Forty-Seventh Annual Allerton Conference
Allerton House, UIUC, Illinois, USA
September 30 - October 2, 2009

Distributed Detection and Localization of Events in Large Ad Hoc Wireless Sensor Networks

K. Premkumar, Anurag Kumar
Department of Electrical Communication Engineering
Indian Institute of Science, Bangalore, 560 012, INDIA
Email: {kprem,anurag}@ece.iisc.ernet.in

Joy Kuri
Center for Electronics Design and Technology
Indian Institute of Science, Bangalore, 560 012, INDIA
Email: kuri@cedt.iisc.ernet.in

***Abstract*—We consider the problem of event detection in wireless sensor networks (WSNs) that are large in the sense that an event affects the statistics of the observations of a small number of sensors in the vicinity of where it occurs. An event occurs at a random time at a random location in the region (called the region of interest, ROI) covered by the WSN. We consider a *distance based* sensing model in which the physical signal sensed by a sensor at a distance d from the event is attenuated by a factor $\rho(d)$. We formulate the problem of detecting an event as early as possible and locating it to a subregion of the ROI under the constraints that the average time to false alarm (TFA) and the average time to false isolation (TFI) are bounded by γ. This formulation is motivated by the change detection/isolation framework introduced by Nikiforov [6].**

We extend the decentralized detection procedures MAX [11] and ALL [9], [5], which are designed for colocated networks, to the case of a large WSN where the event localization is also a critical issue. The extended MAX and ALL detect the change and identify a subregion where the event is located. Sensor noise can make the local decisions of ALL toggle rapidly. Motivated by this fact, we propose a distributed change detection/isolation procedure, HALL (Hysteresis modified ALL). We study the supremum average detection delay (SADD) performance of the change detection/isolation procedures MAX, ALL and HALL for a required $\min\{\mathsf{TFA}, \mathsf{TFI}\} \geqslant \gamma$. We show that as $\gamma \to \infty$, the (asymptotic) $\mathsf{SADD(ALL)} \leqslant \frac{\ln \gamma}{\underline{\omega_0}\widetilde{MI}}$, $\mathsf{SADD(HALL)} \leqslant \frac{\ln(\gamma+1)}{\underline{\omega_0}(1-1/\beta)\widetilde{MI}} + C$, and $\mathsf{SADD(MAX)} \leqslant \frac{\ln \gamma}{\underline{\omega_0} I}$, where $\underline{\omega_0}$, C, β and $\widetilde{M}$ are constants that depend on the sensor deployment, the postchange and the prechange distributions of sensor measurements, and I is the Kullback–Leibler divergence between a worst–case postchange distribution and the prechange distribution of sensor measurements. We also compare the SADD of the distributed procedures with that of the asymptotically optimal centralized procedure given by Nikiforov [6] for a Boolean sensing model. We show that the SADD performance of ALL and HALL is of the same order as that of Nikiforov's. We also provide numerical comparison of SADD and TFA for the centralized asymptotically optimal scheme [6], and the distributed schemes MAX, ALL and HALL.

I. Introduction

Sensor networks are application–specific networks that comprise a large number of tiny, energy–limited, smart sensor devices called nodes or *motes* [1]. The motes make measurements, perform computations and communicate information (e.g., an alarm) to a *fusion center*. In the centralized detection strategy, all nodes send their measurements to the fusion center, which is wasteful of resources (as the events are typically rare), and hence reduces the lifetime and the utility of the network. Thus, it is essential to have energy efficient *distributed* algorithms.

In this paper, we are concerned with the scenario in which a number of motes are deployed in a *large region* (Fig. 1(a)) to detect and locate events. An event occurs at a random time and at a random location in the region. In this paper, we only consider the case of a single stationary event. *We wish to study sequential algorithms for detecting and locating such an event* (also called change detection and isolation [6]) with small detection delay, subject to a constraint on false alarm and false isolation. For the reasons discussed above, *we seek algorithms that are distributed* and require only local information from the neighborhood of each node.

Related literature:

The problem of *change detection and isolation* is introduced by Nikiforov in [6], and by Malladi and Speyer in [4]. Nikiforov [6] extended the classical minimax optimal change detection framework of Lorden to multiple postchange hypotheses and proposed an asymptotically minimax delay optimal procedure which is centralized, parametric and non–recursive. Malladi and Speyer [4] consider a centralized Bayesian change detection/isolation problem and show that a threshold based test is optimal.

In the context of colocated networks (where the event influences all the nodes in the network), considerable work has been done in the decentralized setting. Niu and Varshney [7] studied a simple (non sequential) hypothesis testing problem and proposed a *counting rule* based on the number of alarms. They showed that for sufficiently large number of sensors, the detection performance of the counting rule is close to that of the optimal rule. In the quickest decentralized change detection problem, considerable work has been done by Tartakovsky and Veeravalli [11], Tartakovsky and Kim [9] and by Mei [5]. Tartakovsky and Veeravalli [11] proposed the decentralized detection procedure MAX for a colocated network. In MAX, CUSUM is employed as the local decision rule and a global alarm is raised as soon as all the sensor nodes are in the alarm state. In MAX, once a sensor enters into an alarm state, it stays in the alarm state until a global alarm is raised. Tartakovsky and Kim [9], and Mei [5] proposed a decentralized procedure ALL (again for a colocated network) based on the CUSUM [2] statistic that continues to "run" even after crossing the threshold and a global alarm is raised as soon as the CUSUM

statistic of all the sensor nodes are above a threshold. It has been shown ([9], [5]) that, asymptotically, the minimax SADD performance of ALL is the same as that of the centralized CUSUM procedure [3]. In [10] Tartakovsky and Polunchenko studied the quickest change detection problem in a colocated network in the presence of some unknown parameters, where the unknown parameters take values in some parameter space. Tartakovsky and Polunchenko chooses some reference points in the parameter space and run CUSUM for each of the chosen reference point. The global alarm is raised as soon as one of the CUSUMs crosses a threshold. This procedure is shown to be asymptotically optimal only when the unknown parameters match with one of the reference points.

Our work differs from [11], [9], [5], and [7] in the sense that these works consider detection in a colocated network (where there is single postchange hypothesis) whereas we consider a combined change detection/isolation problem in a large WSN. Also, we propose a distributed algorithm whereas [6] and [4] consider a centralized change detection/isolation problem. Note that the computational complexity of Nikiforov's procedure is prohibitively high. We study distributed procedures based on local CUSUM detectors at the nodes. Also note that centralized procedures ([6], [4]) involve high communication costs. In [10], the authors consider only the change detection problem and the problem of *isolating the correct parameter* is not considered.

Summary of our contributions:

1) To the best of our knowledge, our work is the first on distributed sequential event detection/isolation in a large WSN where the event influences the observations of a few sensor nodes in the vicinity of where it occurs.
2) We consider a practical sensing model which incorporates distance losses, and we provide the resulting performance of the distributed change detection/isolation procedures.
3) We identify subregions in the region of interest ROI, each subregion being influenced by a unique set of sensors. We formulate the quickest detection of an event and locating it in a subregion under the average time to false alarm (TFA) and the average time to false isolation (TFI) requirements. This formulation is motivated by Nikiforov [6].
4) We extend the decentralized procedures MAX ([11]) and ALL ([5],[9]) that are developed for colocated networks to large WSNs. We also propose HALL (**H**ysteresis modified **ALL**), a distributed detection/isolation procedure. Note that the distributed procedures MAX, ALL and HALL are computationally less complex and energy–efficient (the communication cost is very less in distributed procedures compared to centralized procedures).
5) We analyze the supremum average detection delay (SADD) of MAX, ALL and HALL for a required $\min\{\mathsf{TFA}, \mathsf{TFI}\} \geqslant \gamma$. Also, we use Nikiforov's asymptotically optimum centralized procedure ([6]) as a benchmark and show that as $\gamma \to \infty$, the $\mathsf{SADD(ALL)} \leqslant \frac{\ln \gamma}{\underline{\omega_0}\widetilde{M}I}$, $\mathsf{SADD(HALL)} \leqslant \frac{\ln \gamma}{\underline{\omega_0}(1-1/\beta)\widetilde{M}I} + C$, and $\mathsf{SADD(MAX)} \leqslant \frac{\ln \gamma}{\underline{\omega_0}I}$, where $\underline{\omega_0}$, C, β and $\widetilde{M}$ are constants that depend on the sensor deployment, the postchange and the prechange distributions of sensor measurements, and I is the Kullback–Leibler divergence between a worst–case postchange and the prechange distributions of sensor measurements. Our numerical results show that for a given TFA requirement, the SADD of Nikiforov's centralized algorithm is the least and that of the MAX is the highest. Also, we show that $\mathsf{SADD(ALL)} \approx \mathsf{SADD(HALL)}$ and is little larger than $\mathsf{SADD(Nikiforov)}$.

Outline of the paper: The rest of the paper is organized as follows. In Section II, we provide the model for sensing, for the sensor measurements under various hypotheses and for the event. We formulate the problem in Section III. In Section IV, we describe MAX, ALL and HALL and discuss the asymptotic minimax delay optimality of these procedures. In Section V, we provide numerical results and we conclude in Section VI.

II. System Model

A. Sensing Model:

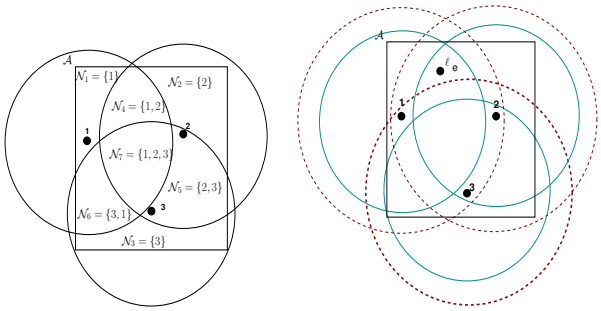
Let $\mathcal{A}$ be the region of interest (ROI) in which a WSN has to be engineered for event detection. n sensor nodes are deployed in the region $\mathcal{A}$. Let $\ell^{(i)} \in \mathcal{A}$ be the location of sensor node i and define $\boldsymbol{\ell} = [\ell^{(1)}, \ell^{(2)}, \cdots, \ell^{(n)}]$. An event is a source of some physical signal that is sensed at a sensor node; for example, the source signal could be acoustic, infrared or magnetic. In this work, we assume that all nodes have only one type of sensor. Let h_e be the signal strength of the event[1]. A sensor at a distance d from the event senses a signal $h_e\rho(d) + W$, where W is random (sensing) noise and $\rho(d)$ is a loss function which is, in general, a decreasing function of the distance d. Note that the mean signal received by all sensors at a distance d (from the event) is the same.

B. Detection Region and Detection Partition:

We define the *detection range* of a sensor as the distance r_s within which the mean signal level from an event is at least μ_1, a quantity to be defined later. We say that a location x in the ROI is *detection–covered* by sensor i, if x lies within the detection range of sensor i, i.e., if $\|\ell^{(i)} - x\| \leqslant r_s$. Let $\mathcal{S}^{(i)} \subseteq \mathcal{A}$ be the detection–coverage region of sensor i, i.e., any point $x \in \mathcal{S}^{(i)}$ is detection–covered by sensor i. Thus, the detection–coverage region of sensor node i is given by $\mathcal{S}^{(i)} = \mathcal{D}(\ell^{(i)}, r_s) := \{x \in \mathcal{A} : \|\ell^{(i)} - x\| \leqslant r_s\}$.

We assume that the sensor deployment is such that every $x \in \mathcal{A}$ is detection–covered by at least one sensor. For each $x \in \mathcal{A}$, define $\mathcal{N}(x)$ to be the largest set of sensors by which x is detection–covered, i.e., $\mathcal{N}(x) := \{i : x \in S^{(i)}\}$. Let $\mathcal{C}(\mathcal{N})$

[1]The signal strength of the event in general lies in an interval $[\underline{h}, \overline{h}]$ and the exact value of h_e is usually unknown. In this case, we can work with $h_e = \underline{h}$ as this corresponds to the least KL divergence between the "*event not occurred*" hypothesis and "*event occurred*" hypothesis. See [10] for change detection with unknown parameters for a colocated network.



(a) Partitioning of $\mathcal{A}$ in a large WSN by Detection regions

(b) Influence and Detection regions

Fig. 1. A simple example of partitioning of $\mathcal{A}$ in a large WSN. The three sensor nodes, in the figure, divide the ROI into regions $\mathcal{A}_1, \cdots, \mathcal{A}_7$ such that region $\mathcal{A}_i$ is detection–covered by a unique set of sensors $\mathcal{N}_i$. In Fig. 1(b), the dotted circles represent the influence region and the thin smaller circles represent the detection region of sensors. Note the location of the event, $\ell_e \in \mathcal{A}(\{1,2\})$.

be the collection of all such sensor–sets, i.e., $\mathcal{C}(\mathcal{N}) = \{\mathcal{N}(x) : x \in \mathsf{ROI}\}$. Note that $\mathcal{C}(\mathcal{N})$ is a finite set and it can have at most $2^n - 1$ elements. Let N be the number of elements of $\mathcal{C}(\mathcal{N})$. For each $\mathcal{N}_i \in \mathcal{C}(\mathcal{N})$, we denote the detection–covered region by $\mathcal{A}_i = \mathcal{A}(\mathcal{N}_i) := \{x \in \mathsf{ROI} : \mathcal{N}(x) = \mathcal{N}_i\}$. Hence, the ROI is *detection–partitioned* into a *minimal number of subregions*, $\mathcal{A}_1, \mathcal{A}_2, \cdots, \mathcal{A}_N$, such that the subregion $\mathcal{A}_i$ is detection–covered by (each sensor in) a unique set of sensors $\mathcal{N}_i$, and $\mathcal{A}_i$ is the maximal detection–covered region of $\mathcal{N}_i$, i.e., $\forall i \neq i'$, $\mathcal{N}_i \neq \mathcal{N}_{i'}$ and $\mathcal{A}_i \cap \mathcal{A}_{i'} = \emptyset$. See Fig. 1(a) for a three sensor example.

C. Event Model:

An event (or change) occurs at a random time T and at an unknown location $\ell_e \in \mathcal{A}$. In this work, we consider only stationary point events, i.e., an event occurs at a point and, having occurred, stays there forever. Examples that would motivate such a model are 1) gas leakage in the wall of a large storage tank, 2) excessive strain at a point in a large flat structure.

D. Measurement Model:

We consider a *discrete time system* and the basic unit of time is one slot. The sensor nodes are assumed to be synchronised (see, for example, [8]). At the beginning of every slot $k \geqslant 1$, each sensor i samples its environment and obtains the observation $X_k^{(i)} \in \mathbb{R}$. Under the hypothesis "*event not occurred*," the observation $X_k^{(i)}$ is just the sensor noise $W_k^{(i)}$. Under the hypothesis "*event occurred*," the observation of sensor i is given by $h_e\rho(d_i) + W_k^{(i)}$ where d_i is the distance of sensor i from the event. The noise processes $\{W_k^{(i)}\}$ are identically distributed and assumed to be independent (i.i.d.) across time and across sensor nodes, the probability density function (pdf) of which is $f_0(\cdot)$. The pdf of $h_e\rho(d) + W_k^{(i)}$ is denoted by $f_1(\cdot; d)$.

E. Influence Region:

In this paper, we propose distributed procedures for change detection/isolation based on CUSUM [2], [3] at *each node*. For each node i, at each time slot k, the driving term in the CUSUM procedure should be the log likelihood–ratio (LLR) of the observation $X_k^{(i)}$ defined as $Z_k^{(i)}(d_i) := \ln \frac{f_1(X_k^{(i)}; d_i)}{f_0(X_k^{(i)})}$ where $d_i := \|\ell_e - \ell^{(i)}\|$. As the location of the event ℓ_e is unknown, the distance of each sensor i from the event, d_i is also unknown. Hence, one can not work with the pdfs $f_1(\cdot; d_i)$. Recall the detection range r_s introduced in Section II-B. We propose to drive the CUSUM at each node with $Z_k^{(i)}(r_s) := \ln \frac{f_1(X_k^{(i)}; r_s)}{f_0(X_k^{(i)})}$. Note that for $d < r_s$, the SADD of CUSUM with $Z(r_s)$ as the driving term is larger than that with $Z(d)$ as the driving term. Thus, having the postchange pdf to be $f_1(\cdot; r_s)$ corresponds to a worst–case assumption on the location of the event (see Section IV-A for a discussion on SADD). We now discuss results concerning the choice of r_s.

A sensor can potentially pick up a signal from an event located at a distance larger than r_s. Thus, we are interested in studying the driving term of the CUSUM at each node for various distances that the event can be from the node.

Lemma 1: For Gaussian noise (i.e., f_0 is Gaussian with mean μ_0 and variance σ^2), for some $r_s > 0$, let $f_1(\cdot; r_s)$ be taken as the postchange pdf in the CUSUM algorithm being used at a sensor. When the distance–loss function is $\rho(d)$, the mean of the driving term in CUSUM is *negative* if the event occurs beyond a distance of $\bar{r} := \min\{d : 2\rho(d) \leqslant \rho(r_s)\}$ from the sensor.

Proof: Let $Z(d)$ be the LLR between pdfs $f_1(\cdot; d)$ and f_0, and let $Z(r_s)$ be the LLR between pdfs $f_1(\cdot; r_s)$ and f_0. It is easy to see for Gaussian f_0 (and hence for Gaussian $f_1(\cdot;\cdot)$) that $\mathbb{E}_{f_1(\cdot;d)}[Z(r_s)] = \frac{(h_e\rho(r_s))^2}{2\sigma^2}\left(\frac{2\rho(d)}{\rho(r_s)} - 1\right)$. Thus, $\mathbb{E}_{f_1(\cdot;d)}[Z(r_s)]$, the mean of the increment that drives CUSUM decreases with d and hits 0 at $\bar{r}$ given by $\bar{r} := \min\{d' : 2\rho(d') \leqslant \rho(r_s)\}$. Thus, $\mathbb{E}_{f_1(\cdot;d)}[Z(r_s)]$ is negative when an event occurs at a distance $d > \bar{r}$. ■

We call $\bar{r}$ as the *sensing–range* of a sensor.

Examples: The Boolean model is defined as

$$\rho(d) = \begin{cases} 1 & \text{if } d \leqslant r_s, \\ 0 & \text{otherwise} \end{cases}$$

In this case, it is clear that the sensing range is just the detection range and hence $\bar{r} = r_s$. *The path–loss model* is given by

$$\rho(d) = d^{-\eta}$$

for some path loss exponent $\eta > 0$, we can show that $\bar{r} = 2^{1/\eta} r_s$. In the case of free–space, $\eta = 2$ and hence $\bar{r} = \sqrt{2} \cdot r_s$. ■

Note that for $d > \bar{r}$, the CUSUM process at a node is a positive recurrent Markov process and can eventually cross

any value. We are interested in finding the average time the CUSUM takes to cross a threshold c when the driving term is $Z(r_s)$ and for $d > \bar{r}$.

Lemma 2: When an event is at a distance $d > \bar{r}$ from a sensor, the average time the CUSUM of that sensor takes to cross a threshold c when $Z(r_s)$ is the driving term is at least $\exp(\omega_0 \cdot c)$, where $\omega_0 = 1 - \frac{2\rho(d)}{\rho(r_s)}$.

Proof: Let τ be the stopping time of CUSUM defined as the earliest time at which the CUSUM statistic crosses a threshold c. From (Eqn. 5.2.79 pg. 177 of) [2], we can show that $\mathbb{E}_{f_1(\cdot;d)}[\tau] \geqslant \exp(\omega_0 c)$ where ω_0 is the solution to the equation $\mathbb{E}_{f_1(\cdot;d)}\left[e^{\omega_0 Z(r_s)}\right] = 0$, which is given by $\omega_0 = 1 - \frac{2\rho(d)}{\rho(r_s)}$. ∎

Note that $\omega_0 > 0$ for $d > \bar{r}$. We would be interested in $\omega_0 \geqslant \underline{\omega_0}$ a parameter which we discuss later. This requirement is achieved if $d > \bar{R} := \min\{d' : 2\rho(d') \leqslant (1 - \underline{\omega_0})\rho(r_s)\}$. We call $\overline{R}$ as the *influence–range* of a sensor. We define the *influence–region* of sensor i as the region $\mathcal{T}^{(i)}$ such that each point $x \in \mathcal{T}^{(i)}$ is within the influence–range of sensor i. Thus, the influence–region of sensor i is $\mathcal{T}^{(i)} = \mathcal{D}(\ell^{(i)}, \bar{R})$. We define the influence region of a set of sensors $\mathcal{N}_i$ as the region $\mathcal{B}_i$ such that each $x \in \mathcal{B}_i$ is within the influence range of all the sensors in $\mathcal{N}_i$, i.e.,

$$\mathcal{B}_i := \mathcal{B}(\mathcal{N}_i) := \{x \in \mathcal{A} : x \in \mathcal{T}^{(j)} \ \forall j \in \mathcal{N}_i\}. \quad (1)$$

Note that $\mathcal{A}(\mathcal{N}_i) \subseteq \mathcal{B}(\mathcal{N}_i)$. In the following Section, we formulate the problem of quickest detection of an event and isolate it to one of the subregions $\mathcal{B}_1, \mathcal{B}_2, \cdots, \mathcal{B}_N$ under a false alarm and false isolation constraint.

F. Discussion and Motivation for Formulation

See Fig. 1(b) for a three sensor example of detection and influence regions. In this example, the larger dotted circles represent the influence region and the smaller continuous circles represent the detection region of each sensor. Consider an example in which an event occurs at ℓ_e as shown in Fig. 1(b). Note that $\ell_e \in \mathcal{A}(\{1,2\})$. Also, note that the event lies in the following influence–regions $\mathcal{B}(\{1\})$, $\mathcal{B}(\{2\})$, and $\mathcal{B}(\{1,2\})$ (and the event does not lie in any of the the following influence regions, $\mathcal{B}(\{3\})$, $\mathcal{B}(\{1,3\})$, $\mathcal{B}(\{2,3\})$ and $\mathcal{B}(\{1,2,3\})$). A global alarm will be raised if all the local detectors in any of the $\mathcal{N}_i$s are triggered. Thus, with the event at ℓ_e, the global alarm can be due to **case 1:** sensor 1 alone or **case 2:** sensor 2 alone or **case 3:** due to both sensors 1 and 2. In case 1, we isolate the event to subregion $\mathcal{B}(\{1\})$, in case 2, we isolate the event to subregion $\mathcal{B}(\{2\})$, and in case 3, we isolate the event to subregion $\mathcal{B}(\{1,2\})$. Thus, in all the three cases, the event is located to a region that contains it. If the global alarm is due to **case 4:** sensors 1 and 3 or **case 5:** sensors 2 and 3, or **case 6:** sensors 1, 2 and 3, or **case 7** due to sensor 3, then in all these cases the isolation region does not contain the event and hence these cases correspond to the *false isolation*.

III. Problem Formulation

We are interested in the problem of distributed detection/isolation for a large WSN. Given a sample node deployment (i.e., given $\boldsymbol{\ell}$), and *having chosen a value of* r_s, we have a detection–partition of the regions $\mathcal{A}_1, \mathcal{A}_2, \cdots, \mathcal{A}_N$ and let $\mathcal{N}_1, \mathcal{N}_2, \cdots, \mathcal{N}_N$ be the corresponding set of sensors. Let $\mathcal{B}_i$ be the influence region of $\mathcal{N}_i$. We define the following set of hypotheses

$$\begin{aligned} \mathbf{H}_0 &: \text{event not occurred,} \\ \mathbf{H}_i &: \text{event occurred in subregion } \mathcal{B}_i, \quad i = 1, 2, \cdots, N. \end{aligned}$$

Note that the location of the event ℓ_e is in exactly one $\mathcal{A}_i$, but it can be in one or more of the $\mathcal{B}_i$s. However, as an event can only influence a few of the $\mathcal{B}_i$ in its vicinity, we are satisfied by isolating it to any of these $\mathcal{B}_i$. We study distributed procedures that detect and locate an event (to any of the $\mathcal{B}_i$s) subject to a false alarm and false isolation constraint. The *false alarm constraint* considered is the average time to false alarm at the region $\mathcal{B}_i$ (TFA_i) defined as the average number of samples taken under the hypothesis $\mathbf{H}_0$ to raise a global alarm and declare the hypothesis $\mathbf{H}_i$. The *false isolation constraint* considered is the average time to false isolation TFI_{ij} defined as the average number of samples taken to raise a global alarm and declare the hypothesis $\mathbf{H}_j$ when the event has not occurred in $\mathcal{B}_j$ but has occurred in $\mathcal{B}_i$. The problem is to obtain a detection/isolation procedure that minimizes the supremum average detection delay SADD for the procedure τ for a given $\min\{\mathsf{TFA}_i, \mathsf{TFI}_{ij}\}$ requirement. The $\mathsf{SADD}(\tau)$ is defined in the same sense as Lorden [3] as the worst–case average number of samples taken under any hypothesis $\mathbf{H}_i,\ i = 1, 2, \cdots, N$, to raise a global alarm, i.e.,

$$\mathsf{SADD}(\tau) := \sup_{1 \leqslant i \leqslant N} \sup_{\{\mathbf{s}: \ell_e \in \mathcal{B}_i\}} \sup_{k \geqslant 1} \operatorname{ess\,sup} \mathbb{E}_k^{(\mathbf{s})} \left[(\tau - k + 1)^+ | \mathbf{X}_{[1,\cdots,k-1]}\right],$$

where $\mathbb{P}_k^{(\mathbf{s})}$ is the probability measure when the change happens at time k and at a point $\ell_e \in \mathcal{B}_i$ such that the distance between sensor j and the event is s_j, with $\mathbf{s} := [s_1, s_2, \cdots, s_n]$, and $\mathbb{E}_k^{(\mathbf{s})}$ is the corresponding expectation operator. Also note that $\mathbb{P}_\infty$ is the probability measure when there is no change and $\mathbb{E}_\infty$ is the corresponding expectation operator.

A. Nikiforov's Setting

Nikiforov [6] considers a multi–hypothesis postchange problem where there are N postchange hypotheses (and one prechange hypothesis) and the joint pdf of the observations under hypothesis $\mathbf{H}_i$ is g_i and is completely known. Since, for each $\mathbf{H}_i$ the corresponding postchange pdf is g_i, this model *corresponds to the Boolean sensing model*. The problem posed in [6] is given by

$$\begin{aligned} \inf \quad & \mathsf{SADD}(\tau) \\ \text{such that} \quad & \mathsf{TFA}_i(\tau) \geqslant \gamma, \quad i = 1, 2, \cdots, N \\ & \mathsf{TFI}_{ij}(\tau) \geqslant \gamma, \quad i \neq j,\ i, j = 1, 2, \cdots, N. \end{aligned} \quad (2)$$

Nikiforov [6] *solved the centralized version* of the change detection/isolation problem (defined in Eqn. 2) for the asymptotic case when $\gamma \to \infty$. The SADD of the asymptotically optimal procedure is given by the following theorem.

Theorem 1 (Nikiforov 1995): For the N–hypotheses change detection/isolation problem (for the Boolean sensing model) defined in Eqn. 2, let g_i be the pdf corresponding to the hypothesis $\mathbf{H}_i$, $i = 0, 1, 2, \cdots, N$. The asymptotically minimax delay optimal detection/isolation procedure τ^* has the property,

$$\begin{aligned} \mathsf{SADD}(\tau^*) &\sim \frac{\ln \gamma}{I^*}, \quad \text{as } \gamma \to \infty, \qquad (3) \\ \text{where} \quad I^* &:= \min_{0 \leqslant i \leqslant N} \min_{1 \leqslant j \neq i \leqslant N} I(g_i, g_j), \end{aligned}$$

and $I(\cdot, \cdot)$ is the KL divergence. ■

Nikiforov also proposed a centralized procedure $\tau^{\mathsf{Nikiforov}}$ that achieves this asymptotic lower bound on SADD.

Nikiforov's model does not capture the distance based sensing and neighborhood effects in a WSN. However, we show Nikiforov's result, as this is one of the first works in change detection/isolation, and provides a benchmark on the asymptotic lower bound on the SADD for centralized procedures. Note that our problem is different from Nikiforov's problem in the following sense. In Nikiforov's setting, the observation under hypothesis $\mathbf{H}_i$ is assumed to follow a pdf g_i which is completely known. In our problem, under hypothesis $\mathbf{H}_i$, the event can be anywhere in the region $\mathcal{B}(\mathcal{N}_i)$ and hence the exact pdf of the observation is not known. Note that in terms of the WSN scenario, Nikiforov's setting corresponds to a Boolean sensing model whereas our work is for a general (distance based) sensing model. Also, we are interested in providing a computationally feasible and distributed solution that is optimal under the TFA and TFI constraints.

In Section IV, we propose a distributed detection procedure HALL and analyze the false alarm (TFA), false isolation (TFI) and the detection delay (SADD) properties. We also discuss the false alarm (TFA), false isolation (TFI) and the detection delay (SADD) properties of the detection/isolation procedures, MAX and ALL.

IV. Distributed Change Detection/Isolation Procedures

In this section, we propose HALL a distributed detection procedure for a large WSN and analyze the TFA, TFI and SADD performance. We also analyze the performance of the procedures MAX and ALL. We provide the procedures MAX and ALL followed by HALL.

The MAX procedure: Tartakovsky and Veeravalli propose a decentralized procedure MAX in a colocated scenario in [11]. We extend the MAX procedure to a large WSN under the TFA and TFI constraints. Here, each sensor node i employs CUSUM [2] for local change detection between pdfs f_0 and $f_1(\cdot; r_s)$. Let $\tau^{(i)}$ be the random time at which the CUSUM of sensor node i crosses the threshold c. Here, the local decision of sensor node i, $D_k^{(i)}$ is defined as

$$D_k^{(i)} := \begin{cases} 0 & \text{for } k < \tau^{(i)}, \\ 1 & \text{for } k \geqslant \tau^{(i)}. \end{cases}$$

The global decision rule is a stopping time τ^{MAX} defined as the earliest time slot k at which all sensor nodes corresponding to a subregion $\mathcal{B}_i$ have crossed the threshold c. Thus,

$$\begin{aligned} \tau^{\mathsf{MAX}}(\mathcal{N}_i) &:= \max\left\{\tau^{(j)},\ \forall j \in \mathcal{N}_i\right\}, \\ \tau^{\mathsf{MAX}} &:= \min\left\{\tau^{\mathsf{MAX}}(\mathcal{N}_i) : 1 \leqslant i \leqslant N\right\}. \end{aligned}$$

The isolation rule is to declare the region $\mathcal{B}(\mathcal{N}_i) \supseteq \mathcal{A}(\mathcal{N}_i)$ corresponding to the set of sensors $\mathcal{N}_i$ which raised the alarm.

The ALL procedure: Mei [5], and Tartakovsky and Kim [9], propose a decentralized procedure ALL for a colocated network. We extend the ALL procedure to a large WSN under the TFA and TFI constraints. Here, each sensor node i employs CUSUM for local change detection between pdfs f_0 and $f_1(\cdot; r_s)$. Let $C_k^{(i)}$ be the CUSUM statistic of sensor node i at time k. The CUSUM in the sensor nodes is allowed to run freely even after crossing the threshold c. Here, the local decision of sensor node i is

$$D_k^{(i)} := \begin{cases} 0 & \text{if } C_k^{(i)} < c, \\ 1 & \text{if } C_k^{(i)} \geqslant c. \end{cases}$$

The global decision rule is a stopping time τ^{ALL} defined as the earliest time slot k at which all the sensor nodes corresponding to a subregion $\mathcal{B}_i$ are above the threshold c, i.e.,

$$\begin{aligned} \tau^{\mathsf{ALL}}(\mathcal{N}_i) &:= \inf\left\{k : D_k^{(j)} = 1,\ \forall j \in \mathcal{N}_i\right\} \\ &= \inf\left\{k : C_k^{(j)} \geqslant c, \forall j \in \mathcal{N}_i\right\} \\ \tau^{\mathsf{ALL}} &:= \min\left\{\tau^{\mathsf{ALL}}(\mathcal{N}_i) : 1 \leqslant i \leqslant N\right\}. \end{aligned}$$

The isolation rule is to declare the region $\mathcal{B}(\mathcal{N}_i) \supseteq \mathcal{A}(\mathcal{N}_i)$ corresponding to the set of sensors $\mathcal{N}_i$ which raised the alarm. Note that ALL declares an alarm only when all the nodes corresponding to a region $\mathcal{B}_i$ are above the threshold.

The HALL Procedure: Motivated by the fact that sensor noise can make the CUSUM statistic fluctuate around the threshold, we propose a local decision rule which is 0 when the CUSUM statistic has visited zero and has not crossed the threshold yet and is 1 otherwise. We explain the HALL procedure below.

Each sensor node i computes a CUSUM statistic $C_k^{(i)}$ based on the LLR of its own observations between the pdfs $f_1(\cdot; r_s)$ and f_0. Define $U_0^{(i)} := 0$. Define $V_1^{(i)}$ as the time at which $C_k^{(i)}$ crosses the threshold c (for the first time) as,

$$V_1^{(i)} := \inf\left\{k : C_k^{(i)} \geqslant c\right\}$$

(see Fig. 2). Note that $\inf \emptyset := \infty$. Next define

$$U_1^{(i)} := \inf\left\{k > V_1^{(i)} : C_k^{(i)} = 0\right\}.$$

Now starting with $U_1^{(i)}$ we can recursively define $V_2^{(i)}, U_2^{(i)}$ etc. in the obvious manner, see Fig. 2. We define the *quiet–times* and the *active–times* of the CUSUM process $C_k^{(i)}$ as

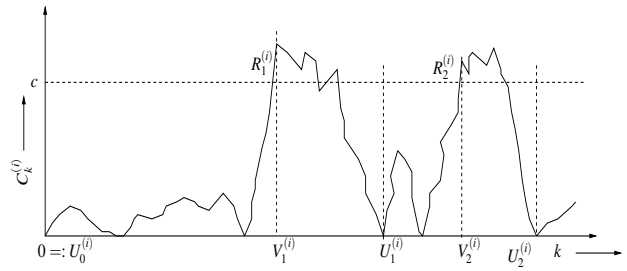


Fig. 2. ALL and HALL: Evolution of CUSUM statistic $C_k^{(i)}$ of node i plotted vs. k. Note that at time $k = V_j^{(i)}$, $R_j^{(i)}$ is the excess above the threshold.

$Q_j^{(i)} := V_j^{(i)} - U_{j-1}^{(i)}$ and $A_j^{(i)} := U_j^{(i)} - V_j^{(i)}$. Each node i computes the local decision $D_k^{(i)}$ based on the CUSUM statistic $C_k^{(i)}$ as follows:

$$D_k^{(i)} = \begin{cases} 1 & \text{if } V_j^{(i)} \leqslant k < U_j^{(i)} \text{ for some } j \\ 0 & \text{otherwise.} \end{cases} \quad (4)$$

The global decision rule[2] is a stopping time τ^{HALL} defined as the earliest time slot k at which all the sensor nodes in a region have a local decision 1, i.e.,

$$\begin{aligned} \tau^{\mathsf{HALL}}(\mathcal{N}_i) &:= \inf\left\{k : D_k^{(j)} = 1,\ \forall j \in \mathcal{N}_i\right\}, \\ \tau^{\mathsf{HALL}} &:= \min\left\{\tau^{\mathsf{HALL}}(\mathcal{N}_i) : 1 \leqslant i \leqslant N\right\}. \end{aligned}$$

The isolation rule is to declare the region $\mathcal{B}(\mathcal{N}_i) \supseteq \mathcal{A}(\mathcal{N}_i)$ corresponding to the set of sensors $\mathcal{N}_i$ which raised the alarm.

For the distributed procedures MAX, ALL and HALL, we analyze the SADD performance in Section IV-A, the TFA_r in Section IV-B and the TFI_{ij} in Section IV-C.

A. Supremum Average Detection Delay (SADD)

In this section, we analyze the SADD performance of the distributed detection/isolation procedures. Note that for any sample path of the observation process, the MAX rule raises an alarm first, followed by the HALL rule and then by the ALL rule. Thus we have, $\tau^{\mathsf{MAX}} \leqslant \tau^{\mathsf{HALL}} \leqslant \tau^{\mathsf{ALL}}$. We recall that each of the stopping times is the minimum of stopping times corresponding to the sets of sensors $\{\mathcal{N}_r : r = 1, 2, \cdots, N\}$. Note that there can be more than one true hypothesis. But there exists a unique true hypothesis $\mathbf{H}_i$ such that $\forall j \in \mathcal{N}_i,\ s_j \leqslant r_s$ (this is possible as $\ell_e \in \mathcal{A}_i$ for exactly one i). Hence, for any rule MAX, HALL or ALL, we have

$$\begin{aligned} \tau^{\mathsf{rule}} &= \min\{\tau^{\mathsf{rule}}(\mathcal{N}_r) : r = 1, 2, \cdots, N\}, \\ &\leqslant \tau^{\mathsf{rule}}(\mathcal{N}_i), \end{aligned}$$

and hence,

$$\begin{aligned} &\sup_{k\geqslant 1} \operatorname{ess\,sup} \mathbb{E}_k^{(\mathbf{s})} \left[\left(\tau^{\mathsf{rule}} - k + 1\right)^+ \mid \mathbf{X}_{[1:k-1]}\right] \\ \leqslant\ &\sup_{k\geqslant 1} \operatorname{ess\,sup} \mathbb{E}_k^{(\mathbf{s})} \left[(\tau^{\mathsf{rule}}(\mathcal{N}_i) - k + 1)^+ \mid \mathbf{X}_{[1:k-1]}\right]. \end{aligned}$$

[2]The procedures HALL, MAX and ALL differ only in their local decision rule; the global decision rule as a function of $D_k^{(i)}$s is the same for HALL, MAX and ALL.

We note from [9], as $c \to \infty$,

$$\begin{aligned} &\sup_{k\geqslant 1} \operatorname{ess\,sup} \mathbb{E}_k^{(\mathbf{s})} \left[(\tau^{\mathsf{ALL}}(\mathcal{N}_i) - k + 1)^+ \mid \mathbf{X}_{[1:k-1]}\right] \\ =\ &\frac{c}{\min_{j\in\mathcal{N}_i} \mathbb{E}_{f_1(\cdot;s_j)}\left[Z\left(s_j\right)\right]}\left(1 + o(1)\right), \\ \leqslant\ &\frac{c}{\mathbb{E}_{f_1(\cdot;r_s)}\left[Z\left(r_s\right)\right]}\left(1 + o(1)\right), \\ =\ &\frac{c}{I(f_1(\cdot;r_s), f_0)}\left(1 + o(1)\right). \end{aligned} \quad (5)$$

The inequality above follows from $\mathbb{E}_{f_1(\cdot;s_j)}\left[Z\left(s_j\right)\right] = \frac{(h_e\rho(s_j))^2}{2\sigma^2} \leqslant \frac{(h_e\rho(r_s))^2}{2\sigma^2} = \mathbb{E}_{f_1(\cdot;r_s)}\left[Z\left(r_s\right)\right]$ as $f_1(\cdot;d)$ is Gaussian with mean $h_e\rho(d) + \mu_0$ and variance σ^2, and f_0 is Gaussian with mean μ_0 and variance σ^2, and $s_j \leqslant r_s$ $\forall j \in \mathcal{N}_i$. Thus, for a given threshold c, asymptotically (as $c \to \infty$), $\mathsf{SADD}(\tau^{\mathsf{ALL}}) \leqslant \frac{c}{I\left(f_1(\cdot;r_s), f_0\right)}(1 + o(1))$ and hence $\mathsf{SADD}(\tau^{\mathsf{MAX}}) \leqslant \mathsf{SADD}(\tau^{\mathsf{HALL}}) \leqslant \frac{c}{I(f_1(\cdot;r_s), f_0)}(1 + o(1))$, where $I(f_1(\cdot;r_s), f_0)$ is the Kullback–Leibler divergence between the pdfs $f_1(\cdot;r_s)$ and f_0. Recall from Section II-B that $\mu_1 = h_e\rho(r_s) + \mu_0$ and the parameter μ_1 is chosen such that a requirement on SADD is met. Thus, to achieve a required SADD, we need to choose r_s appropriately. A small value of r_s gives less detection delay compared to a large value of r_s and requires more sensors to detection–cover the ROI.

B. Average Time to False Alarm in region $\mathcal{B}_r$ (TFA_r)

From [11], we note for a threshold c that

$$\mathsf{TFA}_r\left(\tau^{\mathsf{MAX}}\right) \geqslant \exp(c) \quad (6)$$

and from [5], we have

$$\mathsf{TFA}_r\left(\tau^{\mathsf{ALL}}\right) \geqslant \exp\left(|\mathcal{N}_r|c\right). \quad (7)$$

Thus, for MAX we choose $c = \ln\gamma$ and for ALL we choose $c \geqslant \frac{\ln\gamma}{|\mathcal{N}_r|}$ to achieve a TFA_r of γ. The TFA_r performance of HALL is given by the following theorem.

Theorem 2: $\mathsf{TFA}_r\left(\tau^{\mathsf{HALL}}\right) \geqslant \gamma$ when the threshold c is chosen such that $c\left(1 - \frac{1}{\beta+o(1)}\right) \geqslant \frac{\ln(\gamma+1)}{|\mathcal{N}_r|} + \ln\frac{\beta+o(1)}{I(f_0, f_1(\cdot;r_s))}$, where $\beta > 1$ is a constant that depends upon the distribution of the LLR of the observation.

Proof: See Appendix. ∎

C. Average Time to False Isolation (TFI_{ij})

Here, we consider the scenario in which the hypothesis $\mathbf{H}_i$ is true and the hypothesis $\mathbf{H}_j$ is declared to be true at the time of global alarm and the event does not lie in the region $\mathcal{B}(\mathcal{N}_j)$, i.e., we are interested in finding the time to global alarm due to $\mathcal{N}_j$ when $\ell_e \in \mathcal{B}(\mathcal{N}_i)$ and $\ell_e \notin \mathcal{B}(\mathcal{N}_j)$. This happens when $\exists$ node $j' \in \mathcal{N}_j$ such that $\ell_e \notin \mathcal{T}^{(j')}$, the influence region of node j'. We are interested in obtaining the corresponding average time to false isolation, TFI_{ij} which we show in the following theorem.

Theorem 3: Define the parameter λ_{ij} for the Boolean sensing model, $\lambda_{ij} = |\mathcal{N}_j \setminus \mathcal{N}_i|$ and for the path–loss sensing model, $\lambda_{ij} = 1$. The TFI_{ij} for the change detection/isolation procedures is given by

1) $\mathsf{TFI}_{ij}\left(\tau^{\mathsf{MAX}}\right) \geqslant \gamma$, when the threshold c is chosen such that $c \geqslant \frac{\ln \gamma}{\underline{\omega_0}}$.
2) $\mathsf{TFI}_{ij}\left(\tau^{\mathsf{ALL}}\right) \geqslant \gamma$, when the threshold c is chosen such that $c \geqslant \frac{\ln \gamma}{\underline{\omega_0}\lambda_{ij}}$.
3) $\mathsf{TFI}_{ij}\left(\tau^{\mathsf{HALL}}\right) \geqslant \gamma$, when the threshold c is chosen such that $c\left(1-\frac{1}{\beta+o(1)}\right) \geqslant \frac{\ln(\gamma+1)}{\underline{\omega_0}\lambda_{ij}} + \ln \frac{\beta+o(1)}{\underline{\omega_0} I(f_0, f_1(\cdot; r_s))}$, where $\beta > 1$ is a constant that depends upon the distribution of the LLR of the observation.

Proof: The event lies in the region $\mathcal{B}(\mathcal{N}_i) \setminus \mathcal{B}(\mathcal{N}_j)$. Note that $\mathcal{N}_j = (\mathcal{N}_j \cap \mathcal{N}_i) \bigcup (\mathcal{N}_j \setminus \mathcal{N}_i)$. The sensors in the set $\mathcal{N}_j \cap \mathcal{N}_i$ are within the influence range. For the Boolean sensing model, all the sensors in the set $\mathcal{N}_j \setminus \mathcal{N}_i$ are beyond the influence range from the event and for the path–loss model there exists at least one sensor in $\mathcal{N}_j \setminus \mathcal{N}_i$ that lies beyond the influence range from the event. From sample path argument, it is clear that the time taken to raise a global alarm in region $\mathcal{B}_j$ is at least as large as the time taken when we restrict the set $\mathcal{N}_j$ to those which are beyond the influence range of each of the sensors in $\mathcal{N}_j \cap \mathcal{N}_i$. There are at least λ_{ij} nodes in this restricted set. The observation of any of these sensor nodes is just the sensor noise and hence the theorem follows from Eqns. 6 and 7, from Lemma 2, and from Theorem 2. ∎

In the next subsection, we discuss the asymptotic minimax delay optimality of the distributed procedures in relation to *Theorem 1*.

D. Asymptotic Order Optimality of ALL and HALL

Define the parameter $\widetilde{M}$ as

$$\widetilde{M} \;=\; \min_{1 \leqslant i \leqslant N} \min_{1 \leqslant j \neq i \leqslant N} \min\left\{|\mathcal{N}_i|, \lambda_{ij}\right\}.$$

By choosing $c = \frac{\ln \gamma}{\underline{\omega_0}\widetilde{M}}$ for ALL, $c = \frac{\ln(\gamma+1)}{\left(1-\frac{1}{\beta}\right)\underline{\omega_0}\widetilde{M}} + \frac{1}{1-\frac{1}{\beta}} \ln \frac{\beta}{\underline{\omega_0} I(f_0, f_1(\cdot; r_s))}$ for HALL and $c = \frac{\ln \gamma}{\underline{\omega_0}}$ for MAX, we see from Eqns. 6 and 7, and Theorems 2 and 3, that as $\gamma \to \infty$, $\min\{\mathsf{TFA}_i(\tau^{\mathsf{rule}}), \mathsf{TFA}_{ij}(\tau^{\mathsf{rule}})\} \geqslant \gamma$ for any rule, ALL, HALL or MAX, and from Eqn. 5 that

$$\begin{aligned}
\mathsf{SADD}(\tau^{\mathsf{ALL}}) &\leqslant \frac{\ln \gamma\,(1+o(1))}{\underline{\omega_0}\widetilde{M} \cdot I(f_1(\cdot; r_s), f_0)}, \\
\mathsf{SADD}(\tau^{\mathsf{HALL}}) &\leqslant \frac{\ln(\gamma+1)\,(1+o(1))}{\left(1-\frac{1}{\beta}\right)\underline{\omega_0}\widetilde{M} \cdot I(f_1(\cdot; r_s), f_0)} + C + o(1), \\
\mathsf{SADD}(\tau^{\mathsf{MAX}}) &\leqslant \frac{\ln \gamma\,(1+o(1))}{\underline{\omega_0} \cdot I(f_1(\cdot; r_s), f_0)}, \qquad (8)
\end{aligned}$$

where the constant C is $\frac{\ln(\beta)}{\left(1-\frac{1}{\beta}\right)\underline{\omega_0} \cdot I(f_0, f_1(\cdot; r_s)) \cdot I(f_1(\cdot; r_s), f_0)}$. Note that as we decrease r_s, $I(f_1(\cdot; r_s), f_0)$ and $\underline{\omega_0}$ increases. But $\widetilde{M}$ decreases as r_s decreases. We are interested in the tradeoff between $\widetilde{M}$ and $\underline{\omega_0} I(f_1(\cdot; r_s), f_0)$ which we will study in our future work.

To compare the asymptotic SADD performance of the distributed procedures HALL, MAX and ALL against the optimal centralized scheme of Nikiforov, we consider the Boolean sensing model when $\bar{R} = r_s$ and $\underline{\omega_0} = 1$. From Theorem 1 for Gaussian f_0 and f_1, we can show that

$$\begin{aligned}
\mathsf{SADD}(\tau^*) &\sim \frac{\ln \gamma}{M^* \cdot I(f_1, f_0)} \\
\text{where } M^* &= \min_{1 \leqslant i \leqslant N} \min_{1 \leqslant j \leqslant N, j \neq i} \min\left\{|\mathcal{N}_i|, |\mathcal{N}_i \Delta \mathcal{N}_j|\right\}.
\end{aligned}$$

Note that for the same $\min\{\mathsf{TFA}, \mathsf{TFI}\}$ requirement of γ, the SADD of τ^*, τ^{ALL} and τ^{HALL} scale as $\ln \gamma$. Hence, ALL and HALL are asymptotically order optimal. The factor $1/(1-1/\beta)$ in the SADD of HALL makes this slightly larger than $\mathsf{SADD}(\tau^{\mathsf{ALL}})$. To achieve the desired false alarm performance for the MAX procedure, we choose the threshold $c = \ln \gamma$ and hence $\mathsf{SADD}\left(\tau^{\mathsf{MAX}}\right) \sim \frac{\ln \gamma}{I(f_1(\cdot; r_s), f_0)}$. Thus, the SADD performance of MAX is worse than that of ALL or HALL. Note that $\widetilde{M} \leqslant M^*$. Hence, there is a small price to pay for the distributed procedure compared to the optimum centralized procedure. But, the gain in SADD in the centralized procedure is offset by substantially larger computation and communication costs, particularly in terms of the node energy expenditure, a critical issue in the context of wireless ad hoc sensor networks.

V. Numerical Results

We compute the SADD and the TFA_i performance of MAX, HALL, ALL and the Nikiforov's procedure for the Boolean sensing model with $f_0 \sim \mathcal{N}(0,1)$ and $f_1 \sim \mathcal{N}(1,1)$. We deployed 10 nodes in the ROI in a deterministic manner such that we get $N = 9$ regions and $\mathcal{N}_1 = \{1,2,4\}$, $\mathcal{N}_2 = \{2,4,5\}$, $\mathcal{N}_3 = \{2,3,5\}$, $\mathcal{N}_4 = \{4,6,7\}$, $\mathcal{N}_5 = \{4,5,7\}$, $\mathcal{N}_6 = \{5,7,8\}$, $\mathcal{N}_7 = \{6,7,9\}$, $\mathcal{N}_8 = \{7,9,10\}$, and $\mathcal{N}_9 = \{7,8,10\}$. We obtain the SADD and the TFA_i for the above mentioned setup and plot the SADD against the TFA_i in Fig. 3.We observe from Fig. 3 that SADD increases linearly with $\ln(\mathsf{TFA}_i)$ and that the slope is $\approx 1/I(f_1, f_0)$ for MAX, and $\approx 1/(3 \cdot I(f_1, f_0))$ for HALL, ALL, and Nikiforov's procedure. Also, we observe that for a given TFA_i, Nikiforov's procedure has the smallest SADD and MAX has the largest SADD. For example, for a TFA_i requirement of 2000 slots, the SADD of Nikiforov's procedure is 6.2 slots, that of ALL is 8 slots, HALL is 9 slots, and that of MAX is 17.5 slots. Nikiforov's procedure is an optimal centralized procedure and hence it outperforms ALL, HALL and MAX. Also, SADD of MAX is the largest as it does not scale with $\widetilde{M}$. From Eqn. 8, it is clear that SADD of ALL is little smaller than that of HALL.

VI. Conclusions

We consider the quickest distributed event detection/isolation problem in a large WSN with a practical sensing model which incorporates distance losses. We study the distributed detection/isolation rules, MAX, ALL and HALL in the optimality framework introduced by Nikiforov [6]. We show that for a required $\min\{\mathsf{TFA}, \mathsf{TFI}\}$, the SADD performance of HALL and ALL is asymptotically of the same order and differs from the optimal centralized procedure of Nikiforov [6], by a constant factor. The SADD performance of the MAX procedure is inferior to that of ALL or HALL.

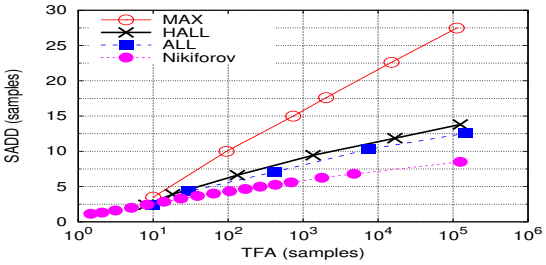


Fig. 3. SADD versus TFA for MAX, HALL, ALL and Nikiforov's procedure. The system parameters are $f_0 \sim N(0,1)$, $f_1 \sim N(1,1)$, $r_s = 1.0$.

APPENDIX
SKETCH OF PROOF OF THEOREM 2

Note that $\{U_j^{(i)}, j \geqslant 1\}$ is a renewal process for any i and let $A_j^{(i)}$ be the reward in the jth cycle. By renewal theory, we can show that the probability of a node in false alarm is given by $\frac{\mathbb{E}_\infty A_1^{(1)}}{\mathbb{E}_\infty A_1^{(1)} + \mathbb{E}_\infty Q_1^{(1)}}$ and because, the CUSUMs of nodes $i \in \mathcal{N}_r$ are independent, probability of false alarm in region $\mathcal{A}_r$ is

$$\mathsf{PFA_r} = \left(\frac{\mathbb{E}_\infty A_1^{(1)}}{\mathbb{E}_\infty A_1^{(1)} + \mathbb{E}_\infty Q_1^{(1)}} \right)^{|\mathcal{N}_r|}. \quad (9)$$

Let $\tau_1 < \tau_2 < \cdots$ be the random times at which the system enters false alarm in region $\mathcal{A}_r$ and let $0 =: \zeta_0 < \zeta_1 < \zeta_2 < \cdots$ be the random times at which the system comes out of the false alarm in region $\mathcal{A}_r$. We define TTF_r, the mean time to false alarms in $\mathcal{A}_r$, as $\mathsf{TTF}_r := \lim_{m\to\infty} \frac{1}{m} \sum_{j=1}^m [\tau_j - \zeta_{j-1}]$. From sample path and coupling arguments, one can show that $\mathsf{TTF}_r \leqslant \mathsf{TFA}_r$. Note that PFA_r, the fraction of time the system is in false alarm in region $\mathcal{A}_r$ is given by

$$\begin{aligned}
\mathsf{PFA}_r &= \lim_{m\to\infty} \frac{\sum_{j=1}^m [\zeta_j - \tau_j]}{\sum_{j=1}^m [(\tau_j - \zeta_{j-1}) + (\zeta_j - \tau_j)]} \\
&\geqslant \lim_{m\to\infty} \frac{m}{\sum_{j=1}^m [\tau_j - \zeta_{j-1}] + m} \\
&= \lim_{m\to\infty} \frac{1}{\frac{1}{m}\sum_{j=1}^m [\tau_j - \zeta_{j-1}] + 1} \\
&= \frac{1}{\mathsf{TTF}_r + 1} \quad (10)
\end{aligned}$$

The inequality in the second step is due to the fact that the time duration the system spends in false alarm state is at least one and the function $g : \{1, 2, 3, \cdots\} \to \mathbb{R}$ defined by $g(x) := \frac{x}{K+x}$ is an increasing function, for a constant K. Therefore, combining Eqn. 9 and Eqn. 10, we have

$$\begin{aligned}
\mathsf{TTF}_r &\geqslant \left(\frac{\mathbb{E}_\infty \left[A_1^{(1)} + Q_1^{(1)} \right]}{\mathbb{E}_\infty \left[A_1^{(1)} \right]} \right)^{|\mathcal{N}_r|} - 1, \\
&\geqslant \left(\frac{e^c}{\mathbb{E}_\infty \left[A_1^{(1)} \right]} \right)^{|\mathcal{N}_r|} - 1 \\
&\geqslant \exp\left(|\mathcal{N}_r| c \left(1 - \frac{\ln \mathbb{E}_\infty A_1^{(1)}}{c} \right) \right) - 1. \quad (11)
\end{aligned}$$

The inequality in the second step follows from $\mathbb{E}_\infty Q_1^{(1)} \geqslant e^c$ (see Eqn. 5.2.80 in [2]) and $\mathbb{E}_\infty A_1^{(1)} \geqslant 0$. In the same way as in Eqn. 5.2.67 in [2], we can show that

$$\mathbb{E}_\infty A_1^{(1)} \leqslant \frac{c + r + o(1) + \alpha}{I(f_0, f_1(\cdot; r_s))},$$

where $r + o(1) = \mathbb{E}_\infty R_j^{(i)}$ is the mean excess above c (in up–crossing) and α is the mean excess in down–crossing. Note that the positive constants, r and α do not depend on the threshold c ([12], [2]). Define $\beta = r + \alpha$. Hence,

$$\begin{aligned}
&\frac{\ln \mathbb{E}_\infty A_1^{(1)}}{c} \\
\leqslant\ & \frac{\ln(c + \beta + o(1)) - \ln I(f_0, f_1(\cdot; r_s))}{c} \\
=\ & \frac{\ln(\beta + o(1))}{c} + \frac{\ln(1 + \frac{c}{\beta + o(1)})}{c} - \frac{\ln I(f_0, f_1(\cdot; r_s))}{c} \\
\leqslant\ & \frac{\ln(\beta + o(1))}{c} + \frac{1}{\beta + o(1)} - \frac{\ln I(f_0, f_1(\cdot; r_s))}{c}. \quad (12)
\end{aligned}$$

Hence, *Theorem 2* follows from Eqns. 11 and 12 as $\mathsf{TTF}_r \leqslant \mathsf{TFA}_r$.

REFERENCES

[1] http://www.xbow.com/.

[2] M. Basseville and I. V. Nikiforov, *Detection of Abrupt Changes: Theory and Application*. Englewood Cliffs, NJ: Prentice Hall, 1993.

[3] G. Lorden, "Procedures for reacting to a change in distribution," *The Annals of Mathematical Statistics*, vol. 42, no. 6, pp. 1897–1908, December 1971.

[4] D. P. Malladi and J. L. Speyer, "A generalized Shiryayev sequential probability ratio test for change detection and isolation," *IEEE Trans. Autom. Control*, vol. 44, no. 8, pp. 1522–1534, Aug 1999.

[5] Y. Mei, "Information bounds and quickest change detection in decentralized decision systems," *IEEE Trans. Inf. Theory*, vol. 51, no. 7, pp. 2669–2681, Jul. 2005.

[6] I. V. Nikiforov, "A generalized change detection problem," *IEEE Trans. Inf. Theory*, vol. 41, no. 1, pp. 171–187, Jan 1995.

[7] R. Niu and P. K. Varshney, "Distributed detection and fusion in a large wireless sensor network of random size," *EURASIP Journal on Wireless Communications and Networking*, vol. 4, no. 7, pp. 462–472, 2005.

[8] R. Solis, V. S. Borkar, and P. R. Kumar, "A new distributed time synchronization protocol for multihop wireless networks," in *45th IEEE Conference on Decision and Control (CDC'06)*, December 2006.

[9] A. G. Tartakovsky and H. Kim, "Performance of certain decentralized distributed change detection procedures," in *9th International Conference of Information Fusion*, July 2006.

[10] A. G. Tartakovsky and A. S. Polunchenko, "Quickest changepoint detection in distributed multisensor systems under unknown parameters," in *11th International Conference of Information Fusion*, Germany, July 2008.

[11] A. G. Tartakovsky and V. V. Veeravalli, "Quickest change detection in distributed sensor systems," in *Sixth International Conference of Information Fusion*, vol. 2, 2003, pp. 756–763.

[12] M. Woodroofe, *Nonlinear Renewal Theory in Sequential Analysis*. Philadelphia: SIAM, 1982.