Automatica 80 (2017) 95–101

Contents lists available at ScienceDirect

Automatica

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

Brief paper

Performance analysis of averaging based distributed estimation algorithm with additive quantization model☆

Shanying Zhu [a,1], Shuai Liu [b,c], Yeng Chai Soh [c], Lihua Xie [c]

[a] Department of Automation, Shanghai Jiao Tong University, and Key Laboratory of System Control and Information Processing, Ministry of Education of China, Shanghai 200240, China

[b] School of Control Science and Engineering, Shandong University, Jinan 250061, China

[c] School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 639798, Singapore

ARTICLE INFO

Article history:
Received 23 December 2014
Received in revised form
20 July 2016
Accepted 26 January 2017

Keywords:
Distributed sensor fusion
Additive quantization model
Directed topology
Log log law

ABSTRACT

In this paper, we consider the distributed sensor fusion problem over sensor networks under directed communication links and bandwidth constraint. We investigate the impact of the additive quantization model on the proposed two-stage averaging based algorithm. Existing works on the effect of the additive model show that convergence can be guaranteed only if the quantization error variances form a convergent series. We show that the proposed algorithm achieves the performance of the optimal centralized estimate even if the quantization error variances are not vanishing. This is guaranteed by establishing a law of the iterated logarithm for weighted sums of independent random vectors. Moreover, an explicit bound of the convergence rate of the proposed algorithm is given to quantify its almost sure performance.

1. Introduction

In a typical sensor fusion problem in sensor networks, nodes make noisy measurements of variables of interest. In this paper, we focus on a specific and simple model of sensor fusion problem, e.g., Ribeiro and Giannakis (2006), Xiao, Boyd, and Lall (2005) and Xiao and Luo (2005), where the common goal is to estimate the scalar[2] using observations from N nodes: $y_i = \theta + n_i,\ i = 1, 2, \ldots, N$, where n_i are zero mean, i.i.d. Gaussian noises. The main concern is how to utilize the samples collected from the nodes to produce a desirable estimate of θ. It is well known that if all $\{y_i\}_{i=1}^N$ are available at a fusion center perfectly, then the best way is to take the average and produce the sample mean estimate $\hat{\theta} \triangleq (1/N)\sum_{i=1}^N y_i$.

In this paper, we focus on distributed solutions of the above sensor fusion problem in ad hoc networks. The procedures of distributed algorithms are generally based on successive refinements of local estimates at nodes, e.g., consensus-based algorithms (Aysal & Barner, 2010; Kar, Moura, & Ramanan, 2012; Xiao et al., 2005; Xie et al., 2012; Zhu, Chen, Ma, Yang, & Guan, 2015; Zhu, Soh, & Xie, 2015), diffusion algorithm (Cattivelli & Sayed, 2010), and learning-based algorithm (Rad & Tahbaz-Salehi, 2010). One major challenge of ad hoc networks is that limitations in bandwidth place tight constraints on the rate of information exchanged between nodes, i.e., data at each node needs to be quantized prior to its transmission to the neighboring nodes.

Recently, much work has been done to examine the effects of several quantizers on distributed algorithms, e.g., probabilistic quantizers (Aysal, Coates, & Rabbat, 2008; Carli, Fagnani, Frasca, & Zampieri, 2010; Kar & Moura, 2010), which is first used in the context of consensus in Aysal et al. (2008), deterministic quantizers (Cai & Ishii, 2011; Chamie, Liu, & Başar, 2014; Kashyap, Başar, & Srikant, 2007; Liu, Li, Xie, Fu, & Zhang, 2013), and dynamic quantization schemes (Li, Fu, Xie, & Zhang, 2011; Li, Liu, Wang, & Lin, 2013; Li & Xie, 2011). A rule of thumb is to choose the algorithm state as an approximation of $\hat{\theta}$. However, the above works show that there is always some gap between the state and $\hat{\theta}$ for static quantizers, which is a function of the quantization resolution. The dynamic encoding/decoding schemes in Li et al. (2011) can

☆ The material in this paper was partially presented at the 54th IEEE Conference on Decision and Control, December 15–18, 2015, Osaka, Japan. This paper was recommended for publication in revised form by Associate Editor Wei Ren under the direction of Editor Ian R. Petersen.

E-mail addresses: shyingzhu@gmail.com (S. Zhu), lius0025@e.ntu.edu.sg (S. Liu), eycsoh@ntu.edu.sg (Y.C. Soh), elhxie@ntu.edu.sg (L. Xie).

1 Fax: +86 2134204022.

2 Our algorithm can be easily extended to the vector case $\boldsymbol{\theta} \in \mathbb{R}^{m_\theta}$ with measurement $\mathbf{y}_i = \mathbf{H}_i\boldsymbol{\theta} + \mathbf{n}_i \in \mathbb{R}^{m_y}$, which is a popular model in power systems (Xie, Choi, Kar, & Poor, 2012). This is done by introducing a second consensus algorithm as in Xiao et al. (2005). Similar convergence results can be established.

http://dx.doi.org/10.1016/j.automatica.2017.02.022

eliminate this gap at the price of global knowledge of some topology-related quantities, which is generally not possible for distributed algorithms. The scheme proposed in Fang and Li (2010) is closely related to our work, where the authors introduced a different form of estimator based on sequence averaging for undirected networks. It was shown that $\hat{\theta}$ can be achieved in the mean square sense even with the basic probabilistic quantizer.

Most of the aforementioned works assume symmetric communication links between nodes. Actually, in ad hoc networks, communication links between certain pairs of nodes may be directed due to non-homogeneous interference, packet collision and so on. To tackle the directed nature of communication links coupled with quantization residues, we proposed a two-stage distributed estimation algorithm using the basic probabilistic quantizer in Zhu, Soh, and Xie (2015). This algorithm can achieve the centralized $\hat{\theta}$ both in the mean square and almost sure senses without any spectral knowledge of the corresponding Laplacian matrix and out-neighbor information. We remark that Zhu, Soh, and Xie (2015) considers the scenario that the nodes take a snapshot measurement of the field, and the consensus strategy is adopted. This can also be formulated within the consensus+innovation framework in Xie et al. (2012), allowing the combination of the two time scales as in Cattivelli and Sayed (2010), Kar and Moura (2011), Kar et al. (2012) and Zhu, Chen et al. (2015) for continuous observations. However, such formulation in Xie et al. (2012) cannot handle the situation with quantized transmissions.

In this paper, we extend our previous work and consider the sensor fusion problem by employing the general additive quantization model. The analysis for the general model is not so straightforward as in Zhu, Soh, and Xie (2015) for the probabilistic quantizer, since no *a priori* knowledge on the boundedness of the quantization errors can be ensured. Hence those approaches heavily relying on such boundedness as in the dithered quantization schemes (Aysal et al., 2008; Carli et al., 2010; Zhu, Soh, & Xie, 2015) are not applicable either. For symmetric communication topologies, consensus problem with additive quantization model has been examined in Yildiz and Scaglione (2008), where the necessary and sufficient conditions for consensus in mean square sense are given. The results indicate, however, that the states of all nodes converge to the same value, with bounded error, with respect to the centralized $\hat{\theta}$ if and only if the quantization noise variances form a convergent series. To achieve this, two dedicatedly designed coding schemes are proposed. Here, we significantly relax the above restrictive conditions on the quantization noise. A preliminary version of this paper was presented at CDC'15 (Zhu, Liu, Xu, Soh, & Xie, 2015). The main contributions of the current paper are twofold: (i) We show that with the two-stage distributed algorithm, the centralized $\hat{\theta}$ can be achieved almost surely even if the quantization noise variances are not vanishing. In our approach, only the existence of slightly higher than the quadratic moment of the quantization noise is needed. This greatly improves the results in Yildiz and Scaglione (2008) and Zhu, Soh, and Xie (2015) and may simplify the design of corresponding coding schemes; (ii) An explicit bound of the convergence rate is provided, which is not available for the consensus algorithms in the literature and cannot be obtained using existing approaches in Aysal et al. (2008), Cai and Ishii (2011), Carli et al. (2010), Cattivelli and Sayed (2010), Chamie et al. (2014), Kar and Moura (2010), Kar et al. (2012), Liu et al. (2013), Yildiz and Scaglione (2008) and Zhu, Soh, and Xie (2015). Our tool is borrowed from probability theory, which establishes a law of the iterated logarithm for weighted sums of random vectors.

The rest of the paper is organized as follows: In Section 2, we present the problem formulation. In Section 3, a law of the iterated logarithm for weighted sums of random vectors is given, based on which the almost sure performance of the proposed algorithm is provided in Section 4. Section 5 discusses the implications of our results. Finally, Section 6 concludes the paper.

Table 1
Two-stage averaging based distributed estimation algorithm.

Initialize $0 < \alpha < 1/\max_i d_i$, $z_{ii}(0) = 1$, $z_{ij}(0) = 0$, $\forall j \neq i$, and $x_i(t_0) = y_i$
Stage 1: Distributed estimation of the left eigenvector $\boldsymbol{\omega}$
(1.1) $\mathbf{z}_i(t+1) = \mathbf{z}_i(t) + \alpha \sum_{j \in \mathcal{N}_i} a_{ij}[\mathcal{Q}(\mathbf{z}_j(t)) - \mathcal{Q}(\mathbf{z}_i(t))]$;
(1.2) $\bar{\mathbf{z}}_i(t+1) = \frac{t}{t+1}\bar{\mathbf{z}}_i(t) + \frac{1}{t+1}\mathbf{z}_i(t+1)$;
Stage 2: Approximation of the centralized estimate $\hat{\theta}$
(2.1) $\epsilon_i(t) \triangleq \begin{cases} \left(\frac{1}{N\bar{z}_{ii}(t_0)} - 1\right)x_i(t_0), & t = t_0, \\ \left(\frac{1}{N\bar{z}_{ii}(t)} - \frac{1}{N\bar{z}_{ii}(t-1)}\right)x_i(t_0), & t > t_0; \end{cases}$
(2.2) $x_i(t+1) = x_i(t) + \epsilon_i(t) + \alpha \sum_{j \in \mathcal{N}_i} a_{ij}[\mathcal{Q}(x_j(t) + \epsilon_j(t)) - \mathcal{Q}(x_i(t) + \epsilon_i(t))]$;
(2.3) $\bar{x}_i(t+1) = \frac{t-t_0}{t-t_0+1}\bar{x}_i(t) + \frac{1}{t-t_0+1}x_i(t+1)$.

Notation: $o(\cdot)$ and $\mathcal{O}(\cdot)$ are the Landau symbols. We simply use $f(k)+o_k$ in the context of $f(k)+o(f(k))$. $\|\cdot\|_2$ and $\|\cdot\|_F$ denote the spectral norm and Frobenius norm for matrices with compatible vector ℓ_2-norm $\|\cdot\|$. $\mathbb{E}\{\mathbf{x}\}$ and $\mathrm{Cov}(\mathbf{x})$ denote its expectation and covariance matrix for a random vector $\mathbf{x}$, respectively.

2. Problem formulation

Consider a sensor network composed of N nodes, which is modeled as a *weighted directed graph* $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathbf{A})$, where $\mathcal{V} = \{1, 2, \ldots, N\}$, $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$ denotes all the unidirectional communication links between nodes and $\mathbf{A} = [a_{ij}]_{N\times N}$ is composed of weights $a_{ij} > 0$ associated with each edge $(j, i) \in \mathcal{E}$. We do not consider self-loops in graph $\mathcal{G}$, i.e., $a_{ii} = 0$, $\forall i$. The directed edge (j, i) means that node i can receive data from node j. We collect all these nodes in the set $\mathcal{N}_i = \{j|(j, i) \in \mathcal{E}\}$ and call them the *neighbors* of node i. Denote by $d_i \triangleq \sum_{j\in\mathcal{N}_i} a_{ij}$ the *in-degree* of node i. Let $\mathbf{L} \triangleq \mathbf{D} - \mathbf{A}$, where $\mathbf{D} \triangleq \mathrm{diag}\{d_1, d_2, \ldots, d_N\}$, represent the Laplacian matrix of graph $\mathcal{G}$. It is clear that $\mathbf{L1} = \mathbf{0}$, and there is a left eigenvector $\boldsymbol{\omega} = [\omega_1, \ldots, \omega_N]^T$ with $\boldsymbol{\omega}^T\mathbf{L} = \mathbf{0}$ and $\mathbf{1}^T\boldsymbol{\omega} = 1$.

Our distributed algorithm involves an estimation of θ at each node with local measurement y_i. In the case of limited bandwidth, each node is only allowed to transmit the quantized data $\mathcal{Q}(\cdot)$ to its neighbors. In this paper, we adopt the additive quantization model as in Gersho and Gray (1992) and Yildiz and Scaglione (2008). Denote by $\mathbf{z}_i(t) \triangleq [z_{i1}, z_{i2}, \ldots, z_{iN}]^T \in \mathbb{R}^N$ or $m_i(t) \in \mathbb{R}$ the unquantized message of node i at iteration t, then the quantized value can be expressed as

$$\mathcal{Q}(\mathbf{z}_i(t)) = \mathbf{z}_i(t) + \mathbf{u}_i(t) \quad \text{or} \quad \mathcal{Q}(m_i(t)) = m_i(t) + v_i(t), \tag{1}$$

where $\mathbf{u}_i(t) \in \mathbb{R}^N$ and $v_i(t) \in \mathbb{R}$ are the quantization errors, which might have unbounded supports (e.g., quantization with finite precision in Kar & Moura, 2010).

The two-stage averaging based distributed estimation algorithm run by node i is summarized in Table 1 (see Zhu, Soh, & Xie, 2015 for some implementation considerations). In the algorithm, Stage 1 first runs for $t_0 \geq 0$ steps, then Stage 2 is triggered. After that, at each iteration, Stage 1 and Stage 2 will be sequentially examined. In examining (1.1) and (2.2), each node incorporates both its unquantized message and the quantized data into updating the local estimate, which is called the compensating updating rule in Carli et al. (2010). The averaging quantity $\bar{x}_i(t)$ rather than $x_i(t)$ is chosen as the approximation of the centralized $\hat{\theta}$ in the algorithm. Besides, the additive quantization model in (1) may result in quantization errors with unbounded supports, which is a big difference from the dithered quantization schemes used in Aysal et al. (2008), Carli et al. (2010) and Zhu, Soh, and Xie (2015). We discuss the communication cost of the proposed algorithm. At each iteration t, each node i transmits its $\mathbf{z}_i(t)$ and $x_i(t)+\epsilon_i(t)$ to its neighbors. Thus, the total number of such transmissions across the entire network at