Asymptotically optimum tests for decentralized change detection

Georgios Fellouris[1] and George V. Moustakides[2]

[1] Department of Statistics, Columbia University,
1255 Amsterdam Avenue,
New York, NY 10027, USA
(e-mail: gaf2106@columbia.edu)

[2] Department of Electrical and Computer Engineering,
University of Patras,
26500 Patras, Greece
(e-mail: moustaki@upatras.gr)

Abstract. Assuming that sensors have full local memory, we introduce a novel test for the problem of decentralized change detection that is asymptotically optimum. According to our proposed scheme, sensors perform locally repeated SPRTs and communicate, asynchronously, their one-bit decisions to a fusion center. The fusion center in turn uses the sequentially acquired information to perform a CUSUM test in order to decide whether a change took place or not. We prove that the average detection delay of the proposed test differs from the optimum centralized CUSUM test only by a constant. This fact suggests order-2 asymptotic optimality as compared to existing schemes that are optimal of order-1.

Keywords. CUSUM, SPRT, Decentralized detection.

1   Introduction

Consider the geometry depicted in Fig. 1 where K sensors observe K statistically independent, *continuous-time* processes $\{\xi_{t,i}\}_{t \geq 0}$, $i = 1, \ldots, K$. At a deterministic but unknown time $\tau \geq 0$ the observed processes experience a
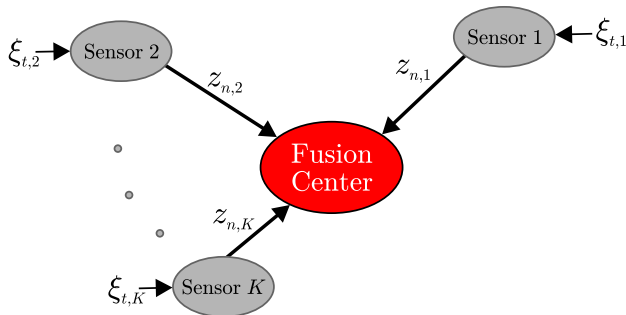

Fig. 1. Geometry of the decentralized change detection problem.

simultaneous change in their statistics from a probability measure $\mathbb{P}_{\infty,i}$ to an alternative $\mathbb{P}_{0,i}$. The goal is to detect this change as soon as possible. Detection must be performed at a Fusion Center which receives, sequentially, *discrete-time* information $\{z_{n,i}\}_{n\geq 0}, n \in \mathbb{Z}$, conveyed from the sensors with the help of standard (wireless) digital communication systems. In order to limit the need in communication bandwidth between sensors and fusion center the communication rate must be kept low. This requirement practically excludes transmission of samples obtained from the continuous-time signals by conventional deterministic canonical sampling and demands for sampling strategies that are more efficient.

Decentralized detection was first introduced by Tsitsiklis (1990) and later in Veeravali et.al. (1993) we find a detailed presentation of the different sampling models that can be defined. We note that a decentralized detection scheme is comprised of two parts 1) the *sampling strategy* at the sensors and 2) the fusion center *detection policy*. Sampling strategies define the type of information to be transmitted from sensors to the fusion center whereas detection policies how this information should be utilized by the fusion center to produce its final decision (whether a change took place or not). We also distinguish the decentralized schemes from the *centralized* detection structures in which the fusion center has complete access to the continuous-time processes $\{\xi_{t,i}\}$. It is clear that the application of a centralized optimum test gives rise to the *ultimate point of reference* in performance.

Current decentralized literature mainly refers to discrete-time signals and to *synchronous* communication between sensors and fusion center, silently assuming the existence of a *global clock*. Transferring this methodology to the continuous-time case requires the processes $\{\xi_{t,i}\}$ to be sampled concurrently, using canonical sampling. The acquired samples $\xi_{nT,i}$ (T being the constant sampling period) need to be further processed with the help of an additional sampler, more accurately quantizer, in order to produce the signals $\{z_{n,i}\}$ to be transmitted. We should mention the work by Tartakovsky and Veeravalli (2002) where one can find (asymptotically) optimum schemes for the models with partial local memory and feedback or with no local memory and no feedback, under a Bayesian setting. The corresponding results under a non-Baysian formulation can be found in Moustakides (2006).

In both cases the proposed sampling/detection strategies result in tests whose performance is *significantly inferior* to the corresponding centralized test, suggesting possibilities for important improvements. The first effort in this direction was offered by Mei (2005). The sampling/detection scheme developed in this work was for the model with *full local memory* where sensors remember all their past acquired samples. The proposed test was shown to be asymptotically optimum, in the sense that the ratio of its performance and the performance of the centralized test tends to 1, as the false alarm period tends to ∞. Comparisons of this scheme with other decentralized and centralized alternatives are reported in Tartakovsky and Kim (2006).

In this work we follow the same model as in Mei, i.e. examine sensors with full local memory. However, the test we develop has performance that falls behind the centralized test only by a finite amount. This should be compared with Mei's scheme whose performance, as reported in Tartakovsky and Kim (2006), diverges from the centralized test, even though the ratio tends to 1.

2   Problem formulation and random sampling

At Sensor-i, let $u_{t,i} = \log \frac{d\mathbb{P}_{0,i}}{d\mathbb{P}_{\infty,i}}(\mathcal{F}_{t,i})$ denote the local log-likelihood ratio, where $\{\mathcal{F}_{t,i}\}$ is the filtration generated by the observed continuous-time signal $\{\xi_{t,i}\}$. Let also $\{t_n^i\}$ be a strictly increasing sequence of sampling instances with $\lim_{n\to\infty} t_n^i = \infty$ a.s. where $t_{n,i}$ are stopping times (s.t.) adapted to $\{\mathcal{F}_{t,i}\}$. Consider now the sampled version $\{u_{t_n^i,i}\}$ of the local log-likelihood ratio and suppose that these values are available at the fusion center at the sampling times $\{t_n^i\}$.

Let $\{\tilde{\mathcal{F}}_t\}$ be the filtration at the fusion center generated by the received samples from all sensors, i.e. $\tilde{\mathcal{F}}_t = \sigma\{u_{t_n^i,i};\ t_n^i \le t; i = 1,\ldots,K\}$ is all the *asynchronously* received information up to time t. We also distinguish the filtration $\{\mathcal{F}_t\}$ where $\mathcal{F}_t = \sigma\{u_{s,i};\ s \le t; i = 1,\ldots,K\}$ is the received information at the fusion center for the continuous-time centralized test. Note that since $\{u_{t,i}\}$ is a sufficient statistics for Sensor-i in the centralized test, transmitting this signal instead of the observed $\{\xi_{t,i}\}$ produces no performance loss for the optimum detection structures.

We now recall a popular test for the solution of the change detection problem under the centralized setup. The Cumulative Sum (CUSUM) test is perhaps the most well-known test for solving the problem of interest. It is defined as follows

$$u_t = u_{t,1} + u_{t,2} + \cdots + u_{t,K};\ m_t = \inf_{0 \le s \le t} u_s;\ y_t = u_t - m_t;$$

$$\mathcal{S} = \inf_{t \ge 0}\{t : y_t \ge \nu\},$$

where $\nu > 0$ is a constant threshold and $\mathcal{S}$ is the first time the test statistics y_t exceeds this quantity. Optimality of CUSUM in continuous-time was first established by Shiryayev (1996), for a Brownian Motion (BM) with constant drift before and after the change. Specifically one can demonstrate that $\mathcal{S}$ solves the following optimization problem due to Lorden (1971)

$$\inf_{\mathcal{T}} \sup_{\tau \ge 0} \operatorname{esssup} \mathbb{E}_\tau[(\mathcal{T} - \tau)^+|\mathcal{F}_\tau];\ \ \text{subject to } \mathbb{E}_\infty[\mathcal{T}] \ge \gamma. \tag{1}$$

Here $\mathbb{E}_\tau[\cdot]$ denotes expectation with respect to the probability measure induced by a change at time τ. More precisely we are looking for the s.t. $\mathcal{T}$ that minimizes the worst average detection delay subject to the constraint that the average period between false alarms is no less than γ. As was indicated, CUSUM solves (1) for BMs provided that ν is selected so that the false alarm constraint is satisfied with equality.

2.1 CUSUM with general sampling

Replacing the continuous-time log-likelihood ratios $\{u_{t,i}\}$ with their sampled versions $\{u_{t_n^i,i}\}$, gives rise to the test statistics $\tilde{u}_t = u_{t_{n_t}^1,1} + \cdots + u_{t_{n_t}^K,K}$, where $t_{n_t}^i = \max_{t_n^i \geq 0}\{t_n^i : t_n^i \leq t\}$ is the last sampling instant before (and including) t at Sensor-i. In other words, at every time t we add the most recently acquired local log-likelihood ratios. The test then continuous as in the case of the continuous-time CUSUM, specifically we define

$$\tilde{m}_t = \inf_{0 \leq s \leq t} \tilde{u}_s; \quad \tilde{y}_t = \tilde{u}_t - \tilde{m}_t; \quad \tilde{\mathcal{S}} = \inf_{t \geq 0}\{t : \tilde{u}_t \geq \tilde{\nu}\}, \quad (2)$$

where again $\tilde{\nu}$ is selected to satisfy the false alarm constraint with equality.

The need for transmitting the samples $\{u_{t_n^i,i}\}$ exhibits the same difficulty as in the case of canonical sampling with $t_n^i = nT$, since these quantities are real numbers. Regarding canonical sampling, we recall that the resulting test applied at the fusion center, is the discrete-time CUSUM which is known to optimize the discrete-time version of Lorden's criterion. Next we are going to introduce a suitable selection of sampling instances $\{t_n^i\}$ that allows for the communication of $\{u_{t_n^i,i}\}$ by simply transmitting *1-bit* information.

3   Proposed sampling/detection strategies

Crucial point in being able to implement the test in (2) is the availability, at the fusion center, of the samples $\{u_{t_n^i,i}\}$ of the local log-likelihood ratios. We observe that we can write

$$u_{t_n^i,i} = [u_{t_n^i,i} - u_{t_{n-1}^i,i}] + [u_{t_{n-1}^i,i} - u_{t_{n-2}^i,i}] + \cdots + [u_{t_1^i,i} - u_{t_0^i,i}] \quad (3)$$

where we define $t_0^i = 0$ and assume that $u_{0,i} = 0$. It is therefore sufficient for Sensor-i to transmit the differences $[u_{t_n^i,i} - u_{t_{n-1}^i,i}]$ between consecutive sampling times. The key idea is to select the sequence of s.t. $\{t_n^i\}$ so that these differences constitute 1-bit information. In fact, as we shall see next, this is not very complicated.

For Sensor-i select before hand two boundaries $A_i < 0 < B_i$ which are also known to the fusion center. Suppose that t_{n-1}^i is already set and define

$$t_n^i = \inf_{t > t_{n-1}^i} \{t : u_{t,i} - u_{t_{n-1}^i,i} \notin (A_i, B_i)\}.$$

If $\{u_{t,i}\}$ has continuous paths we observe that at time t_n^i the difference $u_{t_n^i,i} - u_{t_{n-1}^i,i}$ will hit either A_i or B_i and this information can be transmitted, at time t_n^i, to the fusion center using 1 bit. Indeed if $z_{t_n^i,i}$ is the quantity to be transmitted we define it to be 1 when $u_{t_n^i,i} - u_{t_{n-1}^i,i} \geq B_i$ and 0 when $u_{t_n^i,i} - u_{t_{n-1}^i,i} \leq A_i$. Once t_n^i has been set, we repeat the same process for $t_{n+1}^i, \ldots$, etc. The procedure we just described is simply a *repeated* Sequential Probability Ratio Test (SPRT) where every time the test statistics $u_{t,i} - u_{t_{n-1}^i,i}$ hits one of the two boundaries A_i, B_i we restart the SPRT after

updating the term $u_{t_{n-1}^i, i}$. The interesting point is that by properly selecting the two boundaries A_i, B_i we have complete control over the *average sampling period* since the latter is simply the "detection delay" of the corresponding SPRT. At the fusion center, whenever we receive a new information bit from any sensor, we update the corresponding local log-likelihood ratio from (3), then compute the new test statistics $\tilde{y}_t$ and we apply the test following (2).

4   The Brownian Motion case

In this section we focus on the special case of the signals $\{\xi_{t,i}\}$, $i = 1, \ldots, K$ being standard BMs with drifts equal to 0 and $\mu_i \neq 0$ before and after the change respectively. We can then verify that $u_{t,i} = -0.5\mu_i^2 t + \mu_i \xi_{t,i}$. We first examine the centralized CUSUM test. For this scheme we have the following convenient formulas

$$\mathbb{E}_0[\mathcal{S}] = \frac{2}{\mu_1^2 + \cdots + \mu_K^2}(\nu + e^{-\nu} - 1); \quad \mathbb{E}_\infty[\mathcal{S}] = \frac{2}{\mu_1^2 + \cdots + \mu_K^2}(e^\nu - \nu - 1). \quad (4)$$

Threshold ν can be computed by imposing validity of the false alarm constraint with equality, i.e. $\mathbb{E}_\infty[\mathcal{S}] = \gamma$. This yields $\nu = \log(\gamma) + O(1)$. Regarding now the proposed test, the next theorem demonstrates that our scheme differs from the optimum, only by a bounded quantity.

Theorem 1. *Let $\xi_{t,i}, i = 1, \ldots, K$, be as above; consider the centralized CUSUM test $\mathcal{S}$ and the proposed test $\tilde{\mathcal{S}}$ with fixed local boundaries $A_i < 0 < B_i$ and select both thresholds $\nu, \tilde{\nu}$ to satisfy the false alarm constraint γ with equality. Then, uniformly over γ, we have i) $|y_t - \tilde{y}_t| \leq C < \infty$ and ii) $0 \leq \mathbb{E}_0[\tilde{\mathcal{S}}] - \mathbb{E}_0[\mathcal{S}] \leq D < \infty$.*

Proof. We only highlight the main steps. To show i), because of SPRT sampling we have $A_i < u_{t,i} - u_{t_{n_t}^i, i} < B_i$, suggesting $\sum A_i < u_t - \tilde{u}_t < \sum B_i$. Exactly the same double inequality holds for the difference $m_t - \tilde{m}_t$, yielding the necessary uniform bound C for the difference $y_t - \tilde{y}_t$. For ii), using i) we observe that $\mathcal{S}_l \leq \tilde{\mathcal{S}} \leq \mathcal{S}_u$ where $\mathcal{S}_l = \inf_{t \geq 0}\{t : y_t + C \geq \tilde{\nu}\}$ and $\mathcal{S}_u = \inf_{t \geq 0}\{t : y_t - C \geq \tilde{\nu}\}$. Combining this observation with the formulas in (4), we can show that $\tilde{\nu} = \log(\gamma) + O(1)$ and then that the desired difference is indeed $O(1)$. Our proof, when the local thresholds A_i, B_i are bounded, assures order-2 asymptotic optimality. If we let A_i, B_i increase with γ but at a rate such that $(A_i, B_i)/\log(\gamma) \to 0$ then we can demonstrate only order-1 optimality. Such increase however in the local thresholds will result in an *extremely infrequent communication* between sensors and fusion center.

4.1   Simulations

We examine the case of $K = 2$ sensors with $\mu_1 = \mu_2 = 1$ and apply three different tests. First is the continuous-time centralized CUSUM which serves as a point of reference and its performance is analytically given in (4). Second, we simulate the discrete-time centralized CUSUM, with signals sampled with

a constant period $T = 8\tanh(2)$ and the corresponding samples sent, without quantization, to the fusion center. The latter applies the discrete-time CUSUM test for its final decision. Third is our scheme with local boundaries $B_i = -A_i = 4$ having an *average sampling period* of $2B_i \tanh(B_i/2)/\mu_i^2$ that matches the period T of the centralized discrete-time test. This is necessary for a fair comparison of the two schemes.

Fig. 2 depicts the average detection delay (in linear scale) as a function of the average false alarm period γ (in logarithmic scale). We observe that the proposed test performs better than the discrete-time CUSUM, exhibiting a performance which is very close to the optimum. We recall that in our test, sensors transmit, *asynchronously*, just 1-bit information to the fusion center while the centralized transmits, *synchronously*, real numbers. Mei's scheme also includes 1-bit (but synchronous) transmissions. Its performance on the other hand, as mentioned before, *diverges* from the optimum.
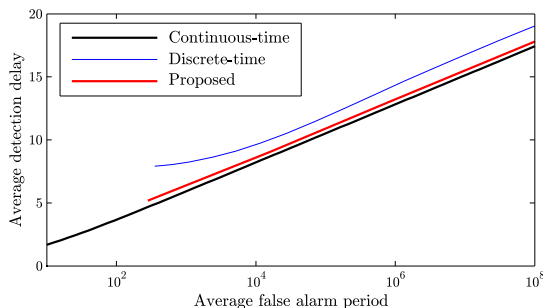

Fig. 2. Relative performance of various change detection schemes.

References

Lorden, G. (1971): Procedures for reacting to a change in distribution. *Ann. Math. Stat., 42*, 1897-1908.

Mei, Y. (2005): Information bounds and quickest change detection in decentralized decision systems. *IEEE Tr. Inf. Th., 51*(7), 2669-2681.

Moustakides, G.V. (2006): Decentralized CUSUM change detection. *Proc. 9th Int. Conf. Inf. Fusion*, Fusion 2006, Florence, Italy.

Shiryayev, A.N. (1996): Minimax optimality of the method of cumulative sums (CUSUM) in the case of continuous time. *Russ. Math. Surv., 51*, 750751.

Tartakovsky, A.G. and Veeravalli, V. (2002): An efficient sequential procedure for detecting changes in multichannel and distributed systems. *Proc. 5th Int. Conf. Inf. Fusion*, Fusion 2002, Annapolis, USA.

Tartakovsky, A.G. and Kim, H. (2006): Performance of certain decentralized distributed change detection procedures. *Proc. 9th Int. Conf. Inf. Fusion*, Fusion 2006, Florence, Italy.

Tsitsiklis, J.N. (1990): Decentralized detection. *Advances in Statistical Signal Processing*, Greenwich, CT: JAI Press.

Veeravalli, V.V., Basar, T. and Poor, H.V. (1993): Decentralized sequential detection with a fusion center performing the sequential test, *IEEE Tr. Inf. Th., 39*(3), 433-442.