# Detection of a Random Sequence of Disorders*

**Aleksandra Ochman-Gozdek** and **Krzysztof Szajowski**

*Wrocław University of Technology*
*Institute of Mathematics and Computer Science, Wrocław, Poland*
*e-mail:* aleksandra.ochman@pwr.wroc.pl; Krzysztof.Szajowski@pwr.wroc.pl

**Abstract**

A random sequence having segments being the homogeneous Markov processes is registered. Each segment has his own transition probability law and the length of the segment is unknown and random. The transition probabilities of each process are known and *a priori* distribution of the disorder moment is given. The former research on such problem has been devoted to various questions concerning the distribution changes when more than one homogeneous segment is expected. The detection of the disorder rarely is precise. The decision maker accepts some deviation in estimation of the disorder moment. In the models taken into account the aim is to indicate the change point with fixed, bounded error with maximal probability. The case with various precision for over and under estimation of this point is analysed. The case when the disorder does not appears with positive probability is also included. The observed sequence, when the change point is known, has the Markovian properties. The results insignificantly extends range of application, explain the structure of optimal detector in various circumstances and shows new details of the solution construction. The motivation for this investigation is the modelling of the attacks in the node of networks. The objectives is to detect one of the attack immediately or in very short time before or after it appearance with highest probability. The problem is reformulated to optimal stopping of the observed sequences. The detailed analysis of the problem is presented to show the form of optimal decision function.

**AMS 2010 subject classifications:** Primary 60G40, ; secondary 60K99; 90D60.
**Keywords and phrases:** Bayesian approach, disorder problem, sequential detection, optimal stopping, Markov process, change point.

## 1. Introduction

Suppose that the process $X = \{X_n, n \in \mathbb{N}\}$, $\mathbb{N} = \{0, 1, 2, \ldots\}$, is observed sequentially. It is obtained from Markov processes by switching between them at a random moment $\theta$ in such a way that the process after $\theta$ starts from the state $X_{\theta-1}$. It means that the state at moment $n \in \mathbb{N}$ has conditional distribution given the state at moment $n-1$, where the formulae describing these distributions have the different form: one for $n < \theta$ and another for $n \geq \theta$. Our objective is to detect the moment $\theta$ based on observation of $X$. There are some papers devoted to the discrete case of such disorder detection which generalize in various directions the basic problem stated by Shiryaev (1961) (see e.g. Bojdecki (1979), Yoshida (1983)).

Such model of data appears in many practical problems of the quality control (see Shewhart (1931) and in the collection of the papers Basseville and Benveniste (1986)), traffic anomalies in networks (see Tartakovsky et al. (2006)), epidemiology models (see Baron (2004)). The aim is to recognize the moment of the change over the one probabilistic characteristics to another of the phenomenon.

Typically, the disorder problem is limited to the case of switching between sequences of independent random variables (see Bojdecki (1979)). Some developments of the basic model can be found in Yakir (1994) where the optimal detection

* The research has been supported by grant S20096/I-18

rule of the switching moment has been obtained when the finite state-space Markov chains is disordered. Moustakides (1998) formulates conditions which help to reduce the problem of the quickest detection for dependent sequences before and after the change to the case for independent random variables. Our result is a generalization of the results obtained by Bojdecki (1979) and Sarnowski and Szajowski (2011). It admits Markovian dependence structure for switched sequences (with possibly uncountable state-space). We obtain an optimal rule under probability maximizing criterion.

Formulation of the problem can be found in Section 2. The main result is presented in Section 3.

## 2. Formulation of the problem

Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a probability space which supports sequence of observable random variables $\{X_n\}_{n\in\mathbb{N}}$ generating filtration $\mathcal{F}_n = \sigma(X_0, X_1, ..., X_n)$. Random variables $X_n$ take values in $(\mathbb{E}, \mathcal{B})$, where $\mathbb{E}$ is a subset of $\mathbb{R}$. Space $(\Omega, \mathcal{F}, \mathbf{P})$ supports also unobservable (hence not measurable with respect to $\mathcal{F}_n$) random variable $\theta$ which has the geometric distribution:

$$\mathbf{P}(\theta = j) = \pi \mathbb{I}_{\{j=0\}} + (1-\pi)p^{j-1}q\mathbb{I}_{\{j\geq 1\}}, \tag{2.1}$$

where $q = 1-p, \pi \in (0,1)$, $j = 1, 2, \ldots$.

We introduce in $(\Omega, \mathcal{F}, \mathbf{P})$ also two time homogeneous and independent Markov processes $\{X_n^0\}_{n\in\mathbb{N}}$ and $\{X_n^1\}_{n\in\mathbb{N}}$ taking values in $(\mathbb{E}, \mathcal{B})$ and assumed to be independent of $\theta$. Moreover, it is assumed that $\{X_n^0\}_{n\in\mathbb{N}}$ and $\{X_n^1\}_{n\in\mathbb{N}}$ have transition densities with respect to a $\sigma$-finite measure $\mu$, i.e., for $i = 0, 1$ and $B \in \mathcal{B}$

$$\mathbf{P}_x^i(X_1^i \in B) \;=\; \mathbf{P}(X_1^i \in B | X_0^i = x) = \int_B f_x^i(y)\mu(dy) = \int_B \mu_x(dy). \tag{2.2}$$

Random processes $\{X_n\}$, $\{X_n^0\}$, $\{X_n^1\}$ and random variable $\theta$ are connected via the rule: conditionally on $\theta = k$

$$X_n = \begin{cases} X_n^0, & \text{if } k > n, \\ X_{n+1-k}^1, & \text{if } k \leq n, \end{cases}$$

where $\{X_n^1\}$ is started from $X_{k-1}^0$ (but is otherwise independent of $X^0$).

Let us introduce the following notation:

$$\begin{aligned} \underline{x}_{k,n} &= (x_k, x_{k+1}, \ldots, x_{n-1}, x_n),\; k \leq n, \\ L_m(\underline{x}_{k,n}) &= \prod_{r=k+1}^{n-m} f_{x_{r-1}}^0(x_r) \prod_{r=n-m+1}^{n} f_{x_{r-1}}^1(x_r), \\ \underline{A}_{k,n} &= \times_{i=k}^n A_i = A_k \times A_{k+1} \times \ldots \times A_n,\; A_i \in \mathcal{B} \end{aligned}$$

where the convention $\prod_{i=j_1}^{j_2} x_i = 1$ for $j_1 > j_2$ is used.

Let us now define functions $S_\cdot(\cdot)$ and $G_\cdot(\cdot,\cdot)$

$$S_n(\underline{x}_{0,n}) = \pi L_n(\underline{x}_{0,n}) + \bar{\pi}\left(\sum_{i=1}^n p^{i-1}qL_{n-i+1}(\underline{x}_{0,n}) + p^nL_0(\underline{x}_{0,n})\right), \tag{2.3}$$

$$\begin{aligned} G_{l+1}(\underline{x}_{n-l-1,n}, \alpha) &= \alpha L_{l+1}(\underline{x}_{n-l-1,n}) + (1-\alpha) \\ &\quad \times \left(\sum_{i=0}^l p^{l-i}qL_{i+1}(\underline{x}_{n-l-1,n}) + p^{l+1}L_0(\underline{x}_{n-l-1,n})\right). \end{aligned} \tag{2.4}$$

where $x_0, x_1, \ldots, x_n \in \mathbb{E}^{n+1}, \alpha \in [0,1], 0 \le n-l-1 < n$.

The function $S(\underline{x}_{0,n})$ stands for the joint density of the vector $\underline{X}_{0,n}$. For any $\underline{D}_{0,n} = \{\omega : \underline{X}_{0,n} \in \underline{B}_{0,n}, B_i \in \mathcal{B}\}$ and any $x \in \mathbb{E}$ we have:

$$\mathbf{P}_x(\underline{D}_{0,n}) = \mathbf{P}(\underline{D}_{0,n}|X_0 = x) = \int_{\underline{B}_{0,n}} S(\underline{x}_{0,n})\mu(d\underline{x}_{0,n})$$

The meaning of the function $G_{n-k+1}(\underline{x}_{k,n}, \alpha)$ will be clear in the sequel.

Roughly speaking our model assumes that the process $\{X_n\}$ is obtained by switching at the random and unknown instant $\theta$ between two Markov processes $\{X_n^0\}$ and $\{X_n^1\}$. It means that the first observation $X_\theta$ after the change depends on the previous sample $X_{\theta-1}$ through the transition pdf $f^1_{X_{\theta-1}}(X_\theta)$. For any fixed $d_1, d_2 \in \{0,1,2,\ldots\}$ (the problem $\mathfrak{D}_{d_1 d_2}$) we are looking for the stopping time $\tau^* \in \mathcal{T}$ such that

$$\mathbf{P}_x(-d_1 \le \theta - \tau^* \le d_2) = \sup_{\tau \in \mathfrak{S}^X} \mathbf{P}_x(-d_1 \le \theta - \tau \le d_2) \tag{2.5}$$

where $\mathfrak{S}^X$ denotes the set of all stopping times with respect to the filtration $\{\mathcal{F}_n\}_{n\in\mathbb{N}}$. Using parameters $d_i$, $i = 1, 2$, we control the precision level of detection. The problem $\mathfrak{D}_{dd}$, *i.e.* the case $d_1 = d_2 = d$, when $\pi = 0$ has been studied in Sarnowski and Szajowski (2011).

## 3. Solution of the problem

Let us denote:

$$\begin{aligned} Z_n^{(d_1,d_2)} &= \mathbf{P}_x(-d_1 \le \theta - n \le d_2 \mid \mathcal{F}_n), \quad n = 0,1,2,\ldots, \\ V_n^{(d_1,d_2)} &= \operatorname*{ess\,sup}_{\{\tau\in\mathfrak{S}^X,\, \tau \ge n\}} \mathbf{P}_x(-d_1 \le \theta - \tau \le d_2 \mid \mathcal{F}_n), \quad n = 0,1,2,\ldots, \\ \tau_0 &= \inf\{n : Z_n^{(d_1,d_2)} = V_n^{(d_1,d_2)}\} \end{aligned} \tag{3.1}$$

Notice that, if $Z_\infty^{(d_1,d_2)} = 0$, then $Z_\tau^{(d_1,d_2)} = \mathbf{P}_x(-d_1 \le \theta - \tau \le d_2 \mid \mathcal{F}_\tau)$ for $\tau \in \mathfrak{S}^X$. Since $\mathcal{F}_n \subseteq \mathcal{F}_\tau$ (when $n \le \tau$) we have

$$V_n^{(d_1,d_2)} = \operatorname*{ess\,sup}_{\tau \ge n} \mathbf{P}_x(-d_1 \le \theta - \tau \le d_2 \mid \mathcal{F}_n) = \operatorname*{ess\,sup}_{\tau \ge n} \mathbf{E}_x(Z_\tau^{(d_1,d_2)} \mid \mathcal{F}_n).$$

The following lemma (see Bojdecki (1979), Sarnowski and Szajowski (2011)) ensures existence of the solution

**Lemma 3.2.** *The stopping time $\tau_0$ defined by formula (3.1) is the solution of problem (2.5).*

By the following lemma we can limit the class of possible stopping rules to $\mathfrak{S}^X_{d_2+1}$ i.e. stopping times equal at least $d_2+1$. Then rule $\tilde{\tau} = \max(\tau, d_1+1)$ is at least as good as $\tau$.

For further considerations let us define the posterior process:

$$\begin{aligned} \Pi_0 &= \pi, \\ \Pi_n &= \mathbf{P}_x\left(\theta \le n \mid \mathcal{F}_n\right), \; n = 1, 2, \ldots \end{aligned}$$

which is designed as information about the distribution of the disorder instant $\theta$. Next lemma transforms the payoff function to the more convenient form.

**Lemma 3.3.** *Let*

$$h(\underline{x}_{0,d_1+1}, \alpha) = \left(1 - p^{d_2} + q \sum_{m=0}^{d_1} \frac{L_{m+1}(\underline{x}_{0,d_1+1})}{p^m L_0(\underline{x}_{0,d_1+1})}\right)(1-\alpha), \tag{3.4}$$

*where* $x_0, \ldots, x_{d_1+1} \in \mathbb{E}$, $\alpha \in (0,1)$, *then*

$$\mathbf{P}_x(-d_1 \leq \theta - n \leq d_2) = \mathbf{E}_x\left[h(\underline{X}_{n-1-d_1,n}, \Pi_n)\right].$$

*Proof.* We rewrite the initial criterion as the expectation

$$\begin{aligned}\mathbf{P}_x(-d_1 \leq \theta - n \leq d_2) &= \mathbf{E}_x\left[\mathbf{P}_x(-d_1 \leq \theta - n \leq d_2 \mid \mathcal{F}_n)\right] \\ &= \mathbf{E}_x\left[\mathbf{P}_x(\theta \leq n + d_2 \mid \mathcal{F}_n) - \mathbf{P}_x(\theta \leq n - d_1 - 1 \mid \mathcal{F}_n)\right]\end{aligned}$$

The probabilities under the expectation can be transformed to the convenient form using the lemmata A1 and A4 of Sarnowski and Szajowski (2011). Next, with the help of Lemma A5 (*ibid*) (putting $l = d_1$) we can express $\mathbf{P}_x(\theta \leq n + d_2 \mid \mathcal{F}_n)$ in terms of $\Pi_n$. Straightforward calculations imply that:

$$\mathbf{P}_x(-d_1 \leq \theta - n \leq d_2 \mid \mathcal{F}_n) = \left(1 - p^{d_2} + q \sum_{m=0}^{d_1} \frac{L_m(\underline{X}_{n-d_1-1,n})}{p^m L_0(\underline{X}_{n-d_1-1,n})}\right)(1 - \Pi_n).$$

This proves the lemma. □

**Lemma 3.5.** *The process* $\{\eta_n\}_{n \geq d_1+1}$, *where* $\eta_n = (\underline{X}_{n-d_1-1,n}, \Pi_n)$, *forms a random Markov function.*

*Proof.* According to Lemma 17 pp. 102–103 in Shiryaev (1978) it is enough to show that $\eta_{n+1}$ is a function of the previous stage $\eta_n$, the variable $X_{n+1}$ and that conditional distribution of $X_{n+1}$ given $\mathcal{F}_n$ is a function of $\eta_n$. Let us consider, for $x_0, \ldots, x_{d_1+2} \in \mathbb{E}, \alpha \in (0,1)$, a function

$$\varphi(\underline{x}_{0,d_1+1}, \alpha, x_{d_1+2}) = \left(\underline{x}_{1,d_1+2}, \frac{f^1_{x_{d_1+1}}(x_{d_1+2})(q + p\alpha)}{G(\underline{x}_{d_1+1,d_1+2}, \alpha)}\right)$$

We will show that $\eta_{n+1} = \varphi(\eta_n, X_{n+1})$. Notice that we get (see Lemma 5 in Sarnowski and Szajowski (2011) ($l = 0$))

$$\Pi_{n+1} = \frac{f^1_{X_n}(X_{n+1})(q + p\Pi_n)}{G(\underline{X}_{n,n+1}, \Pi_n)}. \tag{3.6}$$

Hence

$$\begin{aligned}\varphi(\eta_n, X_{n+1}) &= \varphi(\underline{X}_{n-d_1-1,n}, \Pi_n, X_{n+1}) = \left(\underline{X}_{n-d_1,n}, X_{n+1}, \frac{f^1_{X_n}(X_{n+1})(q + p\Pi_n)}{G(\underline{X}_{n,n+1}, \Pi_n)}\right) \\ &= \left(\underline{X}_{n-d,n+1}, \Pi_{n+1}\right) = \eta_{n+1}.\end{aligned}$$

Define $\hat{\mathcal{F}}_n = \sigma(\theta, \underline{X}_{0,n})$. To see that the conditional distribution of $X_{n+1}$ given $\mathcal{F}_n$ is a function of $\eta_n$, let us consider the conditional expectation of $u(X_{n+1})$ for any Borel function $u : \mathbb{E} \longrightarrow \mathbb{R}$ given $\mathcal{F}_n$. Having Lemma A1 we get:

$$\begin{aligned}\mathbf{E}_x(u(X_{n+1}) \mid \mathcal{F}_n) &= \mathbf{E}_x\left(u(X_{n+1})(1 - \Pi_{n+1}) \mid \mathcal{F}_n\right) + \mathbf{E}_x\left(u(X_{n+1})\Pi_{n+1} \mid \mathcal{F}_n\right) \\ &= \int_{\mathbb{E}} u(y) f^0_{X_n}(y)\mu(dy)\mathbf{P}_x(\theta > n+1 \mid \mathcal{F}_n) + \int_{\mathbb{E}} u(y) f^1_{X_n}(y)\mu(dy)\mathbf{P}_x(\theta \leq n+1 \mid \mathcal{F}_n) \\ &= \int_{\mathbb{E}} u(y)(p(1 - \Pi_n) f^0_{X_n}(y) + (q + p\Pi_n) f^1_{X_n}(y))\mu(dy) = \int_{\mathbb{E}} u(y) G(X_n, y, \Pi_n)\mu(dy)\end{aligned}$$

This is our claim. □

Lemmata 3.3 and 3.5 are crucial for the solution of the posed problem (2.5). They show that the initial problem can be reduced to the problem of stopping Markov random function $\eta_n = (\underline{X}_{n-d_1-1,n}, \Pi_n)$ with the payoff given by the equation (3.4). In the consequence we can use tools of the optimal stopping theory for finding the stopping time $\tau^*$ such that

$$\mathbf{E}_x\left[h(\underline{X}_{\tau^*-d_1-1,\tau^*}, \Pi_{\tau^*})\right] = \sup_{\tau\in\mathfrak{S}^X_{d+1}} \mathbf{E}_x\left[h(\underline{X}_{\tau-d_1-1,\tau}, \Pi_\tau)\right]. \tag{3.7}$$

To solve the reduced problem (3.7) for any Borel function $u : \mathbb{E}^{d_1+2}\times[0,1] \longrightarrow \mathbb{R}$ let us define operators:

$$\begin{aligned}
\mathbf{T}u(\underline{x}_{0,d_1+1}, \alpha) &= \mathbf{E}_x\left[u(\underline{X}_{n-d_1,n+1}, \Pi_{n+1}) \mid \underline{X}_{n-1-d_1,n} = \underline{x}_{0,d_1+1}, \Pi_n = \alpha\right],\\
\mathbf{Q}u(\underline{x}_{0,d_1+1}, \alpha) &= \max\{u(\underline{x}_{0,d_1+1}, \alpha), \mathbf{T}u(\underline{x}_{0,d_1+1}, \alpha)\}.
\end{aligned}$$

By the definition of the operator $\mathbf{T}$ and $\mathbf{Q}$ we get

**Lemma 3.8.** *For the payoff function $h(\underline{x}_{0,d_1+1}, \alpha)$ characterized by (3.4) and for the sequence $\{r_k\}_{k=0}^{\infty}$:*

$$\begin{aligned}
r_0(\underline{x}_{1,d_1+1}) &= p\left[1 - p^{d_2} + q\sum_{m=0}^{d_1}\frac{L_{m-1}(\underline{x}_{1,d_1+1})}{p^m L_0(\underline{x}_{1,d_1+1})}\right],\\
r_k(\underline{x}_{1,d_1+1}) &= p\int_{\mathbb{E}} f^0_{x_{d_1+1}}(x_{d_1+2}) \max\left\{1 - p^{d_2} + q\sum_{m=1}^{d_1+1}\frac{L_m(\underline{x}_{1,d_1+2})}{p^m L_0(\underline{x}_{1,d_1+2})}; r_{k-1}(\underline{x}_{2,d_1+2})\right\}\mu(dx_{d_1+2}),
\end{aligned}$$

*the following formulae hold:*

$$\begin{aligned}
\mathbf{Q}^k h_1(\underline{x}_{1,d_1+2}, \alpha) &= (1-\alpha)\max\left\{1 - p^{d_2} + q\sum_{m=1}^{d_1+1}\frac{L_m(\underline{x}_{1,d_1+2})}{p^m L_0(\underline{x}_{1,d_1+2})}; r_{k-1}(\underline{x}_{2,d_1+2})\right\},\ k\geq 1,\\
\mathbf{T}\,\mathbf{Q}^k h_1(\underline{x}_{1,d_1+2}, \alpha) &= (1-\alpha) r_k(\underline{x}_{2,d_1+2}),\ k\geq 0.
\end{aligned}$$

The following theorem is the main result of the paper.

**Theorem 3.9.** *(a) The solution of the problem (2.5) is given by:*

$$\tau^* = \inf\{n \geq d_1+1 : 1 - p^{d_2} + q\sum_{m=1}^{d_1+1}\frac{L_m(\underline{X}_{n-d_1-1,n})}{p^m L_0(\underline{X}_{n-d_1-1,n})} \geq r^*(\underline{X}_{n-d_1,n})\} \tag{3.10}$$

*where $r^*(\underline{X}_{n-d,n}) = \lim_{k\longrightarrow\infty} r_k(\underline{X}_{n-d,n})$.*

*(b) The value of the problem,* i.e. *the maximal probability for (2.5) given $X_0 = x$, is equal to*

$$\begin{aligned}
\mathbf{P}_x(-d_1 \leq \theta - \tau^* \leq d_2) = p^{d_1+1}\int_{\mathbb{E}^{d_1+1}} &\max\left\{1 - p^{d_2} + q\sum_{m=1}^{d_1+1}\frac{L_m(x, \underline{x}_{1,d_1+1})}{p^m L_0(x, \underline{x}_{1,d_1+1})}; r^*(\underline{x}_{1,d_1+1})\right\}\\
&\times L_0(x, \underline{x}_{1,d_1+1})\mu(dx_1, \ldots, dx_{d_1+1}).
\end{aligned}$$

## 4. Final remarks

The presented analysis of the problem when the acceptable error for stopping before or after the disorder allows to see which protection is more difficult to control. When we admit that it is possible sequence of observation without disorder it is interesting question how to detect not only that we observe the second kind of data but that there were no data of the first kind. It can be verified by standard testing procedure when we stop very early ($\tau \leq \min\{d_1, d_2\}$).

## References

Baron, M. (2004). Early detection of epidemics as a sequential change-point problem. In *Longevity, aging and degradation models in reliability, public health, medicine and biology, LAD 2004. Selected papers from the first French-Russian conference, St. Petersburg, Russia, June 7–9, 2004* (V. Antonov, C. Huber, M. Nikulin and V. Polischook, eds.). *IMS Lecture Notes-Monograph Series* **2** 31–43. St. Petersburg State Politechnical University, St. Petersburg, Russia.

Basseville, M. and Benveniste, A., eds. (1986). *Detection of abrupt changes in signals and dynamical systems. Lecture Notes in Control and Information Sciences* **77** 373. Springer-Verlag, Berlin.

Bojdecki, T. (1979). Probability maximizing approach to optimal stopping and its application to a disorder problem. *Stochastics* **3** 61-71.

Moustakides, G. V. (1998). Quickest detection of abrupt changes for a class of random processes. *IEEE Trans. Inf. Theory* **44** 1965-1968.

Sarnowski, W. and Szajowski, K. (2011). Optimal detection of transition probability change in random sequence. *Stochastics* **83** 569-581.

Shewhart, W. A. (1931). *Economic control of quality of manufactured products.* D. Van Nostrand, Yew York.

Shiryaev, A. N. (1961). The detection of spontaneous effects. *Sov. Math, Dokl.* **2** 740-743. translation from Dokl. Akad. Nauk SSSR 138, 799-801 (1961).

Shiryaev, A. N. (1978). *Optimal Stopping Rules.* Springer-Verlag, New York, Heidelberg, Berlin.

Tartakovsky, A. G., Rozovskii, B. L., Blažek, R. B. and Kim, H. (2006). Detection of intrusions in information systems by sequential change-point methods. *Stat. Methodol.* **3** 252–293. MR2240956

Yakir, B. (1994). Optimal detection of a change in distribution when the observations form a Markov chain with a finite state space. In *Change-point Problems. Papers from the AMS-IMS-SIAM Summer Research Conference held at Mt. Holyoke College, South Hadley, MA, USA, July 11–16, 1992* (E. Carlstein, H. G. Müller and D. Siegmund, eds.). *IMS Lecture Notes-Monograph Series* **23** 346–358. Institute of Mathematical Statistics, Hayward, California.

Yoshida, M. (1983). Probability maximizing approach for a quickest detection problem with complicated Markov chain. *J. Inform. Optimization Sci.* **4** 127-145.