number of nodes of the network, one may evaluate the evaluation of the number of iterations required until convergence, see Figure 4. The complexity seems to be linear with the number of sensors.

7. CONCLUSION

This paper introduces a guaranteed robust bounded-error distributed estimation algorithm. This algorithm is robust to any number $q$ of outliers. It is able to provide at each sensor of the NWS an outer-approximation of the set of all values of the parameter vector which are consistent with all except $q$ measurements, or more, the model structure and the noise bounds. It is not necessary to specify a priori the measurements which are deemed as outliers.

The number of outliers the estimator has to be robust to has to be specified a priori. If labeled trees are used to represent subpavings, themselves used to describe $\Gamma$, exchanged between sensors, the complexity of these structures is not affected by the value of $q$. The complexity of the algorithm has to be evaluated more carefully. The sets $\Gamma$ are quite complex, and their transmission may require some resources. One could imagine an alternative way to provide a robust estimator by exchanging measurements within clusters, and...
exchange estimates between cluster heads to optimize the amount of data to be exchanged within the network.

A. PROOFS

For a node \( i \) of a WSN of \( N \) nodes, consider \( \mathcal{L}_1(i) \subset [1, N] \), the set of all indexes of sensors which are directly connected to \( i \). One chooses \( \mathcal{L}_0(i) = \{i\} \). Then, for any \( k \in [1, N] \) the set \( \mathcal{L}_k(i) = \mathcal{L}_{k-1}(i) \cup \left( \bigcup_{1 \leq i \leq k-1} \mathcal{L}_i(s) \right) \) contains all indexes of sensors which can communicate with \( i \) communication or less. Assuming that the network is fully connected translates into

\[
\forall i \in [1, N] \quad \exists d \in \mathbb{N} \quad \mathcal{L}_d(i) = [1, N]. \tag{12}
\]

A.1. Computation rule

Considering \( J_1 \subset [1, N] \) and \( J_2 \subset [1, N] \), with \( J_1 \cap J_2 = \emptyset \), one aims at proving (10). For that purpose, take some \( q' \in [0, N] \).

First consider \( p \in \bigcup_{q_1 + q_2 = q'} S_{q_1}^{J_1} \cap S_{q_2}^{J_2} \). There exists \( q_1, q_2 \in [1, N] \) with \( q_1 + q_2 = q' \), and with \( p \in S_{q_1}^{J_1} \) and \( p \in S_{q_2}^{J_2} \). One may also find \( I_1 \subset J_1 \) and \( I_2 \subset J_2 \), with \( \text{card}(I_1) = \text{card}(J_1) - q_1 \), \( \text{card}(I_2) = \text{card}(J_2) - q_2 \), and with \( p \in \bigcap_{i \in I_1} \mathcal{P}_i \) and \( p \in \bigcap_{i \in I_2} \mathcal{P}_i \). Since \( J_1 \cap J_2 = \emptyset \), then \( I_1 \cap I_2 = \emptyset \), and \( \text{card}(I_1 \cup I_2) = \text{card}(J_1) - q_1 + \text{card}(J_2) - q_2 = \text{card}(J_1 \cup J_2) - q' \). Now, take \( I = I_1 \cup I_2 \subset J_1 \cup J_2 \), then \( p \in \bigcap_{i \in I} \mathcal{P}_i \) with \( \text{card}(I) = \text{card}(J_1 \cup J_2) - q' \), and one has

\[
p \in \bigcup_{I_1 \subset J_1 \cup J_2} \left( \bigcap_{i \in I_1} \mathcal{P}_i \right) \tag{13}
\]

Finally, \( p \in S_{q_1}^{J_1} \cup S_{q_2}^{J_2} \) and \( \bigcup_{q_1 + q_2 = q'} S_{q_1}^{J_1} \cap S_{q_2}^{J_2} \subset S_{q'}^{J_1 \cup J_2} \).

Second, take \( p \in S_{q_1}^{J_1} \cup S_{q_2}^{J_2} \). There exists \( I \subset J_1 \cup J_2 \) such that \( p \in \bigcap_{i \in I} \mathcal{P}_i \), and \( \text{card}(I) = \text{card}(J_1 \cup J_2) - q' \). Consequently, there exists \( I_1 \subset J_1 \) and \( I_2 \subset J_2 \) with \( I_1 \cup I_2 = I \). Since \( J_1 \cap J_2 = \emptyset \), then \( I_1 \cap I_2 = \emptyset \). Let \( q_1 = \text{card}(J_1) - \text{card}(I_1) \) and \( q_2 = \text{card}(J_2) - \text{card}(I_2) \). One has then \( q_1 + q_2 = \text{card}(J_1) + \text{card}(J_2) - (\text{card}(I_1) + \text{card}(I_2)) = \text{card}(J_1 \cup J_2) - \text{card}(I_1 \cup I_2) = \text{card}(J_1 \cup J_2) - \text{card}(I) = q' \). So

\[
\begin{align*}
p \in \bigcap_{i \in I_1} \mathcal{P}_i & \quad \text{card}(I_1) = \text{card}(J_1) - q_1 \\
p \in \bigcap_{i \in I_2} \mathcal{P}_i & \quad \text{card}(I_2) = \text{card}(J_2) - q_2
\end{align*} \tag{14}
\]

One has then \( p \in S_{q_1}^{J_1} \) and \( p \in S_{q_2}^{J_2} \) and \( q_1 + q_2 = q' \). Finally, \( p \in \bigcup_{q_1 + q_2 = q'} S_{q_1}^{J_1} \cap S_{q_2}^{J_2} \).

B. REFERENCES


