CUSUM Based Distributed Detection in WSNs

Mandar Nadgir
Member of Technical Staff
Computational Reseach Labs
Pune, India
mandar@crlindia.com

K. Premkumar
Department of Electrical
Communication Engineering
Indian Institute of Science
Bangalore, India
kprem@ece.iisc.ernet.in

Anurag Kumar
Department of Electrical
Communication Engineering
Indian Institute of Science
Bangalore, India
anurag@ece.iisc.ernet.in

Joy Kuri
Centre for Electronics Design
and Technology
Indian Institute of Science
Bangalore, India
kuri@cedt.iisc.ernet.in

ABSTRACT

We are concerned with the situation in which a wireless sensor network is deployed in a region, for the purpose of detecting an event occurring at a random time and at a random location. The sensor nodes periodically sample their environment (e.g., for acoustic energy), process the observations (in our case, using a CUSUM-based algorithm) and send a *local decision* (which is binary in nature) to the fusion centre. The fusion centre collects these local decisions and uses a *fusion rule* to process the sensors' local decisions and infer the state of nature, i.e., if an event has occurred or not. Our main contribution is in analyzing two local detection rules in combination with a simple fusion rule. The local detection algorithms are based on the nonparametric CUSUM procedure from sequential statistics. We also propose two ways to operate the local detectors after an alarm. These alternatives when combined in various ways yield several approaches. Our contribution is to provide analytical techniques to calculate false alarm measures, by the use of which the local detector thresholds can be set. Simulation results are provided to evaluate the accuracy of our analysis. As an illustration we provide a design example. We also use simulations to compare the detection delays incurred in these algorithms.

Keywords

event detection in wireless sensor networks, distributed detection of a change in distribution, nonparametric CUSUM technique

1. INTRODUCTION

Sensor networks are used to efficiently monitor the area they are deployed in, and to help in detecting events so that remedial or corrective measures can be taken. This paper focuses on the study of *sequential event detection* or *sequential change detection* in sensor networks. Since the cost of communication, in terms of battery consumption, is large as compared to computation [2], instead of sending raw measurements to be processed by a common *fusion centre*, we consider distributed schemes, where local decisions are made at the sensor nodes. These partial decisions, which may require as little as 1 bit to communicate them, are then fused at the fusion centre, or within the network.

In this paper, we consider the simplest case of a single event occurring at a random time and a random location in the area $\mathcal{A}$. The event remains in the same place for a sufficiently long time, so that we can confine our attention to the problem of detection and localization of the event. Our aim is to develop algorithms both at the sensor level as well as at the fusion centre so that we can detect intrusions as early as possible subject to a constraint on false alarms.

We consider sequential detection procedures in our work. Here, each sensor in the network receives a sequence of observations and sends a sequence of summary messages to the fusion centre where a sequential test is carried out to conclude whether the event occured or not. The design objective is to minimize the detection delay with a constraint on some measure of false alarm rate.

Related Literature: Event detection in sensor networks has been extensively studied, starting from the works of Tenny and Sandell [3]. Further work on event detection in decentralized setup was carried out by Tsitsilklis [4]. Ben-David et al. [6] discuss the use of nonparametric statistics for the event detection in sensor networks. Niu and Varshney [13] discuss the effect of random number of active sensors on the event detection procedures. The CUSUM approach was first discussed in statistics by Page [7]. Moustakides [8] discusses the CUSUM algorithm for a decentralized setup (which is used in sensor networks). Veeravalli [5] discusses an optimal decentralized sequential detection procedure. In [1], Prasanthi and Kumar consider the additional aspect that the measurements need to be transported over a network in which there are multiple access delays. They study the tradeoff between network delay and decision delay.

Our Contributions: Based on the nonparametric CUSUM algorithm, we propose two local detector algorithms. A simple fusion rule is proposed. Two variations are proposed for operating the detection process after an alarm. For three combinations of these approaches we provide analytical techniques for calculating false alarm measures, namely, mean time to false alarm, or the fraction of time in false alarm. Simulation results are provided to validate the analytical techniques that we develop.

2. THE SYSTEM MODEL

n sensors (indexed $i = 1, \ldots, n$) are deployed in an area of interest, $\mathcal{A}$, according to some deployment methodology, e.g., grid based deployment, uniform i.i.d. deployment, etc.

The primary aim of the deployed sensors is to detect events that occur in the area $\mathcal{A}$. Events are modeled as energy sources (e.g., acoustic or vibration energy) that yield a signal strength greater than or equal to $h_{0,min}$ at a reference distance d_0. In the region of deployment there could also exist *clutter* that could be falsely detected as an intruder (e.g., clutter could comprise small animals that live in the area). In our model, a single event occurs at an unknown time, at a random location in the area. We assume that the event stays in the place for a sufficiently long time, so that it appears to be stationary.

A discrete-time model is considered in this work. All the sensors sample their surroundings periodically and obtain a sequence of measurements $\left\{\mathbf{X}_k = \left(X_k^1, X_k^2, \ldots, X_k^n\right); k \geq 1\right\}$ (k denotes the sample index). Prior to the change, the samples are independent and identically distributed (i.i.d.) across the sensors and independently distributed across time, and have 0 mean. After the change, the samples are (conditionally) independent across time and space; the mean of the observations, however, depends on the position of the event.

The samples are processed by the sensors to give local decisions $\{D_k^i, 1 \leq i \leq n, k \geq 1\}$, which can be written as components of the local decision vector, $\{\mathbf{D}_k, k \geq 1\}$. The local decisions are forwarded to a fusion centre which maps the local decision into a global decision $\mathcal{D}_{glob,k} \in \{0,1\}$, with 0 meaning "no change until k" and 1 meaning "a change has occurred at or before k".

2.1 The Motivation For L Coverage

The energy from a source decays with distance homogeneously in every direction, so that a sensor placed at distance u from the source receives

$$h(u) = h_0 \ \rho(u)$$

with $\rho(d_0) = 1$. Also, $\rho(u)$ decreases with u. For example, one model could be

$$\rho(u) = e^{-\alpha_2(u-d_0)} \left(\frac{d_0}{u}\right)^{\alpha_1}$$

with $\alpha_1 > 1$ and $\alpha_2 > 0$.

Since we propose to use the nonparametric CUSUM at each sensor node (see Section 3), the means of the sensor observations become important. The sensor output could be biased in such a way that if the source energy received at the sensor drops below a threshold (say, h_{min}), then the sensor output has a negative mean. If an event of interest has source energy equal to $h_{0,min}$ (at d_0), then the measurements will have a positive mean up to the distance d_{min} given by: $h_{min} = h_{0,min} \ \rho(d_{min})$.

Hence, the effect of an event will persist (in the sense of positive mean of sensor output), up to a distance d_{min} from the source, whereas, clutter (whose radiated energy will be substantially less than that of the event of interest) will only have an effect up to a distance less than d_{min}.

Given d_{min}, we say that a point in $\mathcal{A}$ is L *covered*, if there are at least L sensors at a distance less than or equal to d_{min} from that point. Let the sensors be deployed in $\mathcal{A}$ such that each point is L covered. In the case of random deployment, all the points in $\mathcal{A}$ should be L covered with a high probability.

Thus, if an event occurs at some point in $\mathcal{A}$, at least L sensors will receive measurements with a positive mean. On the other hand, if there is noise or a weak energy source (i.e., clutter) at some point, then less than L sensors will make measurements with a positive mean. If the local detection algorithm at each node is such that the sensors trigger faster (by orders of magnitude) if the mean of the samples is positive, then, in the case of events, more than L sensors will trigger rapidly, else, less than L sensors will be triggered. This motivates L to be included as an important design parameter.

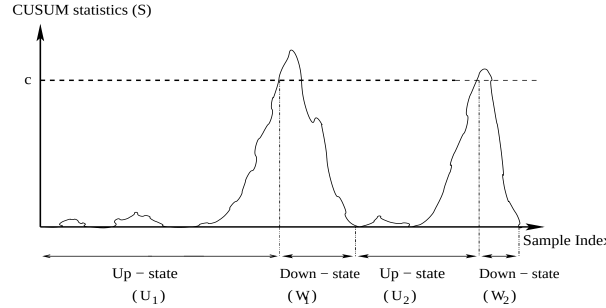

Figure 1: A sample path for the CUSUM statistic with *Local Detector* 1. Note the way the Up-state and the Down-state are defined.

2.2 A Simple Fusion Rule (SFR)

Based on the above idea of L coverage, a simple fusion rule at the fusion centre is to declare a change at k, i.e., $\mathcal{D}_{glob,k} = 1$, if

$$\sum_{i=1}^{n} I_{\{D_k^i = 1\}} \geq L \tag{1}$$

else, $\mathcal{D}_{glob,k} = 0$, and continue to make observations. There are clearly other fusion rules that are easily seen to be better than the simple fusion rule proposed above. An improvement is obtained if we include the sensor location information, or the local decision history. However, the analysis of such fusion rules is complex and in this paper we confine ourselves to the analysis of the simple rule.

3. THE LOCAL DECISION ALGORITHMS

We consider the following nonparametric CUSUM statistic at each node (the superscript indicating the node index has been dropped for notational simplicity)

$$S_k = (S_{k-1} + X_k - b)^+ \text{ with } S_0 = 0 \tag{2}$$

The local decision is a 1 at k if $S_k \geq c$. The local threshold (with which S_k is compared, at each k) is denoted by $c > 0$. The bias parameter b is set so as to obtain a negative mean for the term $(X_n - b)$, prior to the change, and a positive mean for that term after the change. Following the discussion in Section 2.1, if an event occurs at a point in $\mathcal{A}$ then for all nodes in a disk of radius d_{min} around that point, the expectation of $X_n - b$ will become positive, driving the CUSUM statistic at at least L nodes to increase without bound, and thus cross the local CUSUM threshold.

In the classical CUSUM procedure, on crossing the threshold, the test is stopped and a decision is made in favor of a change. However, in our case, the CUSUM algorithm is allowed to run freely, i.e., it is performed at every node in the region of interest $\mathcal{A}$ even after some nodes have crossed the threshold. This stems out of the need for distributed computation in case of sensor networks, where sensors collaborate in the decision process. Although one sensor may cross the threshold early, one needs to wait until one has "sufficient" information for declaring the change. In addition, a sensor may cross the threshold because of noise, in which case the network should continue making observations, while letting the falsely triggered CUSUM return to the "untriggered" state.

In the following sections we describe two approaches for using the CUSUM statistic to carry out the local detection.

3.1 Local Detector 1 (LD 1)

LD1 evolves as follows. $S_0 = 0$ and we say that the node is in the *Up-state*. Let τ_0^c denote the random time at which the CUSUM

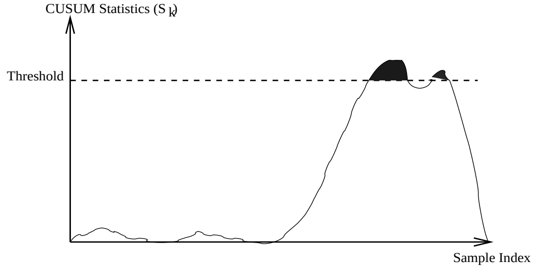

Figure 2: A sample path for the CUSUM statistics under *Local Detector* 2. The shaded region denotes those samples where the local decision is 1.

statistic crosses the threshold c, i.e., $S_{\tau_0^c} \geq c$ while $S_k < c$ for $0 \leq k < \tau_0^c$. We then set $S_{\tau_0^c} = c$ and let the statistic evolve. We say that the node is in the *Down-state*, starting from τ_0^c. If the expectation of $X_k - b$ is positive then with a high probability the statistic will continue to increase and stay above c. On the other hand, if the expectation of $X_k - b$ is negative then the statistic will return to 0 with probability 1. We call the Up-state intervals as *Up-times* and the Down-state intervals as *Down-times*.

For $k \leq \tau_0^c$, i.e., in the Up-state, the node sends the local decision $D_k = 0$, whereas in the Down-state the node sends $D_k = 1$.

Now, in order to analyse the false alarm performance of the system, we consider the situation in which the event never occurs, and, hence, $\mathsf{E}(X_k) = 0$ for all k. It is then clear that the end-points of the Down-times (equivalently, the start-points of the Up-times) are renewal instants. Also, since we reset S_k to c at the ends of Up-times, it is also clear that the alternating sequence of Up-times and Down-times constitute an alternating renewal process. Let the Up-times be denoted by $\{U_k, k \geq 1\}$ and the Down-times be denoted by $\{W_k, k \geq 1\}$ (refer to Figure 1).

3.2 Local Detector 2 (LD2)

In the second local detector, too, the CUSUM algorithm is implemented at each sensor with the statistics S_k being obtained at each sensor using Equation (2). However, in LD2 the CUSUM statistic is allowed to free-run, i.e., it is never reset. The local decision is 1 if S_k is above the threshold, irrespective of whether S_k had crossed the threshold earlier. This is shown in Figure 2. Thus the local decisions in LD2 are described as follows. For $k \geq 0$,

$$D_k = \begin{cases} 0 & S_k < c \\ 1 & S_k \geq c \end{cases} \tag{3}$$

4. FUSION OPTIONS AND FALSE ALARM MEASURES

Once the fusion centre declares a change, we consider the following two options for operating the system.

1. *Fusion Option 1 (FO1): Reset the system.*
   The CUSUM statistic is set to 0 at all the sensors and the count at the fusion centre is reset to 0. In this case, the false alarm measure is taken to be *the mean time to false alarm (TFA)*.

2. *Fusion Option 2 (FO2): Continue the process.*
   The CUSUM statistic is allowed to run without being reset. The false alarm measure is taken to be the fraction of time that the number of sensors with CUSUM statistic greater than 1 is at least L. This is the *fraction of time in false alarm (FROTIFA)*, i.e.,

$$\lim_{m\to\infty} \frac{1}{m} \sum_{k=0}^{m-1} I_{\{(\sum_{i=1}^{n} D_k^i) \geq L\}} \tag{4}$$

where $I\{\cdot\}$ denotes an indicator function.

Table 1: Table of mean Up-time values and the Down-time values under the null hypothesis when each sensor uses LD1. Here the bias $b = 0.75$. Here, Sim stands for the simulation results and NE stands for the numerical eveluation.

Threshold	ARL			
	Up(Sim)	Up(NE)	Down(Sim)	Down(NE)
1.0	18.67	19.03	2.41	2.40
1.25	28.26	28.18	2.72	2.71
1.5	40.78	41.82	3.04	3.03
1.75	62.48	61.92	3.37	3.37
2.0	96.64	91.12	3.72	3.70

5. ANALYSIS OF LOCAL DETECTOR 1

From Section 3, it is seen that the local detector can be viewed in terms of an alternating renewal process $\{(U_k, W_k), k \geq 1\}$ (the U_k's are i.i.d., W_k's are i.i.d., and the U_k's are independent of W_k's). If the distributions of U_k and W_k are approximately exponential, then the means $\mathsf{E}(U_k)$, $\mathsf{E}(W_k)$ are sufficient for the analysis of the Fusion Option 1. The fusion process can then be modeled as a continuous time Markov chain (CTMC).

For the LD1 and FO1 combination, we are interested in evaluating the mean time to false alarm (TFA). Hence, we consider the situation, in which the event occurs at infinity, i.e., the sensor observations are i.i.d. across the sensors and over time, with $\mathsf{E}(X_k) - b < 0$, for all k. The analysis of TFA begins with the study of the average run length (ARL) of the CUSUM algorithm at the node level, followed by relating this ARL to the TFA at the fusion centre.

We note here that the approximate exponentiality of the Up-times is related to the results in [10]. The exponentiality of the Down-time distribution is a heuristic. The results from the approximation match well with simulations.

5.1 Analysis of the Up-time

In terms of the classical CUSUM statistics, the average Up-time under the null hypothesis, i.e., $\mathsf{E}(U_k)$, is the ARL. Let the mean value of the time to cross the threshold with $S_0 = s(< c)$ be denoted by $L_\xi(s)$. The following integral equation is easily written down.

$$L_\xi(s) = 1 + L_\xi(0)\Phi(b-s) + \int_0^c L_\xi(v)\phi(v+b-s)dv \tag{5}$$

where $\Phi(\cdot)$ is the cdf of X_k, and $\phi(\cdot)$ is the pdf of X_k (under the null hypothesis). The ARL is obtained by solving for $L_\xi(s)$ and evaluating it at $s = 0$. Equation (5) is a Fredholm integral equation of the second type. Under the null hypothesis, let $\phi(\cdot)$ be the standard Gaussian for every i and k. With this we can solve the integral equation numerically. We chose the value of $b = 0.75$ and allowed the values of the threshold c to vary from 1.0 to 2.0 in the steps of 0.25. The results are tabulated in Table 1.

5.2 Analysis of Down-time

Let $L_\nu(s)$ be the mean time spent in the Down-state if the initial value of the statistic is $s(> 0)$. Using a renewal argument along

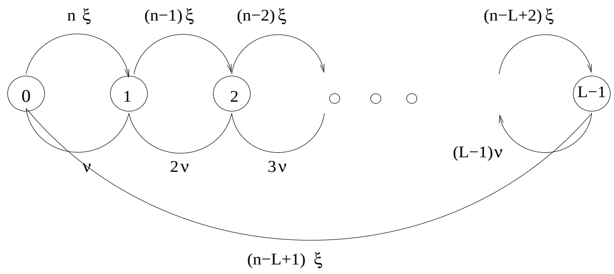

Figure 3: Markov chain for the fusion rule that count L ones in the area $\mathcal{A}$.

with Equation (2), and with a few algebraic manipulations, we obtain:

$$L_\nu(s) = 1 + \int_0^\infty L_\nu(v)\phi(v + b - s)\, dv \qquad (6)$$

This integral equation is used to evaluate the mean time spent in Down-time, i.e., $L_\nu(c)$ which equals $\mathsf{E}(W_k)$. For $\phi(\cdot)$ being the standard Gaussian density, the results are presented in Table 1.

6. LD1 WITH FUSION OPTION 1

In the previous section, we obtained the mean Up and Down times in LD1. With the exponential approximation for the distributions of the Up-times and Down-times, the combination LD1 FO1 can be modeled with a CTMC, the analysis of which yields the mean time to false alarm (TFA).

For notational convenience, define $\xi = \frac{1}{L_\xi(0)}$, $\nu = \frac{1}{L_\nu(0)}$ and T_{FA} as the mean time to false alarm. Let $N(t)$ denote the state of the fusion statistic at time t; i.e., $N(t)$ is incremented by 1 each time a sensor transitions from Up-state to Down-state, and is decremented by 1 is the transition is from Down-state to Up-state. Because of the exponentiality approximation, we can model $N(t)$ is the CTMC who transition diagram is shown in Figure 3. When $N(t) = L - 1$, and another sensor's local decision becomes 1, the fusion centre declares that a change has been detected and the system is reset.

Since we have the null hypothesis, each transition from $L-1$ to 0 is false alarm. Let $\pi(m), 0 \le m \le L-1$, denote the stationary probability distribution of the this CTMC. We observe that the rate of false alarms is $\pi_{L-1}(n-(L-1))\xi$. The instants when these false alarms occur are renewal instants. Hence, by the elementary renewal theorem, we conclude that

$$T_{FA} = \frac{1}{\pi_{L-1}(n-(L-1))\xi} \qquad (7)$$

The term π_{L-1} is calculated by solving the equation $\pi Q = 0$, where the matrix Q is the transition rate matrix.

Solving for π_{L-1}, we obtain

$$\pi_{L-1} = \frac{1}{\sum_{i=0}^{L-1}(\frac{n-(L-1)}{n-i})\gamma_i} \qquad (8)$$

where, $\gamma_{i+1} = 1 + \gamma_i \frac{(i+1)\nu}{(n-i)\xi}$ and $\gamma_0 = 1$.

Numerical Results: $n = 1000$ sensors were deployed in $\mathcal{A}$. L was chosen to be 40. At each sample instant, each sensor receives a sample of the standard Gaussian distribution. The CUSUM algorithm was implemented for two different values of $b = 0.75, 1$. Numerical results were The results are plotted in Figure 4.

Observations:

The analytical and simulation results match well, thus justifying the exponentiality approximation made above. Note that the y-axis

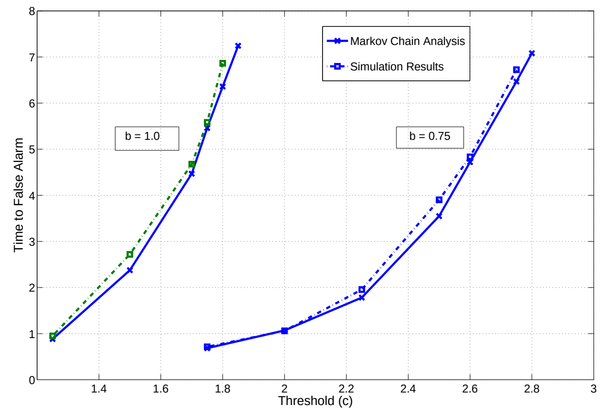

Figure 4: T_{FA} vs. the threshold c for Local Detector 1 at the sensors and with the Fusion Option 1 at the fusion centre, for two bias values, $b = 1.0, 0.75$. Note that the y-axis is $\log_{10}(T_{FA})$.

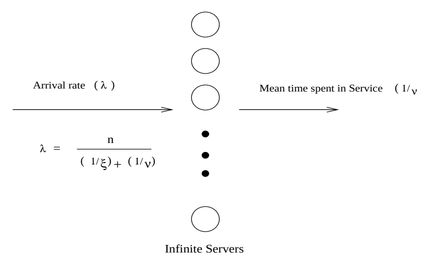

Figure 5: The $M/G/\infty$ model of the fusion count with LD1 and Fusion Option 2.

is in the log scale to the base 10. For $T_{FA} = 10^5$ samples, we can read off the bias and threshold pairs as: $(b = 0.75, c = 2.6)$ and $(b = 1.0, c = 1.75)$. If the sampling interval is 10 seconds, then $T_{FA} = 10^5$ corresponds to a average false alarm rate of 1 every 11 days. For a small change in the value of the threshold, c, there is a large change in the mean time to false alarm. This indicates that the design is quite sensitive to the value of c. In fact, the rate of change is larger for larger values of bias, b. Thus, the user has finer control over T_{FA} with smaller values of the bias. We also see that a small change in the value of the bias (from 0.75 to 1) led to a large change in the value of c (for the same value of TFA). This reflects that the design is sensitive to the bias value.

7. LD1 WITH FUSION OPTION 2

As seen in Section 3, with LD1, there are n independent alternating renewal process. The fusion centre maintains a fusion count $N(t)$ which is incremented at and Up to Down transition at a sensor, and decremented at a Down to Up transition. We are interested in the process $N(t)$. The Up to Down transitions at each sensor yields a point process. Each such transition causes the count $N(t)$ to increase. Since $L \ll n$ the aggregate rate of such transitions does not change much even if some sensors are already in the Down state. Further, we approximate the superposition of these point processes with a Poisson process. Combining these two approximations, we model the aggregate process of Up to Down transitions

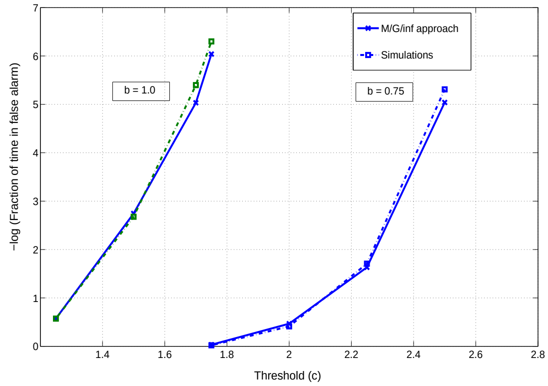

Figure 6: FROTIFA (on a log scale) vs. the threshold, c, for Local Detector 1 at the sensors with Fusion Option 2, for two bias values, $b = 1.0, 0.75$. Note that the y-axis is $-\log(FROTIFA)$.

by a Poisson process of rate

$$\lambda_Q \approx \frac{n}{\frac{1}{\xi} + \frac{1}{\nu}} \tag{9}$$

We then see that $N(t)$ is the same as the number of customers in an $M/G/\infty$ queue, with arrival rate λ_Q and holding time distribution the same at that of the Down time, i.e., $\frac{1}{\nu}$. This model can thus be represented as shown in the Figure 5.

We now seek the *fraction of time in false alarm* (FROTIFA). It is well known that the stationary distribution of $N(t)$ is Poisson with mean $\frac{\lambda_Q}{\nu}$. It then follows that FROTIFA is given by

$$\sum_{i=L}^{\infty} e^{-\frac{\lambda_Q}{\nu}} \frac{(\frac{\lambda_Q}{\nu})^i}{i!} \tag{10}$$

Numerical Results: Again we take $n = 1000$, and $L = 40$. Under the null hypothesis the samples are taken to be standard Gaussian. Each sensor runs the LD1 algorithm. The FROTIFA values were evaluated for the bias values $b = 0.75, 1$. The analysis and simulation results are plotted in Figure 6. We observe that the results are very similar to those for LD1 and Fusion Option 1, with T_{FA} as the false alarm measure.

8. ANALYSIS OF LOCAL DETECTOR 2

In this section we consider the use of LD2 in the sensor, along with Fusion Option 2. As explained earlier, the false alarm measure is FROTIFA. The theory of large deviations helps us relate the FROTIFA measure to the *fraction of time above threshold* (*FOTAT*) measure at each sensor. A target value of FROTIFA yields a target FOTAT. Once a target FOTAT is available, the various design parameters like bias, b, and the threshold, c, can be evaluated.

8.1 Analysis of LD2 to Obtain FOTAT

In this section, we assume that a bound, ϵ, on the FOTAT at a local detector is available to the user (the next section deals with the procedure to obtain ϵ). The problem then is to relate the bias, b, and threshold, c, at a sensor to the target FOTAT value. We use Chernoff's bound and the affine approximation given by Elwalid et al. [11] to carry out the analysis. Since we are concerned mainly with false alarm analysis, it suffices to analyse the CUSUM process at any one sensor. Hence, as before, we drop the sensor index in the notation.

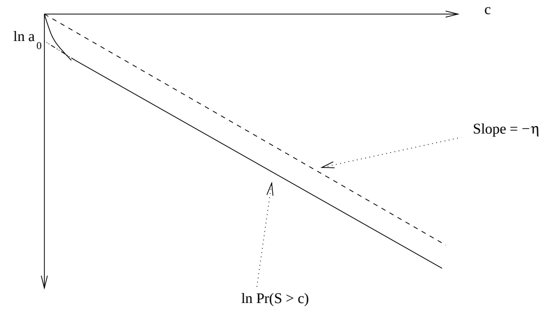

Figure 7: Typical plot of $\ln Pr(S > c)$ along with a linear bound and an affine approximation.

In LD2 the CUSUM process is not stopped but is allowed to freely evolve based on the measurements. Thus we are interested in analyzing the stochastic process $S_k, k \geq 0$, defined by Equation (2), where the $X_k, k \geq 1$, are i.i.d. with $\mathsf{E}(X_k) - b < 0$. By unraveling the CUSUM recursion we easily obtain an alternate expression for the process S_k.

$$S_n = \max_{0 \leq k \leq n} (\sum_{j=(n+1)-k}^{n} X_j - kb) \tag{11}$$

8.1.1 Using Chernoff's Bound

In this section, we use the Chernoff's bound to obtain an upper bound on the fraction of time during which $S_k \geq c$. Using Equation (11) and the union bound, along with Chernoff's bound, and observing that under the null hypothesis the $X_j, j \geq 1$, are i.i.d., for any $\theta > 0$, we have

$$Pr(S_k \geq c) \leq e^{-\theta c} \sum_{k=0}^{n} e^{k(\Gamma(\theta) - \theta b)} \tag{12}$$

where $\Gamma(\theta) = \ln \mathsf{E}\big(e^{\theta X_1}\big)$. If

$$\Gamma(\theta) < \theta b$$

then as the sample index k goes to infinity, it is seen that the summation in Equation (12) converges to a constant, say $\mathcal{K}$. Hence, for large values of k, we can write

$$Pr(S_k \geq c) \leq \mathcal{K} e^{-\theta c}$$

In the steady state, writing the marginal of the stationary version of the process as S, we get the following.

$$Pr(S \geq c) \leq \mathcal{K} e^{-\theta c}$$

Evidently, to obtain the best bound, we can take $\theta = \eta(b)$ such that $\frac{\Gamma(\eta(b))}{\eta(b)} = b$, thus yielding the following bound on FOTAT at any sensor, under the null hypothesis.

$$Pr(S \geq c) \leq \mathcal{K} e^{-\eta(b)c} \tag{13}$$

Now one approximation is to take $\mathcal{K} \approx 1$. Then there exists a linear relationship between $\ln(Pr(S > c))$ and the threshold c. This is plotted as the dashed line in Figure 7.

As before, we let the samples come from a standard Gaussian distribution. Then, $\Gamma(\theta) = \frac{1}{2}\theta^2\sigma^2$. Using this expression and the definition of $\eta(b)$, we obtain $\eta(b) = \frac{2b}{\sigma^2}$. Setting the bound in Equation (13) equal to ϵ, we obtain $\eta(b) = -\frac{\ln \epsilon}{c}$, yielding the following relation between b, c and the the FOTAT bound, ϵ,

$$bc = -\frac{\sigma^2}{2} \ln \epsilon$$

Table 2: Results for Local Detector 2 showing FOTAT obtained for various values of c with $b = 0.75$.

Threshold	Simulations	Chernoff Bound	Affine (Elwalid)
c	ϵ_{sim}	ϵ_{Cher}	$\epsilon_{Aff_{Elw}}$
1.8	0.0285	0.0672	0.0270
1.9	0.0246	0.0578	0.0232
2.0	0.0213	0.0498	0.0200
2.1	0.0181	0.0429	0.0172
2.2	0.0158	0.0369	0.0148
2.3	0.0135	0.0317	0.0127
2.4	0.0115	0.0273	0.0109
2.5	0.0100	0.0235	0.0094

8.1.2 Using an Affine Approximation

As seen from Figure 7 the linear approximation could be quite loose. A more accurate model can be obtained from the affine approximation due to Elwalid et al. [11]. The approximation is

$$Pr(S > c) \approx a_0 e^{-\eta(b)c} \tag{14}$$

where a_0 is approximated as the probability that a single sample exceeds the bias. If the samples come from a standard normal distribution, we have

$$a_0 = Pr(X > b) \approx \frac{\sigma}{b\sqrt{2\pi}} e^{-\frac{b^2}{2\sigma^2}} \tag{15}$$

8.1.3 Simulation Setup and Results

We consider a single sensor receiving a sequence of i.i.d. samples that have the standard Gaussian distribution. The simulations were allowed to run for a long time (typically 10^5 samples or more), as compared to the reciprocal of ϵ values. The results are presented in Table 2.

8.1.4 Observations

It is seen from the table that the results obtained using Chernoff's bound fare poorly and that the affine approximation is quite close to the simulation results. Hence, we make use of only the affine approximation in our work.

8.2 Analysis of the Fraction of Time in False Alarm (FROTIFA)

We now show how obtain the overall FROTIFA for the Fusion Option 2, with LD2 at each sensor. We make use of results from the theory of large deviations. We begin by using Cramer's theorem, following it with the Bahadur-Rao approximation. We conclude this section with simulation results to compare these two approaches.

8.2.1 Using Cramer's Theorem

The sensor level decision processes D_k^i are independent 0-1 for each of which the fraction of time spent in State 1 being ϵ, the FOTAT. Let $D^i, 1 \leq i \leq n$, denote the steady state marginal random variables for the n sensors; these are i.i.d. Bernoulli, with probability of state 1 being ϵ. Since we use the Fusion Option 2, sensor processes are not reset on a global alarm, and the false alarm measure is FROTIFA, which is given by $Pr(\sum_{i=1}^{n} D \geq an)$, where we have defined $a = \frac{L}{n}$.
Define the quantities :

$$M(\theta) = \mathsf{E}\left(e^{\theta D^1}\right) \qquad l(a) = \sup_{\theta}(\theta a - \log M(\theta))$$

Using Cramer's theorem, we can easily show that:

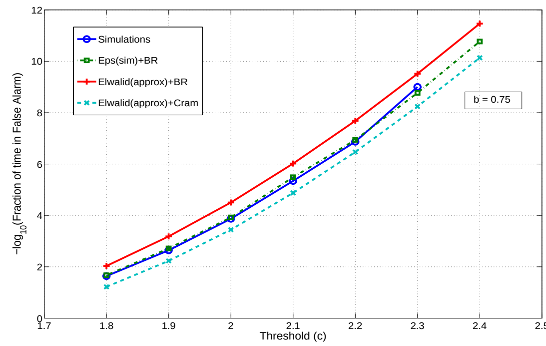

Figure 8: Results for Local Detector 2 with Fusion Option 2. ($-\log_{10}$ FROTIFA) is plotted for different values of the threshold c; $b = 0.75$. Results from simulations and various approximations are shown.

$$Pr(\sum_{i=1}^{n} D^i \geq an) \leq e^{-nl(a)} \tag{16}$$

For Bernoulli random variables, the following expression is well known.

$$l(a) = a\log(\frac{a}{\epsilon}) + (1-a)\log(\frac{1-a}{1-\epsilon})$$

8.2.2 Using the Bahadur-Rao approximation

We will see that conservative results are obtained due to the application of Cramer's theorem. In this section, we investigate a better result, namely, the Bahadur-Rao approximation [12], that can be used instead of the Chernoff's bound. The Bahadur-Rao approximation is given as follows:

$$Pr(\sum_{i=0}^{n} D^i > na) \approx \frac{1}{\theta(a)\sqrt{2\pi n}\sqrt{M''(\theta(a))}} e^{-nl(a)} \tag{17}$$

where $\theta(a)$ achieves the supremum in $\sup_{\theta}(\theta a - \ln M(\theta))$.

8.2.3 Simulation Results

In our simulations, we choose the following parameters: $n = 1000$, $L = 40$, $b = 0.75$. The samples were i.i.d. with standard Gaussian distribution. The simulations were run for a long duration (on the order of 10^9 samples) and compared with the analytical expressions. See Figure 8.

8.2.4 Observations

We observe that the FROTIFA is quite sensitive to the values of threshold. Note the log scale on the y-axis. We observe that when the value of ϵ is available (through simulation) and the Bahadur-Rao approximation is used to analyse the fusion rule, the resultant values of FROTIFA match with those obtained entirely through simulations (middle two curves). This shows that the Bahadur-Rao approximation is excellent for analyzing the process at the fusion centre.

However, if we take the FOTAT (ϵ) from the affine approximations used at the fusion centre, the design turns out to be slightly optimistic (the top curve). This can be attributed to the optimistic design at the sensor level (see the last column in Table 2, where where the epsilon values are smaller than those actually obtained (first column)).

Table 3: Design using LD1 and FO1, for user defined TFA.

TFA_{User}	LD1 with FO1	
Samples	Threshold (c)	Design Result (TFA)
10^4	2.54	2×10^4
5×10^4	2.6	6.3×10^4
10^5	2.63	2×10^5
5×10^5	2.67	6.34×10^5
10^6	2.7	1.48×10^6

9. A DESIGN EXAMPLE

The user defines the objective in terms of the mean time to false alarm, which encapsulates the user's ability to pay a price for false alarms. However, if a very stringent false alarm objective is specified then there could be a large detection delay.

9.1 Converting the User-defined TFA constraint to the FROTIFA constraint

The user defined constraint for false alarms will typically be given in the form of the mean time to false alarm (TFA). However, for analyzing the Fusion Option 2 (Section 4)), it is convenient to pose the false alarm constraint in terms of FROTIFA. If, under the null hypothesis, the mean time spent in false alarm is much smaller than the mean time to false alarm (a reasonable assumption in practice), then the following approximation is obtained

$$FROTIFA \approx \frac{1}{TFA}$$

We use this approximation in our design example.

9.2 A Design Example

The values of n, L and the bias b are related through the event model, the model for propagation of the sensing modality and the sensor response model. We assume these as obtained by other means and provide the design of the CUSUM threshold at the sensors. The analysis procedures discussed in Sections 6, 7 and 8 are used to obtain the value of the threshold c. The values of $n = 1000$, $L = 40$ and $b = 0.75$ were chosen for this example.

9.2.1 *Using LD1 and Fusion Option 1*

The mean time to false alarm is obtained for the algorithm that uses LD1 and Fusion Option 1 using Equation (7). The parameters b, c are implicit in Equation (7) (through π_{L-1} and ξ). However, evaluating the threshold directly from the above equation is not possible as the integral equations involved do not have closed form solutions. Hence, we need to resort to charts (similar to Figure 4). We present the results of the design in Table 3.

9.3 Using the LD1 and Fusion Option 2

This approach too makes use of the parameters ξ and ν for evaluating the false alarm constraint. Hence, we can evaluate the parameter pair b, c using charts like Figure 6. We present the results of the design in Table 4.

9.4 Using the LD2 and Fusion Option 2

We make use of the affine approximation at the sensor level and the Bahadur-Rao approximation at the fusion centre. The affine approximation was used to obtain the sensor level time above the threshold ϵ.

Using the user-defined T_{FA}, we can obtain the parameter ϵ from the Bahadur-Rao approach. For our case, the ϵ and the threshold are related using Eqn(15). Thus, we can directly obtain the value of the threshold. We present the results of the design in Table 5.

Table 4: Design using LD1 and FO2, for user defined TFA.

TFA_{User}	LD1 with FO2		
Samples	FROTIFA-aimed	Threshold (c)	Design(TFA)
10^4	10^{-4}	2.42	2×10^4
5×10^4	2×10^{-5}	2.47	1.1×10^5
10^5	10^{-5}	2.5	2.2×10^5
5×10^5	2×10^{-6}	2.54	1.2×10^6
10^6	10^{-6}	2.58	2.51×10^6

Table 5: Design using LD2 and FO2, for user defined TFA.

TFA_{User}	LD2 with FO2		
Samples	$FROTIFA_{target}$	Threshold	Design TFA
10^4	10^{-4}	1.97	4×10^3
5×10^4	2×10^{-4}	2.01	1.4×10^4
10^5	10^{-5}	2.04	4×10^4
5×10^5	2×10^{-6}	2.07	1.2×10^5
10^6	10^{-6}	2.1	3×10^5

9.5 Discussion

From Tables 3 and 4 we conclude that the analytical approaches we have developed for LD1+FO1 and for LD1+FO2 result in a slightly conservative design. The TFA actually obtained is slightly larger than the user specification. On the other hand from Table 5 we conclude that the affine approximation and Bahadur-Rao approximation based approach for analyzing LD2+FO2 yields a design that provides a smaller TFA than the user requirement. In each case it is seen that the TFA is very sensitive to the threshold c, as observed earlier in the paper.

10. DETECTION DELAY: SIMULATION RESULTS

We have only the simulation results for the detection delays. We deploy $n = 1000$ sensors in the area of interest in a grid fashion. The event occurs at a random place (chosen using the uniform distribution), and at random time (chosen geometrically with the parameter $p = 0.0005$). The coverage parameter was chosen to be $L = 40$. For the post-change distributions, the values of μ_i's were calculated as: $\mu_i = \frac{b * d_{min}}{d_i}$ where, $d_{min} = 0.1$ Km and d_i is the distance of the i^{th} sensor from the event. We chose the values of $b = 0.75$ or 1 for our simulations. The results are presented in Figure 6.

It is seen that the mean detection delay is in the order of a few samples (0.5 to 2) for the values of threshold that gave the mean time to false alarm of the order of $10^5 - 10^7$ samples.

11. CONCLUSION

We formulated an event detection problem with a user-defined objective of mean time to false alarm, TFA. For this problem, we

Table 6: Detection delay. Here, thr stands for threshold, DD stands for Detect Delay, and LFi stands for local decision i with fusion rule i

Bias	Thr(LF1)	DD(LF1)	Thresh (LF2)	DD(LF2)
0.75	2.6	1.13	1.7	1.02
0.75	2.7	1.17	1.75	1.2
0.75	2.75	1.45	1.8	1.39
0.75	2.8	1.54	1.85	1.65
0.75	2.9	1.8	1.9	1.66

explored three different approaches based on CUSUM based local detectors at each sensor, and fusion of these local decisions at a fusion centre. Our main contribution was to develop analytical techniques for setting the CUSUM threshold for achieving the target TFA. Simulation results were presented to demonstrate how our techniques work in a design example.

Our future work includes the study of an optimal parametric distributed detection scheme within the same framework. Also, we will study ways to design the other parameters, i.e., n, L, and b.

12. ACKNOWLEDGEMENTS

This work was supported by a project grant from DRDO 571, IISc.

13. REFERENCES

[1] V. K. Prasanthi and Anurag Kumar, “Optimizing delay in sequential change detection on an adhoc wireless sensor network," *Proc. of IEEE SECON, 2006*.

[2] G. J. Pottie and W. J. Kaiser, “Embedding the internet: Wireless integrated network sensors," *Communications of ACM Vol. 43, No. 5, 51–58, 2000*.

[3] R. Tenney and N. Sandell, “Detction with distributed sensors," *IEEE transactions on Aerospace and Electronics systems Vol. 17, No. 4, 501–510, 1981*.

[4] J. N. Tsitsiklis, “Decentralized detection," *Advances in statistical signal processing Greenwich CT JAI 297–344, 1993*.

[5] V. V. Veeravalli, “Sequential decision fusion: theory and application," *Journal of the Franklin institute 301–322, 1999*.

[6] Shai Ben-David, Ting He and Lang Tong, “Nonparametric approaches to change detectionand estimation in large scale sensor networks," *CISS, 2004*.

[7] E. S. Page, “ Continuous Inspection Schemes," *Biometrika Vol. 41, 100–115, 1954*.

[8] G. Moustakides, “Optimal Procedures for Detecting Changes in Distribution," *Annals Statistics Vol. 14, 1379–1387, 1986*.

[9] M. Basseville and I. Nikiforov, “Detection of abrupt changes:theory and application," *Prentice hall, 1993*.

[10] M. Pollack and A. G. Tartakovsky, “On Asymptotic Exponentiality of the Distribution of First Exit Times for a Class of Markov Processes," *The Annals of Applied Probability, 2006*.

[11] A. Elwalid, D. Heyman, T. V. Lakshman, D. Mitra and A. Weiss, “Fundamental Bounds and Approximations for ATM Multiplexers with Applications to Video Conferencing," *IEEE Journal on Selected Areas in Communications Vol. 13, No. 6, 1004-1016, 1995*.

[12] R. R. Bahadur and R. R. Rao, “On Deviations of Sample Mean," *The Annals of Mathematical Statistics Vol. 31, No. 4, 1015-1027, 1960*.

[13] R. Niu and P. K. Varshney, “Distributed Detection and Fusion in a Large Wireless Sensor Network of Random Size," *EURASIP Journal on Wireless Communication and Networking Vol. 4, 462-472, 2005*