On the effect of low-quality node
Observation on learning over
Incremental adaptive networks

M. Farhid, M. H. Sedaaghi and M. Shamsi
Faculty of Electrical Engineering, Sahand University of Technology, Tabriz, Iran.
{m_farhid, sedaaghi, shamsi}@sut.ac.ir
Corresponding author: M.Farhid

Abstract— In this paper, we study the impact of low-quality node on the performance of incremental least mean square (ILMS) adaptive networks. Adaptive networks involve many nodes with adaptation and learning capabilities. Low-quality mode in the performance of a node in a practical sensor network is modeled by the observation of pure noise (its observation noise) that leads to an unreliable measurement. Specifically, we consider ILMS networks with different number of low-quality nodes and compare their performance in two different cases including (i) ideal and (ii) noisy links in homogeneous and inhomogeneous environments. We show that in the case of ideal links among nodes, one node with low-quality mode degrades the estimation performance significantly and increasing the variance of observation noise does not degrade the performance anymore. Even with increasing node numbers with low-quality mode in network, estimation performance does not divergence. On the other hand, in the presence of noisy links, different behavior is observed and degradation is dependent on variance of noisy links and it may go unstable. Simulation results are provided to illustrate the discussions.

Index Terms— Adaptive network, incremental cooperation, LMS algorithm, mean square deviation (MSD).

I. INTRODUCTION

A wireless sensor network consists of a definite number of sensor nodes distributed over a geographical area. These sensors (nodes) are self-powered and in some cases comprise a local computing mechanism. Distributed estimation problem arises in many applications where a set of nodes are used to estimate a parameter of interest by the data collected at nodes. This problem has first been studied in the context of distributed control, tracking, data fusion [1]-[3], and recently in wireless sensor networks [4]-[5]. A sensor node is a tiny device that includes three essential components: a sensing subsystem for data acquisition from the physical surrounding environment, a processing subsystem for local data processing and storage, and a wireless communication subsystem for data transmission [6]. The estimation problem can be solved by either a centralized or a decentralized approach [7]. In a centralized approach, measurements from all nodes are collected and
On the effect of low-quality node observation on …

processed by a fusion center. This scheme requires extensive amounts of communication between the nodes and the processor. These issues, along with geographical difficulties limit the use of fusion-based solutions. An alternative way is the decentralized solution where the nodes rely solely on their local data and on interactions with their immediate neighbors. The amount of processing and communications is significantly reduced in this scheme [8]. In most of these applications, the statistical information about the process of interest is not available or it varies over the time. The adaptive networks have been introduced in the literature to solve the distributed estimation problem in such cases along with response capability to excitations and tracking [9]-[15]. Having real time response also requires that the nodes must be working continuously whether there is an observation or not. The existing distributed adaptive networks can be roughly classified, based on the cooperation mode between the nodes, into incremental [9]-[12], diffusion [13]-[14] and consensus [15]-[16] algorithms. Our focus in this work is on incremental LMS adaptive networks where at each iteration, each node receives the prior node’s local estimate, updates its local data using LMS algorithm, and finally sends it to the next node with assuming ideal links between nodes [10, 17]. However, in [18], the performance of incremental adaptive network differs considerably in the presence of noisy links. Beside the incremental LMS and RLS, incremental techniques based on the affine projection algorithm, parallel projections, and randomized incremental protocols are the other examples of incremental adaptive networks [11, 19, 20]. The incremental solution suffers from a number of limitations for applications involving adaptation and learning from streaming data [20]. First, the incremental strategy is sensitive to agent or link failures. If an agent or link over the cyclic path fails, then the information flow over the network is interrupted. Second, starting from an arbitrary topology, determining a cyclic path that visits all agents is generally an NP-hard problem. Third, cooperation between agents is limited to an agent which it is allowed to receive data from one preceding agent and to share data with the proceeding one. In addition, processing at the agents needs to be fast enough so that the $N$ update steps can be completed before the next cycle begins. However, this mode of cooperation requires the least amount of communication and power and is also suitable for small-size networks besides offering excellent estimation performance [10]. Another aspect that deserves attention in the distributed algorithms analysis is the case where the observation in each node in the network does not provide the same reliability regarding the estimation of the unknown parameter [27]. In this situation, the estimate provided by the noisy nodes, i.e., the nodes that provide the least reliable parameter estimation, tends to deteriorate the global estimate in the network. This heterogeneous scenario in a WSN can arrive in many situations. For instance, the spectrum sensing in a cognitive radio network in which the channel between each radio and the common primary user presents a different signal-to-noise ratio is an example of such situation [24]. In this case, the noisy cognitive radios would deteriorate the network global estimate. A possible solution to overcome this problem is to provide an adaptive cooperative scheme as in [25] or to use a variable step size LMS as proposed in
As mentioned before, in a realistic sensor network, a sensor may be damaged or attacked, making its measurement unreliable. If this happens, the sensor will only observe the pure noise (observation noise) and may degrade the estimation performance. We refer to this event as low-quality mode of node and investigate the performance of MSD for this issue.

This article is organized as follows: the problem of distributed parameter estimation and incremental solution are explained in Section 2. In Section 3, the effect of low-quality mode is modeled and discussed. Subsequently, Section 4 shows simulation results and the effect of number of nodes with low-quality mode on the accuracy of parameter estimation over network. Finally, conclusions are provided in Section 5.

**Notation:** Throughout the paper, we use boldface letters for random quantities. The * symbol is used for both complex conjugation for scalars and Hermitian transpose for matrices.

II. DISTRIBUTED ESTIMATION AND INCREMENTAL SOLUTION

Consider a network with \( N \) nodes. At time \( i > 0 \), node \( k \) obtains scalar measurement \( d_k (i) \) and regression vector \( u_{k,i} (1 \times M) \) which are the time-realizations of zero mean jointly wide-sense stationary spatial data \( \{d_k, u_k\} \). These quantities are related via:

\[
d_k (i) = u_{k,i} w^0 + v_k (i), \quad k = 1, 2, \ldots, N
\]

where \( M \times 1 \) vector \( w^0 \) is an unknown parameter and \( v_k (i) \) is the observation noise term with variance \( \sigma_{v,k}^2 \). The purpose of the network is to estimate \( w^0 \) from measurements collected from \( N \) nodes. Note that \( w^0 \) is the solution of the following optimization problem:

\[
\min \sum_{k=1}^{N} J_k (w) = \sum_{k=1}^{N} E|d_k (i) - u_{k,i} w|^2 \quad \text{(network objective)}
\]

where \( E \) denotes the statistical expectation. The optimal solution of (1) satisfies the normal equations \([9,10]\):

\[
w^0 = R_u^{-1} R_{du}
\]

Where

\[
R_{du,k} = \sum_{k=1}^{N} E\{u_i d_k\}, \quad R_{u,u,k} = \sum_{k=1}^{N} E\{u_i^* u_k\}
\]

In order to use (4) each node must have access to the global statistical information, which in many applications are not available or change in time and since the optimization problem involves decoupled cost functions as \( J_k (w) = E|d_k - u_k w|^2 \), the incremental methods can be used to seek the solution in a distributed approach. In \([9,10]\) the DILMS is proposed to address the mentioned problems:
On the effect of low-quality node observation on …

Fig. 1. A schematic of DILMS algorithm

\[
\begin{align*}
\psi_i^{(i)} &\leftarrow w_{i-1} \\
\psi_k^{(i)} &\leftarrow \psi_{k-1}^{(i)} + \mu u_{k,i}^* (d_k(i) - u_{k,i} \psi_{k-1}^{(i)}) \\
w_i &\leftarrow \psi_N^{i}
\end{align*}
\]

Where \( \psi_k^{i} \) denotes the local estimate of \( w^o \) at node \( k \) at time \( i \) and \( \mu_k \) is the step size. In the incremental LMS algorithm, the estimated values (i.e., \( \psi_k^{i} \)) are sequentially circulated from node to node as shown in Fig. 1. Assumptions on the statistical properties of the data are as follows: the regression data \( u_{k,i} \) are temporally and spatially independent and identically distributed (i.i.d.) circular white Gaussian random variables with zero mean and diagonal covariance matrix \( \lambda I_M \). The noise signals \( v_k(i) \) are independent of \( d_L(i), u_{L,j} \) for all \( L,j \).

The given algorithm in (5), utilizes both spatial and temporal dimensions of the data. In [10], the mean-square performance of DILMS is studied using energy conservation arguments. In mean-square analysis, we are interested in evaluating the steady-state values of mean-square deviation (MSD), for every node \( k \). This quantity is defined as follows:

\[
\eta_k = E\left(\|\tilde{\psi}_{k-1}^{i}\|^2\right)
\]

where

\[
\tilde{\psi}_{k-1}^{i} = w^o - \psi_{k-1}^{i}
\]

The analysis relies on the linear model (1) and the following assumptions:

(i) \( \{u_{k,i}\} \) are spatially and temporally independent.
(ii) The regressors \{u_{k,i}\} arise from a circular Gaussian distribution with covariance matrix \(R_{u,k}\).

In [9], a complex closed-form expression for MSD has been derived. However, in the case of small step sizes, simplified expressions for the MSD can be described as follows: for each node \(k\), an eigen decomposition is introduced as \(R_{u,k} = U_k \Lambda_k U_k^*\), where \(U_k\) and \(\Lambda_k\) are unitary and diagonal matrices of the eigenvalues of \(R_{u,k}\).

\[ \Lambda_k = \text{diag}\{\lambda_{k,1}, \lambda_{k,2}, \ldots, \lambda_{k,M}\}, \text{(node } k\) \] (8)

Then, according to the results from [10] we have:

\[ \eta_k \approx \frac{1}{2} \sum_{i=1}^{M} \left( \frac{\mu_i^2 \sigma_i^2}{\sum_{j=1}^{N} \mu_j \lambda_{i,j}} \right) \] (9)

Also, in [18] the performance of incremental least mean square adaptive networks was discussed in a more realistic case in which communication links between nodes are considered noisy. In the presence of noisy links the update equation for DILMS changes to:

\[
\left\{ \begin{array}{l}
\psi_i^k = \psi_{k-1}^i + q_{k,i} + \mu_i u_{k,i}^* \left[ d_k(i) - u_{k,i} (\psi_{k-1}^i + q_{k,i}) \right] \\
w_i = \psi_N^i
\end{array} \right. \text{ with } k = 1, \ldots, N \] (10)

Where the \(M \times 1\) vector \(q_{k,i}\) is the channel noise term between nodes \(k-1\) and \(k\). Also, \(q_{k,i}\) is a time-realization of wide-sense stationary random process \(q_{k,i}\) which is assumed to be zero mean with covariance matrix \(Q_k = E[q_k q_k^*]\). Again, by the assumption of constant and small step size and, \(Q_k = \sigma_{v,k}^2 I_M\), \(R_{u,k} = \lambda I_M\) the approximation of MSD can be derived as follows:

\[ \eta_k = \frac{M}{2\mu\lambda N} \sum_{i=1}^{N} \left( \mu_i^2 \sigma_{v,k}^2 \lambda + \sigma_{c,k}^2 (1 - 2\mu\lambda) \right) \] (11)

III. LOW-QUALITY MODEL

In a practical sensor network, a sensor may be damaged or attacked, making its measurement unreliable and does not have any information about \(w^0\). If this happens, the sensor will only observe the pure noise and certainly degrade the estimation performance [22]. We refer to this phenomenon as low-quality mode of node. This is simulated by data model (12). Two ways of low-quality node is studied, namely random and intentional removal (according to SNR).
On the effect of low-quality node observation on …

Fig. 2. A schematic of DILMS algorithm with 2 nodes

\[ d_t(i) = \begin{cases} u_{k,i}w^0 + v_k(i) & \text{usual nodes} \\ v_k(i) & \text{low-quality nodes} \end{cases} \]

\[ d_1(i) = u_{1,i}w^0 + v_1(i) \quad d_2(i) = v_2(i) \]

The mean stability analysis aims to find out the sufficient conditions such that the local estimation at each node converges in the mean to the unknown parameter \( w^0 \).

So we refer to the two nodes as nodes 1 and 2. Node 1 is assumed to measure data that satisfy a linear regression model of the form and node 2 only observe the pure noise (its observation noise):

For node 1, we have:

\[
\begin{align*}
\psi_1' &= \psi_0' + \mu u_{1,i}^* [d_1(i) - u_{1,i} \psi_0'] \\
\psi_2' &= \psi_0' + \mu u_{1,i}^* [u_{1,i}w^0 + v_1(i) - u_{1,i} \psi_0'] \\
\psi_3' &= \mu u_{1,i}^* u_{1,i} w^0 + \mu u_{1,i}^* v_1(i)
\end{align*}
\]

(13)

For node 2, and with assumption small enough step size, we have:

\[
\begin{align*}
\psi_1'' &= \psi_1' + \mu_2 u_{2,i}^* [d_2(i) - u_{2,i} \psi_1'] \\
\psi_2'' &= \psi_1' + \mu_2 u_{2,i}^* [v_2(i) - u_{2,i} \psi_1'] \\
\psi_3'' &= \mu_2 u_{2,i}^* u_{2,i} w^0 + \mu_2 u_{2,i}^* v_1(i) + \mu_2 u_{2,i}^* u_{2,i} w^0 + \mu_2 u_{2,i}^* v_1(i)
\end{align*}
\]

(14)

Now, let’s consider global error vector as:

\[ \tilde{\psi}_k^i = w^0 - \psi_k^i \]

(15)

Using model (12) and Supposing that the regressors \( \{u_{k,i}\} \) are spatially and temporally independent, along with taking expectation value of both sides of (13), we find that the mean relation of \( \tilde{\psi}_k^i \) evolves in time according to the recursion as follows:

\[ E\tilde{\psi}_k^i = w^0 [I - \mu R] \]

(16)

This condition ensures mean stability, as \( i \to \infty \), and again with assumption of enough step size even in the presence nodes with faults is guaranteed. This condition is also right for noisy links with assuming mean of noisy links are zero. The value of this degradation can be achieved with energy conservation method [10].
IV. SIMULATIONS RESULTS

We present a simulation example for the distributed incremental LMS algorithms with ideal links in Figs. 3–6, for a network with 10 nodes. The regressors are zero-mean complex Gaussian, independent of time and space, of size $M=5$ and with covariance matrices $R_{u,k}$. The background noise power is denoted by $\sigma_{v,k}^2$. Fig. 3 shows $\sigma_{v,k}^2$ and $\text{Tr}(R_{u,k})$ for each node. The results are averaged over 100 independent experiments using a step-size $\mu=.01$. Figs. 4-5 show the MSD transient network when two condition (i) homogeneous environment that all of the nodes have same observation noise that are equal .02 and (ii) inhomogeneous environment that observation noise is as shown in Fig.3(bottom left). As it shows, with one low-quality node the performance of the network decreases significantly and this value is more in inhomogeneous network. Note that this low-quality node is chosen randomly. Also, it is resulted for other choices of low-quality node according to the defined SNR values in [23] and observed that MSD value doesn’t change and is shown in Fig.6. For investigating how the value of power noise affects degradation, we simulate default network with different values for homogeneous and inhomogeneous condition. As Figs. 7-8 show, the values of MSD in different values of noise variance. is approximately constant in both homogeneous and inhomogeneous.
environments and also does not divergence. Also we simulate the network with different nodes low-quality mode and as is shown in Fig. 9, with increasing number of low-quality nodes, the degradation is not significant and also until one node in the network works normal, the network does not divergence. In the case of networks with noisy links, all simulation results are averaged over 200 independent Monte Carlo runs. Also we simulate this network with different step sizes as is shown in Fig. 10 for showing that, this behavior is true for all allowable duration of step sizes. We consider the previous example of distributed network again with 10 nodes seeking an unknown filter with $M=5$ taps and with assuming covariance matrix, $Q=10^{-5}I$, the transient MSD of network is shown in Fig. 11. Also as shown in Fig. 11, we have again degradation in this condition but compare to ideal links
this degradation is small and as shown in Fig.12 this degradation is dependent to variance of noisy links and with increasing this value, it may be becomes unstable. In fact this concept comes from simulation in Fig.13 that for different number of node with low-quality mode, with increasing number of low-quality nodes, the degradation is approximately equal and with increasing noisy links variance the network becomes divergence(unlike ideal links).
Fig. 8. Steady state of MSD for usual and one node with low-quality incremental LMS in different inhomogeneous environment.

Fig. 9. Steady state of MSD for usual and different low-quality nodes incremental LMS in inhomogeneous environment.

Fig. 10. Steady state of MSD for usual and different low-quality nodes incremental LMS in inhomogeneous environment for different step sizes.
Fig. 11. Transient network MSD for usual (blue) and one node is low-quality mode (red) incremental LMS with noisy links $Q=10^{-5}$ in inhomogeneous environment.

Fig. 12. MSD for usual and one low-quality node for incremental LMS with different noisy links in inhomogeneous environment.

Fig. 13. Steady state of MSD for usual and different low-quality nodes incremental LMS in inhomogeneous environment.
V. CONCLUSION

In a realistic sensor network, a sensor may be damaged or attacked, making its measurement unreliable. If this happens, the sensor will only observe the pure noise (modeled by its observation noise) and degrade the estimation performance. We refer to this event as low-quality mode of node. In applications where node deployment is controlled, incremental strategies are applicable since they can achieve better performance than diffusion. We considered two different cases, including the ideal and noisy links in homogeneous and inhomogeneous environments. Our results exposed that when the links are ideal, one node failure in the network deteriorates the network’s learning and estimation performance and this degradation does not depend on the value of observation nodes. As to noisy links condition, the performance estimation depends on variance of channel noise and with increasing this value performance of MSD may become worse and divergent. Also, it was revealed that by increasing the number of low-quality nodes, the variation in the percentage of degradation in MSD for each additional node in ideal links is approximately steady and as long as there is at least one properly working node the system will function reliably. Also step size choices are studied and it is concluded that in small step sizes the system is less open to performance degradation caused by increased number of low-quality nodes. On the other hand, for the case of noisy links with a noise level of below some threshold we show that the trend of performance degradation is rapidly increasing one with the number of failed nodes. It should also be noted that link noise of enough intensity can drive the system unstable even in low numbers of low-quality nodes.

REFERENCES


