# An Effective Decentralized Nonparametric Quickest Detection Approach

Dayu Yang, Hairong Qi
*Department of Electrical Engineering and Computer Science*
University of Tennessee
Knoxville, TN 37996
*(dyang, qi@eecs.utk.edu)*

## Abstract

*This paper studies decentralized quickest detection schemes that can be deployed in a sensing environment where data streams are simultaneously collected from multiple channels located distributively to jointly support the detection. Existing decentralized detection approaches are largely parametric that require the knowledge of pre-change and post-change distributions. In this paper, we first present an effective nonparametric detection procedure based on Q-Q distance measure. We then describe two implementations schemes, binary quickest detection and local decision fusion by majority voting, that realize decentralized nonparametric detection. Experimental results show that the proposed method has a comparable performance to the parametric CUSUM test in binary detection. Its decision fusion-based implementation also outperforms the other three popular fusion rules under the parametric framework.*

## 1. Introduction

Quickest detection is the real-time version of abrupt change detection [5] that dates back to 1930s. It detects changes happening at unknown points in time, as rapidly as possible, while maintaining the false alarm rate (FAR) at a given level. In general, quickest detection approaches can be categorized as parametric and nonparametric, where the former assume full knowledge of the pre-change and post-change distributions while the latter do not make this assumption.

As the technology advances, distributed sensing and detection become feasible. Multiple sensors can be deployed to observe the same phenomenon, making decentralized detection an active research topic. There have been two scenarios [4] to perform the decentralized detection task, decision-based fusion and data-based fusion. Decision-based fusion schemes conduct the detection procedure at the local sensor and only transmit detection results to the processing center. On the other hand, in data-based fusion schemes, data (or transformed and compressed version of the data) collected at individual sensor are transmitted to a processing center to conduct quickest detection. In general, data-based fusion schemes usually provide better performance but would sacrifice network bandwidth in transmitting the data. With the limited communication bandwidth, decision-based fusion promotes more efficient usage of the bandwidth but might suffer from performance. The term, "decentralized detection", refers to both fusion schemes.

In this paper, we focus on the nonparametric detection problem in a decentralized manner where the information needed for event detection is distributed and nonparametric. We briefly describe, in Sec. 2, a recently proposed nonparametric scheme from the authors that is based on Q-Q distance measure. In Sec. 3, we develop the decentralized implementation of the Q-Q distance-based detection scheme using both data-based fusion and decision-based fusion. Experimental results are provided in Sec. 4. Finally we conclude our work in Sec. 5.

## 2 Centralized Nonparametric Quickest Detection based on Q-Q Distance

Consider a distributed sensing environment with $N$ geographically deployed sensors. Let $\mathbf{X}(n)$, $n \geq 1$, be the $N$-dimensional data collected at time moment $n$. The component $\mathbf{X}_i$ represents the $i$-th channel of the data stream. At an unknown point in time, $\lambda$ ($\lambda \geq 1$), a change happens and the distributions of channels change accordingly. All channels are assumed to be independent and the observations in each channel are independent and identically distributed (i.i.d.).

Suppose the change occurs at time $k$ such that $\lambda = k$ and in the $i$-th channel the observations

1051-4651/10 $26.00 © 2010 IEEE
DOI 10.1109/ICPR.2010.558

$\{\mathbf{X}_i(1), \mathbf{X}_i(2), \ldots, \mathbf{X}_i(k-1)\}$ follow a distribution $f_i^{(0)}$ while the observations $\{\mathbf{X}_i(k), \mathbf{X}_i(k+1), \ldots\}$ follow a distribution $f_i^{(1)}$. The detection procedure, which is an optimization task, is to use the multi-channel data stream to locate the change point as early as possible, while keeping the rate of false alarm under given level [4]. The well-known CUSUM test [3] has been adopted to examine the solution of the optimization problem assuming both $f_i^{(0)}$ and $f_i^{(1)}$ are known, in which the stopping time of the $i$-th channel is defined as

$$\tau_i(h) = \inf\{n \geq 0 \mid S_{n,i} \geq \omega_i h\} \tag{1}$$

where $h$ is the prescribed threshold and $\omega_i$ is the weight of the threshold. The detection statistics, $S_{n,i}$, can be written in a recursive form as

$$S_{n,i} = \left( S_{n-1,i} + \ln \frac{f_i^{(1)}(\mathbf{X}_i(n))}{f_i^{(0)}(\mathbf{X}_i(n))} \right)^{+} \tag{2}$$

with $S_{0,i} = 0$. Note that $x^+ = \max\{0, x\}$.

Although parametric detection schemes are effective, the requirement of known pre-change and post-change distributions make them less applicable in many applications where such knowledge is not provided. As an alternative, nonparametric detection schemes are studied.

In this section, we describe a nonparametric quickest detection approach proposed by the authors [8] based on the Q-Q distance measure derived from the Quantile-Quantile plot [7]. This method requires no prior assumptions on the nature of the underlying distributions that generate the data stream except for keeping the i.i.d. assumption.

**Definition 2.1** *For any distribution function $F(x)$, the quantile function $Q(t)$ is the inverse of $F$, which is*

$$Q(t) = F^{-1}(t) = \inf\{x : F(x) \geq t\}, 0 < t < 1. \tag{3}$$

Quantile-Quantile plot is a technique for comparing distributions inferred directly from two different data sets. A Q-Q plot is a plot of the quantiles of the first data set against the quantiles of the second data set. If the two batches of data come from a population with the same distribution family, the points should fall approximately along a straight reference line and when two streams are identically distributed, their Q-Q plot should be a straight line with slope 1. So we use the $45°$ line as the reference line. The greater the plot departures from this reference line, the greater the chance that the two data sets have come from populations with different distributions.

In order to quantify the difference between two distributions, that is, the distance between the Q-Q plot and the reference line, we define the *Q-Q distance*. Consider two data sequences $v_0$ and $v_1$ of size $s$, respectively. On their Q-Q plot, for the $j$-th point, $(Q_{v_1}(\frac{j}{s}), Q_{v_0}(\frac{j}{s}))$, the distance from this point to the $45°$ reference line is $\frac{\sqrt{2}}{2} \cdot |Q_{v_1}(\frac{j}{s}) - Q_{v_0}(\frac{j}{s})|$. The Q-Q distance between these two finite sequences can thus be defined as

$$d_{qq} = \frac{1}{s} \cdot \sum_{j=1}^{s} \frac{\sqrt{2}}{2} \left| Q_{v_1}(\frac{j}{s}) - Q_{v_0}(\frac{j}{s}) \right| \tag{4}$$

where $Q_{v_i}$, $i \in \{0, 1\}$ is the quantile of data stream $v_i$.

The Q-Q distance based detection algorithm transforms the problem from the sequential detection into the problem of comparing two static sequences. The detection procedure involves creating two windows with fixed sizes, where Window One is stationary serving as a reference to the pre-change distribution, and Window Two is a moving window always containing the latest observation $\mathbf{X}_i(n)$. Q-Q distance is then calculated between the two windows for each channel. Based on the Q-Q distance, we define the centralized nonparametric detection stopping time as

$$\tau_{qq}(h) = \min\{n \geq 1 \mid \sum_{i=1}^{N} d_{qq,i}(n) \geq h\}. \tag{5}$$

where $d_{qq,i}$ is the Q-Q distance in the $i$-th channel. Once $d_{qq}$ exceeds the threshold $h$, a change is declared. In [8], we have shown this nonparametric scheme has comparable performance to other parametric and nonparametric schemes and it outperforms the CUSUM test in detecting small changes.

## 3 Decentralized Nonparametric Detection

In a detection environment with bandwidth and energy constraints, decentralized implementation of the detection approach is needed. As discussed in Sec 1, there are, in general, two decentralized approaches for quickest detection, binary quickest detection that is basically data-based fusion and decision-based fusion. In this section, we introduce our Q-Q distance based decentralized quickest detection method.

### 3.1 Binary Quickest Detection

Binary quickest detection has been studied for decentralized implementation of only parametric algorithms. It looks for an optimal quantizer such that the original observation data can be coded as binary streams, which are then transmitted as the data compressed representation to processing center for detection. In this section, we present the binary detection implementation for Q-Q distance-based nonparametric algorithm.

In parametric binary detection procedure with known distributions, the optimal quantizer is the one that maximizes the K-L information distance of the induced distributions. But in nonparametric detection the computation of the quantization threshold is not feasible, we can instead use a fixed stationary sensor quantizer, e.g., the estimated pre-change mean, as the threshold.

Let $m_i$ be the estimated mean of the observations under distribution $f_i^{(0)}$, which is used as the quantization threshold. Let $\mathbf{B}_i(n) = 1$ if $X_i(n) \geq m_i$, otherwise, $\mathbf{B}_i(n) = 0$. Then $\mathbf{B}_i$ becomes the to-be-transmitted binary stream, which is a Bernoulli sequence with induced probability mass functions $p_i^{(0)}$ and $p_i^{(1)}$. If we draw the Q-Q plot of the two segments, all drawings fall in only four possible spots, (0,0), (1,1), (0,1) and (1,0). It is easy to see that the greater the difference between $p^{(0)}$ and $p^{(1)}$, the more points fall in (0,1) or (1,0), and the bigger the distance $d_{qb,i}$ between the two binary sets, where $d_{qb}$ is the Q-Q distance between two binary data sets. The stopping time of the binary detection procedure at the processing center can then be defined as

$$\tau_{qb}(h) = \min\{n \geq 1 \mid \sum_{i=1}^{N} d_{qb,i}(n) \geq h\} \tag{6}$$

### 3.2 Decision Fusion

Another decentralized detection scenario is to perform quickest detection locally at each sensor and only send local decisions to the processing center for decision fusion. When parametric detection procedures, such as CUSUM tests, are performed at local sensors, three popular fusion rules [6] have been adopted to generate the global decision. Suppose $\tau_i$, as defined in Eq. 1, is the local stopping time of the $i$-th sensor. At the fusion center, the first fusion rule, $\tau_{max}$, declares a change only when all the local sensors have voted for a change.

$$\tau_{max} = \max_{1 \leq i \leq N} \tau_i. \tag{7}$$

The second rule, $\tau_{min}$, reports a change at the first time any local decision is in favor of the change.

$$\tau_{min} = \min_{1 \leq i \leq N} \tau_i. \tag{8}$$

The third fusion rule, $\tau_{all}$, reports a change at the first time when all sensors send 1's to the fusion center.

$$\tau_{all} = \min\{n \geq 1 \mid \min_{1 \leq i \leq N} (S_{n,i}/\omega_i) \geq h\}. \tag{9}$$

The fusion rules above are widely used in asymptotic performance analysis [2], [6] but these rules actually do not consider the highly possible sensor/channel failures that may lead to incorrect local decisions being transmitted to the fusion center. For example, a malfunctioning sensor with stuck-at faults can continuously report "no change" to the fusion center irrespective of the fact that a change has occurred. This faulty sensor would cause a complete failure of detection when the $\tau_{all}$ or $\tau_{max}$ rule is adopted. It is desired that the errors or uncertainty in some sensors can be corrected by other sensors. Here we use majority voting [1] to aggregate the local decisions generated by the Q-Q distance-based procedure in real time. Majority voting is one of the simplest fusion methods for decision fusion tasks and is as effective as the other more complicated schemes [1]. Assume each sensor reports a local decision

$$\psi_{n,i} = \begin{cases} 1 & \text{if } d_{qq,i} \geq \omega_i h \\ 0 & \text{otherwise} \end{cases} \tag{10}$$

to the center at every time interval $n$. The stopping time at the center is defined as

$$\tau_v = \min\{n \geq 1 \mid \sum_{i=1}^{N} \psi_{n,i} > N/2\} \tag{11}$$

which means a change is declared the first time when more than half of the sensors agree.

## 4 Experimental Results

For decentralized detection, we conduct two experiments that evaluate binary detection and decision-based fusion. We simulate 6 independent channels which have distribution changes at exactly the same time. From channel 1 to channel 6, the changes are Uniform(0,1) to Uniform(0,5), Uniform(-1,1) to Uniform(-1,4), Normal(0,1) to Normal(2,1), Normal(0,1) to Uniform(0,2), Poisson distributions with mean of 8 to 10, and Normal(0,1) to Normal(0.5,1). The results presented are based on Monte Carlo simulations with $10^4$ replications which are sufficient for estimating the average detection delay (ADD) and false alarm rate (FAR). The plot of ADD *vs.* -log(FAR) is used to show the relationship between ADD and FAR. The lower the FAR, the longer the delay. We generate 5000 samples in each channel and choose the window size as 300.

In the first experiment, we compare the performance of binary detection using the Q-Q distance-based nonparametric scheme and the benchmark parametric CUSUM detection scheme. Figure 1 shows the operating characteristics. With complete information of the

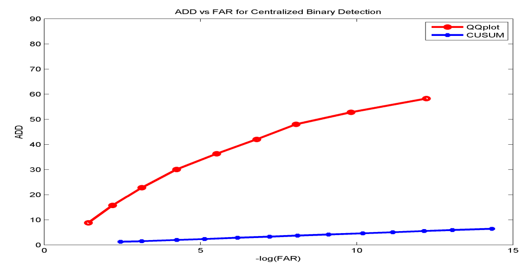


**Figure 1. Operating Characteristics of Centralized Binary Detection Procedures**

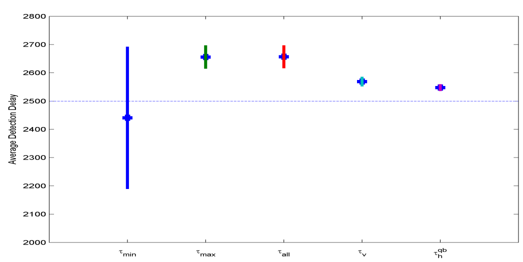


**Figure 2. Results of Decision Fusion Rules**

distributions, the CUSUM scheme outperforms the Q-Q distance-based scheme which is within expectation. However, the Q-Q distance-based scheme still succeeds as a nonparametric detection approach, showing acceptable performance in a sense of small detect delay and low false alarm rate.

In the second experiment, we compare the performance of the four decision-based fusion rules, $\tau_{min}$, $\tau_{max}$, $\tau_{all}$, and $\tau_v$, as well as the result from binary detection, $\tau_{qb}(h)$, as defined in Eq. 6. Note that only the Q-Q distance-based detection procedure is applied. We observe from Fig. 2 that the $\tau_{min}$ rule yields an average stopping time even before the real change time with big variance which means it produces lots of false alarms and the performance is very unstable or unpredictable. The $\tau_{max}$ and $\tau_{all}$ rules generate similar results and introduce fewer false alarms but present much longer delays than the others. $\tau_v$ from the majority voting gives the smallest detection delay with the smallest variance among the four decision-based fusion rules, showing its effectiveness in producing more accurate and reliable results. $\tau_{qb}$ is the average delay from the Q-Q distance-based binary detection. It yields an even smaller delay and variance compared to the $\tau_v$ rule, which is consistent with our previous discussion that although decision-based fusion provides the most effective usage of communication bandwidth and energy, it experiences a little bit degradation in performance in terms of delay and stability because only local decisions are fused. However, the $\tau_v$-based fusion still shows very close performance to binary detection. Fig. 2 data are also tabulated in Table 1 for clarity purpose.

**Table 1. Decision fusion results**

| Rules | $\tau_{min}$ | $\tau_{max}$ | $\tau_{all}$ | $\tau_v$ | $\tau_{qb}$ |
|---|---|---|---|---|---|
| Average Delay | -89 | 156 | 157 | 69 | 48 |
| Variance | 252.5 | 41.0 | 40.1 | 17.2 | 12.2 |
| Minimum Delay | -2071 | 65 | 65 | 28 | 8 |
| Maximum Delay | 42 | 299 | 299 | 133 | 81 |

## 5 Conclusions

In this paper, we described a decentralized nonparametric quickest detection algorithm based on the Q-Q distance measure, using both the binary detection scheme and the decision-based fusion scheme. Experimental results with simulated data showed that this procedure is able to detect the changes with comparable performance as the benchmark CUSUM detection scheme. The majority voting based decision fusion rule also generates better detection decision than other traditional rules from parametric detection.

## References


[1] L. Lam and S. Suen. Application of majority voting to pattern recognition: An analysis of its behavior and performance. *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans*, 27(5):553–568, 1997.

[2] Y. Mei. Information bounds and quickest change detection in decentralized decision systems. *IEEE Transactions on Information Theory*, 51(7):2669–2681, 2005.

[3] E. Page. Continuous inspection schemes. *Biometrika*, 41(1-2):100–115, 1954.

[4] H. Poor and O. Hadjiliadis. *Quickest Detection*. Cambridge University Press, 2009.

[5] W. Shewhart. *Economic Control of Quality of Manufactured Product*. Princeton, NJ: Van Nostrand Reinhold Co., 1931.

[6] A. Tartakovsky and V. Veeravalli. Asymptotically optimal quickest change detection in distributed sensor systems. *Sequential Analysis*, 27(4):441–475, 2008.

[7] M. Wilk and R. Gnanadesikan. Probability plotting methods for the analysis of data. *Biometrika*, 55(1):1–17, 1968.

[8] D. Yang and H. Qi. A new nonparametric quickest detection procedure. *Proceedings of the $35^{th}$ International Conference on Acoustics, Speech, and Signal Processing*, March 2010. Dallas, Texas.