Bayesian Quickest Change Point Detection with Multiple Candidates of Post-Change Models

Samrat Nath and Jingxian Wu

Department of Electrical Engineering, University of Arkansas, U.S.A.

27 November 2018

.

OUTLINE

- Introduction
- Problem Formulation
- Quickest Change Detection (QCD) Algorithm
- Analytical Results
- Simulation Results
- Conclusions

INTRODUCTION (1/4)

- Change-point Detection:
  - Process of identifying time instants at which the distribution of a random process changes.
  - **Applications:** Signal and image processing, quality control engineering, seismology, financial markets, etc.

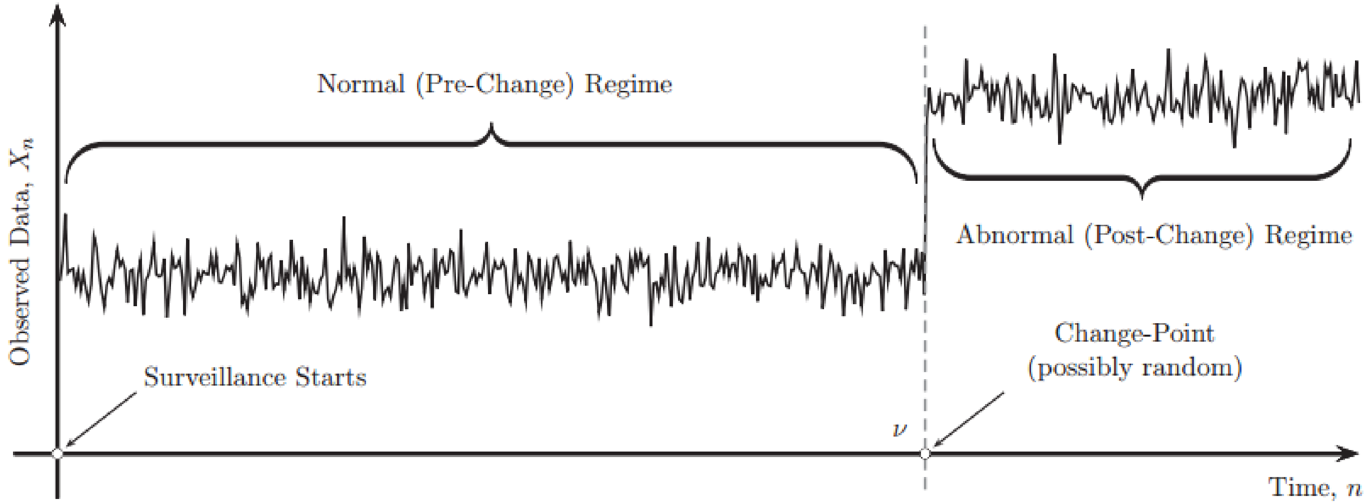

Figure: An example of the behavior observed sequences from a certain process of interest [1]

[1] Polunchenko, Aleksey S., and Alexander G. Tartakovsky. "State-of-the-art in sequential change-point detection." Methodology and computing in applied probability 14, no. 3 (2012): 649-684.

INTRODUCTION (2/4)

- Categories of Detection Procedures:
  - Based on nature of observations:
    - **Offline:** based on the observations of the entire time sequence
    - **Online:** based on currently observed samples (real-time)
  - Based on knowledge of change-point:
    - **Bayesian:** Prior probability of the change-point is known
    - **Non-Bayesian**: Prior is unknown

INTRODUCTION (3/4)

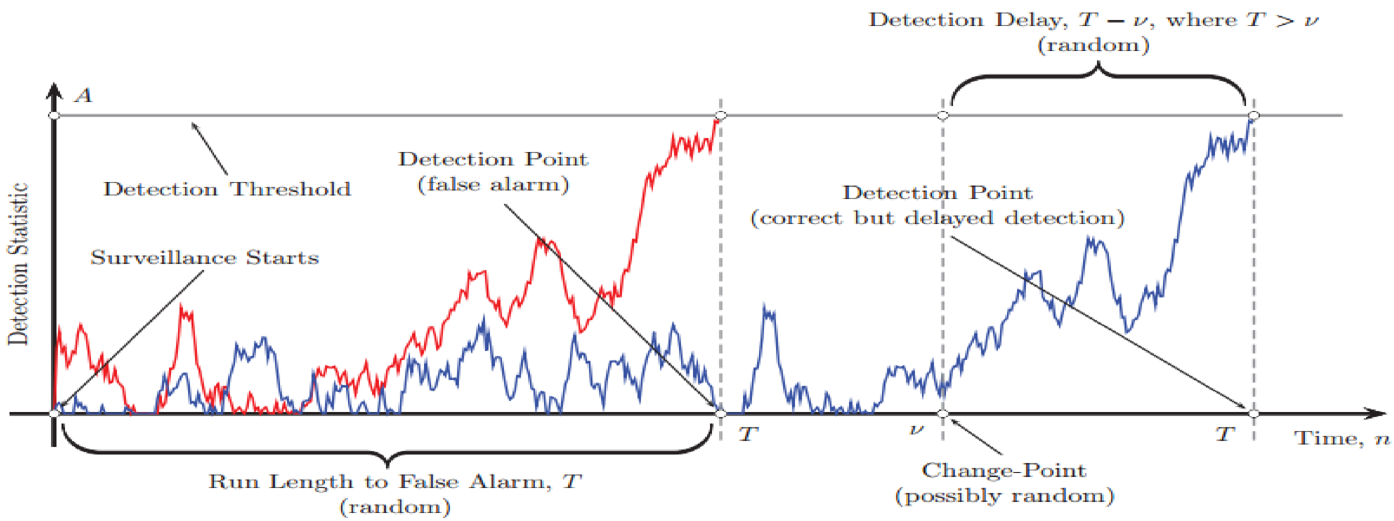

Figure: Two possible scenarios: False alarm (red trajectory), and delayed detection (blue trajectory) [1]

- Quickest Change Detection (QCD):
  - Aims at minimizing detection delay with an upper bound on probability of false alarm [2]

[2] V. V. Veeravalli and T. Banerjee, "Quickest change detection," in Academic Press Library in Signal Processing. Elsevier, 2014, vol. 3, pp. 209–255.

INTRODUCTION (4/4)

- Motivation:
  - There are limited works with unknown or uncertain post-change models.
  - For many applications, the post-change model might be from a finite set of possible models.
  - **Objective:** To design change detection algorithm for systems with multiple post-change models under Bayesian setting.

OUTLINE

- Introduction
- **Problem Formulation**
- QCD Algorithm
- Analytical Results
- Simulation Results
- Conclusions

PROBLEM FORMULATION (1/3)

- Notations:
  - X_n: a sequentially observed random sequence, where $n = 1, 2, \cdots$.
  - $\mathcal{F}_n^{\dot{X}} = \sigma(\mathbf{X}^{1:n})$ is σ −algebra generated by $\mathbf{X}^{1:n} = [X_1, \cdots, X_n]$
  - θ: Change point
  - $f_{0,n}(X_n|\mathbf{X}^{1:n-1})$: Probability density function (pdf) when $n < \theta$.
  - $f_{i,n}(X_n|\mathbf{X}^{1:n-1})$ pdf when $n \geq \theta$ for $i = 1, 2, \cdots, M$
  - M: number of post-change distribution models
  - $\beta \in \{1, \cdots, M\}$: index for true post-change distribution

PROBLEM FORMULATION (2/3)

- Bayesian Setting:
  - Probability mass function (PMF) for change-point θ

$$\mathbf{P}(\theta = k) = \pi_k, \text{ for } k = 1, 2, \cdots.$$

  - PMF for post-change model index:

$$\mathbf{P}(\beta = i) = \omega_i, \text{ for } i = 1, \cdots, M.$$

- Sequential Test (δ):
  - To detect change point θ based on sequentially observed data X_n
  - $\hat{\theta}$: estimate of change point
  - Sequential test can be defined as $\delta : \mathcal{F}_n^X \to \hat{\theta}$
  - Needs to be designed by optimizing with respect to detection delay and false detection

UNIVERSITY OF ARKANSAS

PROBLEM FORMULATION (3/3)

- Performance Metrics:
  - Average Detection Delay:

$$\mathrm{ADD}(\delta) = \mathbb{E}[\hat{\theta} - \theta | \hat{\theta} \geq \theta]$$

  - Probability of False Alarm:

$$\mathrm{PFA}(\delta) = \mathbf{P}(\hat{\theta} < \theta | \mathcal{F}_n^X)$$

- Optimization Problem:

$$\begin{array}{lll} \text{(P1)} & \text{minimize} & \mathrm{ADD}(\delta) \\ & \text{subject to} & \mathrm{PFA}(\delta) < \alpha. \end{array}$$

UNIVERSITY OF ARKANSAS

OUTLINE

- Introduction
- Problem Formulation
- **QCD Algorithm**
- Analytical Results
- Simulation Results
- Conclusions

UNIVERSITY OF ARKANSAS

QCD ALGORITHM (1/2)

- Test Statistic:
  - At moment n, the detector needs to decide between two hypotheses

$$\mathcal{H}_1 : \theta \leq n,$$
$$\mathcal{H}_0 : \theta > n.$$

  - Test statistic is the ratio of posterior probabilities:

$$\Delta(n) = \frac{\mathbf{P}(\mathcal{H}_1|\mathcal{F}_n^X)}{\mathbf{P}(\mathcal{H}_0|\mathcal{F}_n^X)} = \frac{\sum_{k=1}^{n} \pi_k \cdot d\mathbf{P}(\mathbf{x}^{1:n}|\theta = k)}{\Omega_n \cdot d\mathbf{P}(\mathbf{x}^{1:n}|\theta > n)}$$

$$= \sum_{i=1}^{M} \omega_i \sum_{k=1}^{n} \frac{\pi_k}{\Omega_n} \prod_{t=k}^{n} \frac{f_{i,t}(X_t|\mathbf{X}^{1:t-1})}{f_{0,t}(X_t|\mathbf{X}^{1:t-1})},$$

where $\Omega_n = \mathbf{P}(\theta > n) = \sum_{k=n+1}^{\infty} \pi_k.$

UNIVERSITY OF ARKANSAS

QCD ALGORITHM (2/2)

- Proposed Algorithm:

  - **Definition 1**: For a given PFA upper bound α, the change point is detected as

$$\delta_1 : \hat{\theta}_1 = \inf\left\{n \geq 1 : \Delta(n) \geq \frac{1-\alpha}{\alpha}\right\}$$

  - δ_1 can be considered as an extension of the **Shiryaev** procedure [3], which only considers the case of one known post-change model.

[3] A. N. Shiryaev, "On optimum methods in quickest detection problems," Theory of Probability & Its Applications, vol. 8, no. 1, pp. 22–46, 1963.

OUTLINE

- Introduction
- Problem Formulation
- QCD Algorithm
- **Analytical Results**
- Simulation Results
- Conclusions

ANALYTICAL RESULTS (1/4)

- Probability of False Alarm:
  - **Lemma 1**: $\boxed{\mathrm{PFA}(\delta_1) \leq \alpha}$
- Sketch of proof:
  - Let $p(n) = \mathbf{P}(\mathcal{H}_1 | \mathcal{F}_n^X) = \mathbf{P}(\theta \leq n | \mathcal{F}_n^X)$
  - $\Delta(n) = \dfrac{\mathbf{P}(\mathcal{H}_1 | \mathcal{F}_n^X)}{\mathbf{P}(\mathcal{H}_0 | \mathcal{F}_n^X)} = \dfrac{p(n)}{1 - p(n)}$
  - $p(n) = 1 - \frac{1}{1 + \Delta(n)}$
  - Since, $\delta_1 : \hat{\theta}_1 = \inf\left\{ n \geq 1 : \Delta(n) \geq \dfrac{1-\alpha}{\alpha} \right\}$
  - $\Delta(\hat{\theta}_1) \geq \dfrac{1-\alpha}{\alpha} \Longrightarrow p(\hat{\theta}_1) \geq 1 - \alpha$
  - $\mathrm{PFA}(\delta_1) = \mathbf{P}(\hat{\theta}_1 < \theta | \mathcal{F}_n^X) = 1 - p(\hat{\theta}_1) \leq \alpha$

UNIVERSITY OF ARKANSAS

ANALYTICAL RESULTS (2/4)

- Average Detection Delay:

$$\mathrm{ADD}(\delta) = \frac{\mathbb{E}(\hat{\theta} - \theta)^+}{\mathbf{P}(\hat{\theta} \geq \theta)}$$

$$= \frac{1}{\mathbf{P}(\hat{\theta} \geq \theta)} \sum_{k=1}^{\infty} \pi_k \, \mathbf{P}_k(\hat{\theta} \geq k) \, \mathbb{E}_k(\hat{\theta} - k | \hat{\theta} \geq k),$$

  - $x^+ = \max(0, x)$
  - $\mathbf{P}_k$: probability measure when $\theta = k$
  - $\mathbb{E}_k$: corresponding expectation operator

ANALYTICAL RESULTS (3/4)

- Asymptotic Notations:
  - $f(x)$ and $g(x)$ are continuous function such that

$$\lim_{x \to x_0} f(x) = \lim_{x \to x_0} g(x) = \infty.$$

  - Notation 1: $f(x) \underset{x \to x_0}{\preceq} g(x) \Longleftrightarrow \lim\limits_{x \to x_0} \dfrac{f(x)}{g(x)} \leq 1.$
  - If both $f(x) \underset{x \to x_0}{\preceq} g(x)$ and $g(x) \underset{x \to x_0}{\preceq} f(x)$, then two functions are asymptotically equivalent.

$$f(x) \underset{x \to x_0}{\asymp} g(x) \Longleftrightarrow \lim_{x \to x_0} \frac{f(x)}{g(x)} = 1$$

UNIVERSITY OF ARKANSAS

ANALYTICAL RESULTS (4/4)

- Asymptotic Optimality of Proposed QCD Algorithm:
  - **Theorem 1:** $\mathbb{E}_k[(\hat{\theta}_1 - k)^+] \underset{\alpha\to 0}{\preceq} \min_i \left[\frac{\log\left(\frac{1-\alpha}{\alpha}\right) - \log \omega_i}{D_i + |\log(1-\rho)|} \right]$
  - **Theorem 2:** $\mathbb{E}_k[(\hat{\theta}_1 - k)^+] \underset{\alpha\to 0}{\succeq} \min_i \left[\frac{\log\left(\frac{1-\alpha}{\alpha}\right) - \log \omega_i}{D_i + |\log(1-\rho)|} \right]$
  - **Theorem 3:** $\mathrm{ADD}(\delta_1) \underset{\alpha\to 0}{\asymp} \min_i \left[\frac{\log\left(\frac{1-\alpha}{\alpha}\right) - \log \omega_i}{D_i + |\log(1-\rho)|} \right]$
  - Here,
    - $D_i = \mathbb{E}\left[\log \frac{f_i(X)}{f_0(X)}\right]$ is Kullback-Leibler (KL) divergence.
    - $\pi_k = (1-\rho)^{k-1}\rho$ is the prior PMF of change-point θ.

UNIVERSITY OF ARKANSAS

OUTLINE

- Introduction
- Problem Formulation
- QCD Algorithm
- Analytical Results
- **Simulation Results**
- Conclusions

UNIVERSITY OF ARKANSAS

SIMULATIONS (1/2)

- Environment:
  - Two post-change models
  - Models follow Normal distribution,

$$f_i \sim \mathcal{N}(0, \sigma_i^2)$$

  - For pre-change model,

$$\sigma_0^2 = 1$$

  - Geometric dist. $\rho = 0.1$
  - Asymptotic results give good predictions regarding trend of detection delay.

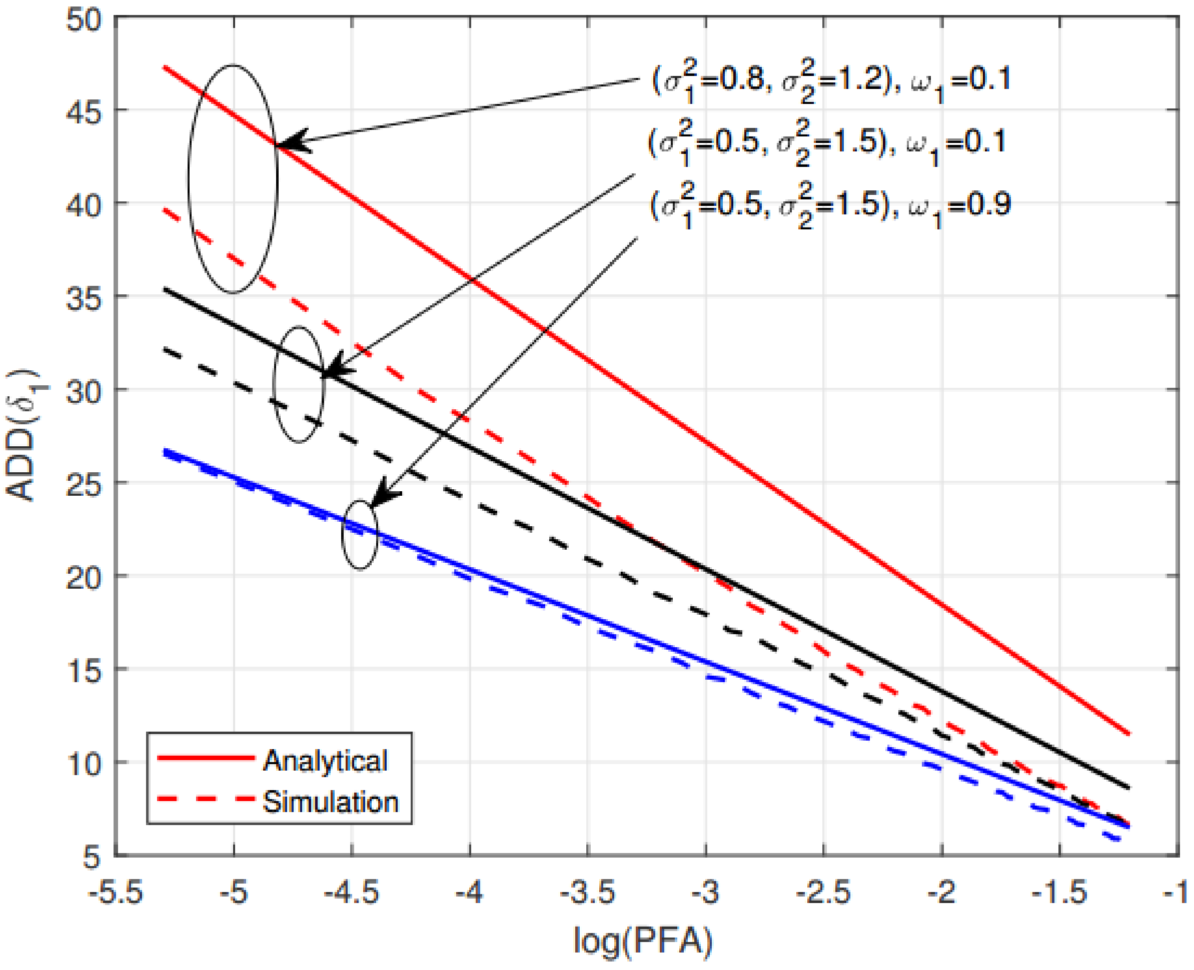

Figure: Average Detection Delay

UNIVERSITY OF ARKANSAS

SIMULATIONS (2/2)

- Environment:
  - Models follow:

    $$f_i \sim \mathcal{N}(\mu_i, 1)$$

  - For pre-change model,

    $$f_0 \sim \mathcal{N}(1, 1)$$

  - Four post-change models

    $$\mu_1 = 0.6, \mu_2 = 0.8,$$
    $$\mu_3 = 1.2, \mu_4 = 1.4$$

  - Mixture model in Shiryaev Procedure [3]:

    $$h(x) = \sum_{i=1} \omega_i f_i(x)$$

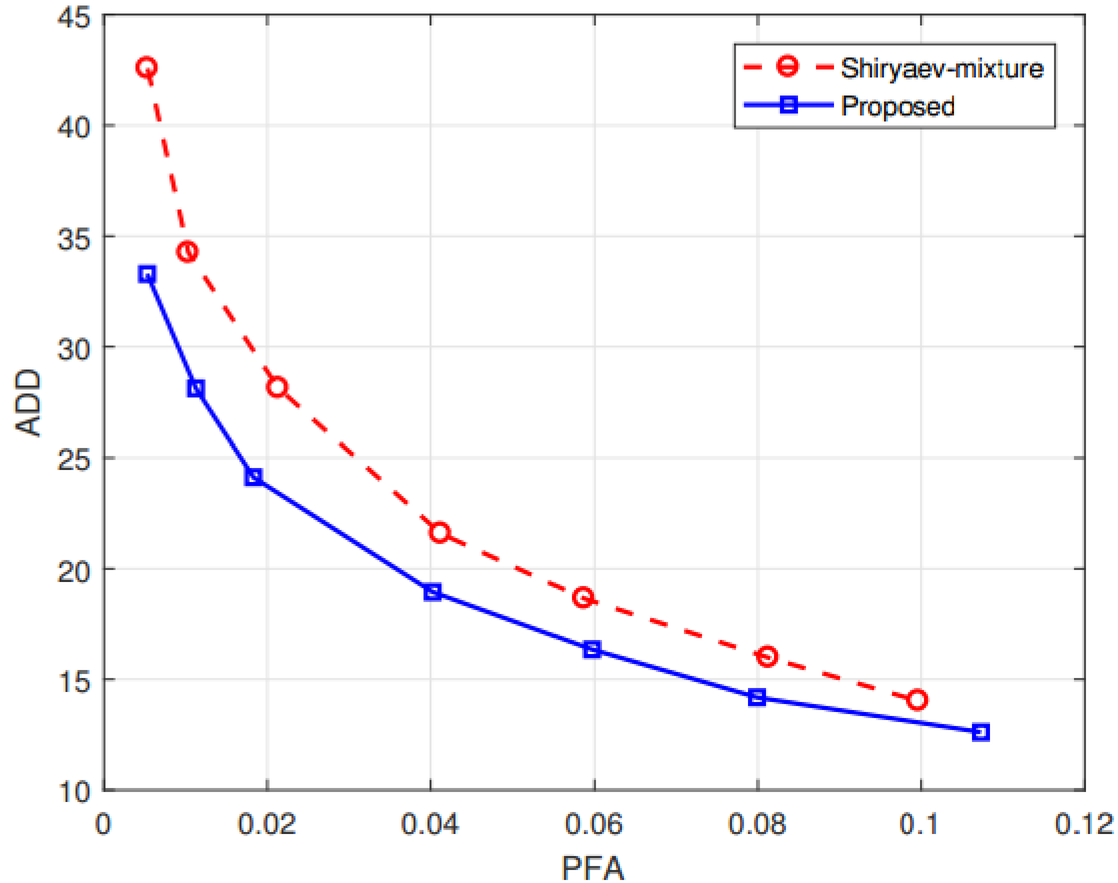

Figure: Tradeoff between ADD and PFA

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

CONCLUSIONS

- Summary:
  - Proposed a threshold-based sequential test for Bayesian QCD when there are multiple possible post-change models.
  - Analytical ADD has been obtained through asymptotic analysis when the PFA is small.
  - Proposed algorithm is asymptotically optimal in terms of ADD.
  - Proposed algorithm outperforms the Shiryaev procedure with a mixture post-change model.

- Future Direction:
  - QCD algorithm in non-Bayesian setting.
  - Detection of multiple change-points.
  - Both change-point detection and identification of post-change model.

REFERENCES

[1] Polunchenko, Aleksey S., and Alexander G. Tartakovsky. "State-of-the-art in sequential change-point detection." Methodology and computing in applied probability 14, no. 3 (2012): 649-684.

[2] V. V. Veeravalli and T. Banerjee, “Quickest change detection,” in Academic Press Library in Signal Processing. Elsevier, 2014, vol. 3, pp. 209–255.

[3] A. N. Shiryaev, “On optimum methods in quickest detection problems,” Theory of Probability & Its Applications, vol. 8, no. 1, pp. 22–46, 1963.

[4] T. L. Lai, “Information bounds and quick detection of parameter changes in stochastic systems,” IEEE Transactions on Information Theory, vol. 44, no. 7, pp. 2917–2929, 1998

Thanks!

for your attention

Questions?