

Distributed Change Detection Based on a Consensus Algorithm

Srdjan S. Stanković,* Nemanja Ilić,* Miloš S. Stanković,**
Karl Henrik Johansson**

* Faculty of Electrical Engineering, University of Belgrade, Belgrade,
Serbia (e-mail: stankovic@etf.rs, in085049p@student.etf.rs)

** School of Electrical Engineering, Royal Institute of Technology
100-44 Stockholm, Sweden (e-mail: milsta@kth.se, kallej@kth.se)

Abstract: In this paper a novel consensus based distributed recursive algorithm is proposed for real time change detection using sensor networks. Convergence of the algorithm to the optimal centralized solution defined by a weighted sum of the results of local signal processing is proved in the cases of constant and time varying forgetting factors of the underlying recursions, assuming correlated data and different local values of the parameter changes. Simulation results illustrate characteristic properties of the algorithms.

1. INTRODUCTION

A great deal of attention has been paid recently to signal processing using *distributed sensors*, having in mind the low cost and increased computational capabilities of sensors, as well as the availability of high speed *networks* connecting the sensors (e.g., Varshney [1996], Vishwanathan and Varshney [1997]). One of the typical tasks of *sensor networks*, which are in the focus of many researchers, is *distributed detection* (e.g., Chamberland and Veeravalli [2003], Vishwanathan and Varshney [1997], Speranzon et al. [2006]). In the classical multi-sensor detection schemes the local sensors send all their data to other sensors, and ultimately to a *fusion center*. This topology has been found to be too restrictive in many applications. *Distributed signal processing* has, in principle, many advantages, consisting of the increased reliability, reduced communication bandwidth requirements and reduced overall cost. However, distribution of functions may result in a certain loss of performance with respect to the optimal centralized system. In the case of detection of changes in the monitored environment, it is often desirable to have a possibility to test the decision variables in real time at any node in the network, and not only at predefined fusion nodes (e.g., Braca et al. [2008, 2010], Chamberland and Veeravalli [2003]).

Consensus techniques have been studied for many years, starting from the early 80's, when important results were obtained in the areas of distributed asynchronous iterations in parallel computation and distributed optimization (Tsitsiklis et al. [1986], Olfati-Saber et al. [2007], Fax and Murray [2004], Jadbabaie et al. [2003], Moreau [2005], Ren and Beard [2005], etc.). There have been some recent attempts to apply consensus techniques to the distributed detection problem, Franco et al. [2006]. However, the underlying assumption is that the dynamic agreement process starts *after all data have been collected*, implying inapplicability to real time change detection problems. In Stanković et al. [2009b,a, 2007] algorithms for distributed state and parameter estimation have been proposed

by combining local overlapping decentralized estimation schemes with a dynamic consensus algorithm. Analogous algorithms for distributed detection based on “running consensus” have been proposed and discussed in Braca et al. [2008, 2010], on the basis of Boyd et al. [2006].

In this paper an algorithm is proposed for *distributed change detection* while monitoring the environment through a wireless sensor network. It is assumed that all the nodes in the network can generate local decision variables by recursive schemes belonging to the *geometric moving average control charts* (Basseville and Nikiforov [1993]). By applying a dynamic consensus scheme with preselected asymmetric communication gains, one obtains an algorithm which asymptotically provides nearly equal behavior of all the nodes, i.e., any node can be selected for testing the decision variable w.r.t. a pre-specified threshold. It is proved that the mean-square error between the optimal centralized decision variable obtained as a general weighted sum of the results of the local signal processing results and the ones generated by the proposed algorithm is of the order of magnitude of $O((1 - \alpha)^2)$, where α is the forgetting factor of the algorithm. The network gains used by the recursive algorithms can be selected by linear programming, starting from the selected weights of local decision variables. The proposed algorithm follows methodologically Stanković et al. [2009a, 2010] and represents a generalization of the algorithm in Braca et al. [2008, 2010] in which symmetric matrix gains connected to simple averaging have been considered. In the case of time varying forgetting factors tending to zero when t tends to infinity, it is proved that the algorithm converges to the optimal centralized scheme in the mean-square sense under general conditions by using stochastic approximation arguments. Some simulation results are given as an illustration of the characteristic properties of the proposed algorithm.

The outline of the paper is as follows. In Section 2 a distributed change detection scheme based on a consensus algorithm is proposed. In Section 3 the convergence analysis is given assuming constant forgetting factors of

the local recursive schemes, while in Section 4 these results are extended to the case of time varying forgetting factors. Section 5 deals with some illustrative simulation results.

2. DISTRIBUTED CHANGE DETECTION ALGORITHM

Consider a sensor network containing n nodes, where each node collects locally available measurements and generates at each discrete time instant t a scalar quantity $x_i(t)$, $i = 1, \dots, n$, directly, or as a result of local signal processing. We shall consider in the sequel $\{x_i(t)\}$ as mutually independent stationary random sequences with means $E\{x_i(t)\} = m_i$ and covariances $r_i(\tau) = E\{(x_i(t) - m_i)(x_i(t + \tau) - m_i)\}$. We shall assume that the network is aimed at change detection purposes.

In order to explain clearly the main line of thought and to make connection with the arguments usual for the domain of hypotheses testing, we shall start the presentation by assuming the following simple model:

$$x = m + \epsilon, \quad (1)$$

where $x = [x_1, \dots, x_n]^T$, $m = [m_1, \dots, m_n]^T$ and $\epsilon = [\epsilon_1, \dots, \epsilon_n]^T$, with $\epsilon \sim \mathcal{N}(0, \Sigma)$, where $\Sigma = \text{diag}\{\sigma_1^2, \dots, \sigma_n^2\}$. Assuming that $m = \theta^0 = 0$ in the case of no change and that $m = \theta^1 = [\theta_1^1, \dots, \theta_n^1]^T$, where $\theta_i^1 > 0$ for some i in the case of change, we can calculate the log likelihood ratio for the data set containing $x(t)$, $t = 1, \dots, N$, and obtain

$$\begin{aligned} L(N) &= \sum_{t=1}^N \log \frac{p_{\theta^1}(x(t))}{p_{\theta^0}(x(t))} \\ &= \theta^{1T} \Sigma^{-1} \sum_{t=1}^N (x(t) - \frac{1}{2} \theta^1) \end{aligned} \quad (2)$$

(Ding [2008], Basseville and Nikiforov [1993]). Starting from (2), one can apply the general methodology for constructing *on-line change detection* algorithms belonging to the *geometric moving average control charts* (Basseville and Nikiforov [1993]) and obtain the global decision function for the whole network, generated recursively by

$$s_c(t+1) = \alpha s_c(t) + (1-\alpha) \sum_{i=1}^n w_i x_i(t+1), \quad s_c(0) = 0 \quad (3)$$

where $0 < \alpha < 1$ is a *forgetting factor* and $w_i = k \theta_i^1 \sigma_i^{-2}$, with $k = (\sum_{i=1}^n \theta_i^1 \sigma_i^{-2})^{-1}$, are the components of the vector $w^T = k \theta^{1T} \Sigma^{-1}$ (Basseville and Nikiforov [1993]).

The global change detection procedure is based on testing the decision function $s_c(t)$ with respect to an appropriately chosen threshold $\lambda_c > 0$, so that a change is detected when $|s_c(t)| > \lambda_c$ (according to Ding [2008], Basseville and Nikiforov [1993], one can take, for example, in the case of model (1), $\lambda_c = \frac{1}{2} k \theta^{1T} \Sigma^{-1} \theta^1$). It is important to notice that the algorithm (3) requires a *fusion center*. It is also important to emphasize that the algorithm (3) can be considered as a representative of a large class of change detection procedures, without claiming optimality in any sense: the weights $w_i \geq 0$ in (3), satisfying $\sum_{i=1}^n w_i = 1$, can result from any *a priori* selection criterion.

The aim of this paper is to propose a consensus based *distributed change detection* algorithm which *does not*

require a fusion center and in which the output of any preselected node can be used as a representative of the whole network and be tested w.r.t. a pre-specified *common threshold*. The basic assumption for this algorithm is that the nodes of the network are connected in accordance with an $n \times n$ matrix $C = [c_{ij}]$ satisfying $c_{ij} \geq 0$, $i \neq j$ and $c_{ii} > 0$, $i, j = 1, \dots, n$, which formally represents the weighted adjacency matrix for the underlying graph representing the network, and that C is *row stochastic* (Horn and Johnson [1985]). We propose in this paper the following algorithm for generating the *vector decision function* of the network, denoted as $s(t) = [s_1(t), \dots, s_n(t)]^T$:

$$s(t+1) = \alpha C s(t) + (1-\alpha) C x(t+1), \quad s(0) = 0. \quad (4)$$

The algorithm is, in general, derived from the *consensus* based state and parameter estimation algorithms proposed in Stanković et al. [2009b,a, 2007]; it is also similar to the detection algorithm based on time averaging proposed in Braca et al. [2008, 2010]. Notice that the consensus matrix C performs “convexification” of the neighboring states for each node and enforces in such a way (under appropriate conditions) consensus between all the nodes in the network. In such a way, after achieving the condition that $s_i(t) \approx s_j(t)$, $i, j = 1, \dots, n$, change detection can be done by testing $s_i(t)$ for any preselected i with respect to a given common threshold λ_c , provided (4) gives a good approximation of $s_c(t)$ generated by (3).

It is to be noticed that $x(t)$ in (4) can take a more general form than the simple model (1). Therefore, both algorithms (3) and (4) can be treated as general representatives of recursive detection procedures with exponential forgetting, in which $x(t)$ is a random vector with general properties.

The error vector between the vector $s(t)$ and the state of the optimal centralized scheme will be defined as

$$e(t) = s(t) - \mathbf{1} s_c(t), \quad (5)$$

where $\mathbf{1} = [1 \dots 1]^T$. Iterating (4) and (3) back to the zero initial conditions, we get

$$s(t) = (1-\alpha) \sum_{i=0}^{t-1} \alpha^i C^{i+1} x(t-i) \quad (6)$$

and

$$s_c(t) = (1-\alpha) \sum_{i=0}^{t-1} \alpha^i w^T x(t-i), \quad (7)$$

wherefrom we obtain

$$e(t) = (1-\alpha) \sum_{i=0}^{t-1} \alpha^i [C^{i+1} - \mathbf{1} w^T] x(t-i) \quad (8)$$

(compare with Braca et al. [2008, 2010], where the special case when $w = \frac{\mathbf{1}}{n}$ is treated, in conjunction with symmetric time varying consensus matrices).

3. CONVERGENCE ANALYSIS

We shall analyze properties of the proposed algorithm (4) starting from the following assumptions:

- A1) C has the eigenvalue 1 with algebraic multiplicity 1;
- A2) $\lim_{i \rightarrow \infty} C^i = \mathbf{1} w^T$.

The first assumption is related to the topology of the underlying multi-agent network, while the second defines

the connection between the matrix C in (4) and the optimal weights in the centralized scheme (3). Namely, in general, under A1), the graph associated with C has a spanning tree and C^i converges when i tends to infinity to a nonnegative row stochastic matrix with equal rows, e.g. Ren and Beard [2005], Olfati-Saber et al. [2007]. Knowing w from the general problem setting based on the centralized detection strategy, we can construct C satisfying A2) by solving for C the linear equation known from the theory of stationary Markov chains (Kumar and Varaiya [1986])

$$w^T C = w^T, \quad (9)$$

under the constraints that: 1) preselected elements of C are equal to zero (indication that there can be no communication between the corresponding nodes) and 2) matrix C is row stochastic, satisfying the given assumptions.

Using A2) one obtains that $s_c(t) = w^T s(t)$, and, therefore, that $e(t) = (I - \mathbf{1}w^T)s(t)$; also,

$$e(t) = (1 - \alpha) \sum_{i=0}^{t-1} \alpha^i \tilde{C}^{i+1} x(t-i), \quad (10)$$

where $\tilde{C} = C - \mathbf{1}w^T$, having in mind that, under A2), we have $(C - \mathbf{1}w^T)^i = C^i - \mathbf{1}w^T$ and $C^i - \mathbf{1}w^T = (I - \mathbf{1}w^T)(C^i - \mathbf{1}w^T)$.

We first realize that $s(t)$ as an estimator of $\mathbf{1}s_c(t)$ is, in general, biased, since from (10) we have

$$m_e(t) = E\{e(t)\} = (1 - \alpha) \sum_{i=0}^{t-1} \alpha^i \tilde{C}^{i+1} m, \quad (11)$$

where $m = E\{x(t)\} = [m_1, \dots, m_n]^T$. Obviously, $E\{e(t)\} = 0$ only when $m_i = m_j$, $i, j = 1, \dots, n$, having in mind that $\tilde{C}m = 0$ for $m = \mu' \mathbf{1}$, where μ' is a given scalar (compare with Braca et al. [2008], where it has been adopted that $m = 0$).

The bias is, obviously, smaller when α is closer to one. Namely, one can directly conclude that in the steady state we have

$$\lim_{t \rightarrow \infty} E\{s(t)\} = (1 - \alpha)(I - \alpha C)^{-1} C m, \quad (12)$$

and, consequently,

$$\lim_{t \rightarrow \infty} E\{s(t)\} \approx [(1 - \alpha)(I + \alpha C + \dots + \alpha^\nu C^\nu) C + \alpha^{\nu+1} \mathbf{1}w^T] m \quad (13)$$

for some ν large enough. For α close to 1, the first term in the brackets is obviously small, and can be neglected, and the second term is approximately equal to $\lim_{t \rightarrow \infty} E\{s_c(t)\}$.

The focus of the analysis is placed on the error covariance matrix $Q(t) = E\{e(t)e(t)^T\} - m_e(t)m_e(t)^T$. Using (10) and (11) one readily obtains

$$Q(t) = (1 - \alpha)^2 \Phi(t)^T \tilde{R}(t) \Phi(t), \quad (14)$$

where $\Phi(t) = [\alpha^{t-1} \tilde{C}^t; \alpha^{t-2} \tilde{C}^{t-1}; \dots; \alpha^0 \tilde{C}]^T$, $\tilde{R}(t) = R(t) - m_X m_X^T$, $R(t) = E\{X(t)X(t)^T\}$, $X(t) = [x(1)^T \dots x(t)^T]^T$ and $m_X = E\{X(t)\}$. Furthermore, $\tilde{R}(t) = [R_{ij}]$, $i, j = 1, \dots, t$, where R_{ij} are constant $n \times n$ block matrices defined as

$$R_{ij} = \text{diag}\{r_1(i-j), \dots, r_n(i-j)\}, \quad (15)$$

according to the adopted assumptions.

Theorem 1. Let assumptions A1) and A2) hold, together with:

$$\text{A3) } \max_i \sum_{\tau=0}^t |r_i(\tau)| \leq K; \quad 0 < K < \infty.$$

Then,

$$\max_{i,j} Q_{ij}(t) \leq O((1 - \alpha)^2),$$

where $Q_{ij}(t)$ are the elements of $Q(t)$ in (14).

Proof: Consider an arbitrary deterministic n -vector y and analyze the following n -dimensional quadratic form

$$y^T Q(t) y = (1 - \alpha)^2 y^T \Phi(t)^T \tilde{R}(t) \Phi(t) y \quad (16)$$

Having in mind the structure of $\tilde{R}(t)$, we have

$$\lambda_{max}(\tilde{R}(t)) \leq \|\tilde{R}(t)\|_\infty \leq K < \infty$$

by assumption A3) ($\|A\|_\infty = \max_i \sum_j |a_{ij}|$, where $A = [a_{ij}]$ is a given matrix).

Coming back to (16), we realize that the expression $y^T \Phi(t)^T \tilde{R}(t) \Phi(t) y$ is in the form of a sum of terms containing $y^T \tilde{C}^i \tilde{C}^{iT} y$, $i = 0, \dots, t-1$. By assumptions A1) and A2), it follows that C and $\mathbf{1}w^T$ have the same eigenvectors. Therefore, C has the same eigenvalues as \tilde{C} , except for the eigenvalue 1 of C which is replaced by the eigenvalue 0 of \tilde{C} . Having in mind that $c_{ii} > 0$, $i = 1, \dots, n$, it follows that the modules of all the eigenvalues of \tilde{C} are strictly less than 1 (Ren and Beard [2005]).

Consider the recursion $P(t+1) = \tilde{C}P(t)\tilde{C}^T$ with $P(0) = I$; obviously, $P(t) = \tilde{C}^t \tilde{C}^{tT}$ for all t . Let $\text{vec}\{P(t)\}$ be an n^2 -vector obtained by concatenating the column vectors of $P(t)$. Then, we have, for $t = 0, 1, \dots$

$$\text{vec}\{P(t+1)\} = (\tilde{C} \otimes \tilde{C}) \text{vec}\{P(t)\},$$

where " \otimes " denotes the Kronecker's product. According to A1) and A2), we have directly that $|\lambda(\tilde{C} \otimes \tilde{C})|_{max} = \lambda_M < 1$, having in mind that the eigenvalues of $\tilde{C} \otimes \tilde{C}$ are equal to the cross products of the eigenvalues of \tilde{C} (Horn and Johnson [1985]); therefore, $\|P(t)\| \leq k_P \lambda_M^t$ for some matrix norm, where k_P is an appropriate constant.

Therefore, we have

$$y^T \tilde{C}^i \tilde{C}^{iT} y \leq k_P \lambda_M^i \|y\|^2,$$

so that we obtain

$$y^T \Phi(t)^T \tilde{R}(t) \Phi(t) y \leq \|y\|^2 k' K \sum_{i=0}^{\infty} \alpha^i \lambda_M^{i+1} \leq \leq K_1 < \infty, \quad (17)$$

where $0 < k' < \infty$ and K_1 does not depend on α . Notice that the term $y^T \tilde{C} \tilde{C}^T y$ can be directly maximized by $\lambda_{max}(\tilde{C} \tilde{C}^T) \|y\|^2$; however, there is no a priori guarantee that $\lambda_{max}(\tilde{C} \tilde{C}^T) < 1$.

Consequently, by choosing $y = e_i$, where e_i denotes the n -vector of zeros with only the i -th entry equal to one, one obtains that

$$Q_{ii}(t) \leq K_1 (1 - \alpha)^2,$$

$i = 1, \dots, n$. Furthermore, $|Q_{ij}(t)| \leq \max_i Q_{ii}(t)$, having in mind elementary properties of positive semidefinite matrices. Thus, the result. Q.E.D.

The meaning of the obtained result becomes clearer after realizing that for α close to 1 the mean square error between the states of the proposed distributed detection

algorithm and the state of the centralized one is much smaller than the mean square value of the state of the centralized system. Namely, it follows from (7) and (17) that

$$\begin{aligned} \text{var}\{s_c(t)\} &= E\{s_c(t)^2\} - E\{s_c(t)\}^2 \leq \quad (18) \\ &\leq K(1-\alpha)^2 \sum_{i=0}^{\infty} \alpha^{2i} \leq O(1-\alpha) \end{aligned}$$

having in mind that $E\{s_c(t)\} = (1-\alpha) \sum_{i=0}^{t-1} \alpha^i w^T m$ and $\sum_{i=0}^{\infty} \alpha^{2i} = \frac{1}{1-\alpha^2}$. On the other hand, one obtains, using the above arguments, that $\|m_e(t)\| \leq O(1-\alpha)$, so that the overall mean square error of the state of the i -th node satisfies $Q_{ii}(t) + m_e(t)_i^2 \leq O((1-\alpha)^2)$.

4. DISTRIBUTED DETECTION BASED ON TIME AVERAGING

The recursive algorithms (3) and (4) with the constant positive coefficient α represent essentially tracking algorithms with exponential forgetting, able to cope with change detection phenomena (Basseville and Nikiforov [1993]). The same form of the algorithms can be used also in the case when detection has to be based on time averaging on infinite intervals. However, in this case we shall assume that α is a function of time tending to 1 when t tends to infinity, requiring that the error between the states of the algorithms (3) and (4) converges to zero in the mean-square sense (in the above case, only upper bounds in the form of functions of the parameter α can be considered). However, the algorithms are then not directly suitable for change detection purposes. Notice that, in general, both algorithms (3) and (4) can be considered, after assuming time varying forgetting factor, as stochastic approximation algorithms, *e.g.* Chen [2002]. Stochastic approximation algorithms with consensus, representing a generalization of (4) to the regression problem, have been analyzed in Stanković et al. [2007, 2010] starting from the basic results presented in Tsitsiklis et al. [1986].

Theorem 2. Let in (3) and (4) the forgetting factor be in the form $\alpha(t+1) = 1 - \gamma(t+1)$, and let the assumptions A1), A2) and A3) be satisfied, together with:

A4) $\gamma(t)$ is a non-increasing sequence satisfying $\gamma(t) > 0$, $\lim_{t \rightarrow \infty} \gamma(t) = 0$; $\sum_{t=1}^{\infty} \gamma(t) = \infty$.

Then,

$$\|Q(t)\| = o(1)$$

($o(1)$ stands for a sequence tending to zero as t tends to infinity).

Proof: Starting from (4) and (6) one obtains (8). Consequently, in the case of time varying gains one obtains

$$e(t) = \sum_{i=0}^{t-1} \pi(t, t+1-i) \tilde{C}^{i+1} \gamma(t-i) x(t-i), \quad (19)$$

where $\pi(i, j) = 1$ for $i < j$ and $\pi(i, j) = \alpha(i) \cdots \alpha(j)$ for $i \geq j$.

Following the above line of thought, we calculate $Q(t) = E\{e(t)e(t)^T\} - m_e(t)m_e(t)^T$ and obtain, similarly as in (16), the following expression

$$y^T Q(t) y = y^T \Psi(t)^T \tilde{R}(t) \Psi(t) y, \quad (20)$$

where

$$\begin{aligned} \Psi(t) &= [\pi(t, 2) \tilde{C}^t \gamma(1); \pi(t, 3) \tilde{C}^{t-1} \gamma(2); \cdots \\ &\quad \cdots; \pi(t, t-1) \tilde{C} \gamma(t)]. \end{aligned}$$

Proceeding like in the proof of Theorem 1, we obtain

$$y^T Q(t) y \leq \|y\|^2 k' K \sum_{i=0}^{t-1} \pi(t, t+1-i)^2 \lambda_M^{i+1} \gamma(t-i)^2. \quad (21)$$

Now it is possible to apply the Kronecker's lemma (*e.g.* Chen [2002]), and to conclude immediately that

$$\lim_{t \rightarrow \infty} \sum_{i=0}^t \pi(t, t+1-i)^2 \lambda_M^{i+1} \gamma(t-i)^2 = 0,$$

wherefrom the result. Q.E.D.

The result of Theorem 2 can be applied, obviously, to the special case when $\gamma(t) = \frac{1}{t}$, treated in Braca et al. [2008, 2010] under the assumption that the consensus matrix is symmetric and that the sequences $\{x_i(t)\}$, $i = 1, \dots, n$, are mutually independent and iid. Formally speaking, the difference between the algorithm (4) with $\alpha(t+1) = 1 - \frac{1}{t+1}$ and the algorithm analyzed in Braca et al. [2008] lies also in the fact that we apply the "convexification" operator not only to the previous detector state, but also to the measurement term, leading to an additional smoothing effect. It is even more important that the essential generalization allowing nonsymmetric matrices C enables obtaining convergence to the centralized change detector based on a general weighting of the available signals.

Corollary 1. Under the assumptions of Theorem 2 and with $\gamma(t) = \frac{1}{t}$ we have

$$\|Q(t)\| = O(t^{-2}), \quad (22)$$

while $\text{var}\{s_c(t)\} = O(t^{-1})$.

Proof: In this special case, we have from (19)

$$e(t) = \frac{1}{t} \sum_{i=0}^{t-1} \tilde{C}^{i+1} x(t-i), \quad (23)$$

and the result (22) immediately follows after applying the methodology of Theorems 1 and 2. Notice that $\text{var}\{s_c(t)\}$ can be easily calculated in this case similarly as in Braca et al. [2008] using

$$s_c(t) = \frac{1}{t} \sum_{i=0}^{t-1} w^T x(t-i), \quad (24)$$

as well as $E\{s_c(t)\} = w^T m$. Q.E.D.

5. SIMULATION RESULTS

Example 1. - Properties of the Distributed Change Detection Algorithm. Let us consider a sensor network with $n = 10$ nodes, where the means m_i and variances σ_i^2 are randomly taken from the intervals $[0, 1]$ and $[0.5, 1.5]$, respectively ($m_i = 0$ in the case of no change). Communication (consensus) gains are obtained by solving the linear programming problem (9) under the constraints that C is row stochastic and possesses a predefined structure (places of zeros). The moment of change is chosen to be $t = 200$. The proposed algorithm effectively achieves a very similar

behavior of all of the nodes, with local decision functions getting closer to the global decision function as $\alpha \rightarrow 1$. In Fig. 1 the global decision function is given by dashed lines (mean \pm one standard deviation), together with the decision function of one randomly selected node (solid lines). The obtained change detector quality is obvious. In addition, it can be seen that the proposed algorithm achieves better performance when compared to the special case algorithm with the weight vector $w = \frac{1}{n}$, and in which at each discrete time instant two randomly selected nodes exchange their data, leading to a symmetric consensus matrix, as in Braca et al. [2008] (Fig. 2).

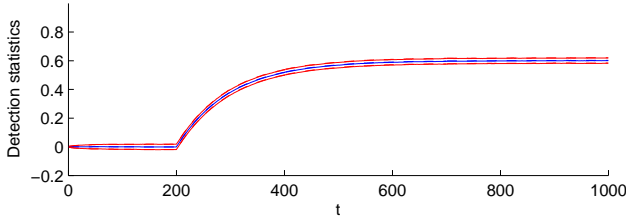
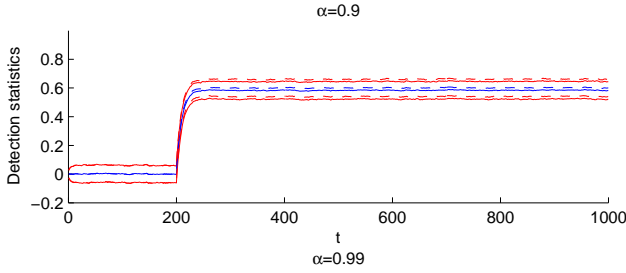


Fig. 1. Decision functions: for one node (solid lines) and global (dashed lines), for asymmetric consensus matrix C

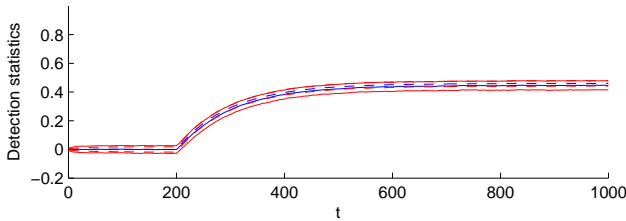
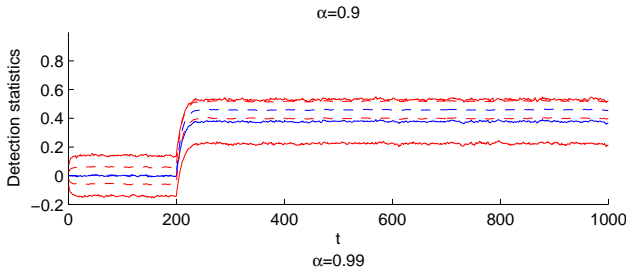


Fig. 2. Decision functions: for one node (solid lines) and global (dashed lines), for symmetric consensus matrix C

Example 2. - Convergence Analysis of the Distributed Change Detection Algorithm. The same sensor network from the previous example is considered. The values of $|E\{e(t)\}|$ and $Q_{ii}(t)$ are estimated for different values of α for $t = 1000$ using 1000 Monte Carlo runs. As can be seen in Fig. 3 and Fig. 4, $|E\{e(t)\}| \sim (1 - \alpha)$ and

$Q_{ii}(t) \sim (1 - \alpha)^2$, which confirms the above convergence analysis.

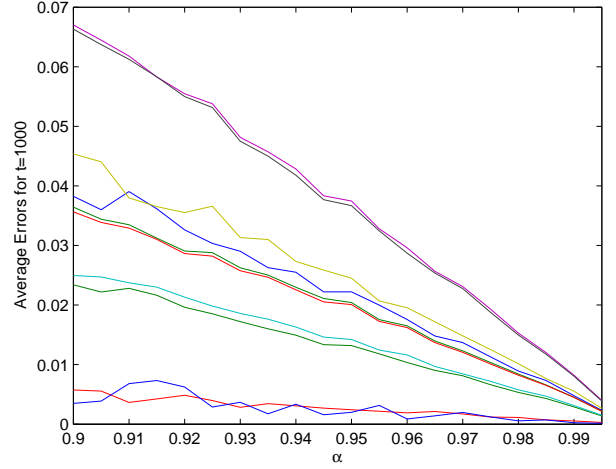


Fig. 3. Average errors as functions of α

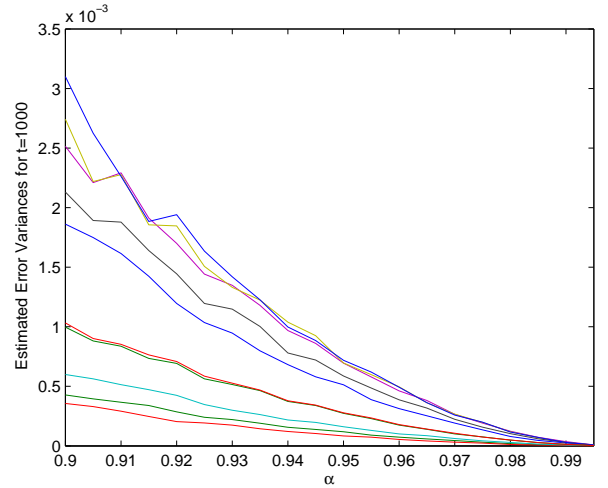


Fig. 4. Estimated error variances as functions of α

Example 3. - Convergence Analysis of the Distributed Detection Algorithm Based on Time Averaging. Consider the sensor network from previous examples. Estimates of $Q_{ii}(t)$ are calculated for $t = 1, \dots, 50$ using 10000 Monte Carlo runs. Fig. 5 shows that $Q_{ii}(t) \sim t^{-2}$, as claimed above.

6. CONCLUSION AND DISCUSSION

In this paper a consensus based distributed recursive algorithm based on geometric moving average control charts is proposed for change detection while monitoring environment by sensor networks. Convergence of the algorithm to the optimal centralized solution is studied assuming temporally correlated data and different local environment. The analysis encompasses constant and time varying forgetting factors in the algorithm. Simulation results give an illustration of the presented theoretical results.

An immediate continuation of the research presented in this paper can be oriented towards two important subjects.

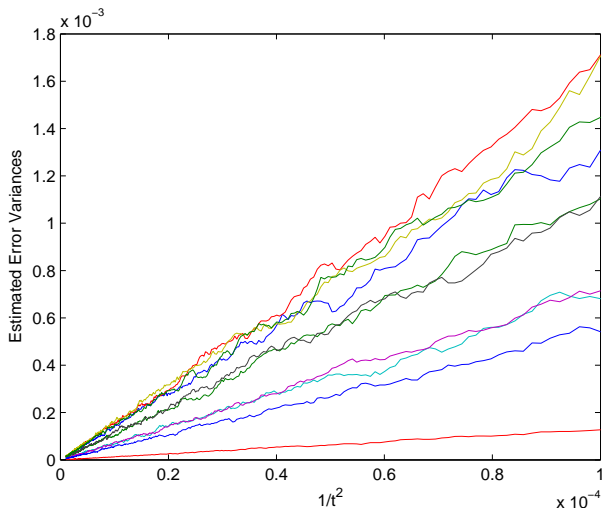


Fig. 5. Estimated error variances as functions of time

The first subject is related to the generalization of the presented results to the case of stochastic time varying consensus matrices in (4), allowing randomized "gossip" algorithms, in which communications between the nodes are unidirectional (not obligatorily bidirectional as in Braca et al. [2008]). Such a case requires an additional care from the point of view of both the asymptotic behavior of the consensus matrix and the convergence to the optimal centralized solution.

The second subject could be related to the application of the proposed consensus based methodology to the recursive algorithms based on the Generalized Likelihood Ratio, which corresponds to the case of unknown parameter jumps (Basseville and Nikiforov [1993], Ding [2008]).

In general, further work can be oriented towards efficient applications of the proposed methodology to complex problems of decentralized overlapping fault detection and isolation in large scale systems.

ACKNOWLEDGEMENTS

This work was partially supported by the EU project PRODI, the Swedish Research Council, the Swedish Strategic Research Foundation, the Swedish Governmental Agency for Innovation Systems and the EU project FeedNetBack.

REFERENCES

M. Basseville and L. V. Nikiforov. *Detection of Abrupt Changes: Theory and Applications*. Prentice Hall, 1993.

S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah. Randomized gossip algorithms. *IEEE Trans. Inf. Theory*, 52: 2508–2530, 2006.

P. Braca, S. Marano, and V. Matta. Enforcing consensus while monitoring the environment in wireless sensor networks. *IEEE Trans. Signal Processing*, 56:3375–3380, 2008.

P. Braca, S. Marano, V. Matta, and P. Willett. Asymptotic optimality of running consensus in testing binary hypotheses. *IEEE Trans. Signal Processing*, 58:814–825, 2010.

J. F. Chamberland and V. Veeravalli. Decentralized detection in sensor networks. *IEEE Trans. Signal Proc.*, 50:407–416, 2003.

H. F. Chen. *Stochastic approximation and its applications*. Kluwer Academic, Dordrecht, the Netherlands, 2002.

S. X. Ding. *Model Based Fault Diagnosis Techniques - Design Schemes, Algorithms and Tools*. Springer Verlag, 2008.

A. Fax and R.M Murray. Information flow and cooperative control of vehicle formations. *IEEE Trans. Automat. Contr.*, 49:1465–1476, 2004.

E. Franco, R. Olfati-Saber, T. Parisini, and M. M. Polycarpou. Distributed fault diagnosis using sensor networks and consensus based filters. In *Proc. 45th IEEE CDC Conf.*, 2006.

R. A. Horn and C. A. Johnson. *Matrix Analysis*. Cambridge Univ. Press, 1985.

A. Jadbabaie, J. Lin, and A. Morse. Coordination of groups of mobile autonomous agents using nearest neighbor rules. *IEEE Trans. Automat. Contr.*, 48:988–1001, 2003.

P.R. Kumar and P. Varaiya. *Stochastic systems: estimation, identification and adaptive control*. Prentice Hall, Englewood Cliffs, New Jersey, 1986.

L. Moreau. Stability of multiagent systems with time-dependent communication links. *IEEE Trans. Automat. Contr.*, 50:169–182, 2005.

R. Olfati-Saber, A. Fax, and R. Murray. Consensus and cooperation in networked multi-agent systems. *Proceedings of the IEEE*, 95:215–233, 2007.

W. Ren and R.W. Beard. Consensus seeking in multi-agent systems using dynamically changing interaction topologies. *IEEE Trans. Autom. Control*, 50:655–661, 2005.

A. Speranzon, C. Fischione, and K. H. Johansson. Distributed and collaborative estimation over wireless sensor networks. In *Proc. IEEE Conf. on Decision and Control*, 2006.

S. S. Stanković, M. S. Stanković, and D. M. Stipanović. Decentralized parameter estimation by consensus based stochastic approximation. In *Proc. 46th IEEE Conference on Decision and Control*, pages 1535–1540, 2007.

S. S. Stanković, M. S. Stanković, and D. M. Stipanović. Consensus based overlapping decentralized estimation with missing observations and communication faults. *Automatica*, 45:1397–1406, 2009a.

S. S. Stanković, M. S. Stanković, and D. M. Stipanović. Consensus based overlapping decentralized estimator. *IEEE Trans. Autom. Control*, 54:410–415, 2009b.

S. S. Stanković, M. S. Stanković, and D. M. Stipanović. Decentralized parameter estimation by consensus based stochastic approximation. *IEEE Trans. Autom. Control*, accepted for publication, 2010.

J. N. Tsitsiklis, D. P. Bertsekas, and M. Athans. Distributed asynchronous deterministic and stochastic gradient optimization algorithms. *IEEE Trans. Autom. Control*, 31:803–812, 1986.

P. K. Varshney. *Distributed Detection and Data Fusion*. Springer, New York, 1996.

R. Vishwanathan and P. Varshney. Distributed detection with multiple sensors: Part I - fundamentals. *Proc. of the IEEE*, 85:54–63, 1997.