DISTRIBUTED ESTIMATION IN SENSOR NETWORKS WITH IMPERFECT MODEL INFORMATION: AN ADAPTIVE LEARNING-BASED APPROACH

Qing Zhou[1], *Soummya Kar*[2], *Lauren Huie*[3], *and Shuguang Cui*[1]

[1]Department of Electrical and Computer Engineering, Texas A&M University, College Station, TX 77843
[2]Department of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, PA 15213
[3]Air Force Research Lab, 525 Brooks Road, Rome, NY 13441

ABSTRACT

The paper considers the problem of distributed estimation of an unknown deterministic scalar parameter (the target signal) in wireless sensor networks (WSNs), in which each sensor receives a single snapshot of the field. The observation or sensing *mode* is only partially known at the corresponding nodes, perhaps, due to their limited sensing capabilities or other unpredictable physical factors. Specifically, it is assumed that the observation process at a node switches stochastically between two modes, with mode one corresponding to the desired signal plus noise observation mode (a *valid* observation), and mode two corresponding to pure noise with no signal information (an *invalid* observation). With no prior information on the local sensing modes (valid or invalid), the paper introduces a learning-based distributed estimation procedure, the mixed detection-estimation (MDE) algorithm, based on closed-loop interactions between the iterative distributed mode learning and estimation. The online learning (or sensing mode detection) step re-assesses the validity of the local observations at each iteration, thus refining the ongoing estimation update process. The convergence of the MDE algorithm is established analytically. Simulation studies show that, in the high signal-to-noise ratio (SNR) regime, the MDE estimation error converges to that of an ideal (centralized) estimator with perfect information about the node sensing modes. This is in contrast with the estimation performance of a naive average consensus based distributed estimator (with no mode learning), whose estimation error blows up with an increasing SNR.

***Index Terms*—** Distributed estimation, distributed learning, adaptive, stochastic switching, sensor networks.

1. INTRODUCTION

A key issue in WSNs concerns the attainment of meaningful network-wide consensus of intelligence based on unreliable locally sensed data [1, 2, 3]. Due to the limited sensing capability and other unpredictable physical factors, these individual observations are usually imperfect. In particular, we assume that the observation process at a node switches stochastically between two modes, i.e., a valid observation mode and an invalid observation mode. This stochastic property of the observation modes causes the unreliable performance of traditional distributed consensus algorithms, e.g., the variance of a naive averaging estimate [4] increases in proportion to the square of SNR.

An MDE algorithm is introduced in this paper, which is a learning-based distributed estimation procedure based on closed-loop interactions between the iterative distributed mode learning and estimation. The main contribution of this MDE algorithm is that the mode learning part detects the validity of the local observation simultaneously when performing the distributed estimation task. In each round of iteration, each node locally detects the observation validity with the maximum a priori probability (MAP) criterion based on the knowledge of the local current estimate of the target together with the local observation. And then the local estimate is refined with the detected validities of the local observations and other exchanged information from their neighbors. By alternatively detecting validity and estimating the target, the sensor network can achieve a global consensus among all nodes. We analytically establish the convergence of the MDE algorithm. With simulations, we show that in the high SNR regime, the MDE estimation error converges to that of an ideal estimator with perfect information about the node sensing modes. The adaptive learning property of the MDE algorithm achieves a reliable estimation performance, in contrast to the estimation performance of a naive average consensus based algorithm in the high SNR regime.

2. NETWORK MODEL

Let $\mathcal{N}_i$ and Ω_i denote sensor node i and the set of its neighbors respectively. The received signal at $\mathcal{N}_i$ is $r_i = a_i x + w_i$, $i \in \{1, 2, ..., N\}$, where $a_i = \{0, 1\}$ is an unknown validity index of the observation at node $\mathcal{N}_i$, i.e., $a_i = 1$ indicates that r_i is a valid observation of the signal + noise form, whereas

This work is supported in part by DoD-AFOSR under Grant FA9550-09-1-0107 and DoD-DTRA under Grant HDTRA1-08-1-0010.

$a_i = 0$ corresponds to the case of pure noise. Also, the w_i's are independent Gaussian white noises with zero mean and variance σ^2. Although the exact instantiations of the a_i's are unknown, we assume that the a_i's are i.i.d. Bernoulli random variables and the probability $p_1 \triangleq \Pr\{a_i = 1\}$ is known *apriori*. We denote the mean and the variance of a_i as $\bar{a}$ and σ_a, respectively, i.e., $\bar{a} = p_1$ and $\sigma_a = p_1(1-p_1)$. We are interested in estimating x using an iterative distributed procedure, in which each node $\mathcal{N}_i$ may only use its neighbors' current state information for updating its estimate (state) at time slot t.

Denoting by $\mathbf{r}$ the network observation vector, i.e.,

$$\mathbf{r} = \mathbf{a}x + \mathbf{w}, \tag{1}$$

with $\mathbf{r} = [r_1, r_2, ..., r_N]^T$, $\mathbf{a} = [a_1, a_2, ..., a_N]^T$, and $\mathbf{w} = [w_1, w_2, ..., w_N]^T$, we note that, a rather straight-forward approach based on naive averaging could be cast as,

$$\hat{x}^N_{\text{Naive}} = \frac{\mathbf{1}^T\mathbf{r}}{N\bar{a}}, \tag{2}$$

which yields an unbiased estimate with the property that $\hat{x}^N_{\text{Naive}} \to x$ as $N \to \infty$. The variance (which coincides with the mean-squared error) of $\hat{x}^N_{\text{Naive}}$ may be expressed as

$$var(\hat{x}^N_{\text{Naive}}) = \frac{1}{N}[\frac{\sigma^2}{\bar{a}^2}(1 + SNR\sigma_a^2)]. \tag{3}$$

Although this naive estimate is quite straight-forward in terms of implementation, we observe from (3) that the precision is poor in the high SNR regime, where in particular, the mean-squared error (MSE) blows up with the SNR. Other than being inaccurate, since the SNR is unknown apriori, the estimate $\hat{x}^N_{\text{Naive}}$ is highly unreliable and, in fact, contradicts with the common intuition that a higher SNR leads to a better performance.

On the other extreme, if we assume that $\mathbf{a}$ is perfectly known, we may generate an *ideal* estimate $\hat{x}^N_{Ideal}$ of x by eliminating the invalid observations, i.e.,

$$\hat{x}^N_{Ideal} = \frac{\sum_{\{i:a_i=1\}} r_i}{\sum_{\{i:a_i=1\}} a_i} = \frac{\sum_{i=1}^N a_i r_i}{\sum_{i=1}^N a_i}, \tag{4}$$

for which a robust version for distributed implementation of $\hat{x}^N_{Ideal}$ was studied recently in [5]. The above estimate is also unbiased, and $\hat{x}^N_{Ideal} \to x$ as $N \to \infty$, and its variance may be expressed as

$$var(\hat{x}^N_{Ideal}) = E(var(\hat{x}^N_{Ideal}|\mathbf{a})) + var(E(\hat{x}^N_{Ideal}|\mathbf{a})) \tag{5}$$

$$= \psi\sigma^2, \tag{6}$$

where $\psi = \sum_{k=0}^N \frac{1}{k}\binom{N}{k} p_1^k p_0^{N-k}$. When $p_1 = 0.5$, we have $\psi \approx \frac{2-2^{-N}}{N+1}$. A key difference from the naive estimate in (2) is that the variance of the ideal estimate stays constant over SNR, i.e., the estimation error does not scale up with the SNR.

From the MSE viewpoint, the ideal estimate is in fact optimal as long as the observation noise is Gaussian. However, such a scheme may not be implementable as it requires perfect knowledge of $\mathbf{a}$, which is unknown apriori. In Section 3, we will introduce a learning-based distributed estimation procedure, the MDE algorithm, based on the simultaneous iterative detection of $\mathbf{a}$ and the estimate refinement of x. Our results will indicate that not only $\mathbf{a}$ may be detected with high accuracy by the MDE algorithm, but also the estimation performance (in terms of MSE) approaches that of the ideal estimate $\hat{x}^N_{Ideal}$ in the high SNR regime (see Section 4).

3. DISTRIBUTED ALGORITHM

In this section, we present the MDE algorithm for the problem of interest. The first subsection introduces the MDE algorithm, and the second one analyzes its convergence.

3.1. MDE Algorithm

In each iteration of the MDE algorithm, each node first locally detects the value of a_i by using its current local estimate of x and its local observation. The new observation validity index is subsequently forwarded to the neighboring nodes, leading to an estimate refinement process. The algorithm is presented as follows.

Step 1. Initialization

$$\tilde{r}_i = \frac{r_i}{\bar{a}}, \quad \hat{x}_i(1) = \sum_{j=1}^N W_{ij}\tilde{r}_j. \tag{7}$$

Step 2. Detection of a_i

$$\hat{x}_i^2(t) - 2r_i\hat{x}_i(t) \underset{\hat{a}_i(t)=1}{\overset{\hat{a}_i(t)=0}{\gtrless}} 2\sigma^2 \ln\frac{p_1}{p_0}. \tag{8}$$

Step 3. Estimation of x

$$y_i(t) = (1-\beta(t))\sum_{j=1}^N W_{ij}y_j(t-1) + \alpha(t)\sum_{j=1}^N W_{ij}r_j\hat{a}_j(t), \tag{9}$$

$$v_i(t) = (1-\beta(t))\sum_{j=1}^N W_{ij}v_j(t-1) + \alpha(t)\sum_{j=1}^N W_{ij}\hat{a}_j(t), \tag{10}$$

$$\hat{x}_i(t) = \frac{y_i(t)}{v_i(t)}, \tag{11}$$

where $y_i(0) = r_i a_i(0)$, $v_i(0) = a_i(0)$, $\alpha(t) = 1/t$, and $\beta(t) = 1/t^{1-\varepsilon}$, with $\varepsilon \in (0,1)$. In the above, the W_{ij}'s denote the Metropolis weights [6] given by

$$W_{ij} = \begin{cases} 1/(1+max\{|\Omega_i|, |\Omega_j|\}) & j \in \Omega_i \\ 1 - \sum_{k\in\Omega_i} W_{ik} & i = j \\ 0 & \text{otherwise} \end{cases}, \tag{12}$$

where $|\cdot|$ denotes the cardinality of a set.

Step 4. Repeat steps 2 and 3 until $|\hat{x}_i(t) - \hat{x}_i(t-1)| < \epsilon$, where ϵ is a predefined small positive error tolerant parameter.

In practice, each node needs to know the degrees of their neighbors across the network topology in order to calculate the local weights W_{ij}'s. Intuitively, in step 2, each node locally detects (re-assesses) the value of a_i using the current local estimate of x and r_i. The validity indices, thus obtained, are forwarded to the neighboring nodes, leading to the state update in step 3, in which each node refines its local variables, i.e., $y_i(t)$ and $v_i(t)$ based on its neighbors' information variables. Finally, a new estimate is obtained from $y_i(t)$ and $v_i(t)$. The next subsection investigates the convergence of this iterative procedure.

3.2. Convergence of the MDE Algorithm

The main convergence result concerning the MDE algorithm is stated as follows.

Theorem 1 *Let the inter-sensor communication network be connected*[1]*, and assume that* $\alpha(t)$ *and* $\beta(t)$ *in* (9)-(10) *satisfy the following three conditions:*

- $\alpha(t) \to 0$, $\beta(t) \to 0$,
- $\beta(t)/\alpha(t) \to \infty$,
- $\sum_{t=0}^{\infty} \alpha(t) = \infty$, $\sum_{t=0}^{\infty} \beta(t) = \infty$.

Then, the estimate sequence $\{\hat{x}_i(t)\}$ *at each node* $\mathcal{N}_i$ *converges almost surely (a.s.), and the limiting value is given by*

$$\lim_{t\to\infty} \hat{x}_i(t) = \frac{\sum_{i=1}^{N} \hat{a}_i r_i}{\sum_{i=1}^{N} \hat{a}_i}, \qquad \forall i, \tag{13}$$

with $\hat{a}_i \in \{0,1\}^N$ *denoting the limiting value of the convergent sequence* $\{\hat{a}_i(t)\}$*. In particular, the local sensor estimates reach consensus. Note that* $\hat{a}_i$ *is, in general, random given the stochasticity of* a_i*.*

We emphasize that the conditions on $\alpha(t)$ and $\beta(t)$ listed above are not hard to satisfy. For example, $\alpha(t) = 1/t$ and $\beta(t) = 1/t^{1-\varepsilon}, \varepsilon \in (0,1)$ satisfy desired conditions.

The proof of Theorem 1 is omitted here due to space limitation and will appear in the journal version of this paper. In the following, we provide an intuitive sketch of the proof by stating the various intermediate propositions leading to Theorem 1.

The first result, Lemma 3.1, shows that as long as the sequences $\{\alpha(t)\}$ and $\{\beta(t)\}$ satisfy the conditions in Theorem 1, the processes $y_i(t)$ and $v_i(t)$ at each node $\mathcal{N}_i$ converge. More importantly, the corresponding limiting values are same for all the nodes, which leads to global consensus.

Lemma 3.1 *Let the conditions in Theorem 1 hold. Then, the sequences* $\{y_i(t)\}$ *and* $\{v_i(t)\}$ *reach the global consensus as* $t \to \infty$*, i.e., for each* i*,* $y_i(t)$ *and* $v_i(t)$ *converge a.s. to* y^* *and* v^**, the limiting values independent of* i*.*

[1]The network is said to be connected if there exists a path (possibly multi-hop) between any pair of nodes.

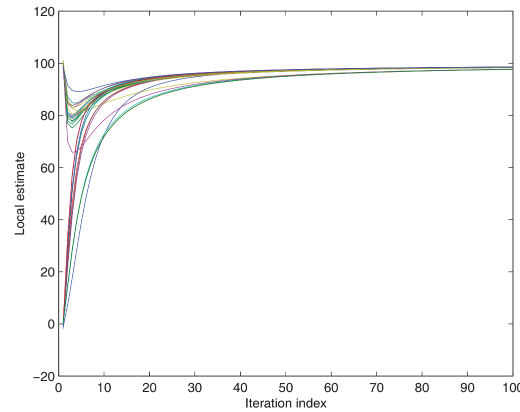

Fig. 1. The convergence of the MDE algorithm, $x = 100$.

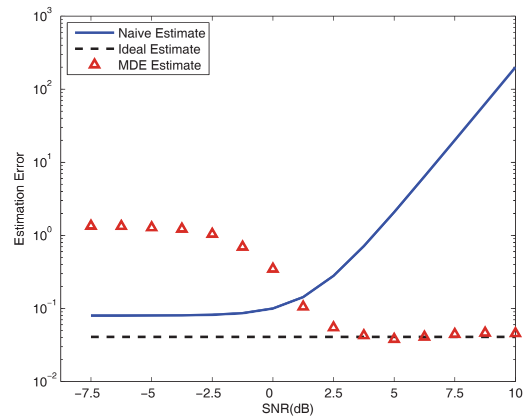

Fig. 2. The performance comparison of the MDE algorithm, the naive averaging algorithm and the ideal estimate.

The limiting consensus values y^* and v^* in Lemma 3.1 may be characterized as follows:

Lemma 3.2 *Under the conditions in Theorem 1, the limiting consensus values* y^* *and* v^* *are given by* $y^* = \frac{\sum_{i=1}^{N} \hat{a}_i r_i}{N}$ *and* $v^* = \frac{\sum_{i=1}^{N} \hat{a}_i}{N}$*, respectively, where* $\hat{a}_i$ *is the limiting value of the convergent sequence* $\{a_i(t)\}$*.*

Sketched proof of Theorem 1 By the convergence of $y_i(t)$ and $v_i(t)$ established in Lemma 3.1 and Lemma 3.2, the convergence of $\hat{x}_i(t)$ follows naturally, and the limiting value may be calculated as

$$\lim_{t\to\infty} \hat{x}_i(t) = \frac{\lim_{t\to\infty} y_i(t)}{\lim_{t\to\infty} v_i(t)} = \frac{y^*}{v^*} = \frac{\sum_{i=1}^{N} \hat{a}_i r_i}{\sum_{i=1}^{N} \hat{a}_i}. \tag{14}$$

4. SIMULATION RESULTS

In the last section, the a.s. convergence of the MDE algorithm has been established analytically, where the limiting value is shown to be a function of $\{\hat{a}_i\}$. Given the stochasticity of a_i, the analytical quantification of the MDE performance is not trivial, which is left for the journal version of this paper. In order to show the characterization of the limiting value, we here present simulation studies implemented in Matlab that demonstrate the estimation performance of the MDE algorithm. In our network setting, 50 nodes are uniformly distributed over a unit square where two nodes are connected by an edge if their distance is less than 0.3, a predefined transmission range. In this section, the a_i's are independently generated with $p_1 = 0.5$, w_i's are independent white Gaussian noises with zero mean and unit variance, and the other parameter values are specified in the description of each figure.

In Fig. 1, we demonstrate the convergence (see Theorem 1) of the proposed algorithm. A realization of the local estimates at the 50 nodes for 100 rounds of iterations, i.e., $\hat{x}_i(t), i \in [1..50], t \in [1..100]$ is plotted. The signal value x is 100 in this simulation, i.e., $SNR = 20$. In the figure, about half of the nodes start around the initial value 100 and the rest start around 0 indicating that the former ones correspond to valid observations and the latter ones are the nodes with invalid observations. We observe that the local estimates of both types of nodes converge to the same limiting value after about 50 rounds of iteration and this consensus is close to the real target value 100 as desired.

In Fig. 2, we compare the performance of the proposed algorithm with the naive averaging algorithm and the ideal algorithm (see Section 2). In the figure, the estimation error of these three estimates are plotted (both numerically and analytically) with SNR ranging from -15 dB to 20 dB, i.e., x from 0.03 to 100. For each SNR, we generated 500 realizations of the MDE algorithm, the limiting consensus value of the local estimates for each realization being taken to be the estimate in the first node at the end of the 3000-th iteration round. The estimation error plotted in the figure is the average (numerical) squared deviation of the limiting consensus value from the true value x over these 500 realizations, i.e., $(\sum(\hat{x}_1(3000) - x)^2)/500$. The topology of the communication graph (given by the random node placement) and the observation values at each nodes are independently generated for each realization. We make several observations from this figure. First, the numerical result of the naive averaging algorithm matches the theoretical results, and the estimation error grows exponentially with SNR (now in dB) as derived in (3); second, the numerical result of the ideal algorithm matches the theoretical results and the estimation error of the ideal algorithm is the lowest among the three algorithms; and third, although the estimation error of MDE is higher than that of the naive averaging in the lower SNR regime (SNR<4), it performs much better in the mid and high SNR regimes (SNR>10), where it approaches the performance of the optimal estimator.

In the following we provide some intuitive explanation of the observed simulation behavior: 1). In the low SNR regime, the target value is relatively small as compared with the Gaussian noise, which leads to a high detection error in (8) of Subsection 3.1. Some invalid observations are wrongly detected as valid ones and wrongly incorporated into the estimate update process, whereas, some valid observations are discarded as invalid ones. Thus, the estimate is largely distorted from the ideal estimate, which leads to the poor estimation performance; 2). In the high SNR regime, the detection error in (8) is very small and almost every observation is correctly detected as valid or invalid. Therefore, the MDE estimate is quite close to the ideal estimate and the MSE of the MDE algorithm approaches this (ideal algorithm) lower bound;

5. CONCLUSIONS AND FUTURE WORK

We studied algorithms for distributed estimation of a scalar target signal with observations of imperfect mode information in sensor networks. The MDE algorithm was presented, in which an online learning step re-assesses the validity of the local observations at each iteration, and thus refining the ongoing estimation update process. We analytically established the convergence of the MDE algorithm. From the results of the simulation, we have shown that in the high SNR regime, the MDE estimation error converges to that of an ideal estimator with perfect information about the node sensing modes. Ongoing research concerns an analytical characterization of the MDE performance. Observing that the estimation error of MDE is higher than that of the naive average consensus in the low SNR regime, an auto-switching mechanism between the naive estimator and the MDE estimator will be developed in the future work.

6. REFERENCES

[1] J. N. Tsitsiklis, "Problems in Decentralized Decision Making and Computation," *Ph. D. Disseration*, Deptartment of Electrical Engineering and Computer Science, Massachusetts Institute Technology, Cambridge, MA, Nov. 1984.

[2] Reza Olfati-Saber and J. Alex Fax, and Richard M. Murray, "Consensus and Cooperation in Networked Multi-Agent Systems," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 215 - 233, Jan. 2007.

[3] Alexandros G. Dimakis, Soummya Kar, José M. F. Moura, Michael G. Rabbat, and Anna Scaglione, "Gossip Algorithms for Distributed Signal Processing," *Proceedings of the IEEE*, vol. 98, no. 11, pp. 1847 - 1864, Nov. 2010.

[4] L. Xiao, and S. Boyd, "Fast Linear Iterations for Distributed Averaging," *IEEE Conference on Decision and Control*, vol. 5, pp. 4997 - 5002, March 2004.

[5] Q. Zhou, S. Kar, L. Huie, H. V. Poor, and S. Cui, "Robust Distributed Least-Squares Estimation in Sensor Networks with Node Failures," *to appear in Proceeding of the Globecom*, Houston, TX, Dec. 2011.

[6] L. Xiao, and S. Boyd, "Distributed Average Consensus with Time-Varying Metropolis Weights," *Unpublished Manuscript, http://www.stanford.edu/ boyd/papers/avg_metropolis.html*, June 2006.