# Crowdsensing based prompt emergency discovery: a sequential detection approach

Rawan F. El Khatib[1], Nizar Zorba[2], and Hossam S. Hassanein[3]

[1]Department of Electrical and Computer Engineering, Queen's University, Canada
[2]Department of Electrical Engineering, Qatar University, Qatar
[3]School of Computing, Queen's University, Canada
18rek@queensu.ca, nizarz@qu.edu.qa, hossam@cs.queensu.ca

***Abstract*—The growth in the number of smart devices has mobilized the rise of CrowdSensing (CS) as an enabler of smart cities, where state-of-the-art technologies are utilized to improve citizens' quality of life. CS is a novel sensing paradigm that leverages data collected from smart devices to support a wide range of services. Particularly, smart emergency management systems are attracting increasing attention due to their potential to save lives, as they accelerate the delivery of emergency services including detection, mitigation and recovery. In this paper, we study the problem of the detection of an abnormal change in a monitored sensory variable, where the change is suggestive of an emergency situation. Specifically, we formulate our problem as a sequential change-point detection problem, where the underlying distribution of the variable changes at an unknown time. Our aim is to detect the change-point with minimal delay, subject to certain performance constraints. We utilize Shiryaev's optimal solution in two variants of the problem depending on the mobility behaviour of the participants, and conduct simulation experiments to show the performance of our schemes.**

## I. Introduction

The upsurge of sensor-enhanced smart devices has facilitated the rise of *CrowdSensing* (CS) as an enabling paradigm for smart cities. A smart city aims to offer its citizens an improved quality of life by providing intelligent, efficient and sustainable environments. Service providers in a smart city exploit various state-of-the-art technologies and paradigms to collect and analyze data to progress towards prosperous urban communities [1].

In essence, a CS framework consists of a central server which we refer to as the *CS organizer*, and a crowd of smart device owners, referred to as the *participants*. Participants contribute sensory data by leveraging the sensing, computing and communications capabilities of their smart devices [2]. This allows the acquisition of real-time sensory data used by service providers in a multitude of applications. Examples include smart transportation, smart healthcare, and safe smart cities; which is the focus of this work [3].

In a smart city, CS can be employed to improve the functionalities of emergency services by increasing situational awareness. In particular, CS improves core management operations, including a) the prompt detection of irregularities associated with an imminent emergency situation, b) provision of real-time data for continuous situation monitoring and c) the mitigation of damages by dynamic recovery planning [4].

In this work, we are concerned with the detection of a change in some given observed phenomena via sensor-enhanced smart devices, indicating a forthcoming or an existing emergency situation. As a motivational example, consider the entrance of a shopping mall where the CS organizer continuously monitors the noise levels by utilizing the microphones in participants' smart devices. Here, an abnormally high noise measurement may be indicative of high density of shoppers in that particular area, which may be a sign of a large gathering that warrants special attention from concerned authorities. Another example involves monitoring the temperature value in a park, where an unexpectedly high value could signify a risk of a fire in that particular location. In either case, it is desirable that the CS organizer detects the change in the reported measurements as soon as it occurs and alerts the concerned authorities to take appropriate courses of action.

Intuitively, we want the CS organizer to implement a scheme that detects the change in the monitored variable with high reliability, measured by the probability of an erroneous decision. Specifically, if the CS organizer incorrectly declares that a change has happened, it is said that a false alarm has occurred. The false alarms rate is a key performance metric in emergency management systems, since a false alarm causes unnecessary calling of service providers and responders, which in turn leads to waste of resources. On the other hand, if the CS organizer successfully recognizes a change in the monitored variable, then there exists a decision delay, which is equal to the difference between the time of the detection and the actual time the change happened. Due to the context of the scenario in emergency management systems, it is trivial that the detection delay is also pivotal to the performance of the scheme.

Therefore, a trade-off emerges between the quality of the detection procedure (as measured by the probability of a false alarm), and the delay until the CS organizer raises an alarm due to a detected change. Our goal is to devise a scheme that minimizes the decision delay within a given false alarm probability threshold. Towards this end, we adopt a *sequential detection* approach [5]. In sequential detection, data samples are observed sequentially until enough samples have been collected to stop further data acquisition and declare a decision. Specifically, we formulate our problem as a sequential change-

978-1-7281-0962-6/19/$31.00 ©2019 IEEE

point detection problem, and present two variants depending on the mobility behaviour of the participants contributing the sensory data.

The remainder of the paper is organized as follows. In section II, we discuss related work. In section III, we introduce some preliminaries of the discussion, followed by the problem formulation and solution in section IV. Section V tackles the simulation results. Finally, section VI presents our conclusions.

## II. Related Work

The notion of a smart city revolves around the engagement and integration of various technologies to enable safer and more sustainable environments. Hence, the concept of smart safe cities which incorporate smart emergency management schemes has gained increasing attention recently.

The work in [6] surveys state-of-the-art IoT based protocols appropriate for use in disaster management scenarios. The authors discuss schemes specific to natural and man-made disasters, as well as service-oriented and post-disaster management frameworks. Relevant issues include early warning and notification, remote monitoring, real-time analytics, and victim localization. The authors in [4] propose a vision of an information infrastructure that supports emergency responders in large-scale unstructured crisis situations. The proposed infrastructure aims to transform the operations of emergency management by exploiting large-scale CS, heterogeneous data integration and analytics.

Another work [7] underlines that a system based solely on municipal IoT components might suffer in crisis situations due to breakages in connectivity. Hence, the authors propose a novel end-to-end architecture that relies on CS for continuous data acquisition. The infrastructure implements a data processing component for timely knowledge discovery from multimedia big data. This will aid the service providers in the prediction and detection of a disaster, and collaborate in the response and recovery missions.

While the above works deliver ambitious perspectives on the realization of smart emergency management schemes, other works focus on the realization of core operations related to the prediction and management of a hazardous situation. The scheme introduced in [8] evaluates the value of the data contributions made by participants in a fire event. The scheme exploits the physical attributes of a fire spread to assess whether the temperature readings provided by the participants are credible. This is achieved by partitioning the affected area into sectors and implementing inter- and intra-sector evaluation algorithms, where the final output is a reputation score for each participant.

The work presented in [9] presents energy-efficient algorithms for personal state recognition based on data collected from smartphones. Specifically, their scheme aims to detect cases when the sensory data indicates the owner has fallen to the ground as opposed to walking or running, and alerts first responders once a fall has been confirmed.

The work in [10] introduces a cost-efficient fine-tuned fire detection convolutional neural network architecture for data drawn from surveillance videos. Experimentation is performed on real-world datasets confirm the effectiveness of the proposed scheme compared to other traditional ones. However, we note here that such schemes need prior training using appropriate relevant data for effective performance, making them suitable in existing infrastructure environments like Wireless Sensor Networks (WSNs) or Closed-Circuit Television (CCTV) networks. However, in CS schemes, this may not always be feasible. Hence, in this work we construct a framework that allows for the detection of irregularities associated with emergencies based on data generated from CS participants.

## III. Preliminaries

We are concerned with the detection of a change in a given variable which we monitor through CS-generated data, where the change signifies a possible emergency situation. We establish a mathematical framework through which we can minimize the time required to detect the change constrained by a certain reliability condition. The framework is based on opportunistic sensing [**?**], where the organizer collects sensory data from the smart devices without direct interference by the participants. This is convenient in an emergency situation where people are less likely to pay attention to their smart devices. We assume that the sensing process includes the sampling of a single sensor in the smart device.

In reference to the shopping mall and park examples mentioned previously, there are two possible manners in which the organizer can administer the sensing process. The first approach is to take advantage of participants transiting the place of interest at every time slot to collect data. In other words, the organizer recruits a new set of participants at the beginning of each time slot to perform sensing. This approach is appropriate in situations where crowd members are available in the place of interest for a period of time that approximates one time slot. Consequently, this is suitable when the places of interest are spatially limited and are characterized by high mobility flow. For example, crowd members walking by the main entrance of a shopping mall will usually be available there for a few seconds, allowing the organizer to exploit their presence for a finite number of time slots. Therefore, the CS-contributed data is from a heterogeneous set of participants, which can be renewed at every time slot.

The second approach is to consider a fixed number of stationary participants in the place of interest until a change is detected. In this manner, the same set of participants periodically generates sensory data according to the organizer's schedule. Contrary to the former approach, this is suitable in scenarios where crowd members are inclined to stay at a certain place for longer durations sufficient to span a large number of sensing slots. Moreover, the place of interest can span a larger spatial area, where the considered participants are sparsely spread. For example, people visiting a park for a picnic averaging for an hour or more.

We observe that the fundamental difference between the two approaches is whether the source of sensory data might

TABLE I: Table of Notations

| Notation | Explanation |
|---|---|
| $\mathbf{x}_i$ | The set of sensory measurements received at time $i$ |
| $N_i$ | Number of participants at time $i$ |
| $X_i$ | Single data instance at time $i$ |
| $\boldsymbol{X}_j$ | Sequence of data instances until time $j$ |
| $\tau$ | The change-point |
| $f_0$ | The pre-change distribution of the sequence |
| $f_1$ | The post-change distribution of the sequence |
| $t_d$ | The time at which the organizer detects a change |
| $\Gamma$ | The time delay until a change is declared |
| $ADD(t_d)$ | The average detection delay |
| $P_{FA}$ | The probability of a false alarm |
| $\alpha$ | The threshold of $P_{FA}$ |
| $\lambda$ | The parameter of the geometric distribution |
| $p_j$ | The posterior probability that a change has occurred |
| $t_s$ | The Shiryaev optimal stopping time |
| $w_{n,i}$ | The normalized LOF weight |
| $\rho^k_{n\to m}$ | The $k$-reachability distance of $x_{n,i}$ |
| $LRD^k_n$ | The $k$ local reachability density of $x_{n,i}$ |
| $LOF^k_n$ | The $k$-LOF score of $x_{n,i}$ |

change at the beginning of each sensing slot. Moreover, the participants in the first approach are confined to a smaller physical space, compared to the second approach where they are distributed to cover a larger area. Consequently, we refer to the former approach as the *Centralized Expedient Sensing* (CES) scheme, and the latter as the *Distributed Recurring Sensing* (DRS) scheme. These subtle differences call for some distinctions in the mathematical treatment, as will be shown in the shortly.

## IV. Problem formulation and solution

In this section, we begin by formulating the sequential detection problem for the CES scheme and present the solution. Subsequently, we extend the formulation in accordance with the DRS scheme and discuss its solution. Table I shows notations employed in the paper.

### A. The Centralized Expedient Sensing (CES) Scheme

In reference to the shopping mall entrance example, we have established that the organizer expects to receive sensory data at each time slot from a different set of participants. As a result, the size of the set might change at each time slot. Hence, we need to introduce a mathematical framework that handles the heterogeneous and varying-size sets, such that we are able to detect the change in the monitored variable under certain performance constraints.

Specifically, let $i = 1, 2, ..$ denote the index of the current time slot. At $i$, the organizer receives the set $\mathbf{x}_i$ which consists of data measurements collected from $N_i$ participants, i.e. $\mathbf{x}_i = \{x_{1,i}, x_{2,i}, ..., x_{N_i,i}\}$. Our CES framework consists of two components. The first component utilizes a pre-processing technique that transforms the set $\mathbf{x}_i$ into a single data instance denoted by $X_i$. The second component, which is the core of our scheme, implements a solution that detects a change with minimal delay. For clarity, we first introduce the solution to the detection problem for the case when the organizer obtains sensory data from a single participant throughout the entire sensing process (i.e., for the case when $N_i = 1 \forall i$), and then introduce the pre-processing technique that enables us to extend the solution for all values of $N_i$.

We formulate our problem as a sequential quickest change-point detection problem [5]. Broadly speaking, detection problems involve the observation of data pertaining to a certain phenomenon in order to make a decision about it. Contrary to traditional detection problems, in sequential detection we do not know beforehand the number of data instances used in making a decision. Rather, there is an unknown number of data instances, and the decision when to stop observing is part of the detection procedure. Sequential detection is appropriate for scenarios when both the reliability of the decision and the decision delay are key to the performance. Moreover, the quickest change-point detection problem, is a special case of sequential problems. In such a problem, the distribution of the monitored variable changes at an unknown time, and we aim to raise an alarm as soon as the change occurs, hence minimizing the decision delay.

Let $\boldsymbol{X}_j = \{X_i : i = 1, 2, ..., j\}$ be a sequence of real random variables observed sequentially from a single participant in time slots $i = 1, 2, ..., j$. Here, $X_i$ resembles the value of the monitored variables obtained via a sensor in the smart device at time slot $i$. Initially, the sequence follows a distribution $f_0$, until a change occurs at an unknown time $\tau \in \mathbb{Z}^+$. Following the change, the random variables $X_\tau, X_{\tau+1}, ...$ follow a different distribution, denoted by $f_1$. At time slot $j$, we must choose between the two hypotheses:

$$\begin{array}{ll} H_0: & \boldsymbol{X}_j = \{X_i \sim f_0, i = 1, 2, ..., j\} \\ H_1: & \exists \tau \in \mathbb{Z}^+, \text{s.t.:} \\ & \boldsymbol{X}_j = \begin{cases} X_i \sim f_0, i = 1, 2, ..., \tau - 1 \\ X_i \sim f_1, i = \tau, \tau + 1, ..., j \end{cases} \end{array} \quad (1)$$

Let $t_d$ denote the time that a change is detected. If $t_d \geq \tau$, then there exists a detection delay $\Gamma = t_d - \tau$, where $\Gamma$ is a discrete random variable. We define the average detection delay as the conditional expectation of $\Gamma$, written as:

$$ADD(t_d) = \mathbb{E}\left[\Gamma \mid t_d > \tau\right] = \sum_{i=1}^{\infty} P(\tau = i)\, \mathbb{E}_i\left[t_d - i \mid t_d > i\right] \quad (2)$$

where $\mathbb{E}_i$ is the expectation when the change occurs at $i$. On the other hand, if a detection is incorrectly declared before a change to the variable actually happens, then a false alarm has occurred. In such a case, $t_d < \tau$, and the probability of a false alarm ($P_{FA}$) is:

$$P_{FA} = P(t_d < \tau) = \sum_{i=1}^{\infty} P(\tau = i) P_i(t_d < i) \quad (3)$$

where $P_i$ is the probability measure when the change occurs at $i$. Our objective is to devise a detection policy that identifies the change as soon as it occurs, while restricting the probability of making an erroneous decision. Therefore, we can formally present our optimization problem as:

$$\begin{array}{rl} \min & ADD(t_d) \\ \text{s.t.} & P_{FA} \leq \alpha \end{array} \quad (4)$$

where $0 < \alpha < 1$ is a threshold limiting $P_{FA}$. The above formulation is formally known as the quickest change-point detection problem [5]. While no general solutions have been found for the problem, there exists an explicit solution in the *Bayesian formulation*, where the change-point $\tau$ is assumed to be random with a known distribution. Specifically, $\tau$ is modelled as a geometric random variable with parameter $0 < \lambda < 1$. Thus, the probability that a change occurs at $i$ is:

$$P(\tau = i) = \lambda(1-\lambda)^{(i-1)}, i = 1, 2, ... \tag{5}$$

Given the above, and under the assumption that the random variables in $\boldsymbol{X}_j$ are i.i.d., the problem can be solved as follows. Let $p_j$ denote the *a posteriori* probability that a change has occurred before time $j$, given the sequence $\boldsymbol{X}_j$. Equivalently, $p_j = P(\tau \leq j \mid \boldsymbol{X}_j)$. It can be shown through Bayes' rule that $p_j$ can be recursively calculated as:

$$p_j = \frac{\big[p_{j-1} + (1-p_{j-1})\lambda\big]L_j}{\big[p_{j-1} + (1-p_{j-1})\lambda\big]L_j + (1-p_{j-1})(1-\lambda)} \tag{6}$$

where $L_i$ is the likelihood ratio between the post-change and pre-change distributions, found as $L_j = \frac{f_1(\boldsymbol{X}_j)}{f_0(\boldsymbol{X}_j)}$. Consequently, it is shown that a Lagrangian relaxation of (4) can be solved through dynamic programming to yield [5]:

$$t_s = \inf\{j \geq 1 : p_j \geq A_\alpha\} \tag{7}$$

where $\inf$ denotes the infimum of a set, $t_s$ is the optimal stopping time at which the organizer halts sensing process, and $0 < A_\alpha < 1$ is an appropriately chosen threshold that satisfies $P_{FA}(t_s) = \alpha$. This result is known as the optimal *Shiryaev test* for the quickest change-point problem [5]. In general, it is not trivial to find $A_\alpha$ that satisfies the condition on $P_{FA}$. However, it has been shown in [11] that setting $A_\alpha = 1 - \alpha$ provides a guarantee that $P_{FA}(t_s \mid A_\alpha) \leq \alpha$, which satisfies the constraint in Eq. (4).

Thus far, we assumed that the sequence $\boldsymbol{X}_j$ is generated from a single participant throughout the sensing process until a change is detected. However, as noted earlier, this is not the case in the CES scheme, where transiting participants provide a set of sensory measurements $\mathbf{x}_i$, whose size might vary for different $i$. Hence, we need to propose a pre-processing technique that is conducted on the set $\mathbf{x}_i$, such that it is transformed from $N_i$ measurements into a singular data instance $X_i$.

In particular, we wish that our pre-processing technique reflects the truth of the monitored variable at the place of interest. We achieve this via a weighted average as follows:

$$X_i = \sum_{n=1}^{N_i} w_{n,i} x_{n,i} \tag{8}$$

where $w_{n,i} \in [0, 1]$ is the Local Outlier Factor (LOF) score [12]. The LOF is a consensus-based outlier detection technique, whose output is a measure of the consistency of the sample $x_n$ among other samples. The LOF is an attractive choice because it is computationally efficient, and it does not require the knowledge of ground-truth to assess the data measurements. Rather, the LOF gauges the distance-based deviation of any data measurement compared to its neighbours.

Specifically, let $d(n,m)$ represent the euclidean distance between the two sensory data measurements $x_{n,i}$ and $x_{m,i}$. Additionally, let $d_n^k$ denotes the distance between $x_{n,i}$ and its $k$th neighbour. Then, we define $E_n^k$ as the set of all data measurements in $\mathbf{x}_i$ whose distance to $x_{n,i}$ is less than or equal to $d_n^k$. Based upon which, we define the $k$-reachability distance of $x_{n,i}$ as:

$$\rho_{n \to m}^k = max\{d(n,m), d_m^k\} \tag{9}$$

Additionally, let the $k$-Local Reachability Density (LRD) of $n$ be the inverse of the average reachability distances in $x_{n,i}$'s neighbourhood, found as:

$$LRD_n^k = \frac{|E_n^k|}{\sum_{\forall m \in E_n^k} \rho_{n \to m}^k} \tag{10}$$

Then, we can find the LOF of $x_{n,i}$ as the average of the ratio of $k$-local reachability densities of $x_{n,i}$ and its $k$ neighbourhood. Mathematically:

$$LOF_n^k = \frac{\sum_{\forall m \in E_n^k} LRD_m^k}{LRD_n^k \cdot |E_n^k|} \tag{11}$$

where $LOF_n^k$ is an outlierness measure of $x_{n,i}$ in $[0, \infty)$. Particularly, an LOF measure of 1 indicates an inlier that is perfectly consistent with the rest of the measurements in $\mathbf{x}_i$. On the hand, $LOF_n^k >> 1$ indicates that $x_{n,i}$ is an outlier. To ensure reliable performance, we repeat this procedure for multiple values of the $k$ parameter in $\mathbf{k}$ and take the maximum, as shown in Algorithm 1, line 1 to line 9.

To be able to use the LOF score in Eq. (8), we must convert the LOF measure into the range $[0, 1]$. We achieve this via the normalization and regularization procedure presented in [13]. Specifically, this procedure normalizes the LOF value while improving the contrast between inliers and outliers. This is achieved by projecting each LOF value onto a Gaussian distribution $F_G(\mu, \sigma)$. Here, $\mu$ and $\sigma$ are the mean and variance of all the LOF scores, respectively. Through Algorithm 1, line 10 to line 12, we obtain $w_{n,i}$ whose value is around 0 for outliers and 1 for inliers.

It is vital to emphasize that the LOF pre-processing component is important in the CES scheme because of the transiting nature of the participants. Even if we restrict $N_i$ to be constant over all time slots, the sensory data in $\mathbf{x}_i$ is unlikely to be generated by the same set of participants. Hence, using the LOF algorithm to project the set onto a single-dimensional variable allows us to overcome this challenge without restrictions on the set or number of participants.

In summary, the organizer is to execute Algorithm 1 for the set $\mathbf{x}_i$ to obtain $X_i$ at each time slot. Then, the organizer recursively evaluates $p_j$ as per Eq. (6) until the optimal stopping time is found according to the optimal Shiryaev test in Eq. (7).

**Algorithm 1** The LOF algorithm

**Input:** $\mathbf{x}_i$, $\mathbf{k}$

**Output:** $w_{n,i}\ \forall n \in \mathbf{x}_i$

**for all** $x_{n,i} \in \mathbf{x}_i$ **do**
  **for all** $k \in \mathbf{k}$ **do**
    **for all** $m \in \mathbf{x}_i, m \neq n$ **do**
      $\rho^k_{n\to m} = max\{d(n,m), d^k_m\}$
      $LRD^k_n = \frac{|E^k_n|}{\sum_{\forall m \in E^k_n} \rho^k_{n\to m}}$
    **end for**
    $LOF^k_n = \frac{\sum_{\forall m \in E^k_n} LRD^k_m}{LRD^k_n \cdot |E^k_n|}$
  **end for**
  $LOF_n = max\{LOF^k_n, \forall k \in \mathbf{k}\}$
  $LOF_n = max\{0, LOF_n - 1\}$
  $\hat{w}_{n,i} = max\{0, 0.5F_G(LOF_n) - 1\}$
  $w_{n,i} = 1 - \hat{w}_n$
**end for**

### B. The Distributed Recurring Sensing (DRS) Scheme

Contrary to the CES scheme explained in the previous section, in the DRS scheme the organizer recruits a fixed number of participants who are committed to their location until a decision is made. As noted before, this is convenient in scenarios where the place of interest stretches over a relatively larger space, such that the organizer requires multiple participants distributed over different sub-locations. This implies that merging the data measurements contributed from each sub-location using Algorithm 1 will cause loss of information.

In other words, our question in the CES scheme is whether a change has been detected in the place of interest (time domain). While in DRS scheme, our question is whether a change has been detected in any of the sub-locations of the place of interest (spatial domain). Thus, the CES scheme carries the detection policy on a single observation location over several times, while the DRS carries the detection policy on every stream originating from each participant at a specific time.

The CES scheme conducts the detection procedure at the central entity, i.e., the organizer. On the other hand, the DRS scheme performs the detection procedure in a distributed manner across the scattered participants. In general, it is shown that the sequential change-point detection problem in (4) in a distributed environment is computationally intractable even in its simplest form [5].

Let us assume that the place of interest is divided into $L$ sub-locations. In the beginning of the sensing process, the organizer recruits one participant who will stay at his designated sub-location until the organizer halts the sensing process in return for monetary incentives. If a participant leaves the assigned sub-location, the organizer may choose another appropriate replacement. Furthermore, the organizer may recruit more than one participant at each sub-location and combine their sensory measurements using Algorithm 1, to ensure the sensing process is not interrupted by participants leaving their respective sub-locations. Without loss of generality, we assume that the organizer recruits $L$ participants, where participant $l$ is dedicated to stay in his/her sub-location until the change is detected.

Following with the notation in the previous section, at time slot $j$, participant $l$ has generated the sequence $Y^l_j = \{Y^l_1, Y^l_2, ..., Y^l_j\}$ of real random variables, where each element corresponds to a single sensory data measurement. Each participant leverages the computing abilities in his smart device to locally perform the optimal detection procedure in Eq. (7) on his own sequence $Y^l_j$. Our interest now shifts to construct a fusion rule which combines the local decisions from each participant. We consider two rules from the detection theory literature, as follows:

- $t_{min}$: the organizer halts the sensing process for all participants and declares that a change is detected as soon as one of the $L$ participants achieves the threshold on $P_{FA}$.
- $t_{max}$: the organizer halts the sensing process at an individual participant as soon as he achieves the threshold on $P_{FA}$, and declares a change is detected once all participants have achieved their respective thresholds.

We observe that the two fusion rules are applicable in slightly different scenarios. Specifically, $t_{min}$ is appropriate for scenarios where a detection from a single participant is worthy of raising an alarm, whereas $t_{max}$ is suitable for scenarios where unanimity of the detection must be reached. Furthermore, it is noteworthy that $t_{max}$ is shown to be globally first order asymptotically optimal, given that an appropriate threshold on $P_{FA}$ is chosen. On the contrary, the fusion rule $t_{min}$ does not possess this asymptotic optimality property [5]. Inevitably, $t_{max}$ entails longer delays than $t_{min}$.

## V. Performance Evaluation

We first introduce our simulation environment and parameters, then present performance evaluation results.

### A. Simulation Environment

We conduct simulations experiments to evaluate the proposed schemes. The CES scheme collects sensory measurements from a set of participants whose number $N_i$ varies in $\{1, 2, ..., 10\}$. These $N_i$ measurements are combine into a single data instance using the LOF algorithm where $\mathbf{k}$ is varied in $\{2, 3, 4\}$ when $N_i > 3$. If $N_i \leq 3$, then we set $w_{n,i} = 1/N_i$. In the DRS scheme, we assume we have $L = 10$ sub-locations and corresponding participants. We assume the pre- and post-change distributions are Gaussian with $f_0 \sim \mathcal{N}(0, 1)$ and $f_1 \sim \mathcal{N}(0.5, 1)$, respectively. Furthermore, we set the geometric distribution parameter $\lambda$ to 0.1, and we vary the false alarm threshold $\alpha$ in $(0, 0.2]$.

### B. Simulation Results

We begin by plotting the evolution of the Shiryaev's test statistic $p_j$ in Fig. 1 for a sequence observed from a single participant. We set the change-point $\tau$ to 100, where the

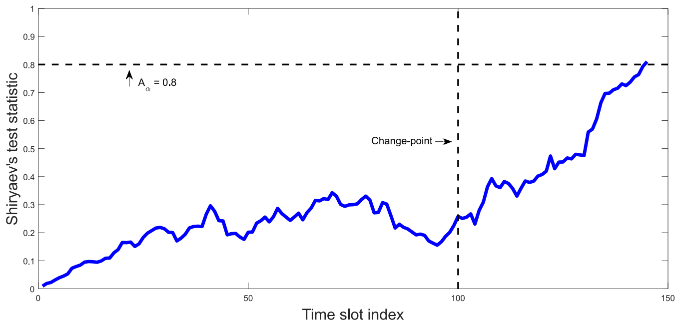


Fig. 1: The evolution of Shiryaev's statistic with $i$.

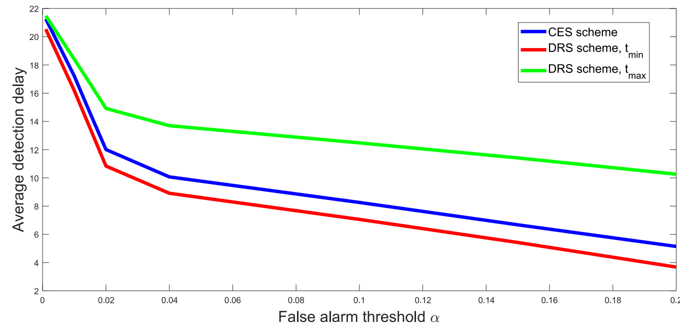


Fig. 2: ADD vs. $\alpha$.

underlying distribution changes from $f_0$ to $f_1$. Moreover, we set the Shiryaev's threshold $A_\alpha = 0.8$, indicating that $P_{FA} = 0.2$. By examining the figure, we note that the statistic stays relatively low until the change-point $\tau$. For subsequent time slots, the statistic $p_j$ grows until it exceeds the threshold $A_\alpha$ at $j = 145$. In other words, the organizer had a delay of $\Gamma = 145-100 = 45$ additional time slots to detect the change. From this figure, it is obvious that choosing a higher value for $A_\alpha$ achieves a lower false alarm rate, but at the expense of longer delay until the Shiryaev's statistic exceeds $A_\alpha$.

In Fig. 2, we perform Monte Carlo simulations to study the behaviour of the ADD versus the false alarm threshold $\alpha$ for both our proposed schemes; CES and DRS. For DRS, we tackle both options of $t_{max}$ and $t_{min}$. By examining the figure, we notice that all schemes follow the same behaviour, as false alarm threshold constraint becomes more relaxed, lesser delay is achieved to detect the change. In addition, we note that the DRS scheme following the $t_{max}$ fusion rule has the largest delay. This is expected since this scheme waits for all the $L$ participants to achieve the threshold on $A_\alpha$. On the other hand, the DRS scheme following the $t_{min}$ fusion rule and the CES scheme achieve comparable performance, with the $t_{min}$ scheme attaining slightly lower delays. This can be attributed to the fact that with $L = 10$ participants, it is more probable that one of these participants will detect the change faster than a single-stream data sequence observation as in the CES scheme.

## VI. Conclusions

In this paper, we studied the detection of a change in an observed variable based on CrowdSensing (CS) generated data, where the change signifies a forthcoming emergency situation. Our problem was formulated as a sequential change-point detection problem, where the distribution of the variable changes at an unknown time. Two variants of the problem were introduced based on the mobility behaviour of the CS participants: Centralized Expedient Sensing (CES) and Distributed Recurring Sensing (DRS). In both variants, we utilized Shiryaev's test to minimize the average detection delay under false alarm constraints. We conducted Monte Carlo simulation experiments to show the performance of our two proposals in balancing the trade-off between the detection delay and the false alarm constraint.

## Acknowledgment

This work was made possible by NPRP grant NPRP 9-185-2-096 from the Qatar National Research Fund (a member of The Qatar Foundation). The statements made herein are solely the responsibility of the authors.

## References

[1] A. Gharaibeh et al., "Smart Cities: A Survey on Data Management, Security, and Enabling Technologies," in *IEEE Communications Surveys & Tutorials,* vol. 19, no. 4, pp. 2456-2501, 2017.

[2] R. F. ElKhatib, N. Zorba and H. S. Hassanein, "A Fair Reputation-based Incentive Mechanism for Cooperative Crowd Sensing", in *Proc. of IEEE Global Communications Conf. (GLOBECOM)*, UAE, Dec., 2018.

[3] H. S. Hassanein, N. Zorba, S. Han, S. Kanhere and M. Shukair, "Crowd Management," in *IEEE Communications Magazine,* vol. 57, no. 4, pp. 18-19, 2019.

[4] M. Abu-Elkheir, H. S. Hassanein and S. M. A. Oteafy, "Enhancing Emergency Response Systems through Leveraging Crowdsensing and Heterogeneous data," in *Proc. of Int. Wireless Communications and Mobile Computing Conf. (IWCMC)*, Cyprus, Sep., 2016.

[5] V. V. Veeravalli and T. Banerjee, "Quickest Change Detection" in *Academic Press Library in Signal Processing: Array and Statistical Signal Processing,* 1st ed. USA: Academic Press, 2014, ch. 6, pp. 209-252.

[6] P. P. Ray, M. Mukherjee and L. Shu, , "Internet of Things for Disaster Management: State-of-the-art and Prospects," in *IEEE Access*, vol. 5, pp.818–835, 2017.

[7] A. Boukerche and R. W. L. Coutinho, "Smart Disaster Detection and Response System for Smart Cities," in *Proc. of IEEE Symposium on Computers and Communications (ISCC)*, Brazil, Jun., 2018.

[8] M.A. Alswailim, H.S. Hassanein, and M. Zulkernine, "A Participant Contribution Trust Scheme for Crisis Response Systems," in *IEEE Global Communications Conf. (GLOBECOM)*, Singapore, Dec., 2017.

[9] L. García et al., "System for Detection of Emergency Situations in Smart City Environments Employing Smartphones," in *Proc. Int. Conf. on Advances in Computing, Communications and Informatics (ICACCI)*, India, Sep., 2018.

[10] K. Muhammad et al., "Convolutional Neural Networks Based Fire Detection in Surveillance Videos," in *IEEE Access*, vol. 6, pp. 18174-18183, 2018.

[11] A. G. Tartakovsky and V. V. Veeravalli, "Quickest Change Detection in Distributed Sensor Systems," in *Proc. of Sixth Int. Conf. of Information Fusion*, Australia, May, 2003.

[12] M. M. Breunig et al., "LOF: identifying density-based local outliers," in *Proc. of the ACM SIGMOD Int. Conf. on Management of Data*, USA, May, 2000.

[13] H.-P. Kriegel et al., "Interpreting and unifying outlier scores," in *Proc. of the SIAM Int. Conf. on Data Mining*, USA, Apr., 2011.