Sign CUSUM Algorithm for Change-Point Detection of the MMPP Controlling Chain State

Yulia Burkatovskaya[1,2,3], Tatiana Kabanova[2(✉)], and Olga Tokareva[1]

[1] Institute of Cybernetics, Tomsk Polytechnic University,
30 Lenin Prospekt, 634050 Tomsk, Russia
tracey@tpu.ru, ostokareva@gmail.com

[2] Department of Applied Mathematics and Cybernetics,
Tomsk State University, 36 Lenin Prospekt, 634050 Tomsk, Russia
tvk@bk.ru

[3] International Laboratory of Statistics of Stochastic Processes
and Quantitative Finance, Tomsk State University,
36 Lenin Prospekt, 634050 Tomsk, Russia
http://www.tpu.ru, http://www.tsu.ru

Abstract. The authors consider the Markov modulated Poisson process with two states of the Markovian controlling chain. The flow intensity of the observed process depends on the unobserved controlling chain state. All the process parameters are supposed to be unknown. The paper develops a new sequential change-point detection method based on the cumulative sum control chart approach to determine the switching points of the flow intensity. Usage of special sign statistics allows the obtaining of theoretical characteristics of the proposed algorithm.

Keywords: Markov modulated Poisson process · Jump intensity · Change-point detection · Sign statistics · CUSUM algorithm

1 Introduction

The Markov-modulated Poisson process (MMPP) has been extensively used for modeling Poisson processes whose arrival intensities vary randomly over time. It qualitatively models the time-varying arrival rate and captures some of the important correlations between the interarrival times while still remaining analytically tractable [1]. It can be described as a Poisson process whose intensity is determined by a controlling chain state. Transition between the states occurs in unknown random instants, and the time of being in a state is distributed exponentially.

MMPP arise in many applications of interest, such as Web-servers, multimedia traffic, call-centers, cell phone call activity, product demand, etc. Mingrui Zou and Jianqing Liu in [2] use MMPP to investigate an unsaturated IEEE 802.16 network with the contention-based access mechanism. The authors model packet arrivals at each subscriber station as a MMPP) and derive analytical expressions

A. Dudin et al. (Eds.): ITMM 2016, CCIS 638, pp. 18–33, 2016.
DOI: 10.1007/978-3-319-44615-8_2

for the network throughput and packet delay subject to the MMPP parameters, i.e., the steady-state probabilities and the average arrival rate.

Nogueira et al. in [3] use a superposition of discrete time MMPP model (dMMPP) for the modeling of network traffic on multiple time scales. Two Markovian models are proposed: the fitting procedure of the first model matches the complete distribution of the arrival process at each timescale of interest, while the second proposed model is constructed using a hierarchical procedure that decomposes each MMPP state into new MMPPs that incorporate a more detailed description of the distribution at finer time scales. The traffic process is then represented by a MMPP equivalent to the constructed hierarchical structure. Both approaches use estimators of the characterizing parameters of each MMPP, that is, the matrices corresponding to the transition probabilities and the Poisson arrival rates for each state.

Giacomazzi in [4] develops a method for using traffic sources modeled as a MMPP in the framework of the bounded-variance network calculus, a novel stochastic network calculus framework for the approximated analysis of end-to-end network delay. The mean and the variance of the cumulative traffic are analyzed for two traffic envelopes, the first, the two-moment envelope, is an approximation of traffic with the same first two moments of the actual source traffic. The second, the linear envelope, provides a less precise approximation but it permits the closed-form analysis of single-node and end-to-end delay with several types of important schedulers.

Choi et al. in [5] consider MMPP as model for traffic streams with bursty characteristic and time correlation between interarrival times. Traffics such as voice and video in ATM networks have these properties. By using the embedded Markov chain method, the authors derive the queue length distribution at departure epochs. They also obtain the queue length distribution at an arbitrary time by the supplementary variable method. The authors apply the results for preventive congestion control in telecommunication networks.

There exist a number of rather complicated flow models based on the MMPP. Vasil'eva and Gortsev in [6] study an asynchronous double stochastic flow of events where each event results in a dead time period when other events cannot be observed. The authors determine the Laplace transformation of the event-event interval probability density in the observed flow and derive the equations of moments for the estimation of dead time and initial event flow parameters. Gortsev and Nezhelskaya in [7] study the stationary mode of an asynchronous double stochastic flow with initiation of superfluous events. The authors determine important properties of the flow studied as interval probability density and joint probability density of neighboring intervals lengths. Also this work specifies conditions in which the flow either becomes recursive or degenerates to an elementary one.

The papers surveyed above describe some applications of the MMPP process but of course not all of them. To take decisions concerning process behavior and to develop dispatching rules one needs to fit a model and to evaluate the model parameters. There are two classical approaches to the MMPP parameter estimation problem: maximum likelihood estimation and its implementation via the EM (expectation-maximization) algorithm and matching moment method. A detailed survey of former methods is given in [8]. The survey [9] with a huge bibliography is focused on the latter methods. These approaches are connected with complicated numerical calculations. It implies difficulties in their theoretical investigations and, hence, the necessity of their study via simulation. That stresses the urgency of the MMPP parameter estimation problem and necessity of simple efficient estimation algorithms, whose non-asymptotic properties can be investigated theoretically.

In [10], we a the sequential analysis approach to parameter estimation for MMPP process. The time intervals between the observed flow events were considered as the values of a stochastic process. The mean of the process was supposed to change in some unknown instants, or change points.

The problem of sequential change-point detection can be formulated as follows. A stochastic process is observed. Several parameters of the process change in random points. The problem is to detect the change points when the process is observed online. Sequential methods include a special stopping rule that determines a stopping time. At this instant, a decision on the change point can be made. There are two types of errors typical for sequential change-point detection procedures: a false alarm, when one makes a decision that change has occurred before a change point (type 1 error), and delay, when the change is not detected (type 2 error).

At the first stage of the algorithm, these change points were detected by using CUSUM (cumulative sum control chart) algorithms. After that the intensity parameters were estimated under the assumption that the intensity was constant between detected change points. The quality of the proposed algorithms was studied via simulation.

In this paper, a modification of the CUSUM algorithm is proposed. Instead of the lengths of the interval between events, special sign statistics are used at the change-point detection procedure. It allows us to investigate theoretically the properties of the algorithm and give recommendations concerning the parameter choice.

2 Problem Statement

We consider a Markov-modulated poisson process, i.e. a flow of events controlled by a Markovian chain with a continuous time. The chain has two states and transition between the states happens at random instants. The time of sojourn of the chain in the i-th state is exponentially distributed with the parameter α_i, $i = 1, 2$.

The flow of events has exponential distribution with the intensity parameter λ_1 or λ_2 subject to the state of the Markovian chain. We suppose that the intensity of the switch between the controlling chain states is sufficiently smaller then the intensity of the arrival process, i.e. $\alpha_i << \lambda_i$. In this case, several events commonly occur before the change of the controlling chain state. The parameters of the system λ_1, λ_2 and the instants of switching between the states are supposed to be unknown. The sequence of instants of arriving events is observed. The problem is to estimate the parameters λ_1, λ_2, α_1, α_2.

3 Sign CUSUM Algorithm

Consider the process $\{\tau_i\}_{i\geq 1}$, where $\tau_i = t_i - t_{i-1}$ is the length of the i-th interval between arriving events in the observed flow. Figure 1 demonstrates the construction of the sequence $\{\tau_i\}$.

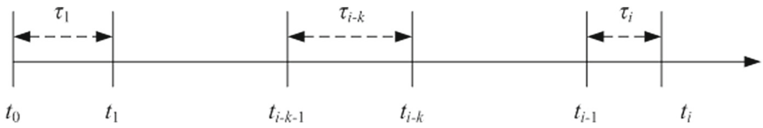

Fig. 1. Construction of the sequence $\{\tau_i\}$.

If the controlling chain is in the l-th state then the mean length between events is equal to $1/\lambda_l$. So at the first stage of our procedure we try to detect the instants of the chain transition from one state to another as the instants of change in the mean of the process $\{\tau_i\}_{i\geq 1}$ using the Sign CUSUM procedure. At the first time such algorithm was proposed in [11] and developed in [12] for the case of the single change point. In this paper we consider the multiple change-point detection problem which has some special features.

Now we describe the procedure. Let the parameters λ_1, λ_2 satisfy the condition

$$\begin{aligned} &0 < \lambda_2 < \lambda_1; \\ &\frac{1}{\lambda_2} - \frac{1}{\lambda_1} > \Delta, \end{aligned} \tag{1}$$

where Δ is a certain known positive parameter. Choose then an integer parameter k describing the memory depth. The idea is to compare the values τ_i and τ_{i-k}. If there are no changes in the controlling chain state within the interval $[t_{i-k-1}, t_i]$ then the variables τ_i and τ_{i-k} have the identical exponential distribution with the mean $1/\lambda_1$ or $1/\lambda_2$. If the chain state changes within the interval $[t_{i-k-1}, t_i]$ then the expectations of the values τ_i and τ_{i-k} are different.

For the case of the single change-point detection the parameter k is taken commonly as large as possible in order to guarantee that the decision is taken while the means of the variables are different. For our case of multiple change-point detection the parameter k should not be too large so as to contain more than one chain state change within the interval $[t_{i-k-1}, t_i]$. In paper [10] we recommended choosing $k \approx r/2$, where $\lambda_i \geq r\alpha_i$. Further we consider the choice of the algorithm parameters in more detail.

As the initial state of the chain is unknown, we shall consider two CUSUM procedures simultaneously. The first procedure is set up to detect an increase in the mean of the process and, hence, decrease of the intensity, and the second procedure is set up to detect an decrease in the mean and, hence, increase of the intensity. For the first procedure we introduce the sequence of the statistics

$$z_i^{(1)} = n\left(\mathrm{sign}(\tau_i - \tau_{i-k}) - \delta\right), \quad i > k. \tag{2}$$

For the second procedure we introduce the sequence of the statistics

$$z_i^{(2)} = n\left(\mathrm{sign}(\tau_{i-k} - \tau_i) - \delta\right), \quad i > k. \tag{3}$$

These statistics are calculated at the instant t_i. Here $\delta = m/n$, m and n are integers, $m < n$, and the fraction m/n is irreducible.

Consider then four hypothese concerning the state of the controlling chain:

- $H_1(t_{i-k-1}, t_i)$ – the intensity of the arrival process on the interval $[t_{i-k-1}, t_i]$ is constant and equal to λ_1;
- $H_2(t_{i-k-1}, t_i)$ – the intensity of the arrival process on the interval $[t_{i-k-1}, t_i]$ is constant and equal to λ_2;
- $H_{1,2}(t_{i-k}, t_{i-1})$ – the intensity of the arrival process on the interval $[t_{i-k}, t_{i-1}]$ changed once from λ_1 to λ_2, i.e., decreased;
- $H_{2,1}(t_{i-k}, t_{i-1})$ – the intensity of the arrival process on the interval $[t_{i-k}, t_{i-1}]$ changed once from λ_2 to λ_1, i.e., increased.

Note that in the conditions of the hypothesis $H_l(t_{i-k-1}, t_i)$ the random variables τ_i and τ_{i-k} have the same mean and the functions $\mathrm{sign}(\tau_i - \tau_{i-k})$ and $\mathrm{sign}(\tau_{i-k} - \tau_i)$ take the values 1 and –1 with the equal probabilities 1/2. So, in this case the expectation of the statistics $z_i^{(l)}$ is negative. To implement CUSUM procedures it is necessary to provide positive expectations of the statistics $z_i^{(1)}$ and $z_i^{(2)}$ after an increase or decrease in the mean of τ_i correspondingly. So, introducing the notations

$$\begin{aligned} p = P\{\tau_i \geq \tau_{i-k} | H_{1,2}(t_{i-k}, t_{i-1})\} = P\{\tau_{i-k} \geq \tau_i | H_{2,1}(t_{i-k}, t_{i-1})\}; \\ q = P\{\tau_i < \tau_{i-k} | H_{1,2}(t_{i-k}, t_{i-1})\} = P\{\tau_{i-k} < \tau_i | H_{2,1}(t_{i-k}, t_{i-1})\}, \end{aligned} \tag{4}$$

where $p > 1/2$, $q < 1/2$, $p + q = 1$ we can obtain the following result.

Theorem 1. *If the parameter δ satisfies the condition*

$$\delta < 2p - 1, \tag{5}$$

then the statistics $z_i^{(j)}$, $j \in \{1, 2\}$ (2) and (3) have the following properties:

$$\begin{aligned} E\left[z_i^{(1)} \middle| H_l(t_{i-k-1}, t_i) \right] &= -m < 0, \quad l = 1, 2; \\ E\left[z_i^{(1)} \middle| H_{2,1}(t_{i-k}, t_{i-1}) \right] &= -n(2p-1) - m < 0; \\ E\left[z_i^{(1)} \middle| H_{1,2}(t_{i-k}, t_{i-1}) \right] &= n(2p-1) - m > 0; \\ E\left[z_i^{(2)} \middle| H_l(t_{i-k-1}, t_i) \right] &= -m < 0, \quad l = 1, 2; \\ E\left[z_i^{(2)} \middle| H_{1,2}(t_{i-k}, t_{i-1}) \right] &= -n(2p-1) - m < 0, \\ E\left[z_i^{(2)} \middle| H_{2,1}(t_{i-k}, t_{i-1}) \right] &= n(2p-1) - m > 0. \end{aligned} \tag{6}$$

Proof. Using (4) and (5) one obtains

$$\begin{aligned}
E\left[z_i^{(1)} \middle| H_l(t_{i-k-1}, t_i) \right] &= n\left(E\left[\text{sign}(\tau_i - \tau_{i-k}) \middle| H_l(t_{i-k-1}, t_i)\right] - \delta\right) \\
&= n((1/2 - 1/2) - \delta) = -m < 0; \\
E\left[z_i^{(1)} \middle| H_{2,1}(t_{i-k}, t_{i-1}) \right] &= n\left(E\left[\text{sign}(\tau_i - \tau_{i-k}) \middle| H_{2,1}(t_{i-k}, t_{i-1})\right] - \delta\right) \\
&= n((q-p) - \delta) = -n(2p-1) - m < 0; \\
E\left[z_i^{(1)} \middle| H_{1,2}(t_{i-k}, t_{i-1}) \right] &= n\left(E\left[\text{sign}(\tau_i - \tau_{i-k}) \middle| H_{1,2}(t_{i-k}, t_{i-1})\right] - \delta\right) \\
&= n((p-q) - \delta) = n(2p-1) - m > 0; \\
E\left[z_i^{(2)} \middle| H_l(t_{i-k-1}, t_i) \right] &= n\left(E\left[\text{sign}(\tau_{i-k} - \tau_i) \middle| H_l(t_{i-k-1}, t_i)\right] - \delta\right) \\
&= n((1/2 - 1/2) - \delta) = -m < 0; \\
E\left[z_i^{(2)} \middle| H_{1,2}(t_{i-k}, t_{i-1}) \right] &= n\left(E\left[\text{sign}(\tau_{i-k} - \tau_i) \middle| H_{1,2}(t_{i-k}, t_{i-1})\right] - \delta\right) \\
&= n((q-p) - \delta) = -n(2p-1) - m < 0; \\
E\left[z_i^{(2)} \middle| H_{2,1}(t_{i-k}, t_{i-1}) \right] &= n\left(E\left[\text{sign}(\tau_{i-k} - \tau_i) \middle| H_{2,1}(t_{i-k}, t_{i-1})\right] - \delta\right) \\
&= n((p-q) - \delta) = n(2p-1) - m < 0
\end{aligned}$$

So the average values of statistics (2) and (3) change from a negative value to positive when the intensity of the process changes. Besides, the statistic $z_i^{(1)}$ reacts to decrease the intensity, i.e., to increase the mean length of the interval between events; the statistic $z_i^{(2)}$ reacts to increase the intensity, i.e., to decrease the mean length of the interval between events. These properties determine the construction of the procedures. We introduce positive values h_1 and h_2 as the procedures thresholds and construct the cumulative sums $S_i^{(1)}$ and $S_i^{(2)}$ which are recalculated at the instants t_i. For the first procedure it is defined as follows

$$\begin{gathered}
S_k^{(1)} = m + n; \\
S_i^{(1)} = \max\{m+n, S_{i-1}^{(1)} + z_i^{(1)}\}, \quad i > k; \\
S_i^{(1)} = m + n, \quad \text{if} \quad S_i^{(1)} \geq h_1.
\end{gathered} \tag{7}$$

For the second procedure the cumulative sum is defined as follows

$$\begin{gathered}
S_k^{(2)} = m + n; \\
S_i^{(2)} = \max\{m+n, S_{i-1}^{(2)} + z_i^{(2)}\}, \quad i > k; \\
S_i^{(2)} = m + n, \quad \text{if} \quad S_i^{(2)} \geq h_2.
\end{gathered} \tag{8}$$

If the cumulative sum $S_i^{(1)}$ reaches the threshold h_1 then we make the decision that the mean time between events increased and hence the intensity of the process decreased, i.e., it changed from λ_1 to λ_2. If the cumulative sum $S_i^{(2)}$ reaches the threshold h_2 then we make the decision that the mean time between events decreased and hence the intensity of the process increased, i.e., it changed from λ_2 to λ_1. Once the sum reaches the threshold it is reset to $m + n$ and the corresponding procedure is restarted.

In connection with sequential change-point detection procedures two types of errors are considered: the false alarm and the skip of the change. A false alarm occurs when one of the cumulative sums reaches the corresponding threshold in the case of the constant intensity of the arrival process. These events can be described as follows:

$$\begin{aligned} F_1 &= \left\{ S_i^{(1)} \geq h_1 \middle| H_l(t_{i-k-1}, t_i) \cup H_{2,1}(t_{i-k}, t_{i-1}) \right\}; \\ F_2 &= \left\{ S_i^{(2)} \geq h_2 \middle| H_l(t_{i-k-1}, t_i) \cup H_{1,2}(t_{i-k}, t_{i-1}) \right\}. \end{aligned} \tag{9}$$

A skip of the change occurs when the change of the parameter occurs but the corresponding cumulative sum does not reach its threshold. These events can be described as follows:

$$\begin{aligned} G_1 &= \left\{ S_i^{(1)} < h_1 \middle| H_{1,2}(t_{i-k}, t_{i-1}) \right\}; \\ G_2 &= \left\{ S_i^{(2)} < h_2 \middle| H_{2,1}(t_{i-k}, t_{i-1}) \right\}. \end{aligned} \tag{10}$$

4 Characteristics of the Algorithm

4.1 Probability *p*

First, we calculate the probability p (4) as

$$p = P\{\tau_i \geq \tau_{i-k} | H_{1,2}(t_{i-k}, t_{i-1})\}.$$

In the conditions of the hypothesis $H_{1,2}(t_{i-k}, t_{i-1})$ the variable τ_{i-k} is distributed exponentially with the parameter λ_1, and the variable τ_i is distributed exponentially with the parameter λ_2, hence

$$p = \int_0^\infty P\{\tau_i > t\} dP\{\tau_{i-k} < t\} = \int_0^\infty e^{-\lambda_2 t} \lambda_1 e^{-\lambda_1 t} dt = \frac{\lambda_1}{\lambda_1 + \lambda_2}.$$

According to (1), $p > 1/2$, and we obtain

$$p = \frac{\lambda_1}{\lambda_1 + \lambda_2}, \quad q = \frac{\lambda_2}{\lambda_1 + \lambda_2}. \tag{11}$$

4.2 Average Delay

Then, we investigate the characteristics of the change-point detection procedure if the change occurs more then once. Figure 4 demonstrates an example of the multiple change point.

In the case of the multiple change-point detection problem any parameter change should be detected before the next change occurs. Without loss of generality, we consider the first procedure and suppose that at the instant θ_1 the intensity parameter changes from λ_1 to λ_2, and then at the instant θ_2 it changes from λ_2 to λ_1, so, our procedure should detect the first change. The first change occurs within the interval $[t_{i-1}, t_i]$, and the second change occurs within the

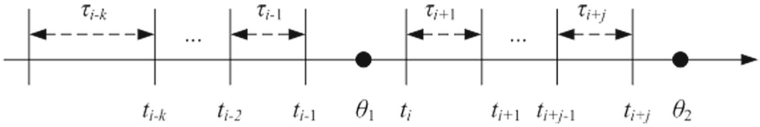

Fig. 2. Multiple change point

interval $[t_{i+j}, t_{i+j+1}]$. It means that the first change should be detected at the interval $[t_i, t_j]$, i.e., the cumulative sum should reach the threshold no later than at the instant t_j. On the other hand, the expectation of the statistics $z_{i+a}^{(1)}$ is positive if and only if $a \leq j$, $a \leq k-1$, so, the change should be detected not more than in $k-1$ steps. Consequently, the most important characteristics of any multiple change-point detection algorithm are those connected with the delay in the detection (Fig. 2).

Note that in the conditions of the hypothesis $H_{1,2}(t_{i-k}, t_i)$, at the instant t_{i+1}, the cumulative sum can increase by $n-m$ with the probability p, and it can decrease by $n+m$ with the probability q. Let us introduce the following notations

$$n + m = N, \quad n - m = M.$$

To simplify the further calculations, we suppose that $M > 1$.

The cumulative sum is recalculated at the moments when the flow events occur. We will call every such recalculation a step. We denote the mean number of steps for the sum $S_i^{(1)}$ necessary to reach the threshold $h^{(1)}$ if at the moment i the sum is equal to j as $T^{(1)}(j)$. Taking into account that N is the minimal value of the cumulative sum one obtains that the mean delay for the first procedure is

$$T_{delay}^{(1)} = T^{(1)}(N). \tag{12}$$

The values $T^{(1)}(j)$ satisfy the following set of equations [11,12]

$$T^{(1)}(j) = 1 + pT^{(1)}(j+M) + qT^{(1)}(j-N) \tag{13}$$

with the initial conditions

$$\begin{gathered} T^{(1)}(h^{(1)}) = ... = T^{(1)}(h^{(1)} + M - 1) = 0; \\ T^{(1)}(0) = ... = T^{(1)}(N-1) = T^{(1)}(N). \end{gathered} \tag{14}$$

The decision of this system is obtained at [11,12] and can be written as follows

$$T^{(1)}(j) = \frac{j}{qN - pM} + A_1\mu_1^j + ... + A_{N+M}\mu_{N+M}^j, \tag{15}$$

where $\mu_1, ..., \mu_{N+M}$ are the roots of the characteristic polynomial of system (13)

$$P_1(\mu) = p\mu^{N+M} - \mu^N + q. \tag{16}$$

The exact formulas for the decision (15) are given in [11,12]. For the second procedure one can obtain the same result. These imply the following lower bound for the memory depth parameter k

$$k \geq T^{(1)}(N) + 1. \tag{17}$$

Table 1 demonstrates some values of the mean delay $T^{(1)} = T^{(1)}(N)$ subject to the parameters λ_1, λ_2 and $h^{(1)}$. Here $\delta = 1/5$, i.e., $M = 4$, $N = 6$.

Table 1. Average delay

λ_1	λ_2	p	q	h	$T^{(1)}$	h	$T^{(1)}$	h	$T^{(1)}$	h	$T^{(1)}$
2	0.4	0.83	0.17	42	12.12	62	19.01	77	24.12	97	30.9
2	0.6	0.77	0.23	37	13.27	52	20.24	72	29.23	87	35.94
2	0.8	0.71	0.29	32	15.85	47	26.04	67	39.34	82	49.23
2	1	0.67	0.33	32	19.07	47	33.04	62	47.23	77	61.57

Let us calculate now the mean delay not in terms of steps, but in terms of real time between the change point and the instant of its detection. If the change point is detected in b steps, i.e., at the instant t_{i+b} then the time necessary to detect the change can be expressed as follows

$$t_{i+b} - t_i = \sum_{j=1}^{b} \tau_{i+j}.$$

So the mean time of the change-point detection can be written in the form

$$Q^{(1)}_{delay} = E \sum_{j=1}^{b} \tau_{i+j}.$$

If $\theta_1 \in [t_{i-1}, t_i]$, $k > b$ and $\theta_2 > t_{i+b}$ then the random variables τ_{i+j} and b satisfy the following conditions

(a) $\{\tau_{i+j}\}$ are all finite-mean random variables having the same expectations;
(b) $E\tau_{i+j}\mathbf{1}_{b \geq i+j} = E\tau_{i+j}\mathbf{1}_{S^{(1)}_{i+j-1} < h^1} = E\tau_{i+j}P(S^{(1)}_{i+j-1} < h^1)$;
(c) b has a finite expectation.

So, using the Wald's identity [13] one can obtain

$$Q^{(1)}_{delay} = E\tau_i Eb = \frac{1}{\lambda_2} T^{(1)}_{delay}. \tag{18}$$

The change point θ_1 should be detected earlier then the next change occurs, so, the necessary condition for this is $t_{i+b} < \theta_2$, or $t_{i+b} - \theta_1 < \theta_2 - \theta_1$. Rewriting the left side of the expression

$$t_{i+b} - \theta_1 = \sum_{j=1}^{b} \tau_{i+j} + (t_i - \theta_1)$$

and taking into account that $t_i - \theta_1$ has exponential distribution with the parameter λ_2, we obtain the following condition

$$E(t_{i+b} - \theta_1) = Q^{(1)}_{delay} + E(t_i - \theta_1) \leq Q^{(1)}_{delay} + \frac{1}{\lambda_2} = \frac{1}{\lambda_2}\left(T^{(1)}_{delay} + 1\right).$$

The variable $\theta_2 - \theta_1$ is distributed exponentially with the parameter α_2. This and (18) implies the condition

$$\frac{1}{\lambda_2}\left(T^{(1)}_{delay} + 1\right) < \frac{1}{\alpha_2} \tag{19}$$

4.3 Average Time Between False Alarms

A false alarm occurs when the cumulative sum reaches the corresponding threshold in the conditions of the hypothesis $H_l(t_{i-k-1}, t_{i+1})$. At the instant t_{i+1}, the cumulative sum can increase by M or decrease by N with the probability 1/2.

Let us denote the mean number of steps for the sum $S^{(1)}_i$ necessary to reach the threshold $h^{(1)}$ if at the moment i the sum is equal to j as $R^{(1)}(j)$. Taking into account that N is the minimal value of the cumulative sum one obtains that the mean number of steps between false alarms for the first procedure is

$$T^{(1)}_{alarm} = R^{(1)}(N). \tag{20}$$

The values $R^{(1)}(j)$ satisfy the following set of equations [11,12]

$$R^{(1)}(j) = 1 + \frac{1}{2}R^{(1)}(j+M) + \frac{1}{2}R^{(1)}(j-N) \tag{21}$$

with the initial conditions

$$\begin{gathered} R^{(1)}(h^{(1)}) = ... = R^{(1)}(h^{(1)} + M - 1) = 0; \\ R^{(1)}(0) = ... = R^{(1)}(N-1) = R^{(1)}(N). \end{gathered} \tag{22}$$

The decision of this system is obtained at [11,12] and can be written as follows

$$R^{(1)}(j) = \frac{2j}{N-M} + C_1\mu_1^j + ... + C_{N+M}\mu_{N+M}^j, \tag{23}$$

where $\mu_1, ..., \mu_{N+M}$ are the roots of the characteristic polynomial of system (21)

$$P_1(\mu) = \mu^{N+M} - 2\mu^N + 1. \tag{24}$$

The exact formulas for the decision (23) are given in [11,12]. For the second procedure one can obtain the same result.

Table 2 demonstrates some values of the mean number of steps between false alarms $R^{(1)} = R^{(1)}(N)$ subject to the parameter $h = h^{(1)}$.

As for the mean delay, we can obtain the formula for the mean time between the false alarms

$$Q^{(1)}_{false} = \frac{1}{\lambda_2}T^{(1)}_{alarm}. \tag{25}$$

Table 2. Average time between false alarms

h	$R^{(1)}$	h	$R^{(1)}$
20	12,31	50	195,17
30	37,99	60	385,13
40	91,86	70	727,26

According to G. Lorden [14], a sequential change-point detection procedure is optimal if both the average delay and the logarithm of the average time between false alarms grow linearly with an increase in the parameter h. Figures 3 and 4 demonstrate this property. Figure 4 indicates that the growth rates of the average delay increase with the decreasing in the probability p.

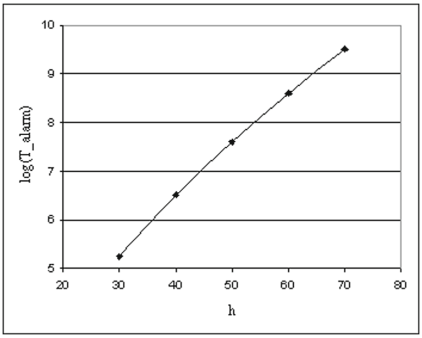

Fig. 3. Logarithm of the average time between false alarms

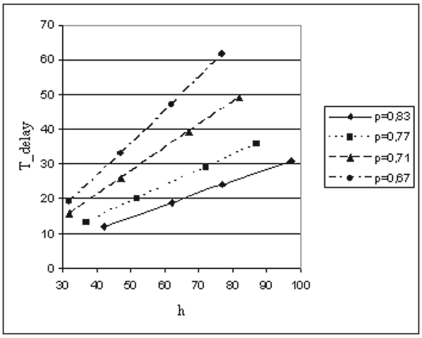

Fig. 4. Average delay

4.4 Memory Depth

Now we consider the event $B_j^{(1)} = \left\{ \sum_{a=1}^{j} \tau_{i+a} + (t_i - \theta_1) < \theta_2 - \theta_1 \right\}$, which means that the next change of the controlling chain state occurs after the instant t_{i+j}.

All the variables τ_a and $t_i - \theta_1$ are independent and identically distributed with the parameter λ_2. Their sum has the gamma distribution Gamma$(1/\lambda_2, j+1)$:

$$f(x) = \begin{cases} \lambda_2^{j+1} x^j \dfrac{e^{-\lambda_2 x}}{j!}, & \text{if } x \geq 0; \\ 0, & \text{if } x < 0. \end{cases}$$

The difference $\theta_2 - \theta_1$ has the exponential distribution with the parameter α_2; hence, we obtain

$$P\left(\sum_{a=1}^{j} \tau_{i+a} + (t_i - \theta_1) < \theta_2 - \theta_1\right) = \int_0^\infty P\left(\theta_2 - \theta_1 > x\right) \lambda_2^{j+1} x^j \frac{e^{-\lambda_2 x}}{j!} dx$$
$$= \frac{\lambda_2^{j+1}}{j!} \int_0^\infty e^{-\alpha_2 x} x^j e^{-\lambda_2 x} dx = \frac{\lambda_2^{j+1}}{j!} \frac{j!}{(\lambda_2 + \alpha_2)^{j+1}} = \frac{\lambda_2^{j+1}}{(\lambda_2 + \alpha_2)^{j+1}}.$$

For the change of the controlling chain state from the second to the first, we have the same result. Finally, we obtain

$$P\left(B_j^{(1)}\right) = \frac{\lambda_2^{j+1}}{(\lambda_2 + \alpha_2)^{j+1}}, \; P\left(B_j^{(2)}\right) = \frac{\lambda_1^{j+1}}{(\lambda_1 + \alpha_1)^{j+1}} \tag{26}$$

On the one hand, the next change point should be detected in not more than k steps. On the other hand, the next change should occur later than the previous is detected. So, the probability of the event B_k should exceed the prescribed probability Q which is close to one. This implies the upper bound for the value of the parameter k

$$\frac{\lambda_l^{k+1}}{(\lambda_l + \alpha_l)^{k+1}} \geq Q, \quad l = 1, 2. \tag{27}$$

Before we supposed that $\lambda_l \geq r\alpha_l$, where $r >> 1$. Table 3 demonstrates the maximum values of k satisfying condition (27) subject to the values of Q (in the rows) and r (in the columns). In practice, the values of $k < 10$ lead to frequent skips of changes; hence, the empty cells correspond to such values of k.

Table 3. Upper bound for the memory depth

$Q \backslash r$	40	60	80	100	120	160	200
0,7	13	22	27	34	41	56	70
0,75	11	16	22	27	33	45	56
0,8		12	16	21	24	34	43
0,85			12	15	18	25	31
0,9					11	16	20

4.5 Error Probabilities

Now we consider the event $A_j = \bigcup_{a=1}^{j} S_{i+a}^{(1)} < h^{(1)}$ which means that the first change point is not detected in j steps. In every step, the cumulative sum can increase by M with the probability p and can decrease by N with the probability q. So, for the event $S_{i+j}^{(1)} < h^{(1)}$, if the condition A_{j-1} holds true, we have

$$P\left(S_{i+j}^{(1)} < h^{(1)}\right) = p\mathbf{1}_{h^{(1)}-M>N} P\left(S_{i+j-1}^{(1)} < h^{(1)} - M\right) \\ +q\mathbf{1}_{h^{(1)}>N} P\left(S_{i+j-1}^{(1)} < h^{(1)}\right).$$

By using this formula again for its right side, we obtain

$$P\left(S_{i+j}^{(1)} < h^{(1)}\right) = p\mathbf{1}_{h^{(1)}-M>N}\left[p\mathbf{1}_{h^{(1)}-2M>N} P\left(S_{i+j-2}^{(1)} < h^{(1)} - 2M\right)\right. \\ \left. + \, q\mathbf{1}_{h^{(1)}-M>N} P\left(S_{i+j-2}^{(1)} < h^{(1)} - M\right)\right] \\ +q\mathbf{1}_{h^{(1)}>N}\left[p\mathbf{1}_{h^{(1)}-M>N} P\left(S_{i+j-2}^{(1)} < h^{(1)} - M\right)\right. \\ \left. + \, q\mathbf{1}_{h^{(1)}>N} P\left(S_{i+j-2}^{(1)} < h^{(1)}\right)\right] = p^2\mathbf{1}_{h^{(1)}-2M>N} P\left(S_{i+j-2}^{(1)} < h^{(1)} - 2M\right) \\ +2pq\mathbf{1}_{h^{(1)}-M>N} P\left(S_{i+j-2}^{(1)} < h^{(1)} - M\right) + q^2\mathbf{1}_{h^{(1)}>N} P\left(S_{i+j-2}^{(1)} < h^{(1)}\right).$$

By using the mathematical induction method, finally we obtain

$$P\left(A_j\right) = \sum_{a=0}^{j} C_a^j p^a q^{j-a} \mathbf{1}_{h^{(1)}-aM>N} P\left(S_i^{(1)} < h^{(1)} - aM\right)$$

If we choose the threshold as $h^{(1)} = N + cM$, where c is an integer, then

$$P\left(A_j\right) = \sum_{a=0}^{\min\{j,c-1\}} C_a^j p^a q^{j-a} P\left(S_i^{(1)} < h^{(1)} - aM\right)$$

As the minimum value of the sum $S_i^{(1)}$ is N, we can bound this probability from above

$$P\left(A_j\right) \le \sum_{a=0}^{\min\{j,c-1\}} C_a^j p^a q^{j-a}.$$

This formula is used for $c < j$, because the probability should be small enough. So, we obtain

$$P\left(A_j\right) \le \sum_{a=0}^{c-1} C_a^j p^a q^{j-a}, \tag{28}$$

where $c = (h^{(1)} - N)/M$.

This formula gives a way to determine the value of the threshold. Let P_1 be the desired value of the probability of the skip of the change, then the value c can be calculated as the maximum integer satisfying the following conditions

$$\sum_{a=0}^{c-1} C_a^j p^a q^{j-a} \le P_1. \tag{29}$$

Note that the probability does not depend on the exact values of the input flow intensities, it includes only the probabilities p and q given by formulas (4). One can use instead of p its minimum desired value p^*, and, instead of q the value $q^* = 1 - p^*$. As a result, the value of the threshold for $p \geq p^*$ will be obtained.

Let $S_i^{(1)} = N$, and the process satisfy the hypothesis $H_l(t_{i-k-1}, t_{i+j})$. Let us consider the event $C_j = \bigcup_{a=1}^{j} S_{i+a}^{(1)} < h^{(1)}$ which means that no false alarms occur in j steps if the initial value of the sum is N. By using the same reasoning as for a skip of the change, we obtain

$$P\left(C_j\right) = \frac{1}{2^j} \sum_{a=0}^{c-1} C_a^j, \tag{30}$$

where $c = (h^{(1)} - N)/M$.

Table 4 gives some results of calculations. Here $P_1 = 0,1$, $m = 1$, $n = 6$ (so, $M = 5$, $N = 7$); λ_1 and λ_2 are the process parameters, j is the number of steps, p and q are calculated by formulas (4); c is the maximum value satisfying condition (29), $h = N + cM$, $\hat{P}_1$ is the upper bound for the probability of a skip of the change calculated by Eq. (28), $\hat{P}_0$ is the probability of a false alarm calculated by Eq. (30).

The error probabilities depend on the probability p: an increase in p leads to a decrease in the probability of a skip of the change: hence, the value of the threshold can be increased for the fixed value of P_1. In turn, this implies an decrease in the false alarm probability. Both error probabilities are sufficiently small for $p > 0.75$. On the other hand, the change point should be detected in k steps; so, we can decrease the false alarm probability by increasing k and, consequently, the threshold.

Table 5 demonstrates the results of calculations for the values of k taken from Table 3. The parameter h was chosen to provide the false alarm probability less then 0.15. Then the probability of a skip of the change was calculated for different values of λ_2 and, hence, p and q.

Table 4. Error probabilities

λ_1	λ_2	p	q	j	c	h	$\hat{P}_1$	$\hat{P}_0$	j	c	h	$\hat{P}_1$	$\hat{P}_0$
2	0.4	0.83	0.17	10	7	42	0.07	0.172	15	11	62	0.09	0.059
2	0.6	0.77	0.23	10	6	37	0.058	0.377	15	9	52	0.038	0.304
2	0.8	0.71	0.29	10	5	32	0.038	0.623	15	8	47	0.038	0.5
2	1	0.67	0.33	10	5	32	0.077	0.623	15	8	47	0.088	0.5
2	0.4	0.83	0.17	20	14	77	0.037	0.058	25	18	97	0.044	0.022
2	0.6	0.77	0.23	20	13	72	0.069	0.132	25	16	87	0.044	0.115
2	0.8	0.71	0.29	20	12	67	0.087	0.252	25	15	82	0.072	0.212
2	1	0.67	0.33	20	11	62	0.092	0.412	25	14	77	0.092	0.345

Table 5. Error probabilities

λ_1	k	h	$\hat{P}_0$	λ_2	p	q	$\hat{P}_1$	λ_2	p	q	$\hat{P}_1$
2	31	107	0.075	0.4	0.83	0.17	0.0028	0.6	0.77	0.23	0.0037
2	43	137	0.111	0.4	0.83	0.17	0.00008	0.6	0.77	0.23	0.0048
2	56	182	0.07	0.4	0.83	0.17	0.00001	0.6	0.77	0.23	0.002
2	70	212	0.094	0.4	0.83	0.17	0.0000002	0.6	0.77	0.23	0.00019
2	31	107	0.075	0.8	0.71	0.29	0.0016	1	0.67	0.33	0.323
2	43	137	0.111	0.8	0.71	0.29	0.0043	1	0.67	0.33	0.153
2	56	182	0.07	0.8	0.71	0.29	0.03	1	0.67	0.33	0.139
2	70	212	0.094	0.8	0.71	0.29	0.0074	1	0.67	0.33	0.061

Greater values of k provide adequate error probabilities in the most of cases.

Then we compare the results presented in Tables 1 and 4 for the same values of the threshold h and the intensities λ_i. Table 1 demonstrates the mean delay $T^{(1)}$ in the change-point detection, and Table 4 contains the number of steps j used to detect a change point with the probability 0.9 and greater.

For $p = 0.83$, the values $T^{(1)}$ and j differ insignificantly (not more than by 25 %), but the decrease in p leads to the increase of that difference, and for $p = 0.67$ the value $T^{(1)}$ exceeds the value of j twice or more. So, in spite of the big values of the mean delay for small values of p, the change in this case will be detected most probably much earlier than in the mean delay time.

Acknowledgements. Y. Burkatovskaya is supported by The National Research Tomsk State University Academic D.I. Mendeleev Fund Program (NU 8.1.55.2015 L) in 2014–2015 and by RFBR Grant 16-01-00121.

The authors are grateful to Prof. Sergey Vorobeychikov from Tomsk State University for useful comments.

References

1. Fischer, W., Meier-Hellstern, K.: The Markov-modulated Poisson process (MMPP) cookbook. Perform. Eval. **18**(2), 149–171 (1993)
2. Zou, M., Liu, J.: Performance analysis of IEEE 802.16 networks with MMPP arrivals. Perform. Eval. **69**, 492–509 (2012)
3. Nogueira, A., Salvador, P., Valadas, R., Pacheco, A.: Modeling self-similar traffic over multiple time scales based on hierarchical Markovian and L-System models. Comput. Commun. **33**, S3–S10 (2010)
4. Giacomazzi, P.: Closed-form analysis of end-to-end network delay with Markov-modulated Poisson and fluid traffic. Comput. Commun. **32**(4), 640–648 (2009)
5. Choi, D.I., Kim, T.-S., Lee, S.: Analysis of an MMPP/G/1/K queue with queue length dependent arrival rates, and its application to preventive congestion control in telecommunication networks. Eur. J. Oper. Res. **187**(2), 652–659 (2008)

6. Vasil'eva, L.A., Gortsev, A.M.: Estimation of the dead time of an asynchronous double stocastic flow of events under incomplete observability. Autom. Remote Control **64**(12), 1890–1898 (2003)
7. Gortsev, A.M., Nezhelskaya, L.A.: An asynchronous double stochastic flow with initiation of superfluous events. Discrete Math. Appl. **21**(3), 283–290 (2011)
8. Asmussen, S.: Phase-type distributions and related point processes: fitting and recent advances. In: Chakravarthy, S., Alfa, A.S. (eds.) Matrix-Analytic Methods in Stochastic Models. Lecture Notes in Pure and Applied Mathematics, vol. 183, pp. 137–149. Marcel Dekker, New York (1997)
9. Gerhardt, I., Nelson, B.L.: On capturing dependence in point processes: matching moments and other techniques. Technical report, Northwestern University (2009)
10. Burkatovskaya, Y., Kabanova, T., Vorobeychikov, S.: CUSUM algorithms for parameter estimation in queueing systems with jump intensity of the arrival process. In: Dudin, A., Nazarov, A., Yakupov, R. (eds.) ITMM 2015. CCIS, vol. 564, pp. 275–288. Springer, Heidelberg (2015). doi:10.1007/978-3-319-25861-4_24
11. Vorobeychikov, S.E.: On the detection of a change in the mean of a sequence of random variables. Autom. Remote Control **59**(3), 50–56 (1998)
12. Vorobeychikov, S.E., Kabanova, T.V.: Detection of the change point in a sequence of independent random variables (in Russian). J. Commun. Technol. Electron. **47**(10), 1198–1203 (2002)
13. Wald, A.: Sequential Analysis. Wiley, New York (1947)
14. Lorden, G.: Procedures for reacting to a change in distribution. Annals. Math. Statist. **42**, 1897–1971 (1971)

Information Technologies and Mathematical Modelling: Queueing Theory and Applications
15th International Scientific Conference, ITMM 2016, named after A.F. Terpugov, Katun, Russia, September 12-16, 2016. Proceedings
Dudin, A.; Gortsev, A.; Nazarov, A.; Yakupov, R. (Eds.)
2016, XI, 397 p. 102 illus., Softcover
ISBN: 978-3-319-44614-1