CUSUM Algorithms for Parameter Estimation in Queueing Systems with Jump Intensity of the Arrival Process

Yulia Burkatovskaya[1(✉)], Tatiana Kabanova[2], and Sergey Vorobeychikov[3]

[1] Institute of Cybernetics, Tomsk Polytechnic University, 30 Lenin Prospekt, 634050 Tomsk, Russia
tracey@tpu.ru
http://www.tpu.ru

[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
sev@mail.tsu.ru
http://www.tsu.ru

Abstract. The problem of Markov-modulated Poisson process intensities estimating is studied. A new approach based on sequential change point detection method is proposed to determine switching points of the flow parameter. Both the intensities of the controlling Markovian chain and the intensities of the flow of events are estimated. The results of simulation are presented.

Keywords: Markov-modulated poisson process · Jump intensity · CUSUM algorithm

1 Introduction

Markovian arrival processes form a powerful class of stochastic processes introduced in [1,2] and thereafter they are widely used now as models for input flows to queueing systems where the rate of the arrival of customers depends on some external factors. MAP is a counting process whose arrival rate is governed by a continuous-time Markov chain.

Queueing systems with jump intensity of customer arrivals is one of the examples of applying MAP. In such models the intensity is supposed to be piecewise constant function depended on the state of random environment. Particulary, this model can be used as a model of a call-center or http-server customers (see [3,4]), healthcare systems (see [5]), etc. Usually the stationary probabilities of system states, sojourn and waiting time distributions, mean length of the queue and other parameters are investigated. To solve such problems there is a need to estimate parameters of customer arrivals.

A. Dudin et al. (Eds.): ITMM 2015, CCIS 564, pp. 275–288, 2015.
DOI: 10.1007/978-3-319-25861-4_24

The typical property of observing time series derived from a MAP is that only the arrivals but not the states of the controlling Markovian chain can be seen. The problem is to estimate both the controlling Markovian chain parameters and parameters of the intensity of the arrival process. A survey of estimation methods is given in [6]. Its emphasis is on maximum likelihood estimation and its implementation via the EM (expectation-maximization) algorithm. The EM iteration alternates between performing an expectation (E) step, which creates a function for the expectation of the log-likelihood evaluated using the current estimator for the parameters, and a maximization (M) step, which computes parameters maximizing the expected log-likelihood found on the E step. These parameter estimators are then used to determine the distribution of the latent variables in the next E step. This approach is developed for different conditions in [7,8], etc. The survey [9] with a huge bibliography is focused on matching moment method which is also widely used for parameter estimation in MAP because of its simplicity. This method is used, for example, in [10]. Bayesian approach based on the a posteriori probability of the controlling chain state is developed in [11]. It provides estimators with the minimum mean square error.

In this paper we propose a different approach. We use the sequential analysis methods for parameter estimation in queueing system with jump intensity of the arrival process. The key idea is to consider time intervals between arrivals as a stochastic process which parameters change in random points. First we are going to detect these points using sequential change point detection methods. Then we are going to estimate the intensity parameters under the assumption that the intensity is constant between detected 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 point. The problem is to detect this change point when the process is observed online. Sequential methods include a special stopping rule that determines a stopping time. At this instant a decision on change point can be made. There are two types of errors typical for sequential change point detection procedures: false alarm, when one makes a decision that change is occurred before a change point (type 1 error), and delay, when the change is not detected (type 2 error).

The CUSUM (or cumulative sum control chart) algorithm was proposed by E.S. Page in [12] and since then it is widely used for online detecting changes in parameters for different time series both with independent and with dependent observations, even for autoregressive type processes. Usually the change in the mean is considered. As far as a change of the state of the controlling Markovian chain causes a jump of the mean length of an interval between arrivals hence the lengths of intervals form a sequence of dependent random variables and it is possible to apply the CUSUM algorithm to this situation. G. Lorden in [14] established that the CUSUM procedure is optimal in a sense that it provides minimum mean time of delay in change detecting when mean time between false alarms is fixed. In this paper we use the CUSUM procedure to determine intervals of the constant intensity of the observed flow of events. After that parameter estimators are constructed.

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, 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, where $i = 1, 2$.

The flow of events has the exponential distribution with the intensity parameter λ_1 or λ_2 subject to the state of the Marcovian chain. The parameters of the system λ_1, λ_2 and the instants of switching between the states are supposed to be unknown. We also suppose that $\lambda_i \ll \alpha_i$, i.e., changes of the controlling chain states occur more rarely than observed events. Thus some events occur between switchings of the controlling chain states. This situation is typical for real processes such as call-center or http-server because one of the states can be interpreted as a "usual" state of the system and another state as a "peak-time" state and during each of these states several customers are supposed to arrive. Besides processes having this property are often used for simulation study of algorithms for processes with jump intensity of customer arrivals (for example, see [8,13]).

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 Algorithm 1

Let the process $\{t_i\}_{i\geq 0}$ be the sequence of the instants when events of the observed flow occur. 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 as it is shown at the diagram (Fig. 1).

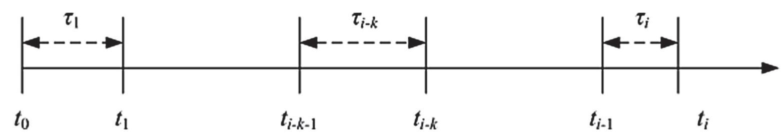

Fig. 1. Construction of the sequence $\{\tau_i\}$.

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 CUSUM procedures.

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 $k > 1$ describing the memory depth. The idea is to compare the values τ_i and τ_{i-k}. If there are no changes of the controlling chain state within the interval $[t_{i-k-1}, t_i]$ then the values 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 $[t_{i-k}, t_{i-1}]$ then the expectations of the values τ_i and τ_{i-k} are different.

On one hand the parameter k should allow us to detect changes with minimal delay, on the other hand it should not be too large to contain more than one chain state change within the interval $[t_{i-k}, t_{i-1}]$. Further we consider the choice of the parameter k in 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 increase in the mean of the process and hence, decrease of the intensity, and the second procedure is set up to detect decrease in the mean and hence, increase of the intensity. For the first procedure we introduce the sequence of the statistics

$$z_i^{(1)} = \tau_i - \tau_{i-k} - \Delta, \quad i > k. \tag{2}$$

For the second procedure we introduce the sequence of the statistics

$$z_i^{(2)} = \tau_{i-k} - \tau_i - \Delta, \quad i > k. \tag{3}$$

This statistics are calculated at the instant t_i.

Consider then four hypothesis 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;
- $H_{2,1}(t_{i-k}, t_{i-1})$ – the intensity of the arrival process on the interval $[t_{i-k}, t_{i-1}]$ changed once from λ_2 to λ_1.

Theorem 1. *If the parameter Δ satisfies condition (1) then the statistics $z_i^{(j)}$, $j \in \{1, 2\}$ (2), (3) have the following properties:*

$$\begin{gathered} E\left[z_i^{(1)} \middle| H_l(t_{i-k-1}, t_i) \right] < 0, \quad l = 1, 2; \\ E\left[z_i^{(1)} \middle| H_{1,2}(t_{i-k}, t_{i-1}) \right] > 0; \\ E\left[z_i^{(2)} \middle| H_l(t_{i-k-1}, t_i) \right] < 0, \quad l = 1, 2; \\ E\left[z_i^{(2)} \middle| H_{2,1}(t_{i-k}, t_{i-1}) \right] > 0. \end{gathered} \tag{4}$$

Proof. Using (1) one obtains

$$
\begin{aligned}
E\left[z_i^{(1)}\middle| H_l(t_{i-k-1},t_i)\right] &= E\left[\tau_i - \tau_{i-k} - \Delta\right| H_l(t_{i-k-1},t_i)] \\
&= \frac{1}{\lambda_l} - \frac{1}{\lambda_l} - \Delta < 0; \\
E\left[z_i^{(1)}\middle| H_{1,2}(t_{i-k},t_{i-1})\right] &= E\left[\tau_i - \tau_{i-k} - \Delta\right| H_{1,2}(t_{i-k},t_{i-1})] \\
&= \frac{1}{\lambda_2} - \frac{1}{\lambda_1} - \Delta > 0; \\
E\left[z_i^{(2)}\middle| H_l(t_{i-k-1},t_i)\right] &= E\left[\tau_{i-k} - \tau_i - \Delta\right| H_l(t_{i-k-1},t_i)] \\
&= \frac{1}{\lambda_l} - \frac{1}{\lambda_l} - \Delta < 0; \\
E\left[z_i^{(2)}\middle| H_{2,1}(t_{i-k},t_{i-1})\right] &= E\left[\tau_{i-k} - \tau_i - \Delta\right| H_{2,1}(t_{i-k},t_{i-1})] \\
&= \frac{1}{\lambda_2} - \frac{1}{\lambda_1} - \Delta > 0.
\end{aligned}
$$

So the means of statistics (2), (3) change from negative value to positive when the intensity of the process changes. 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

$$
\begin{gathered}
S_0^{(1)} = \Delta; \\
S_i^{(1)} = \max\{0, S_{i-1}^{(1)} + z_i^{(1)}\}, \quad i > k; \\
S_i^{(1)} = 0, \quad \text{if} \quad S_i^{(1)} \geq h_1.
\end{gathered} \tag{5}
$$

For the second procedure the cumulative sum is defined as follows

$$
\begin{gathered}
S_0^{(2)} = \Delta; \\
S_i^{(2)} = \max\{0, S_{i-1}^{(2)} + z_i^{(2)}\}, \quad i > k; \\
S_i^{(2)} = 0, \quad \text{if} \quad S_i^{(2)} \geq h_2.
\end{gathered} \tag{6}
$$

If the cumulative sum $S_i^{(1)}$ reaches the threshold h_1 then the decision is made 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 the decision is made 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 a sum reaches threshold it is reset to zero and the corresponding procedure is restarted.

Let the sequence $\left\{\sigma_m^{(l)}\right\}_{m\geq 0}$ be the sequence of the instants when the cumulative sum in the l-th procedure reaches the threshold h_l, i.e.

$$
\begin{gathered}
\sigma_0^{(l)} = 0; \\
\sigma_m^{(l)} = \min\left\{t_j > \sigma_{m-1}^{(l)} : S_j^{(l)} \geq h_l\right\}.
\end{gathered} \tag{7}
$$

Consider a sequence $\left\{n_i^{(l)}\right\}_{i\geq 0}$ associated with the sequence $\left\{\sigma_m^{(l)}\right\}_{m\geq 0}$ as follows

$$n_0^{(l)} = 0;$$
$$n_m^{(l)} = \max\left\{t_j \leq \sigma_m^{(l)} : S_j^{(l)} > 0, S_{j-1}^{(l)} = 0\right\}. \tag{8}$$

Thus the instant $n_m^{(l)}$ is the first instant when the cumulative sum becomes positive to reach then the threshold. The construction of the sequences are illustrated at Fig. 2. The instants of occurrences t_i are marked by vertical dotted lines. At the diagram above an example of the sum $S_j^{(1)}$ behavior is presented and the instants $\sigma_m^{(1)}$ and $n_m^{(1)}$ are marked out. At the diagram in the middle a similar example for the sum $S_j^{(2)}$ is shown.

We consider the instants $n_i^{(1)}$ as the estimators for the instants when the mean length between the events increases. They are pointed by up arrows at the diagram below. In turn the instants $n_i^{(2)}$ are considered as the estimators for the instants when the mean length between the events increases. They are pointed by down arrows at the diagram below.

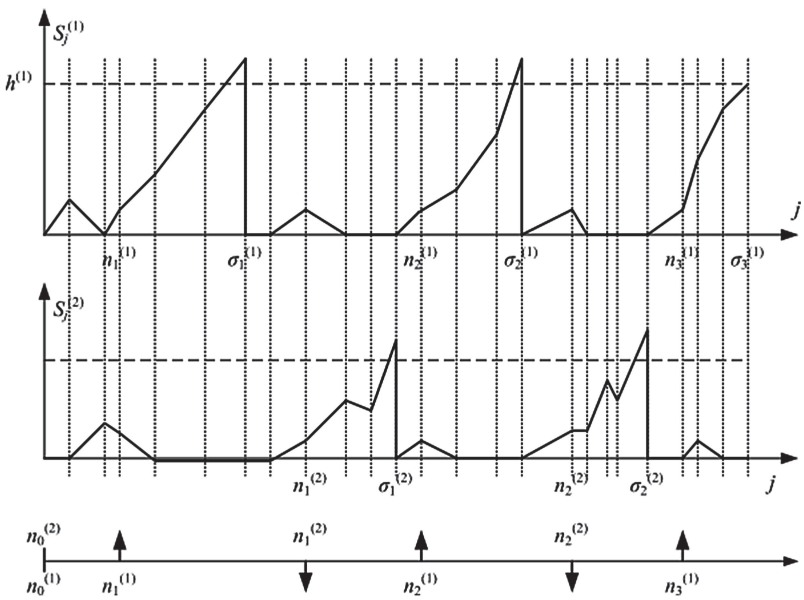

Fig. 2. Construction of the sequences $\left\{\sigma_m^{(l)}\right\}$, $\left\{n_m^{(l)}\right\}$.

In connection with sequential change point detection procedures two type 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. A skip of the change occurs when the change of the parameter occurs but the corresponding cumulative sum does not reach its threshold.

When implementing the procedure it is possible to encounter false alarm situations. We shall record all the exceeding the thresholds by either first or the second cumulative sum. If the same sum reaches threshold several times in a row, we only record the first occurrence.

Thus the procedure for estimation of instants of intensity switching is described as follows. Calculate two cumulative sums given by Eqs. (5), (6). Then construct the sequences $\left\{\sigma_m^{(l)}\right\}$, $\left\{n_m^{(l)}\right\}$ defined by Eqs. (7), (8). Let $n_1^{(1)} < n_1^{(2)}$, then the initial value of the intensity is equal to λ_1. Define the sequence

$$\begin{aligned} q_0 &= 0; \\ q_{2l+1} &= \min\left\{n_i^{(1)} : n_i^{(1)} > q_{2l}\right\}, \quad l \geq 0; \\ q_{2l+2} &= \min\left\{n_i^{(2)} : n_i^{(2)} > q_{2l+1}\right\}, \quad l \geq 0. \end{aligned} \tag{9}$$

The values $q_1,\ q_2, \ldots$ are calculated using formula (9) while the set

$$\left\{n_i^{(2)} : n_i^{(1)} > q_{2l}\right\} \neq \emptyset;$$
$$\left\{n_i^{(1)} : n_i^{(2)} > q_{2l+1}\right\} \neq \emptyset.$$

If
$$\left\{n_i^{(2)} : n_i^{(1)} > q_{2l}\right\} = \emptyset \quad \left(\left\{n_i^{(1)} : n_i^{(2)} > q_{2l+1}\right\} = \emptyset\right)$$
then we set $q_{2l+1} = N$ $(q_{2l+2} = N)$, where N is the instant of the last occurrence. Here the odd instants q_{2l+1} are the estimators of the instants when the intensity changes from λ_1 to λ_2, and the even instants q_{2l+2} are the estimators of the instants when the intensity changes from λ_2 to λ_1.

An example of the sequence construction is illustrated at Fig. 3. The sequences n_i^l are shown at the diagram above. The instants of switching the controlling chain state from 1 to 2 are pointed by up arrows, the instants of switching the controlling chain state from 2 to 1 are pointed by down arrows at the diagram below. The intervals are marked by the numbers of the states of the controlling chain.

Define estimators for parameters λ_1, λ_2

$$\hat{\lambda}_1 = \frac{N_1}{T_1}, \quad \hat{\lambda}_2 = \frac{N_2}{T_2}, \tag{10}$$

where N_1 is the total number of events occurred at the intervals $[q_{2l}, q_{2l+1}]$, $q_{2l+1} \leq N$ and T_1 is the total length of these intervals; N_2 is the total number of events occurred at the intervals $[q_{2l+1}, q_{2l+2}]$, $q_{2l+2} \leq N$ and T_2 is the total length of these intervals; $l \geq 0$ (Fig. 3).

Define estimators for parameters α_1, α_2

$$\hat{\alpha}_1 = \frac{L_1}{T_1}, \quad \hat{\alpha}_2 = \frac{L_2}{T_2}, \tag{11}$$

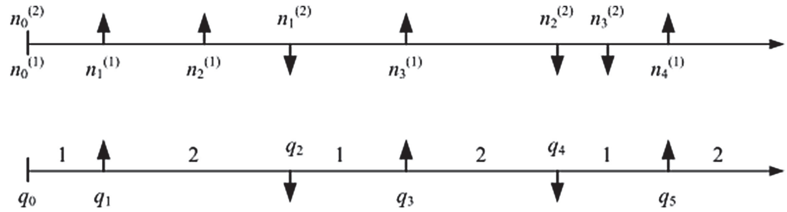

Fig. 3. Construction of the sequences $\{q_m\}$.

where L_1 is the total number of the switching points $q_{2l+1} \leq N$, L_2 is the total number of the switching points $q_{2l+2} \leq N$, $l \geq 0$.

If $n_1^{(2)} < n_1^{(1)}$, then the initial value of the intensity is equal to λ_2 the procedure is similar. Define the sequence

$$\begin{aligned} q_0 &= 0; \\ q_{2l+1} &= \min\left\{n_i^{(2)} : n_i^{(1)} > q_{2l}\right\}, \quad l \geq 0; \\ q_{2l+2} &= \min\left\{n_i^{(1)} : n_i^{(2)} > q_{2l+1}\right\}, \quad l \geq 0. \end{aligned} \tag{12}$$

Here the odd instants q_{2l+1} are the estimators of the instants when the intensity changes from λ_2 to λ_1, and the even instants q_{2l+2} are the estimators of the instants when the intensity changes from λ_1 to λ_2. Estimators for the parameters λ_1, λ_2 are calculated using formula (10), where N_1 is the total number of events occurred at the intervals $[q_{2l+1}, q_{2l+2}]$ and T_1 is the total length of these intervals; N_2 is the total number of events occurred at the intervals $[q_{2l}, q_{2l+1}]$ and T_2 is the total length of these intervals; $l \geq 0$. Estimators for the parameters α_1, α_2 are calculated using formula (11), where L_1 is the total number of the switching points $q_{2l+2} < N$, L_2 is the total number of the switching points $q_{2l+1} < N$, $l \geq 0$.

4 Choice of the Algorithm Parameters

In this section the problem of choice of the parameters k, Δ and h_l is discussed.

We suppose that changes of the controlling chain states occur more rarely than observed events. First, we consider the memory depth parameter k. Let n be a lower bound of the mean number of events between switchings of the controlling chain states. For the model under consideration it means that $n\alpha_i \leq \lambda_i$. It means that it is not effective to choose the memory depth $k \geq n$ or close to n because in this case there can be many situations when more than one chain state change occur within the interval $[t_{i-k}, t_{i-1}]$. On the other hand, the sum $S_i^{(l)}$ should reach the corresponding threshold h_i after switching of the controlling chain state, i.e. some statistics $z_i^{(l)}$ should be positive. It follows from these considerations and numerical calculations that a good choice of the parameter k is

$$k \approx \frac{n}{2}. \tag{13}$$

Then, turn to the parameters Δ and h_l. Condition (1) provides properties (4). The properties make it possible to construct CUSUM procedures. Thus the parameter Δ can be chosen from the interval $(0, 1/\lambda_2 - 1/\lambda_1)$, i.e., it is positive and does not exceed the difference between the mean lengths of the intervals τ_i when the controlling chain is in different states. Let this difference be not less than some $d > 0$:

$$\frac{1}{\lambda_2} - \frac{1}{\lambda_1} \geq d. \tag{14}$$

The parameter Δ and h_l affects the characteristics of the CUSUM procedure, i.e., the mean delay and the mean time between false alarms (see [14]). If the parameter h_l is fixed then increase of the parameter Δ results in decrease of the mean of the statistic $z_i^{(l)}$ and hence the sum $S_i(l)$ reaches the threshold h_l more slowly and hence, a switching of the controlling chain state from the state l can be skipped. Consequently, the number of false detection of the controlling chain state switchings decreases but on the other hand the number of skips of the controlling chain state switchings increases. If the parameter Δ is fixed then increase of the parameter h_l results in the same effect. Vice versa, decrease of the parameter Δ or the parameter h_l while the other parameter is fixed result in increase of the number of false detection of the controlling chain state switchings and decrease of the number of skips of the controlling chain state switchings.

If there are no additional conditions then the procedure is considered to be optimal when the probabilities of the false detection and the skip of the change are equal. It can be guaranteed by the following conditions

$$\begin{aligned} E\left[z_i^{(1)}\middle| H_l(t_{i-k-1}, t_i)\right] &= -E\left[z_i^{(1)}\middle| H_{1,2}(t_{i-k}, t_{i-1})\right]; \\ E\left[z_i^{(2)}\middle| H_l(t_{i-k-1}, t_i)\right] &= -E\left[z_i^{(2)}\middle| H_{2,1}(t_{i-k}, t_{i-1})\right]. \end{aligned} \tag{15}$$

It results in the equations (see Theorem 1)

$$-\Delta = -\frac{1}{\lambda_2} + \frac{1}{\lambda_1} + \Delta.$$

Hence the best choice of the parameter Δ is

$$\Delta = \frac{d}{2} \tag{16}$$

where d is defined by the Eq. (14), i.e., Δ is the half of the difference between the mean lengths of the intervals τ_i when the controlling chain is in different states. If the difference is unknown then one has to choose as d a lower bound of the difference. In other words, one has to define the minimal difference that should be detected by the algorithm.

Consider now the parameter h_l. If the memory depth is equal to k then the sum $S_i^{(l)}$ to reach the threshold h_l in not more then k steps (while $Ez_i^{(l)} > 0$).

If the parameter $\varDelta$ satisfies the condition (16) then using (14) and Theorem 1 one obtains

$$E\left[z_i^{(1)}\middle| H_{1,2}(t_{i-k},t_{i-1})\right] = \frac{1}{\lambda_2} - \frac{1}{\lambda_1} - \varDelta \geq d - \frac{d}{2} = \frac{d}{2};$$
$$E\left[z_i^{(2)}\middle| H_{2,1}(t_{i-k},t_{i-1})\right] = \frac{1}{\lambda_2} - \frac{1}{\lambda_1} - \varDelta \geq d - \frac{d}{2} = \frac{d}{2}.$$

So if the sum $S_i^{(l)}$ starts from zero it reaches the threshold h_l on the average in $2h_l/d$ steps. Hence it is supposed to choose the threshold h_l from the condition $2h_l/d < k$, i.e.

$$h_l < \frac{kd}{2} \approx \frac{nd}{4}. \tag{17}$$

Note that the parameter h_l should not be significantly than its upper bound because it can increase the number of false alarms.

In general the choice of the CUSUM parameters is a rather difficult problem requiring further theoretical investigations. Nevertheless, numerical simulations demonstrated a good quality of the proposed algorithm with the parameters (13), (16), (17).

5 Algorithm 2

The second algorithm is very similar to the first except of the definition of the statistics $z_i^{(l)}$.

Let we have a certain period of observation $[0,T]$ and N is the number of occurrences at the interval. First, we calculate the mean of the length between occurrences using the usual formula

$$\hat{\tau} = \frac{T}{N}. \tag{18}$$

The value $\hat{\tau}$ exceeds the mean length of the interval τ_i when the controlling chain is in the first state, and vice versa, the mean length of the interval τ_i exceeds the value $\hat{\tau}$ when the controlling chain is in the second state, i.e.

$$\frac{1}{\lambda_1} < E\hat{\tau} < \frac{1}{\lambda_2}. \tag{19}$$

Using this property we can construct statistics as follows. For the first procedure we introduce the sequence of the statistics

$$z_i^{(1)} = \tau_i - \hat{\tau}. \tag{20}$$

For the second procedure we introduce the sequence of the statistics

$$z_i^{(2)} = -\tau_i + \hat{\tau}. \tag{21}$$

Consider then two hypothesis concerning the state of the controlling chain:

– $H_1(t_{i-1}, t_i)$ – the intensity of the arrival process on the interval $[t_{i-1}, t_i]$ is constant and equal to λ_1;
– $H_2(t_{i-1}, t_i)$ – the intensity of the arrival process on the interval $[t_{i-1}, t_i]$ is constant and equal to λ_2;

Theorem 2. *The statistics $z_i^{(j)}$, $j \in \{1, 2\}$ (20), (21) have the following properties:*

$$\begin{aligned} E\left[z_i^{(1)} \middle| H_1(t_{i-1}, t_i) \right] &< 0; \\ E\left[z_i^{(1)} \middle| H_2(t_{i-1}, t_i) \right] &> 0; \\ E\left[z_i^{(2)} \middle| H_1(t_{i-1}, t_i) \right] &> 0; \\ E\left[z_i^{(2)} \middle| H_2(t_{i-1}, t_i) \right] &< 0. \end{aligned} \tag{22}$$

Proof. Using (19) one obtains

$$\begin{aligned} E\left[z_i^{(1)} \middle| H_1(t_{i-1}, t_i) \right] &= E\left[\tau_i - \hat{\tau} \middle| H_1(t_{i-1}, t_i) \right] = \frac{1}{\lambda_1} - E\hat{\tau} < 0; \\ E\left[z_i^{(1)} \middle| H_2(t_{i-1}, t_i) \right] &= E\left[\tau_i - \hat{\tau} \middle| H_2(t_{i-1}, t_i) \right] = \frac{1}{\lambda_2} - E\hat{\tau} > 0; \\ E\left[z_i^{(2)} \middle| H_1(t_{i-1}, t_i) \right] &= E\left[-\tau_i + \hat{\tau} \middle| H_l(t_{i-1}, t_i) \right] = -\frac{1}{\lambda_1} + E\hat{\tau} > 0; \\ E\left[z_i^{(2)} \middle| H_2(t_{i-1}, t_i) \right] &= E\left[-\tau_i + \hat{\tau} \middle| H_2(t_{i-1}, t_i) \right] = -\frac{1}{\lambda_2} + E\hat{\tau} < 0. \end{aligned}$$

So one can see that the statistics $Z_i^{(l)}$ change their means when the intensity of the arrival process changes. Using in Algorithm 1 statistics (20), (21) instead of (2), (3) we obtain Algorithm 2.

Consider now the choice of the parameters h_l. If n is a lower bound of the mean number of events between switchings of the controlling chain states then the sum $S_i^{(l)}$ should reach the threshold h_l on the average less then in n steps, for example, in $n/2$ steps. For Algorithm 2 we can not estimate the mean of the statistic $E\left[z_i^{(1)} \middle| H_2(t_{i-1}, t_i) \right]$ if the parameters α_i are unknown because we can not calculate $E\hat{\tau}$. Hence, we use a rather crude estimator

$$\begin{aligned} E\left[z_i^{(1)} \middle| H_2(t_{i-1}, t_i) \right] &\approx \frac{d}{2}; \\ E\left[z_i^{(2)} \middle| H_1(t_{i-1}, t_i) \right] &\approx \frac{d}{2}. \end{aligned}$$

So we come to the inequality

$$h_l < \frac{nd}{4} \tag{23}$$

which is the same as in Algorithm 1.

6 Numerical Simulation

The model for the considered flow and the suggested algorithms was implemented with varying parameters. The results are presented in the tables below (Tables 1 and 2).

Table 1. The results of the simulation for Algorithm 1.

T	λ_1	λ_2	α_1	α_2	h_1	h_2	k	Δ	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{\alpha}_1$	$\hat{\alpha}_2$
1000	5	1	0,4	0,2	1	1	5	0,2	4,2750	1,0951	0,2458	0,1641
1000	5	1	0,4	0,2	1,8	1,8	5	0,2	3,8424	1,0621	0,1687	0,1226
1000	5	1	0,4	0,2	1,8	1,8	8	0,2	4,3561	1,1852	0,1686	0,1180

Here we use the following notations:

- T is the time of simulation;
- λ_1 and λ_2 are the intensities of the arrival process in the first and the second state, correspondingly;
- α_1 and α_2 are the switching intensities from the first to the second state and vise versa, correspondingly;
- h_1 and h_2 are the CUSUM thresholds;
- k is the parameter of the algorithm, the difference between the numbers of the compared intervals at the statistics (2), (3);
- Δ is the parameter of the algorithm;
- $\hat{\lambda}_1$ and $\hat{\lambda}_2$ are the estimators of the parameters λ_1 and λ_2;
- $\hat{\alpha}_1$ and $\hat{\alpha}_2$ are the estimators of the parameters α_1 and α_2.

Table 2. The results of the simulation for Algorithm 2.

T	λ_1	λ_2	α_1	α_2	h_1	h_2	$\hat{\tau}$	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{\alpha}_1$	$\hat{\alpha}_2$
1000	5	1	0,3	0,2	0,5	0,5	0,3883	5,2336	1,2276	0,322	0,1732
1000	5	1	0,3	0,2	0,8	0,8	0,3929	5,0591	1,2370	0,2573	0,1322
1000	5	1	0,3	0,2	1	1	0,4355	4,7475	1,1972	0,2587	0,1144
1000	5	1	0,1	0,2	0,5	0,5	0,2668	5,6804	2,0053	0,2912	0,2604
1000	5	1	0,1	0,2	0,8	0,8	0,2924	5,1544	1,6825	0,1180	0,1880
1000	5	1	0,1	0,2	1	1	0,2501	5,2498	2,6283	0,1207	0,1297
1000	5	2	0,1	0,2	0,5	0,5	0,2351	6,1085	2,8854	0,3632	0,2656
1000	5	2	0,1	0,2	0,8	0,8	0,2564	5,1785	2,8092	0,1652	0,1289
1000	5	2	0,1	0,2	1	1	0,2486	5,1949	2,8831	0,1219	0,1162
10000	5	1	0,3	0,2	0,8	0,8	0,3830	4,8379	1,3439	0,2316	0,1318
10000	5	1	0,3	0,2	1	1	0,3766	4,6783	1,4326	0,1917	0,1157

Here we use the same notations as above, $\hat{\tau}$ is the mean length of the interval between occurrences calculated by (18).

First, the quality of the proposed algorithms on the threshold parameters h_i was studied. Increasing of h_i leads to decreasing of probability for the cumulative sums to reach the thresholds and hence an intensity change can be undetected. It causes increasing of error of the estimators $\hat{\lambda}_l$ because of not correct estimation of the controlling chain current state.

On the other hand, increasing of h_i results in decreasing the total number of false alarms. These theoretical conclusions are supported by the simulation results. As the thresholds increase the estimators of the switching parameters $\hat{\alpha}_l$ decrease because less switching points are detected on the first stage of the procedures. In Table 2 for $h_1 = h_2 = 1$ one can see that the estimators $\hat{\alpha}_i$ considerably less the real values of the parameters α_i. The best results are obtained for $h_1 = h_2 = 0,8$ for all intensity parameter values. It supports our considerations concerning the parameter h_l. According to (23) for $\lambda_1 = 5$ and $\lambda_2 = 1$ minimal difference between the mean length of the intervals τ_i is $d = 1/1 - 1/5 = 0.8$ and the recommended choice of h_l is $h_l < (0.8 \times 5)/4 = 1$, but it should not be significantly less.

Increase of the simulation time from 1000 to 10000 does not influence significantly the estimators quality. This result stresses the fact that the proposed algorithms can be used for small sample size.

7 Conclusion

Markovian arrival processes serve as models for real processes, particularly, for call-centers or http-server customers, healthcare systems, etc. Input flow intensity estimation and pertinent model setup is necessary to develop dispatching rule, to calculate optimal number of servers, etc. The suggested algorithms do not use the distribution function of the observing flow and, hence, can be applied to parameter estimation of other types of flows.

Acknowledgements. This paper is supported by The National Research Tomsk State University Academic D.I. Mendeleev Fund Program (NU 8.1.55.2015 L) in 2014–2015.

References

1. Neuts, M.F.: A versatile Markovian point process. J. Appl. Probab. **16**, 764–774 (1979)
2. Lucantoni, D.M., Meier-Hellstern, K.S., Neuts, M.F.: A single server queue with server vacations and a class of non-renewal arrival processes. Adv. Appl. Probab. **22**, 676–705 (1990)
3. Dudina, O., Dudin, S.: Queueing system $MAP/M/N/N + K$ operating in random environment as a model of call center. In: Dudin, A., Klimenok, V., Tsarenkov, G., Dudin, S. (eds.) BWWQT 2013. CCIS, vol. 356, pp. 83–92. Springer, Heidelberg (2013)

4. Dudin, S.: The servicing system MAP(PH) plus MAP/PH/N/R as a model of optimizing an HTTP server with blockings. Autom. Remote Control **1**, 28–38 (2010)
5. Fackrell, M.: Health Care Manage. Sci. Modelling healthcare systems with phase-type distributions **12**, 11–26 (2008). Springer Science + Business Media, LLC
6. Asmussen, S.: Phase-type distributions and related point processes: fitting and recent advances. In: Chakravarthy, S.R., Alfa, A.S. (eds.) Matrix-analytic Methods in Stochastic Models. Lecture Notes in Pure and Applied Mathematics, pp. 137–149. Marcel Dekker, New York (1997)
7. Breuer, L., Kume, A.: An EM algorithm for Markovian arrival processesobserved at discrete times. In: Müller-Clostermann, B., Echtle, K., Rathgeb, E.R. (eds.) MMB&DFT 2010. LNCS, vol. 5987, pp. 242–258. Springer, Berlin (2010)
8. Okamura, H., Dohi, T., Trivedi, K.S.: Markovian arrival process parameter estimation with group data. IEEE/ACM Trans. Netw. **17**(4), 1326–1340 (2009)
9. Gerhardt, I., Nelson, B.L.: On capturing dependence in point processes: matching moments and other techniques. Technical report, Northwestern university (2009)
10. Duffie, D., Glynn, P.: Estimation of continuous-time markov processes sampled at random time intervals. Econometrica **72**, 1773–1808 (2004)
11. Gortsev, A.M., Zuevich, V.L.: Optimal estimation of parameters of an asynchronous doubly stochastic flow of events with arbitrary number of the states. Tomsk State Univ. J. Control Comput. Sci. **4**(17), 25–40 (2011)
12. Page, E.S.: Continuous inspection schemes. Biometrica **42**(1), 100–115 (1956)
13. Gortsev, A.M., Kalyagin, A.A., Nezhelskaya, L.A.: Optimum estimation of states in generalized semi-synchronous flow of events. Tomsk State Univ. J. Control Comput. Sci. **2**(11), 66–81 (2010)
14. Lorden, G.: Procedures for reacting to a change in distribution. Ann. Math. Stat. **42**, 1897–1908 (1971)