QUICKEST DETECTION OF DYNAMIC EVENTS IN SENSOR NETWORKS

Shaofeng Zou　*Venugopal V. Veeravalli*

University of Illinois at Urbana-Champaign
Email: szou3@illinois.edu, vvv@illinois.edu

ABSTRACT

We consider the problem of quickest detection of dynamic events in sensor networks. After an event occurs, a number of sensors are affected and undergo a change in the statistics of their observations. We assume that the event is dynamic and can propagate with time, i.e., different sensors perceive the event at different times. The goal is to design a sequential algorithm that can detect when the event has affected no less than η sensors as quickly as possible, subject to false alarm constraints. We design a computationally efficient algorithm that is adaptive to unknown propagation dynamics, and demonstrate its asymptotic optimality as the false alarm rate goes to zero. We also provide numerical simulations to validate our theoretical results.

***Index Terms*—** asymptotic optimality, dynamic event, quickest change detection, spartan CuSum

1. INTRODUCTION

Suppose a system is monitored in real time by a set of L sensors. At an unknown time, an event occurs in the system, and causes a change in the observations of an arbitrary, unknown subset of sensors. Moreover, if an event occurs, it can dynamically propagate over the sensor network with time, i.e., different sensors perceive the event at different times. We are interested in detecting a "large" event, i.e., we would like to raise an alarm if more than $\eta \geq 1$ sensors are affected quickly and reliably. Applications of this model can be found in epidemic detection [1, 2], remote sensing [3], etc.

The problem in this paper is closely related to the multichannel sequential change detection setup, in which one or multiple unknown sensors perceive a change simultaneously [4–8], or alternatively, at different times [9–11]. The major differences from these previous works lie in that: (i) we are interested in detecting when the event has affected as least η sensors, whereas previous works focus on the special case with $\eta = 1$; and (ii) instead of considering the worst case of detection delay over all possible perceiving times of the sensors [10] or taking a Bayesian approach [9, 11], we are interested in designing algorithms that are adaptive to unknown propagation patterns.

This research was supported in part by the NSF under grants CCF 16-18658 and ECCS 14-62311, and by the AFOSR under grant FA9550-16-1-0077, through the University of Illinois at Urbana-Champaign.

Our problem is also closely related to the problem of quickest change detection (QCD) under transient dynamics [12–15], in which the pre-change distribution does not change to the persistent post-change distribution instantaneously, but after a number of transient phases, where each of the transient phases is associated with a distribution that is distinct from the pre-change and persistent post-change distributions. In this paper, as the event affects more and more sensors over time, the system also goes through multiple phases. However, the event propagation pattern is unknown and has multiple possibilities.

In this paper, we reformulate this QCD problem as a dynamic hypothesis testing problem, i.e., one where we have to distinguish between two hypotheses at each time instant. The null hypothesis corresponds to the case with less than η affected sensors, and the alternative hypothesis corresponds to the case with no less than η affected sensors. We take the generalized log-likelihood ratio of the two composite hypotheses as the detection statistic, and compare it to a threshold to make a decision about the event. We show that such a test is equivalent to one that compares the sum of the smallest $L - \eta + 1$ local Cumulative Sum (CuSum) statistics [16] to a threshold. This test, which we call spartan CuSum (S-CuSum), is computationally efficient with $O(L)$ complexity. The S-CuSum algorithm is shown to guarantee false alarm constraints for all scenarios with less than η affected sensors, and to adapt to unknown propagation patterns. We also establish the asymptotic optimality of the S-CuSum algorithm up to a first-order approximation as the false alarm rate goes to zero.

2. PROBLEM MODEL

Consider a sensor network consisting of L sensors. Before an event occurs, sensor i receives independent and identically distributed (i.i.d.) samples from distribution f_i, $\forall 1 \leq i \leq L$. If an event occurs, and sensor i is affected by the event at an unknown time v_i, then it starts to receive i.i.d. samples from distribution g_i. More specifically, if we denote the observation received by sensor i at time k by $X_i[k]$, then

$$X_i[k] \sim \begin{cases} f_i, \text{ if } k < v_i, \\ g_i, \text{ if } k \geq v_i. \end{cases} \tag{1}$$

We consider a centralized setting in which all the samples are available to the fusion center. Note that we are only interested in detecting a "large" event, i.e., if an alarm is triggered at a time when less than η sensors are affected, it is considered as a false alarm.

Let $\boldsymbol{v} = \{v_1, \ldots, v_L\}$ denote the set of all change-points which are unknown in advance. Without loss of generality, we assume that $v_1 \leq v_2 \leq \cdots \leq v_L$, with the ordering being unknown in advance. Then, v_η is the first time when at least η sensors are affected. Thus, our problem is to detect the change at v_η as quickly as possible subject to false alarm constraints.

We let $d_i = v_{i+1} - v_i$ denote the time it takes the event to propagate from i affected sensors to $i+1$ affected sensors, and let $\boldsymbol{d} = \{d_1, \ldots, d_{L-1}\}$. Then, given $\boldsymbol{v}$, v_η and $\boldsymbol{d}$ are known. We use $\mathbb{P}_{\boldsymbol{v}}$ to denote the probability measure of the samples with change-points being $\boldsymbol{v}$, and let $\mathbb{E}_{\boldsymbol{v}}$ denote the corresponding expectation. It is clear that for a set of change-points $\boldsymbol{v}$, if $\sum_{i=1}^{L} \mathbb{1}_{\{v_i<\infty\}} < \eta$, i.e., $v_\eta = \infty$, then there are less than η affected sensors under $\mathbb{P}_{\boldsymbol{v}}$. We define the worst-case average run length (WARL) to false alarm as follows:

$$\mathrm{WARL}(\tau) = \inf_{\boldsymbol{v}: v_\eta = \infty} \mathbb{E}_{\boldsymbol{v}}[\tau]. \tag{2}$$

For any fixed $\{d_\eta, d_{\eta+1}, \ldots, d_{L-1}\}$, we further define the worst-case average detection delay (WADD) under Pollak's criterion [17] as follows:

$$J_{\mathrm{P}}[\tau] = \sup_{\boldsymbol{v}: v_1 \leq \cdots \leq v_\eta < \infty} \mathbb{E}_{\boldsymbol{v}}[\tau - v_\eta | \tau \geq v_\eta]. \tag{3}$$

In this paper, we assume for our asymptotic analysis that for each $i \geq \eta$, d_i is deterministic, unknown and either finite or infinite. We note that even for a fixed $\{d_\eta, d_{\eta+1}, \ldots, d_{L-1}\}$, the distribution of the samples after v_η is still composite, since without the knowledge that $v_1 \leq \cdots \leq v_L$, the order of affected sensors after time v_η is unknown.

We denote $\mathcal{F}_k$ as the σ-algebra generated by the observations of all the sensors up to time k, for $k = 1, 2, \ldots$. We wish to find a $\{\mathcal{F}_k\}_{k\in\mathbb{N}}$-stopping time that achieves "small" detection delay, while controlling the false alarm rate. More specifically, the goal is to minimize $J_{\mathrm{P}}[\tau]$ subject to a constraint on the WARL:

$$\inf_{\tau: \mathrm{WARL}(\tau) \geq \gamma} J_{\mathrm{P}}(\tau). \tag{4}$$

In this paper, we denote $\boldsymbol{X}[k] = \{X_1[k], \ldots, X_L[k]\}$, and $\boldsymbol{X}[k_1, k_2] = \{\boldsymbol{X}[k_1], \ldots, \boldsymbol{X}[k_2]\}$. We further denote

$$Z_i[k_1, k_2] = \sum_{k=k_1}^{k_2} \log \frac{g_i(X_i[k])}{f_i(X_i[k])}. \tag{5}$$

We define $\sum_{j=k_1}^{k_2} A_j = 0$ and $\prod_{j=k_1}^{k_2} A_j = 1$ if $k_1 > k_2$. We use X^+ to denote the positive part of X, i.e., $X^+ = \max\{X, 0\}$. We denote the Kullback-Leibler (KL) divergence between g_i and f_i as

$$I_i = \int g_i \log \frac{g_i}{f_i}, \tag{6}$$

which is assumed to be positive and finite, for $1 \leq i \leq L$.

3. THE S-CUSUM ALGORITHM

In this section, we present the design of the S-CuSum algorithm, and show that it can be implemented efficiently with complexity that is linear in L.

We reformulate the quickest detection problem in Section 2 as a dynamic hypothesis testing problem, i.e., to distinguish the following two hypotheses at each time k:

$$\begin{aligned} &\mathcal{H}_0[k] : \sum_{i=1}^{L} \mathbb{1}_{\{v_i \leq k\}} < \eta, \\ &\mathcal{H}_1[k] : \sum_{i=1}^{L} \mathbb{1}_{\{v_i \leq k\}} \geq \eta. \end{aligned} \tag{7}$$

It is clear that the null hypothesis corresponds to the scenario in which there are less than η affected sensors at time k, the alternative hypothesis corresponds to the scenario in which there are no less than η affected sensors at time k, and both the null and alternative hypotheses are composite.

This hypothesis testing procedure stops once a decision in favor of the alternative hypothesis is reached; otherwise, it takes a new sample from each sensor. We take the generalized log-likelihood ratio for this composite hypothesis testing problem as the detection statistic for the S-CuSum algorithm:

$$W[k] = \log \left(\frac{\max_{\boldsymbol{v}: \sum_{i=1}^{L} \mathbb{1}_{\{v_i \leq k\}} \geq \eta} \mathbb{P}_{\boldsymbol{v}}(\boldsymbol{X}[1,k])}{\max_{\boldsymbol{v}: \sum_{i=1}^{L} \mathbb{1}_{\{v_i \leq k\}} < \eta} \mathbb{P}_{\boldsymbol{v}}(\boldsymbol{X}[1,k])} \right). \tag{8}$$

The corresponding stopping time is then given by comparing $W[k]$ against a pre-determined positive threshold:

$$\tilde{\tau}(b) = \inf\{k \geq 1 : W[k] > b\}. \tag{9}$$

Next, we will derive an equivalent but cleaner form for (9), which can be computed efficiently.

Let $\mathbb{P}_\infty$ denote the probability measure with $v_i = \infty$, for $1 \leq i \leq L$, i.e., none of the sensors will be ever affected. It is easily shown that

$$\begin{aligned} W[k] = &\max_{\boldsymbol{v}: \sum_{i=1}^{L} \mathbb{1}_{\{v_i \leq k\}} \geq \eta} \log \left(\frac{\mathbb{P}_{\boldsymbol{v}}(\boldsymbol{X}[1,k])}{\mathbb{P}_\infty(\boldsymbol{X}[1,k])} \right) \\ &- \max_{\boldsymbol{v}: \sum_{i=1}^{L} \mathbb{1}_{\{v_i \leq k\}} < \eta} \log \left(\frac{\mathbb{P}_{\boldsymbol{v}}(\boldsymbol{X}[1,k])}{\mathbb{P}_\infty(\boldsymbol{X}[1,k])} \right). \end{aligned} \tag{10}$$

Now, due to the fact that

$$\begin{aligned} &\log \left(\frac{\mathbb{P}_{\boldsymbol{v}}(\boldsymbol{X}[1,k])}{\mathbb{P}_\infty(\boldsymbol{X}[1,k])} \right) \\ &= \log \left(\prod_{i=1}^{L} \frac{\prod_{j=1}^{\min\{v_i-1,k\}} f_i(X_i[j]) \prod_{j=v_i}^{k} g_i(X_i[j])}{\prod_{j=1}^{k} f_i(X_i[j])} \right) \\ &= \sum_{i=1}^{L} \sum_{j=v_i}^{k} \log \frac{g_i(X_i[j])}{f_i(X_i[j])}, \end{aligned} \tag{11}$$

the first term in (10) is equivalent to

$$\max_{\boldsymbol{v}:\sum_{i=1}^{L}\mathbb{1}_{\{v_i\le k\}}\ge\eta}\sum_{i=1}^{L}\sum_{j=v_i}^{k}\log\frac{g_i(X_i[j])}{f_i(X_i[j])}. \tag{12}$$

Similarly, the second term in (10) is equivalent to

$$\max_{\boldsymbol{v}:\sum_{i=1}^{L}\mathbb{1}_{\{v_i\le k\}}<\eta}\sum_{i=1}^{L}\sum_{j=v_i}^{k}\log\frac{g_i(X_i[j])}{f_i(X_i[j])}. \tag{13}$$

If we denote the individual CuSum [16] statistic at sensor i (testing a change from f_i to g_i) at time k as

$$W_i[k]=\max_{1\le v_i\le k}\sum_{j=v_i}^{k}\log\frac{g_i(X_i[k])}{f_i(X_i[k])}, \tag{14}$$

and define a permutation $\mu(\cdot)$ such that

$$W_{\mu(1)}[k]\ge W_{\mu(2)}[k]\ge\cdots\ge W_{\mu(L)}[k], \tag{15}$$

then, $\tilde{\tau}(b)$ is equivalent to

$$\hat{\tau}(b)=\inf\left\{k\ge 1:\sum_{i=\eta}^{L}\left(W_{\mu(i)}[k]\right)^{+}\ge b\right\}. \tag{16}$$

Such an equivalence can be established as follows.

1. If $W_{\mu(\eta)}[k]\ge 0$, then (12) is equal to $\sum_{i=1}^{L}\left(W_{\mu(i)}[k]\right)^{+}$, and (13) is equal to $\sum_{i=1}^{\eta-1}W_{\mu(i)}[k]$. It then follows that $W[k]=\sum_{i=\eta}^{L}\left(W_{\mu(i)}[k]\right)^{+}$.
2. If $W_{\mu(\eta)}[k]<0$, then (12) is equal to $\sum_{i=1}^{\eta}W_{\mu(i)}[k]$, and (13) is equal to $\sum_{i=1}^{\eta-1}\left(W_{\mu(i)}[k]\right)^{+}$. In this case, $W[k]$ is non-positive, and $\sum_{i=\eta}^{L}\left(W_{\mu(i)}[k]\right)^{+}=0$. Since b is positive, the test in (9) is equivalent to comparing $\sum_{i=\eta}^{L}\left(W_{\mu(i)}[k]\right)^{+}$ to b.

The test in (16) can be implemented efficiently. First of all, for each i, $W_i[k]$ can be updated recursively, i.e.,

$$W_i[k]=(W_i[k-1])^{+}+\log\frac{g_i(X_i[k])}{f_i(X_i[k])}. \tag{17}$$

Second, finding the smallest $L-\eta+1$ numbers from L numbers can be solved in $O(L)$ time using the algorithm in [18]. Then, the total computational cost at each time k is linear in the number of sensors, i.e., $O(L)$.

4. PERFORMANCE ANALYSIS

4.1. Universal Lower Bound on WADD

We first study the universal lower bound on the WADD for any stopping rule with the WARL no less than γ.

For any $\boldsymbol{v}$, denote $\mathcal{C}=\{i:v_i<\infty\}$ as the set that contains all the indices of the affected sensors, and define a permutation $\lambda(\cdot)$ such that $I_{\lambda(1)}\le I_{\lambda(2)}\le\cdots\le I_{\lambda(|\mathcal{C}|)}$. Assume that $|\mathcal{C}|\ge\eta$. Let $\mathcal{S}=\{\lambda(1),\lambda(2),\ldots,\lambda(|\mathcal{C}|-\eta+1)\}$, i.e., $\mathcal{S}$ contains indices of those sensors with the smallest $|\mathcal{C}|-\eta+1$ KL numbers in $\mathcal{C}$, and set $\tilde{I}=\sum_{i\in\mathcal{S}}I_i$.

Theorem 1. *For any event with all affected sensors in set $\mathcal{C}$, and $|\mathcal{C}|\ge\eta$, as $\gamma\to\infty$,*

$$\inf_{\tau:\mathrm{WARL}(\tau)\ge\gamma}J_P(\tau)\ge(1-o(1))\frac{\log\gamma}{\tilde{I}}. \tag{18}$$

The proof is based on a change-of-measure argument and the Law of Large Numbers for log-likelihood ratio statistics similar to those in [19], which is omitted due to space limitations. A major difference in the change-of-measure argument relative to [19] is that the "pre-change" mode is composite, i.e., there are multiple possible scenarios with less than η affected sensors. Furthermore, the application of the Law of Large Numbers requires a decomposition of the sum of the log-likelihood ratio.

4.2. Upper Bound on the WADD

Theorem 2. *For any event with all affected sensors in set $\mathcal{C}$, and $|\mathcal{C}|\ge\eta$, as $b\to\infty$,*

$$J_P[\hat{\tau}(b)]\le(1+o(1))\frac{b}{\tilde{I}}. \tag{19}$$

Proof. It is clear that

$$\begin{aligned}\sum_{i=\eta}^{L}\left(W_{\mu(i)}[k]\right)^{+}&=\min_{\mathcal{D}:|\mathcal{D}|=L-\eta+1}\sum_{i\in\mathcal{D}}(W_i[k])^{+}\\&\overset{(a)}{\ge}\min_{\mathcal{D}\subseteq\mathcal{C}:|\mathcal{D}|=|\mathcal{C}|-\eta+1}\sum_{i\in\mathcal{D}}(W_i[k])^{+}\\&\overset{(b)}{\ge}\min_{\mathcal{D}\subseteq\mathcal{C}:|\mathcal{D}|=|\mathcal{C}|-\eta+1}\sum_{i\in\mathcal{D}}Z_i[v_{|\mathcal{C}|},k],\end{aligned} \tag{20}$$

where $Z_i[v_{|\mathcal{C}|},k]$ is defined in (5), (a) is obtained by the setting $W_i[k]=0$ for $i\notin\mathcal{C}$, since $(W_i[k])^{+}$ is always positive; and (b) is due to the definition of $W_i[k]$.

Define the stopping rule $N(b)$:

$$N(b)=\inf\left\{k:\min_{\mathcal{D}:|\mathcal{D}|=L-\eta+1}\sum_{i\in\mathcal{D}}Z_i[v_{|\mathcal{C}|},k]>b\right\}. \tag{21}$$

It then follows that $\hat{\tau}(b)\le N(b)$. By [20, Theorem 3], as $b\to\infty$, $\mathbb{E}_{\boldsymbol{v}}[N(b)-v_{|\mathcal{C}|}]\le b/\tilde{I}(1+o(1))$. By the assumption that $v_{|\mathcal{C}|}-v_\eta$ is finite and the fact that $N(b)$ is independent from the event $\{\hat{\tau}(b)\ge v_\eta\}$, as $b\to\infty$,

$$\mathbb{E}_{\boldsymbol{v}}[N(b)-v_\eta|\hat{\tau}(b)\ge v_\eta]\le\frac{b}{\tilde{I}}(1+o(1)). \tag{22}$$

□

4.3. Lower Bound on the WARL

Theorem 3. *For any $\boldsymbol{v}$ such that $\sum_{i=1}^{L} \mathbb{1}_{\{v_i \le k\}} < \eta$,*

$$\mathbb{E}_{\boldsymbol{v}}[\hat{\tau}(b)] \ge \frac{e^b}{poly(b)}, \tag{23}$$

where $poly(b)$ denotes a polynomial in b.

Proof. For any $t \in \mathbb{N}$ and $b > 0$, it follows that

$$\begin{aligned}\mathbb{P}_{\boldsymbol{v}}(\hat{\tau}(b) \le t) &= \mathbb{P}_{\boldsymbol{v}}\left(\max_{1\le k\le t} \sum_{i=\eta}^{L} \left(W_{\mu(i)}[k]\right)^+ > b\right)\\ &\le \sum_{k=1}^{t} \mathbb{P}_{\boldsymbol{v}}\left(\sum_{i=\eta}^{L} \left(W_{\mu(i)}[k]\right)^+ > b\right).\end{aligned} \tag{24}$$

Since $\sum_{i=\eta}^{L} \left(W_{\mu(i)}[k]\right)^+ \le \sum_{i=\eta}^{L} (W_i[k])^+$,

$$\mathbb{P}_{\boldsymbol{v}}\left(\sum_{i=\eta}^{L} \left(W_{\mu(i)}[k]\right)^+ > b\right) \le \mathbb{P}_{\boldsymbol{v}}\left(\sum_{i=\eta}^{L} (W_i[k])^+ > b\right). \tag{25}$$

By [6, Lemma B1], it follows that

$$\mathbb{P}_{\boldsymbol{v}}\left(\sum_{i=\eta}^{L} (W_i[k])^+ > b\right) \le \mathrm{poly}(b)e^{-b}. \tag{26}$$

Therefore, $\mathbb{P}_{\boldsymbol{v}}(\hat{\tau}(b) \ge t) \le t\,\mathrm{poly}(b)e^{-b}$, which implies that

$$\mathbb{E}_{\boldsymbol{v}}[\hat{\tau}(b)] \ge \frac{e^b}{\mathrm{poly}(b)}. \tag{27}$$

□

To guarantee that that $\mathrm{WARL}(\hat{\tau}(b)) \ge \gamma$, it then suffices to choose $b \sim \log\gamma$.

Theorem 4 (Asymptotic Optimality)**.** *Let the threshold $b \sim \log\gamma$ so that WARL$(\hat{\tau}(b)) \ge \gamma$. Then for any event with all affected sensors in set $\mathcal{C}$, and $|\mathcal{C}| \ge \eta$,*

$$J_P[\hat{\tau}(b)] \sim \frac{b}{\tilde{I}}. \tag{28}$$

Proof. This result follows from Theorems 1, 2 and 3. □

5. NUMERICAL RESULTS

In this section, we provide some numerical results. We use a simple example with $L = 3$ and $\eta = 2$.

In Fig. 1, we first plot the evolution path of the S-CuSum algorithm. We assume that $f_1 = f_2 = f_3 = \mathcal{N}(0, 1)$, and $g_1 = g_2 = g_3 = \mathcal{N}(1, 1)$. We set $\boldsymbol{v} = \{1, 40, 80\}$. As we can see from Fig. 1, when there is only one affected sensor, i.e., $k < 40$, the detection statistic of the S-CuSum algorithm stays close to 0. Then, when there are two affected sensors, i.e., $40 \le k < 80$, and later when there are three affected sensors, i.e., $k \ge 80$, the detection statistic gradually grows but with different drifts in these two scenarios. Thus, the S-CuSum algorithm is adaptive to the unknown propagation dynamics.

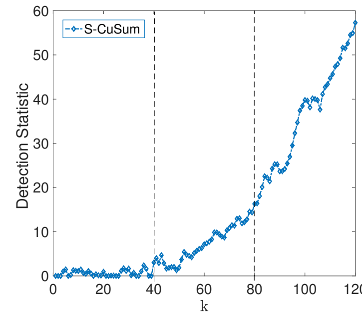

Fig. 1. A sample evolution path of the S-CuSum algorithm.

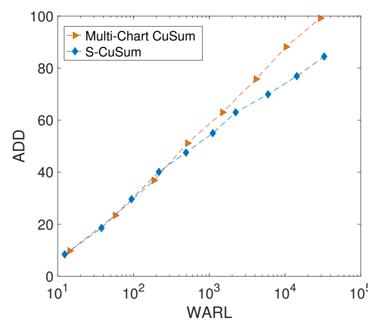

Fig. 2. Comparison between the S-CuSum and multichart CuSum algorithms.

We next study the performance of the S-CuSum algorithm and compare it with a generalization of the multichart algorithm in [21] which stops when at least η local CuSum algorithms have crossed their individual thresholds. We assume that $f_1 = f_2 = f_3 = \mathcal{N}(0, 1)$, and $g_1 = g_2 = g_3 = \mathcal{N}(0.4, 1)$. We set $\boldsymbol{v} = \{1, 1, 40\}$. Here, we consider the average detection delay (ADD) defined as $\mathbb{E}_{\boldsymbol{v}}[\hat{\tau} - v_\eta | \hat{\tau} \ge v_\eta]$ without taking sup over $v_1, \ldots, v_\eta$. We plot ADD versus WARL for the S-CuSum and multichart CuSum algorithms in Fig. 2.

As we can see from Fig. 2, the drift of the S-CuSum algorithm gradually changes after ADD$= v_3 = 40$, which demonstrates that it is adaptive to the unknown $\boldsymbol{v}$. Furthermore, the S-CuSum algorithm has a better performance compared to the multichart CuSum algorithm, especially when ADD$\ge v_3$. This is because when ADD$\ge v_3$, the samples of sensor 3 also contain information about whether there are no less than η affected sensors (although the local CuSum statistic at sensor 3 is not large), and this information is used by the S-CuSum algorithm but not the multichart CuSum algorithm.

6. REFERENCES

[1] N. A. Christakis and J. H. Fowler, "Social network sensors for early detection of contagious outbreaks," *PLOS ONE*, vol. 5, no. 9, pp. e12948, 2010.

[2] F. Pervaiz, M. Pervaiz, N. A. Rehman, and U. Saif, "Flubreaks: Early epidemic detection from google flu trends," *Journal of Medical Internet Research*, vol. 14, no. 5, 2012.

[3] J. B. Campbell and R. H. Wynne, *Introduction to Remote Sensing*, Guilford Press, 2011.

[4] A. G. Tartakovsky and V. V. Veeravalli, "Change-point detection in multichannel and distributed systems," *Applied Sequential Methodologies: Real-World Examples with Data Analysis*, vol. 173, pp. 339–370, 2004.

[5] A. G. Tartakovsky, B. L. Rozovskii, R. B. Blazek, and H. Kim, "A novel approach to detection of intrusions in computer networks via adaptive sequential and batch-sequential change-point detection methods," *IEEE Trans. Signal Proc.*, vol. 54, no. 9, pp. 3372–3382, 2006.

[6] Y. Mei, "Efficient scalable schemes for monitoring a large number of data streams," *Biometrika*, vol. 97, no. 2, pp. 419–433, 2010.

[7] Y. Xie and D. Siegmund, "Sequential multi-sensor change-point detection," *The Annals of Statistics*, pp. 670–692, 2013.

[8] G. Fellouris and G. Sokolov, "Second-order asymptotic optimality in multisensor sequential change detection," *IEEE Trans. Inform. Theory*, vol. 62, no. 6, pp. 3662–3675, 2016.

[9] V. Raghavan and V. V. Veeravalli, "Quickest change detection of a Markov process across a sensor array," *IEEE Trans. Inform. Theory*, vol. 56, no. 4, pp. 1961–1981, 2010.

[10] O. Hadjiliadis, H. Zhang, and H. V. Poor, "One shot schemes for decentralized quickest change detection," *IEEE Trans. Inform. Theory*, vol. 55, no. 7, pp. 3346–3359, 2009.

[11] M. Ludkovski, "Bayesian quickest detection in sensor arrays," *Sequential Analysis*, vol. 31, no. 4, pp. 481–504, 2012.

[12] S. Zou, G. Fellouris, and V. V. Veeravalli, "Asymptotic optimality of D-CuSum for quickest change detection under transient dynamics," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Aachen, Germany, 2017, IEEE, pp. 156–160.

[13] G. Rovatsos, S. Zou, and V. V. Veeravalli, "Quickest change detection under transient dynamics," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, New Orleans, USA, Mar. 2017, IEEE, pp. 4785–4789.

[14] G. Rovatsos, J. Jiang, A. D. Domínguez-García, and V. V. Veeravalli, "Statistical power system line outage detection under transient dynamics," *IEEE Trans. Signal Proc.*, vol. 65, no. 11, pp. 2787–2797, June 2017.

[15] G. V. Moustakides and V. V. Veeravalli, "Sequentially detecting transitory changes," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*. IEEE, 2016, pp. 26–30.

[16] E. S. Page, "Continuous inspection schemes," *Biometrika*, vol. 41, pp. 100–115, 1954.

[17] M. Pollak, "Optimal detection of a change in distribution," *The Annals of Statistics*, pp. 206–227, 1985.

[18] M. Blum, R. W. Floyd, V. Pratt, R. L. Rivest, and R. E. Tarjan, "Time bounds for selection," *Journal of Computer and System Sciences*, vol. 7, no. 4, pp. 448–461, 1973.

[19] T. L. Lai, "Information bounds and quick detection of parameter changes in stochastic systems," *IEEE Trans. Inform. Theory*, vol. 44, no. 7, pp. 2917–2929, 1998.

[20] R. H. Farrell, "Limit theorems for stopped random walks," *The Annals of Mathematical Statistics*, pp. 1332–1343, 1964.

[21] A. G. Tartakovsky, "Asymptotic performance of a multichart CUSUM test under false alarm probability constraint," in *Proc. 44th IEEE Conference on Decision and Control*. IEEE, 2005, pp. 320–325.