Sequential Change Detection: Overview & Recent Results

George V. Moustakides
Rutgers University &
University of Patras, Greece
Outline

- Problem definition: Detectors and Change generation mechanisms
- Formulations involving expected delays
- Formulations involving hard delay constraints
- Decentralized detection (sensor networks)
- Intrusion detection in wireless networks
Problem definition

Change in statistics

Detect change as soon as possible

Specify: a) Detector form
b) Change generation mechanisms

G.V. Moustakides: Sequential Change-Detection, Princeton, January 2014
Applications

Quality monitoring of manufacturing process (1930’s)
Biomedical Engineering
Electronic Communications
Econometrics
Seismology
Speech & Image Processing (segmentation)
Vibration monitoring (Structural health monitoring)
Security monitoring (fraud detection)
Spectrum monitoring
Scene monitoring
Network monitoring (router failures, intrusion detection)
Epidemic detection ..... 

CUSUM: 2,280 hits in 2013. Google Scholar

G.V. Moustakides: Sequential Change-Detection, Princeton, January 2014
We observe **sequentially** a process $\{x_t\}$ that has the following statistical properties:

$$
\begin{align*}
  x_t &\sim \begin{cases} 
    f_0 & \text{for } 0 < t \leq \tau \\
    f_1 & \text{for } t > \tau
  \end{cases}
\end{align*}
$$

Detect occurrence of $\tau$ as soon as possible:

At every time $t$ consult available data: $x_1, \ldots, x_t, x_{t+1}$

- Change did not take place before $t$
  
  Continue sampling

- Change took place before $t$
  
  **Stop** sampling!

Sequential Detector $\leftrightarrow$ Stopping time
Structural health monitoring

Change mechanism independent from data

Amplitude of oscillations overly large

Change mechanism dependent on data
Formulations with expected delays

We are looking for a stopping time $T$.

General criterion:

$$J(T) = E_1[T - \tau \mid T > \tau]$$

Change mechanism independent from data

Shiryaev (1963): $\tau$ is random with **known** prior.

$$\inf_T J(T) \quad \text{subject to: } P_0(T \leq \tau) \leq \alpha$$

If prior is exponential: $$P(\tau = t) = p(1 - p)^t$$
Define the statistic: \[ \pi_t = P(\tau < t \mid x_1, \ldots, x_t) \]
\[ T_S = \min\{t > 0 : \pi_t \geq \nu\} \]

Threshold \( \nu \in (0,1) \) such that the false alarm constraint is satisfied with equality.

In discrete time when \( \{x_t\} \) are i.i.d. before and after the change.

In continuous time when \( \{x_t\} \) is a Brownian motion with constant drift before and after the change.

In continuous time when \( \{x_t\} \) is Poisson with constant rate before and after the change.

Time variation? Dependence? Multiple pre- and/or post-change possibilities?
\[ J(T) = E_1[T - \tau \mid T > \tau] \]

Changetime \( \tau \) is random with \textbf{unknown} prior.


\[ J_P(T) = \sup_{\text{all priors}} E_1[T - \tau \mid T > \tau] \]

We can show:

\[ J_P(T) = \sup_{t \geq 0} E_1[T - t \mid T > t] \]

\[ \inf_T J_P(T) \quad \text{subject to: } E_0[T] \geq \gamma \]
Discrete time: i.i.d. data before and after the change with pdfs $f_0, f_1$.

Compute recursively the following statistic:

$$ S_t = (1 + S_{t-1}) \frac{f_1(x_t)}{f_0(x_t)}; \quad \text{Pollak (1985):} \quad S'_0 \text{ if specially designed, then} $$

$$ T_P = \inf \{ t > 0 : S_t \geq \nu \} $$

$$ [J_P(T_P) - \inf_T J_P(T)] \to 0; \quad \text{as } \gamma \to \infty $$

Order-3 Asymptotic optimality


Continuous-time? Time variation? Dependence? Multiple pre- and/or post-change possibilities?
Change mechanism dependent on data.

Lorden (1971): \( \tau \) unknown dependence.
Follow a worst-case analysis.

\[
J_L(T) = \sup_{\text{data dependent } \tau} \mathbb{E}_1[T - \tau \mid T > \tau]
\]

\[
J_L(T) = \sup_{t \geq 0} \sup_{x_1, \ldots, x_t} \mathbb{E}_1[(T - t)^+ \mid x_1, \ldots, x_t]
\]

\[
\inf_T J_L(T) \text{ subject to } : \mathbb{E}_0[T] \geq \gamma
\]

G.V. Moustakides: Sequential Change-Detection, Princeton, January 2014
CUSUM stopping time:

\[ u_t = \log \left( \frac{f_1(x_1, \ldots, x_t)}{f_0(x_1, \ldots, x_t)} \right); \quad \text{running LLR} \]

\[ m_t = \inf_{0 < s \leq t} u_s; \quad \text{running minimum} \]

\[ S_t = u_t - m_{t-1}; \quad \text{CUSUM statistic} \]

\[ T_C = \inf\{t > 0 : S_t \geq \nu\}; \quad \text{CUSUM stop. time} \]

For i.i.d. \[ S_t = (S_{t-1})^+ + \log \left( \frac{f_1(x_t)}{f_0(x_t)} \right) \]
Discrete time: i.i.d. before and after the change
Lorden (1971) asymptotic optimality (order-1).
Moustakides (1986) strict optimality.
Poor (1998) strict optimality for exponential delay penalty.

Continuous time
Shiryaev (1996), Beibel (1996) strict optimality for Brownian Motion
Moustakides (2004) strict optimality for Ito processes
Moustakides (under review) strict optimality for Poisson processes.

Time variation? Dependence? Multiple pre- and/or post-change possibilities?
Formulations with hard constraints

\[ J(T) = \mathbb{E}_1[T - \tau \mid T > \tau] \]

Detection delay can be arbitrarily large!

Several applications require detection delay at most \( m \).

\[ \tau < T \leq \tau + m \]

If \( \tau + m < T \), this is regarded as failure.
\[ \mathcal{J}(T) = P_1(\tau < T \leq \tau + m \mid T > \tau) \]

Interested in detection probability

Change mechanism independent from data

\( \tau \) random with known prior. (Shiryaev-like)

\[
\sup_{T} \mathcal{J}(T) \text{ subject to } P_0(T \leq \tau) \leq \alpha
\]

\( \tau \) random with unknown prior. (Pollak-like)

\[ \mathcal{J}_P(T) = \inf_{t \geq 0} P_1(t < T \leq t + m \mid T > t) \]

\[
\sup_{T} \mathcal{J}_P(T) \text{ subject to } E_0[T] \geq \gamma
\]

G.V. Moustakides: Sequential Change-Detection, Princeton, January 2014
Change mechanism dependent on data.

\( \tau \) unknown dependence. (Lorden-like)

\[
\mathcal{J}_L(T) = \inf_{t \geq 0} \inf_{x_1, \ldots, x_t} P_1(t < T \leq t + m \mid x_1, \ldots, x_t)
\]

\[
\sup_T \mathcal{J}_L(T) \text{ subject to } : E_0[T] \geq \gamma
\]

Exact solution only for \( m=1 \) (detect the change with the first sample under the alternative regime).

\[
T_{Sh} = \inf \left\{ t > 0 : \frac{f_1(x_t)}{f_0(x_t)} \geq \nu \right\}
\]

If there are two possible changes?

1) \( f_0 \rightarrow f_1^1 \)

2) \( f_0 \rightarrow f_1^2 \)

Run two separate CUSUMs in parallel (2-CUSUM).

Dragalin (1997); Hadjiliadis, Moustakides (2006); Hadjiliadis, Poor (2009): Asymptotic optimality (orders-1,2,3).

\[
J_L(T) = \sup_i \sup_{t \geq 0} \sup_{x_1, \ldots, x_t} E^i_1[(T - t)^+ | x_1, \ldots, x_t]
\]

\[
\inf_T J_L(T) \text{ subject to } E_0[T] \geq \gamma
\]

**Theorem:** If \( \gamma_0 \geq \gamma \geq 1 \), then the Shewhart test

\[
T_{Sh} = \inf \left\{ t > 0 : (1 - q) \frac{f_1(x_t)}{f_0(x_t)} + q \frac{f_2(x_t)}{f_0(x_t)} \geq \nu \right\}
\]

is optimum. **2-CUSUM is not strictly optimum.**

G.V. Moustakides: Sequential Change-Detection, Princeton, January 2014
Decentralized detection

Sensor 1

Sensor 2

Sensor 3

Fusion Center

Challenge: Data quantization.

Sensor K

$\mathbf{x}_t$ $\mathbf{z}_n^1$ $\mathbf{t}_n^1$ $\mathbf{z}_n^2$ $\mathbf{t}_n^2$ $\mathbf{z}_n^3$ $\mathbf{t}_n^3$ $\mathbf{x}_t^1$ $\mathbf{x}_t^2$ $\mathbf{x}_t^3$

Veeravalli (1999,2001)
Dayanik, Poor, Sezer (2008)
Tartakovsky, Veeravalli (2008)

G.V. Moustakides: Sequential Change-Detection, Princeton, January 2014
If more than 1 bits, **quantize overshoot**!
Communication with Fusion Center is:
- at random times
- asynchronous
- control over **average** communication period with $A_i, B_i$

If sensor $i$ sends a bit at time $t$, the Fusion Center updates an estimate of the global log-likelihood ratio:

$$\hat{u}_t = \begin{cases} 
\hat{u}_{t-} + B_i & \text{if bit is 1} \\
\hat{u}_{t-} + A_i & \text{if bit is 0}
\end{cases}$$

and performs a CUSUM test using the estimate of the global log-likelihood ratio.
Number of sensors=5; Communication period=6

Detection delay vs False alarm period for different CUSUM algorithms.

- 1-bit Q-CUSUM
- 2-bit Q-CUSUM
- inf-bit Q-CUSUM
- 1-bit D-CUSUM
- 2-bit D-CUSUM
- inf-bit D-CUSUM
- Centralized CUSUM
Long-term deployment setup at the Chevron-Richmond refinery. The result of this test was a detection rate of 100% with no false alarms. The sensors withstood strong winds and rainy weather.
Intrusion detection with Radosavac and Baras

MAC Layer: If the channel is not in use, nodes wait a random (back-off) time and then ask to reserve the channel.

- The node with the smallest back-off time reserves the channel.
- Back-off times of legitimate users are uniformly distributed. So $f_0 = U[0, W]$.
- Intruder’s goal is to reserve the channel more often than a legitimate user. Back-off distribution $f_1$ is unknown.

Use back-off time measurements to detect intruder.
We would like to apply CUSUM on the back-off times for intruder detection. But we do not know $f_1$!

**Intruder characterization**

- $N$ legitimate nodes have probability $1/N$ of reserving the channel.
- A node is characterized as “intruder” if its probability to reserve the channel is at least $\eta/N$ where $\eta > 1$.

Example: If $\eta = 1.1$ this means I can tolerate illegitimate behavior provided it is no larger than 10% of the legitimate one!
\[ \Pr_{1}(\text{Reserve channel}) \geq \frac{\eta}{N} \Leftrightarrow \int_{0}^{W} x f_{1}(x) \, dx \leq \epsilon \frac{W}{2} \]

Defines a class \( \mathcal{F} \) of possible pdfs

\[ J_{L}(T, f_{1}) = \sup_{t \geq 0} \sup_{x_{1}, \ldots, x_{t}} \mathbb{E}_{1}[(T - t)^{+} | x_{1}, \ldots, x_{t}] \]

\[ \inf_{T} \sup_{f_{1} \in \mathcal{F}} J_{L}(T, f_{1}) \text{ subject to } : \mathbb{E}_{0}[T] \geq \gamma \]

CUSUM with \( f_{1}^{*}(x) = \begin{cases} Ce^{-\mu x} & \text{for } 0 \leq x \leq W \\ 0 & \text{otherwise.} \end{cases} \)