Bandwidth-Constrained Distributed Quickest Change Detection in Heterogeneous Sensor Networks: Anonymous vs Non-Anonymous Settings

Wen-Hsuan Li and Yu-Chih Huang
Institute of Communications Engineering
National Yang Ming Chiao Tung University, Taiwan
E-mail: {vincent, jerryhuang}@nycu.edu.tw

Abstract—The heterogeneous distributed quickest change detection (HetDQCD) problem with 1-bit feedback is studied, in which a fusion center monitors an abrupt change through a bunch of heterogeneous sensors via anonymous 1-bit feedbacks. Two fusion rules, one-shot and voting rules, are considered. We analyze the performance in terms of the worst-case expected detection delay and the average run length to false alarm for the two fusion rules. Our analysis unveils the mixed impact of involving more sensors into the decision and enables us to find near optimal choices of parameters in the two schemes. Notably, it is shown that, in contrast to the homogeneous setting, the first alarm rule may no longer lead to the best performance among one-shot schemes. The non-anonymous setting is then investigated where a novel weighted voting rule is proposed that assigns different weights to votes from different types of sensors. Simulation results show that the proposed scheme is able to outperform all the above schemes and the mixture CUSUM scheme for the anonymous HetDQCD, hinting at the price of anonymity.

I. INTRODUCTION

Quickest change detection (QCD) investigates efficient detection of an abrupt change under a tolerant false alarm constraint. Such a classical problem in statistical analysis and signal processing has numerous applications [1]–[3], and several fundamental results have been proposed from early works [4]–[6]. In particular, Page in [4] proposed the well-known cumulative sum (CUSUM) test. For the criterion in [5], Lorden [5] and Moustakides [6] proved the asymptotic and exact optimality of the CUSUM test, respectively.

With the rise of Internet of Things (IoT) and distributed learning/computing, the requirement of efficient anomaly detection has become exceptionally crucial, and QCD has found its applications in such distributed systems [7]. For example, the smart grid is suffered from the risk of cyberattack, and in order to avoid the damage due to critical voltage variation, one can detect the line outage with low latency by applying QCD algorithm on certain small-signal power model [8]. However, the task of detection is usually carried out by a fusion center through distributed sensors who feedback their observations/decisions via band-limited links. To overcome bandwidth constraints, Mei [9] proposed a family of schemes where the sensors feedback the binary messages that whether the local CUSUM statistics reach a certain predefined threshold. When the fusion center applied consensus rule to combine feedback messages, Mei [9] verified the asymptotic optimality of this stopping rule. However, the simulations in [9] and [10] showed that the consensus rule is not exactly optimal. In [11], Banerjee and Fellouris considered the same problem and further investigated the performance of the one-shot and voting rules with various number of alarms. The second order analysis in [11] indicated that the first alarm is the best choice among one-shot schemes, while the majority vote rule is a reasonably good strategy among the voting rules.

Another challenge in distributed QCD is the difference among the quality of information sent from the sensors. This discrepancy could originate from several sources such as transmission error/noise, sensing abilities, cyberattacks, etc. Such a problem with Byzantine attacks being the source of discrepancy has been investigated in [12] and [13]. The heterogeneous distributed QCD (HetDQCD) problem in [14] modeled the heterogeneity by assuming each sensor’s observations are drawn from a possibly non-identical but independent distribution. Sun et al. in [14] showed that the mixture CUSUM procedure is an optimal anonymous rule when the observations are directly available at the fusion center.

Two questions naturally arise from this premise: 1) How to anonymously detect the change efficiently in HetDQCD under bandwidth constraints? 2) Does knowing the identities of the sensors (i.e., non-anonymous setting) help us achieve a better performance? In this paper, we partially answer the above two questions for HetDQCD with 1-bit feedback. To address the first question, both the one-shot and voting rules in the anonymous HetDQCD are considered and analyzed. Our first order analysis reveals appropriate choices of the ratios between thresholds of different sensors in both the one-shot and voting schemes. Furthermore, based on our second order analysis, the (asymptotically) optimal choice of the number of alarms we should wait in the one-shot scheme and that of the number of votes in the voting rule are derived. Our results indicate that among the one-shot schemes, the first alarm scheme may no longer be the best, which is in sharp contrast to...
the homogeneous setting. To address the second question, two schemes which exploit the knowledge of sensors’ identities are introduced. The first scheme only accept alarms from the most informative sensors, while the second one further generalizes
the idea to weight the votes according to the informativeness
of the observations. Simulation results show that the proposed
schemes which exploit the knowledge of sensors’ identities are
introduced. The first scheme only accept alarms from the most
informative sensors, while the second one further generalizes
the idea to weight the votes according to the informativeness
of the observations. Simulation results show that the proposed
schemes which exploit the knowledge of sensors’ identities are
introduced. The first scheme only accept alarms from the most
informative sensors, while the second one further generalizes

Finally, Section VI concludes the paper.

IV, which are verified via computer simulations in Section V.

The analysis results and some discussions are shown in Section
II states the network model and problem formulation. Section

Anonymous HetDQCD and the mixture CUSUM, hinting at
weighted voting rule (even the former simple special scheme
of the observations. Simulation results show that the proposed
schemes which exploit the knowledge of sensors’ identities are
introduced. The first scheme only accept alarms from the most
informative sensors, while the second one further generalizes

A. Problem formulation

We consider the heterogeneous sensor network as illustrated
in Fig. 1, where \( N \) sensors assist the fusion center in detecting
an abrupt change occurring at time \( \nu \). There are total \( L \)
heterogeneous types and thereby the \( N \) sensors can be divided
into \( L \) groups, namely \( G_1, \ldots, G_L \), according to their types.
Let \( N_l \) be the number of sensors of type \( l \in [L] \) and we use
the pair \((k, l)\) to denote the index of sensor \( k \in [N_l] \) in group
\( l \in [L] \). The sensor \((k, l)\) will observe \( X_{k,l}^t \) at time \( t \) that is
drawn independently according to

\[
X_{k,l}^t \sim \begin{cases} f_l, & t \leq \nu, \\ g_l, & t > \nu, \end{cases}
\]

where \( f_l \) and \( g_l \) are the pre-changed and post-changed distributions, respectively, and they have the positive Kullback-Leibler divergence (KLD) \( \mathcal{I}_l > 0 \). At time \( t \), each sensor can communicate an 1-bit message with the fusion center noiselessly. A fusion rule \( \rho \) is then adopted by the fusion center to form a decision according to the feedbacks. Specifically, let \( \rho \) be an \( \mathcal{F}_t \)-stopping time, in which

\[
\mathcal{F}_t \equiv \sigma(X_{k,l}^s; s \in [t], k \in [N_l], l \in [L]),
\]

is the \( \sigma \)-algebra from all possible observations of all sensors
up to time \( t \). We denote by \( P_\nu \) the underlying probability measure that the event happens at time \( \nu \). In particular, if
no event occurs, we set \( \nu = \infty \) and denote by \( P_\infty \) the underlying probability measure. Two metrics are defined to
measure the performance of \( \rho \), namely the average run length
(ARL) to false alarm and the worst-case expected detection
delay (EDD), as

\[
\text{EDD}(\rho) = \sup_{\nu \geq 0} \mathbb{E}_\nu[(\rho - \nu)^+ | \mathcal{F}_\nu],
\]

respectively, where \( \mathbb{E}_\nu \) is the expectation under \( P_\nu \). Following
Lorden’s setting [5], the problem of QCD is to minimize
EDD(\rho) subject to a constraint ARL(\rho) > \gamma.

B. Local CUSUM

Similar to most of the work in the literature such as [9],
[11], [12], each sensor first computes the CUSUM statistic locally. That is, let \( Z_{k,l}^t \) and \( W_{k,l}^t \) be the log-likelihood ratio (LLR) and the CUSUM statistic of the sensor \((k, l)\) at time \( t \), respectively, each sensor recursively computes

\[
W_{k,l}^t = \max\{0, W_{k,l}^{t-1} + Z_{k,l}^t\}, \quad W_{k,l}^0 \triangleq 0.
\]

To meet the bandwidth constraint, this CUSUM statistic cannot
be losslessly sent to the fusion center and some process must
be carried out. Due to the 1-bit feedback constraint, each
decision is made when at least \( h_l \) alarms received at time
\( t \), where \( h_l \) is the threshold for sensors in type \( l \in [L] \).

C. Anonymous fusion rules

At each time \( t \), based on the feedback signals it has received,
the fusion center makes a global decision about whether there is
a change or not. Let \( \mathbf{h} = [h_1, \ldots, h_L]^T \) be the threshold
vector, two families of fusion rules are considered based on
anonymous feedbacks [11]:

- Anonymous \( M \)-th alarm: Each sensor sends a local alarm
at the first time (hence, one-shot) its local CUSUM statistic
exceeds a predefined threshold and a global decision is made when at least \( M \) alarms are received. Mathematically, we have

\[
\rho^{(M)}(\mathbf{h}) = \inf\{t : \max_{(k, l)} |W_{k,l}^t - h_l| \geq M\}.
\]

- Anonymous \( M \) voting rule: Each sensor constantly sends
the local decision and a global decision is made at the
first time when there are at least \( M \) alarms received
simultaneously. Mathematically, we have

\[
\rho_M(\mathbf{h}) = \inf\{t : \max_{(k, l)} W_{k,l}^t > h_l \geq M\}.
\]

III. PROPOSED WEIGHTED FUSION RULE

The two fusion rules in Section II-C treat every sensor
equally, regardless of its type, and are thus suitable for the
anonymous setting. Once the identities of the sensors are
known to the fusion center, more potential fusion rules can
be applied. A naive strategy is selecting the top informative
sensors to avoid triggering the false alarm by other sensors.
Given the predetermined set \( D \) of selecting sensors, two fusion rules are proposed:

- **M-th alarm within \( D \):**
  \[
  \tilde{\rho}^{(M)}(h, D) = \inf \{ t : \{(k, l) \in D : \max_{i} W_{t}^{k,l} > h_{t}\} \geq M \}. \tag{7}
  \]

- **M voting rule within \( D \):**
  \[
  \rho_{M}(h, D) = \inf \{ t : \{(k, l) \in D : W_{t}^{k,l} > h_{t}\} \geq M \}. \tag{8}
  \]

Notice that the sensor selection scheme can be applied even if the fusion center does not have the prior knowledge of distributions but the order of informative sensors. If furthermore, the fusion center obtains the information of KLD of each group, design of various weights for selecting sensor can further enhance the performance. Given \( 0 \leq \alpha_{k,l} \leq 1 \) as the weight for the \( k \)th sensors in the \( l \)th group and let \( \alpha = [\alpha_{1,1}, ..., \alpha_{N_{L},L}]^{T} \), we can generalize the previous M voting rule to the weighted version:

- **Weighted M voting rule:**
  \[
  \tilde{\rho}_{M}(h, \alpha) = \inf \{ t : \sum_{k,l} \alpha_{k,l} 1\{ W_{t}^{k,l} > h_{t}\} \geq M \}, \tag{9}
  \]

in which \( 1\{ \cdot \} \) is the indicator function.

By setting \( \alpha_{k,l} = 1 \{(k, l) \in D\} \), the proposed weighted voting rule subsumes the selection scheme in (8) as a special case. As a heuristic, in our simulations, we propose setting the weights to be proportional to the KLDs, i.e., \( \alpha_{k,l} \propto \mathcal{I}_{l} \) for \( 1 \leq l \leq L \).

### IV. Analysis of Anonymous Fusion Rules

In this section, we investigate the performance of anonymous fusion rules. We first analyze the first order asymptotic performance of \( \rho^{(M)}(h) \) and \( \rho_{M}(h) \), which guide us to choose appropriate thresholds \( h_{t} \)'s that lead to the best first order asymptotic scaling. Under the threshold setting, a second order analysis is then investigated, which enables us to choose the suitable \( M \) in each family of anonymous fusion rules. Finally, using our analysis, the benefit of knowing the identities of sensors and the price of anonymity are discussed. The detailed analysis of non-anonymous fusion rules is left as future works.

Before proceeding to the analysis, we note that according to [12, Lemma 3], any fusion rule \( \rho \) based on the local CUSUM statistics (as those above) admits a simpler expression of the EDD as

\[
\text{EDD}(\rho) = E_{0}[\rho]. \tag{10}
\]

#### A. First Order Analysis

Our first main result is the upper and lower bounds on the first order asymptotic of the expected detection delay as a function of the mean time to a false alarm for both \( \rho^{(M)}(h) \) and \( \rho_{M}(h) \).

**Theorem 1.** Let \( h = [c_{1}h, \cdots, c_{L}h]^{T} \) with \( c_{1}/\mathcal{I}_{1} \leq c_{2}/\mathcal{I}_{2} \leq \cdots \leq c_{L}/\mathcal{I}_{L} \), \( c_{l} \in \mathbb{R}^{+} \) for all \( 1 \leq l \leq L \). Suppose that \( h = h_{\gamma,1} \) and \( h = h_{\gamma,2} \) lead to \( E_{\infty}[\rho^{(M)}(h)] = \gamma \) and \( E_{\infty}[\rho_{M}(h)] = \gamma \), respectively. Let \( c^{(1)} \leq c^{(2)} \leq \cdots \leq c^{(L)} \) be the rearrangement of \( c_{1}, c_{2} \cdots, c_{L} \), and let \( N^{(l)} \) and \( \mathcal{T}^{(l)} \) represent the amount of the sensors in the group and KLD corresponding to \( c^{(l)} \) respectively, \( 1 \leq l \leq L \). Define \( \lambda : [N] \rightarrow [L] \) by

\[
\lambda(m) = l \text{ if } \sum_{j=1}^{l-1} N^{(j)} \leq m \leq \sum_{j=1}^{l} N^{(j)}.
\]

Then as \( \gamma \rightarrow \infty \),

\[
h_{\gamma,1} \sim \frac{\log \gamma}{c^{(\lambda(M))}}, \quad h_{\gamma,2} \lesssim \sum_{m=1}^{M} c^{(\lambda(m))}(1 + o(1)). \tag{12}
\]

Moreover, we have

\[
E_{0}[\rho^{(M)}(h)] \in \left[ \frac{c_{1} \log \gamma}{\mathcal{I}_{1} c^{(\lambda(M))}}(1 + o(1)), \frac{c_{L} \log \gamma}{\mathcal{I}_{L} c^{(\lambda(M))}}(1 + o(1)) \right], \tag{14}
\]

\[
E_{0}[\rho_{M}(h)] \in \left[ \frac{c_{1} \mathcal{T}_{\lambda}(M) \log \gamma(1 + o(1))}{\mathcal{I}_{1} \sum_{m=1}^{M} c^{(\lambda(m))}}, \frac{c_{L} \mathcal{T}_{\lambda}(M) \log \gamma(1 + o(1))}{\mathcal{I}_{L} \sum_{m=1}^{M} c^{(\lambda(m))}} \right]. \tag{15}
\]

**Proof:** See Appendix in [15].

We notice that in Theorem 1, the ratio of thresholds between each group give a great effect in the upper and lower bound of detection delays in equations (14) and (15), and one suboptimal policy for choosing the thresholds is to minimize these upper bounds. This policy turns out to make the thresholds of each group be proportional to the KLD of the pre-changed and post-changed distributions in this group. From now on, we will choose \( c_{l} = \mathcal{I}_{l} \) for \( 1 \leq l \leq L \), and assume that \( \mathcal{I}_{1} \leq \cdots \leq \mathcal{I}_{L} \) for simplifying notation. This induces the following corollary.

**Corollary 2.** Let \( h = [\mathcal{I}_{1}h, \cdots, \mathcal{I}_{L}h]^{T} \). Define \( h_{\gamma,1} \) and \( h_{\gamma,2} \) as in Theorem 1. Then as \( \gamma \rightarrow \infty \),

\[
h_{\gamma,1} \sim \frac{\log \gamma}{\mathcal{I}_{\lambda}(M)}, \quad h_{\gamma,2} \sim \frac{\log \gamma}{\sum_{m=1}^{M} \mathcal{I}_{\lambda}(m)}. \tag{16}
\]

Furthermore,

\[
E_{0}[\rho^{(M)}(h)] \sim \frac{\log \gamma}{\mathcal{I}_{\lambda}(M)}, \quad E_{0}[\rho_{M}(h)] \sim \frac{\log \gamma}{\sum_{m=1}^{M} \mathcal{I}_{\lambda}(m)}. \tag{17}
\]

**Proof:** Setting \( c_{l} = \mathcal{I}_{l} \) for \( 1 \leq l \leq L \) completes the proof.

#### B. Second Order Analysis

The first order analysis in Theorem 1 has enabled us to set the thresholds for obtaining the best asymptotic scaling as shown in Corollary 2. However, in order to compare the performance of different choices of \( M \) under small to medium \( \gamma \), we need a finer analysis. Here, we present the second order analysis which will later help us choose the best \( M \).
Theorem 3. Let \( h = [I_1 h, \cdots, I_L h]^T \). Assume that
\[
\sigma_i^2 \equiv E_0 \left[ (Z_{i}^{k,l} - I_i)^2 \right] < \infty \quad \text{for all } k, l.
\] (19)

Then as \( h \to \infty \),
\[
E_0[\rho^{(M)}(h)] = h + \xi_M \sqrt{h} (1 + o(1)) \quad \text{and} \quad E_0[\rho_M(h)] \leq h + \xi_M \sqrt{h} (1 + o(1)),
\] (20)
\[
\text{where } \xi_M \text{ is the expected value of } M\text{-th order statistics of independent (but not necessarily identical) Gaussian random variables } G_{k,l} \sim N(0, \sigma_i^2 / I_i^2) \text{ for } 1 \leq k \leq N_i \text{ and } 1 \leq l \leq L.
\]

Proof: See Appendix in [15].

C. Discussions

We discuss the appropriate choice of \( M \) suggested by our analysis. With the asymptotic setting in (16), (20) and (21) become
\[
E_0[\rho^{(M)}(h)] = \left( \frac{\log \gamma}{I_x(M)} + \xi_M \sqrt{I_x(M)} \right) (1 + o(1)),
\] (22)
\[
E_0[\rho_M(h)] \leq \left( \frac{\log \gamma}{\sum_{m=1}^M I_x(M)} + \xi_M \sqrt{\sum_{m=1}^M I_x(M)} \right) (1 + o(1)),
\] (23)
respectively.

We first focus on the voting rule \( \rho_M(h) \). For sufficiently large \( \gamma \), the first order in (23) will eventually dominate and one should set \( M = N \) that minimizes the first term. On the other hand, for small to medium \( \gamma \), due to the impact of \( \xi_M \), one really has to solve for \( M \) that minimizes (23). For example, we can prove that \( \xi_M \) is increasing in \( M \) and is negative if \( M < N/2 \) and non-negative otherwise; thereby, the second order term in (22) and (23) will have negative impact on the detection delay when \( M \) increases. The analysis of \( \xi_M \) in the heterogeneous setting will allow us to quickly determine an appropriate \( M \) for a finite \( \gamma \) and is left for future work.

For the \( M \)-th alarm rule \( \rho^{(M)}(h) \), things become very interesting even for large \( \gamma \). Specifically, due to the piece-wise nature of \( \lambda(\cdot) \) in (11), any choice of \( 1 + \sum_{j=1}^{l-1} N^{(j)} \leq M \leq \sum_{j=1}^{l} N^{(j)} \) for \( 1 \leq l \leq L \) leads to the same first order term in (22). Within this region, one should pick
\[
M = 1 + \sum_{j=1}^{l-1} N^{(j)},
\] (24)
that results in the smallest \( \xi_M \). In particular, the best choice among all \( M \) occurs at \( M = N - N^{(L)} + 1 \). However, the best choice of \( M \) for any fixed \( \gamma \) again requires detailed analysis of \( \xi_M \). We summarize our observations as follows.

Corollary 4. For large \( \gamma \), appropriate choices of \( M \) for \( \rho^{(M)}(h) \) and \( \rho_M(h) \) according to (22) and (23) are \( M = N - N^{(L)} + 1 \) and \( M = N \), respectively.

In the sequel, we discuss the non-anonymous setting with the 1-bit bandwidth constraint, where the fusion center knows the identity of each sensor. The following result compares \( \hat{\rho}^{(1)} \) with \( \rho^{(M)} \).

Corollary 5. When \( N_L < N/2 \), for large \( \gamma \), the stopping rule \( \hat{\rho}^{(1)}(h, G_L) \) that accepts the first alarm from the most informative group \( G_L \) achieves a better tradeoff than \( \rho^{(M)}(h) \) for any \( M \).

Proof: Following the suggestion in Theorem 4, the minimal detection delay for \( M\)-th alarm is achieved by setting \( M = N - N_L + 1 \), which implies that as \( \gamma \) large enough,
\[
E_0[\rho^{(M)}(h)] \geq \left( \frac{\log \gamma}{I_x} + \xi_N - N_L + 1 \sqrt{\frac{\log \gamma}{I_x}} \right) (1 + o(1))
\]
\[
> \frac{\log \gamma}{I_x} (1 + o(1)) = E_0[\hat{\rho}^{(1)}(h, G_L)] \text{ when } N_L < N/2,
\] (25)
where (a) holds since \( \xi_M > 0 \) for \( M > N/2 \).

For \( \hat{\rho}_M \) and \( \rho_M \), we note that precise comparison still requires computation of \( \xi_M \). However, since \( \xi_M \) is an increasing function, it is likely that \( \hat{\rho}_M \) focusing on several most informative groups would outperform \( \rho_M \) for any \( M \) which blindly cooperate among all the sensors in the network.

V. SIMULATION RESULTS

We first examine the discussion in Corollary 4. Here we set \( L = 3 \) and \( N_i = 3 \) for each \( i \in \{1, 2, 3\} \). We first consider the \( M\)-th alarm rule \( \rho^{(M)}(h) \) with \( M \in \{1, 4, 7\} \) satisfying (24) and \( M = 9 \) involving all the sensors. In this simulation, the pre-changed distributions remain standard normal for all sensors and the post-changed distributions of group \( l \) are set to be \( g_l \sim N(2L/3, 1) \). As shown in Fig. 2, when \( \gamma \) is large enough, as predicted in Corollary 4, the best two curves correspond to \( M = 4 \) and \( M = 7 \) and \( M = 7 \) leads to the best performance, indicating a major departure from the homogeneous setting where it was shown in [12] that the first alarm is the best. Moreover, involving all the sensors (i.e., \( M = 9 \)) is also not a good choice. We then study the performance of the \( M \) voting rules \( \rho_M(h) \) with the
pre-changed distributions being the standard normal and the post-changed distributions $g_l \sim N(l/3, 1)$ for $l \in \{1, 2, 3\}$. We show in Fig. 3 that when $\gamma$ is sufficiently large, $M = 9$ is the best. However, when $\gamma$ is not large enough, $M = 9$ is not always optimal, exhibiting the impact from $\xi_M$ in (23).

The comparison between anonymous and non-anonymous settings is then investigated in Fig. 4. In this simulation, there are two groups consisting of 8 and 1 sensors, respectively. The pre-change distributions are again the standard normal distribution, while the post-change distributions for group 1 and group 2 are set to be $g_1 \sim N(0.35, 1)$ and $g_2 \sim N(2, 4)$, respectively. In this figure, the performance of $\rho^{(1)}(h)$, $\rho^{(2)}(h)$, and $\rho^{(3)}(h, G_L)$ are plotted. We note that we have tested $\rho^{(M)}(h)$ and $\rho_M(h)$ for all the choices of $M$ and $M = 9$ is the best for both the families in this setting. In addition, the performance of the mixture CUSUM procedure is also provided. We note that this procedure has been shown to be optimal for the anonymous setting when the observations are directly available at the fusion center [14]. Moreover, the performance of $\tilde{\rho}^{(1)}(h, G_2)$ (equivalent to $\tilde{\rho}_1(h, G_2)$ in this case as $N_2 = 1$) is plotted to reveal the benefit of non-anonymity. It is demonstrated that with the knowledge of sensors’ identities, by simply ignoring the sensors in $G_1$ and focusing solely on the sensor in $G_2$, $\tilde{\rho}^{(1)}(h, G_2)$ outperforms all the anonymous schemes considered, including the mixture CUSUM. This hints at the price of anonymity.

To further examine the performance enhancement of weighted voting and threshold setting, we demonstrate the simulation as shown in Fig 5. There are two groups with 6 and 4 sensors respectively. The pre-changed distribution still remains standard normal, and the post-changed distribution of group 1 and group 2 are $g_1 \sim N(0.55, 1)$ and $g_2 \sim N(1, 1)$ respectively. Following Section III, the weights for weighted voting rule are set to be proportional to the KLDs, and the largest one is set to be unity. We can see that both the threshold and the weight settings improve the performance. In particular, the proposed weight give a better performance than equal weight (i.e. anonymous voting) or binary weight (i.e. group selection), which shows the potential of cooperation.

VI. CONCLUSIONS

In this paper, we investigated the HetDQCD problem with 1-bit feedback. For the anonymous setting, we rigorously analyzed two families of stopping rules, the $M$-th alarm rule and the $M$ voting rule. Using our analysis, an appropriate choice of $M$ for each scheme was suggested. The non-anonymous setting had also been discussed, where it was shown that focusing on the most informative sensors could outperform all the sensors. A weighted voting rule based on KLDs was proposed and simulation results were provided to demonstrate that it can outperform all the schemes based on anonymous reports considered in this paper, including the mixture CUSUM. Potential future works include a thorough investigation of the proposed $\tilde{\rho}^{(M)}(h, D)$ and $\tilde{\rho}_M(h, D)$ for any $D$ in the non-anonymous setting, the optimal weight design for $\tilde{\rho}_M(h, \alpha)$, design of quantizers for HetDQCD with $b$-bit bandwidth constraint for $b > 1$, and taking into account the presence of compromised sensors.
REFERENCES

[1] L. Lai, Y. Fan, and H. V. Poor, “Quickest detection in cognitive radio: A sequential change detection framework,” in IEEE GLOBECOM 2008, 2008, pp. 1–5.

[2] A. Lakhina, M. Crovella, and C. Diot, “Diagnosing network-wide traffic anomalies,” ACM SIGCOMM Comput. Commun. Rev., vol. 34, no. 4, pp. 219–230, 2004.

[3] F. Nizam, S. Chaki, S. A. Mamun, and M. S. Kaiser, “Attack detection and prevention in the cyber physical system,” in Proc. Int. Conf. Comput. Commun. Informat. (ICCCI), 2016, pp. 1–6.

[4] E. S. Page, “Continuous inspection schemes,” Biometrika, vol. 41, no. 1, pp. 100–115, 1954.

[5] G. Lorden, “Procedures for reacting to a change in distribution,” Ann. Math. Statist., vol. 42, no. 6, pp. 1897–1908, Dec. 1971.

[6] G. V. Moustakides, “Optimal stopping times for detecting changes in distributions,” Ann. Math. Statist., vol. 14, no. 4, pp. 1379–1387, Dec. 1971.

[7] L. Xie, S. Zou, Y. Xie, and V. V. Veeravalli, “Sequential (quickest) change detection: Classical results and new directions,” IEEE Journal on Selected Areas in Information Theory, vol. 2, no. 2, pp. 494–514, 2021.

[8] Y. C. Chen, T. Banerjee, A. D. Dominguez-Garcia, and V. V. Veeravalli, “Quickest line outage detection and identification,” IEEE Trans. Power Syst., vol. 31, no. 1, pp. 749–758, Jan. 2016.

[9] 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.

[10] A. G. Tartakovsky and V. V. Veeravalli, “Asymptotically optimum quickest change detection in distributed sensor systems,” IEEE Trans. Inf. Theory, vol. 27, no. 4, pp. 441–475, 2008.

[11] S. Banerjee and G. Fellouris, “Decentralized sequential change detection with ordered CUSUMs,” in 2016 IEEE International Symposium on Information Theory (ISIT), Jul. 2016, pp. 36–40.

[12] G. Fellouris, E. Bayraktar, and L. Lai, “Efficient Byzantine sequential change detection,” IEEE Trans. Inf. Theory, vol. 64, no. 5, p. 3346–3360, Jul. 2018.

[13] Y.-C. Huang, Y.-J. Huang, and S.-C. Lin, “Asymptotic optimality in Byzantine distributed quickest change detection,” IEEE Trans. Inf. Theory, vol. 67, no. 9, pp. 5942–5962, Sep. 2021.

[14] Z. Sun, S. Zou, and Q. Li, “Quickest change detection in anonymous heterogeneous sensor networks,” in Proc. IEEE ICASSP, 2020, pp. 5925–5929.

[15] W.-H. Li and Y.-C. Huang, “Bandwidth-constrained distributed quickest change detection in heterogeneous sensor networks: Anonymous vs non-anonymous settings,” arXiv:2202.02697v2, May 2023.