Int. J. Electron. Commun. (AEÜ) 67 (2013) 263–268

Contents lists available at SciVerse ScienceDirect

International Journal of Electronics and Communications (AEÜ)

journal homepage: www.elsevier.com/locate/aeue

An optimum step-size assignment for incremental LMS adaptive networks based on average convergence rate constraint

Azam Khalili [a], Amir Rastegarnia [a,*], Jonathon A. Chambers [b], Wael M. Bazzi [c]

[a] Department of Electrical Engineering, Malayer University, Malayer 95863-65719, Iran
[b] Department of Electronic and Electrical Engineering, Loughborough University, Loughborough, Leics LE11 3TU, UK
[c] Electrical and Computer Engineering Department, American University in Dubai, P.O. Box 28282, Dubai, United Arab Emirates

ARTICLE INFO

Article history:
Received 2 June 2011
Accepted 20 August 2012

Keywords:
Adaptive networks
Distributed estimation
Least-mean square (LMS)
Step-size assignment

ABSTRACT

This paper presents an optimum step-size assignment for incremental least-mean square adaptive networks in order to improve its robustness against the spatial variation of observation noise statistics over the network. We show that when the quality of measurement information (in terms of observation noise variances) is available, we can exploit it to adjust the step-size parameter in an adaptive network to obtain better performance. We formulate the optimum step-size assignment as a constrained optimization problem and then solve it via the Lagrange multipliers approach. The derived optimum step-size for each node requires information from other nodes, thus with some justifiable assumptions, near-optimum solutions are derived that depend only on local information. We show that the incremental LMS adaptive network with near-optimal step sizes has improved robustness and steady-state performance. Simulation results are also presented to support the theoretical results.

1. Introduction

In many WSN applications, the ultimate goal is to obtain an accurate estimate of an unknown parameter, based on the temporal data acquired by spatially distributed sensors [1,2]. This estimation problem can be solved by either a centralized approach (with fusion center) or a decentralized approach (see e.g. [3] and references therein for a brief review of distributed estimation algorithms). In many applications, however, sensors need to perform estimation in a constantly changing environment without having available a (statistical) model for the underlying processes of interest [4]. This issue motivated the development of distributed adaptive estimation algorithms (or adaptive networks). An adaptive network is a collection of adaptive nodes that observe space–time data and collaborate, according to some cooperation protocol, in order to estimate a parameter [5–14]. Using cooperative processing in conjunction with adaptive filtering per node enables the entire network (and also each individual node) to track not only the variations of the environment but potentially also the topology of the network.

Depending on the manner by which the nodes communicate with each other, adaptive networks may be referred to as incremental networks (algorithms) or diffusion networks (algorithms). In the incremental mode, a cyclic (Hamilton) path through the network is required, and nodes only communicate with neighbors within this path [5–10]. In diffusion based adaptive networks, each node combines estimates form its neighbors using some combiner methodology and then performs adaptation on this combined estimate. Finally, the new (updated) estimate is then diffused into the network [11–14].

The performance of existing incremental adaptive networks [5–10] deteriorates when the measurement quality at some nodes are lower than others; because the poor estimates of such nodes pervades into the entire network due to incremental cooperation [15]. In this paper we consider the observation quality information to design an incremental LMS adaptive network with improved robustness against the spatial variation of observation noise statistics over the network. The step-sizes are allotted so that nodes presenting poor performance are assigned with small step-sizes and vice versa. We formulate the step-size assignment as a constrained optimization problem and then solve it via the Lagrange multipliers approach. Since the optimum step size for each node requires information from other nodes, with some justifiable assumptions, near-optimal solutions are derived that depend only on local information. Simulation results show that the incremental LMS adaptive network with the proposed step-sizes outperforms those with existing static step-sizes.

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. $\|x\|_{\Sigma}^2$ denotes weighted norm for a column vector x, which is given by $x^*\Sigma x$; $\mathbb{1}_N$

* Corresponding author.
E-mail addresses: a.khalili@ieee.org (A. Khalili), amir.rastegarnia@gmail.com (A. Rastegarnia), wbazzi@aud.edu (W.M. Bazzi).

1434-8411/$ – see front matter © 2012 Elsevier GmbH. All rights reserved.
http://dx.doi.org/10.1016/j.aeue.2012.08.010

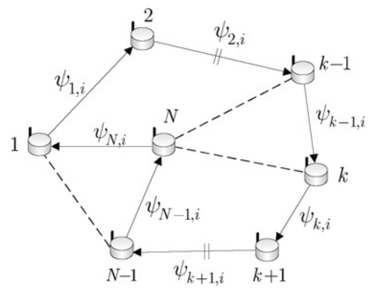

Fig. 1. The structure of incremental LMS adaptive network.

denotes an $N \times 1$ vector with unit entries; while E represents the statistical expectation operator.

2. The incremental LMS adaptive network

We denote by set $\mathcal{N} = \{1, \cdots, N\}$, a distributed network (e.g. a WSN) with N nodes which communicate according to the incremental protocol. At time $i > 0$, each node $k \in \mathcal{N}$ has access to scalar measurement $d_k(i)$ and $1 \times M$ regression vector $u_{k,i}$ that are related via

$$d_k(i) = u_{k,i} w^o + v_k(i) \tag{1}$$

where the $M \times 1$ vector $w^o \in \mathbb{R}^M$ is an unknown parameter and $v_k(i)$ is the observation noise term with variance $\sigma^2_{v,k}$. Note that $\{d_k(i), u_{k,i}\}$ are time-realizations of zero-mean spatial data $\{\boldsymbol{d}_k, \boldsymbol{u}_k\}$. The objective of the network is to estimate w^o from measurements collected at N nodes. Note that w^o is the solution of the following optimization problem

$$\arg\min_w J(w) \quad \text{where } J(w) = E\{\|\boldsymbol{d} - \boldsymbol{U}w\|^2\} \tag{2}$$

where

$$\boldsymbol{U} \triangleq \begin{bmatrix} \boldsymbol{u}_1 \\ \boldsymbol{u}_2 \\ \vdots \\ \boldsymbol{u}_N \end{bmatrix}_{N\times M}, \quad \boldsymbol{d} \triangleq \begin{bmatrix} \boldsymbol{d}_1 \\ \boldsymbol{d}_2 \\ \vdots \\ \boldsymbol{d}_N \end{bmatrix}_{N\times 1} \tag{3}$$

The solution of (2) (i.e. w^o) is given by [5,6]

$$w^o = R_u^{-1} R_{du} \tag{4}$$

where

$$R_{du} = E\{\boldsymbol{U}^* \boldsymbol{d}\}, \text{ and } R_u = E\left\{\boldsymbol{U}^* \boldsymbol{U}\right\} \tag{5}$$

In order to use (4) each node must have access to the global statistical information $\{R_u, R_{du}\}$ which in many applications are not available or change in time. To address this issue and moreover, to enable the network to respond to changes in statistical properties of data in real time, the incremental LMS adaptive network is proposed in [5]. The update equation for incremental LMS is given by

$$\psi_{k,i} = \psi_{k-1,i} + \mu_k u^*_{k,i}(d_k(i) - u_{k,i}\psi_{k-1,i}) \tag{6}$$

where $\psi_{k,i}$ denotes the local estimate of w^o at node k at time i and μ_k is the step size. In the incremental LMS algorithm, the calculated estimates (i.e. $\psi_{k,i}$) are sequentially circulated from node to node as shown in Fig. 1.

A good measure of the adaptive network performance is the steady-state mean-square deviation (MSD) which for each node k is defined as follows

$$\eta_k \triangleq E(\|\widetilde{\boldsymbol{\psi}}_{k-1,\infty}\|^2) = E(\|\widetilde{\boldsymbol{\psi}}_{k-1,\infty}\|^2_I) \tag{7}$$

where

$$\widetilde{\boldsymbol{\psi}}_{k-1,i} \triangleq w^o - \boldsymbol{\psi}_{k-1,i} \tag{8}$$

The mean-square performance of incremental LMS algorithm is studied in [6] using energy conservation arguments. The analysis relies on the linear model (1) and the following assumptions

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

In [6], 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, introduce the eigen decomposition $R_{u,k} = U_k \Lambda_k U_k^*$ where U_k is unitary and Λ_k is a diagonal matrix of the eigenvalues of $R_{u,k}$

$$\Lambda_k = \text{diag}\{\lambda_{k,1}, \lambda_{k,2}, \cdots, \lambda_{k,M}\}, \quad (\text{node } k) \tag{9}$$

Then, according to the results from [5,6]:

$$\eta_k \approx \frac{1}{2}\sum_{j=1}^{M}\left(\frac{\sum_{\ell=1}^{N}\mu_l^2\sigma^2_{v,l}\lambda_{l,j}}{\sum_{\ell=1}^{N}\mu_l\lambda_{l,j}}\right) \tag{10}$$

In the next section we use (10) to derive optimum step-sizes for the incremental LMS algorithm.

2.1. Assumption

To develop our proposed scheme we consider the following assumptions

(**A.1**) For $k = 1, 2, \cdots, N$ we have $R_{u,k} = \gamma_u I_M$ where γ_u is a real positive constant.
(**A.2**) The observation noise variances $\{\sigma^2_{v,k}\}$ are uniformly distributed over $\sigma^2_{v,k} \in [a, b]$.
(**A.3**) The observation noise variance $\sigma^2_{v,k}$ is available to node k.

The assumption (A.1) is valid when regressors $\{\boldsymbol{u}_{k,i}\}$ arise from an independent Gaussian distribution and is a simplifying assumption that makes the expression (10) more mathematically tractable. Also, in (A.2), the lower bound is determined by the sensor measurement accuracy and the upper bound can be chosen to provide robustness in learning. Finally, the assumption (A.3) can be achieved by some training data [16].

2.2. Motivation

Let us assume that there are some nodes with low quality of measurement in the network. We denote these nodes by $\mathcal{N}_s$ (note that $\mathcal{N}_s \subset \mathcal{N}$). We have shown in [15] that this set of nodes can drastically decrease the steady-state performance of an incremental LMS adaptive network. Thus we need to reduce the effects of theses nodes. On the other hand, as it is shown in [6], η_k is a monotonically increasing function of $\{\mu_k\}$. Thus we cannot find any set $\{\mu_1, \mu_2, \cdots, \mu_N\} \in \mathbb{R}^N$, in terms of λ_k and $\sigma^2_{v,k}$ that minimize the MSD. This means that, if the step sizes $\{\mu_k\}$ are large, the convergence rate of the incremental LMS algorithm will be rapid, but the steady-state MSD will increase and vice versa.[1] Therefore, we need to minimize the η_k in terms of $\{\mu_1, \mu_2, \cdots, \mu_N\}$, subject to a suitable

[1] It also must be noted that as we have shown in [17,18] this is not the case in adaptive networks with noisy links, i.e. reducing the adaptation step-size may actually increase the MSD.