

Choice of the Parameters of the CUSUM Algorithms for Parameter Estimation in the Markov Modulated Poisson Process

Yulia Burkatovskaya

Tomsk Polytechnic University,
Institute of Cybernetics
Tomsk State University,

Faculty of Applied Mathematics and Cybernetics and
International Laboratory of Statistics of Stochastic
Processes and Quantitative Finance
Tomsk, Russia
tracey@tpu.ru

Tatiana Kabanova

Tomsk State University,
Faculty of Applied Mathematics and Cybernetics
Tomsk, Russia
tvk@bk.ru

Pavel Khaustov

Tomsk Polytechnic University,
Institute of Cybernetics
Tomsk, Russia
khaustovpavel@gmail.com

Abstract—CUSUM algorithm for controlling chain state switching in the Markov modulated Poisson process was investigated via simulation. Recommendations concerning the parameter choice were given subject to characteristics of the process. Procedure of the process parameter estimation was described.

Keywords—Markov modulated Poisson process, parameter estimation, sequential change point detection, CUSUM algorithm.

I. INTRODUCTION

Markovian arrival processes (MAP) form a powerful class of stochastic processes introduced in [1] and [2] and thereafter they are widely used now as models for input flows to queuing systems where the rate of the arrival of customers depends on some external factors. For example, a queuing system can have two possible states: “stable state” and “peak-time state”. The former state can be regarded as the usual one, i.e., the rate of the customers’ arrival is not very high. The latter is characterized by the sharply increasing rate of the customers.

To develop dispatching rules for such type of systems it is necessary to estimate the process characteristics, such as the stationary probabilities of the system states, the sojourn time and the waiting time distributions, the mean length of the queue and others. In turn, to calculate these characteristics there is a need to estimate the customer’s arrival intensities and the intensities of the switching between the system states if they are unknown.

The main problem of MAP parameter estimation is that the theoretical investigation of an algorithm quality is hardly ever

possible when both all the parameters and the switching instants are unknown. The reason is the presence of correlation between the observations.

There are two main theoretical approaches to MAP parameter estimation. The first one is the 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. Paper [3] contains a detailed survey of the EM algorithms with a huge number of references. The estimators are supposed to maximize the log-likelihood function, but the exact solution of the derived set of equations cannot be obtained. Consequently, at each step the equations are solved numerically, and this, together with the iterative structure of algorithm, implies a big number of calculations.

The second approach is based on the matching moment method. A number of such algorithms are considered in detail in survey [4]. These methods are also connected with numerical solution of a set of equations, but as a rule algorithm are not iterative. Consequently, the volume of calculations is less than at the first method, but obtaining theoretical characteristics of the algorithms quality is hardly ever possible.

These approaches are connected with complicated numerical calculations. It implies difficulties in their theoretical investigations and, hence, the necessity of their study via simulation. Thus, simple heuristic methods are also

This research carried out in 2015-2016 was supported in part by “The Tomsk State University Academic D.I. Mendeleev Fund Program” under grant No 8.1.55.2015 and by RFBR Grant 16-01-00121 A

can be used for MAP parameter estimation. In our paper a method based on the sequential analysis principles is developed.

We consider the Markov-modulated Poisson process (MMPP) where the intensity of the flow of events is controlled by the continuous time Markovian chain with two states. Transitions between states happen at random instants. The sojourn time 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 Markovian chain. We also suppose that $\alpha_i \ll \lambda_i$, i.e., changes of the controlling chain state occur more rarely than observed events. Both all the parameters and the moments of switching of the controlling chain state are supposed to be unknown.

There are a number of papers devoted to the MMPP applications. Sh. Lu in [5] introduces a special type of the MMPP and applies it to characterize the occurrence patterns of New Zealand deep earthquakes. The author uses the EM algorithm for parameter estimation. The model fitting is performed by using Akaike information criteria.

S. R. Mahabhashyam and N. Gautam in [6] presented two applications of the Markov modulated processes in computer systems: Web server with multi-class requests and CPU with multiple processes. Authors investigated some characteristics of the systems such as the first and second moments of the service time of customers in system, the average waiting time in system and the average queue length. The results are illustrated with the help of numerical examples.

T. Yoshihara, Sh. Kasahara and Yu. Takahashi in [7] studied self-similar multimedia traffic characterized by the fact that the correlation never vanishes in a large time-scale. They used superposing two-state MMPP to describe this process and consider the condition on the preliminarily required parameters for the fitting. A number of numerical results were shown in order to verify the usefulness of the method.

S. L. Scott and P. Smyth in [8] considered the MMPP as a superposition of unobserved Poisson processes that were activated and deactivated by an unobserved Markov process, or as the Hidden Markov model. They introduced Markov-Poisson cascade (MPC) is an MMPP which underlying Markov process obeys certain restrictions. Then they applied the MPC to a data set containing click rate data for individual computer users browsing through the WorldWideWeb. P. Yan, T. Schoenharl, A. Pawling, and G. Madey in [9] developed these ideas in application to cell phone call activity. They focused on anomaly detection in the WIPER system that helps to detect possible emergencies from cell phone data, provided the corresponding information to emergency planners and responders, and suggested possible actions to mitigate the emergency. Recursive procedures were used to calculate the parameters and a posteriori distribution of the anomalous behavior of cell phone calls. Authors also described a simulation algorithm for the process under consideration.

H. S. Abhyankar and S. C. Graves in [10] investigated a product demand following a cyclic pattern in which low-demand intervals alternate with high-demand intervals. They used the MMPP to model this type of non-stationary behavior. The research was motivated by Teradyne that is subject to highly variable demand, driven by capital investment cycles in the semi-conductor industry. Authors developed the model to characterize the inventory requirements and service level for a hedging policy. Then they formulated an optimization problem to determine the location for the inventory hedge and conducted some computational experiments. The results provided some insight into the behavior and benefit from the hedging policy.

The papers surveyed above describe some applications of the MMPP process but of course not all of them. To take decisions concerning the process behavior and to develop dispatching rules one needs to fit a model and to evaluate the model parameters. That stresses the urgency of the MMPP parameter estimation problem and necessity of simple efficient estimation algorithms. Our paper [11] develops the sequential analysis approach to the MMPP parameter estimation. The intervals between the customer's arrivals are considered as a stochastic process. The mean of the process changes at the instants of the switching of the controlling chain state. The cumulative sum control chart (CUSUM) approach is used to detect the switching instants of the controlling chain state. After that the parameters of the process are estimated by using usual formulas.

In this paper, we study quality characteristics of our algorithm subject to the choice of the CUSUM algorithm parameters. Because of correlation between the observation and non-stationary structure of the process its theoretical investigation is hardly ever possible. So, the characteristics are studied via simulation as at the papers mentioned above.

This paper is organized as follows. Section II presents the algorithm proposed in [11]. Section III contains the description of the simulation procedure and presents the results of the simulation. Our recommendations concerning the parameter choice are given at the Section IV.

II. CUSUM ALGORITHM FOR THE MMPP PARAMETER ESTIMATION

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. Fig. 1 shows the principle of the sequence τ_i construction.

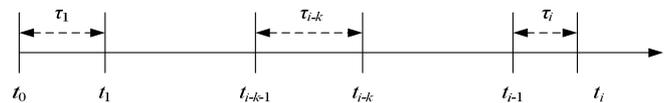


Fig. 1. Construction of the sequence $\{\tau_i\}$.

If the controlling chain is in the i -th state then the mean length between events is equal to $1/\lambda_i$. 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 τ_i by using CUSUM procedures.

Let the parameters λ_1, λ_2 satisfy the condition

$$0 < \lambda_2 < \lambda_1, \quad 1/\lambda_2 - 1/\lambda_1 > \Delta, \quad (1)$$

where Δ is a certain known positive parameter. In fact, the parameter Δ describes the minimum change of the mean length of the interval between events that should be detected by the algorithm. Choose then an integer parameter k describing the memory depth. The idea is to compare the values τ_i and τ_{i-k} , $i > k$. 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 $[t_{i-k-1}, t_i]$ then the expectations of the values τ_i and τ_{i-k} are different.

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

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

For the second procedure we introduce the sequence of the statistics

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

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

– $H_l(t_{i-k-1}, t_i)$ – the intensity of the arrival process on the interval $[t_{i-k-1}, t_i]$ is constant and equal to λ_l , $l = 1, 2$;

– $H_{l,m}(t_{i-k}, t_{i-1})$ – the intensity of the arrival process on the interval $[t_{i-k-1}, t_i]$ changed once from λ_l to λ_m , where $l=1, m=2$ or $l=2, m=1$.

The statistics $z_i^{(1)}$ expressed by formula (2) and $z_i^{(2)}$ expressed by formula (3) have the following properties:

$$\begin{aligned} E[z_i^{(1)} | H_l] < 0, \quad l = 1, 2, \quad E[z_i^{(1)} | H_{1,2}] > 0; \\ E[z_i^{(2)} | H_l] < 0, \quad l = 1, 2, \quad E[z_i^{(2)} | H_{2,1}] > 0. \end{aligned} \quad (4)$$

So the expectations of statistics (2), (3) change from negative values 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 $S_j^{(1)}$ and $S_j^{(2)}$ which are recalculated at the instants t_i . The sums are defined as follows

$$\begin{aligned} S_0^{(l)} &= \Delta; \\ S_j^{(l)} &= \max\{S_{j-1}^{(l)} + z_j^{(l)}, \Delta\}; \\ S_j^{(l)} &= \Delta, \quad \text{if } S_{j-1}^{(l)} + z_j^{(l)} \geq h^{(l)}; \\ & \quad l = 1, 2. \end{aligned} \quad (5)$$

Reaching the threshold h_l by the sum $S_j^{(l)}$ results in a decision considering the parameters changes; $l=1$ indicates decision on increase of the mean length of the interval between the events and decrease of the process intensity, $l=2$ indicates the opposite. In these case the corresponding sums are set Δ .

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

$$\begin{aligned} \sigma_0^{(l)} &= 0; \\ \sigma_m^{(l)} &= \min\{t_j > \sigma_{m-1}^{(l)} : S_{j-1}^{(l)} + z_j^{(l)} \geq h^{(l)}\}; \\ & \quad l = 1, 2. \end{aligned} \quad (6)$$

These are the instants of the detecting the controlling chain state change. Consider a sequence $\{n_m^{(l)}\}_{m \geq 0}$ associated with the sequence $\{\sigma_m^{(l)}\}$ as follows

$$\begin{aligned} n_0^{(l)} &= 0; \\ n_m^{(l)} &= \min\{t_j > \sigma_{m-1}^{(l)} : S_{j-1}^{(l)} = \Delta, S_j^{(l)} > \Delta\}; \\ & \quad l = 1, 2. \end{aligned} \quad (7)$$

Thus the instant $n_m^{(l)}$ is the first instant when the cumulative sum exceeds Δ to reach then the threshold. Fig. 2 illustrates the construction of the sequences.

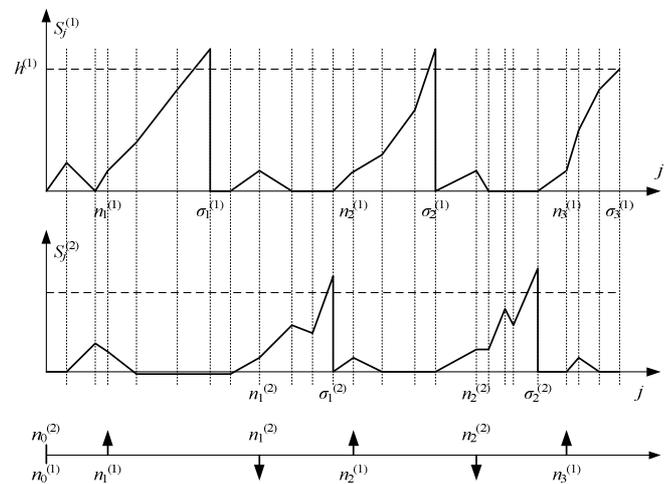


Fig. 2. Construction the sequences $\{\sigma_m^{(l)}\}$ and $\{n_m^{(l)}\}$.

The instants of occurrences t_j are marked by vertical dotted lines. The diagram above demonstrates an example of the sum $S_j^{(1)}$ behavior with the instants $\sigma_m^{(1)}$ and $n_m^{(1)}$ marked out. The diagram in the middle shows a similar example for the sum

$S_j^{(2)}$. We consider the instants $n_m^{(1)}$ 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 $n_m^{(2)}$ 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.

When implementing the procedure it is possible to encounter false alarm situations, i.e. the situations when the cumulative sum exceeds the threshold while the intensity does not change. 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 and consider the next instants in the row as false alarms.

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

$$\begin{aligned} q_0 &= 0; \\ q_{2m+1} &= \min\{n_j^{(1)} : n_j^{(1)} > q_{2m}\}; \\ q_{2m+2} &= \min\{n_j^{(2)} : n_j^{(2)} > q_{2m+1}\}; \\ m &\geq 0. \end{aligned} \quad (8)$$

The values q_1, q_2, \dots , are calculated using formula (8) while it is possible. If

$$\{n_j^{(1)} : n_j^{(1)} > q_{2m}\} = \emptyset \quad (\{n_j^{(2)} : n_j^{(2)} > q_{2m+1}\} = \emptyset)$$

then we set $q_{2m+1} = N$ ($q_{2m+2} = N$), where N is the instant of the last occurrence. Here the odd instants q_{2m+1} are the estimators of the instants when the intensity changes from λ_1 to λ_2 , and the even instants q_{2m+2} are the estimators of the instants when the intensity changes from λ_2 to λ_1 .

Fig. 3 demonstrates the construction of the sequence $\{q_j\}$. The diagram above presents the sequence $n_m^{(l)}$ constructed as it is described earlier.

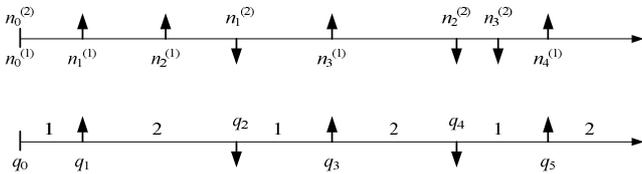


Fig. 3. Construction the sequences $\{q_j\}$.

The diagram below presents the sequence $\{q_j\}$. The estimators for the instants of the control chain states switching from λ_2 to λ_1 are marked by up arrows (because it

means the increase of the intensity), and the estimators for the instants of the control chain states switching from λ_1 to λ_2 are marked by down arrows (because it means the decrease of the intensity). The number of the supposed state of the control chain is also shown above the axis.

The estimators for the parameters λ_1, λ_2 are defined as follows

$$\hat{\lambda}_1 = \frac{N_1}{T_1}, \quad \hat{\lambda}_2 = \frac{N_2}{T_2} \quad (9)$$

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

The estimators for the parameters α_1, α_2 are defined as follows

$$\hat{\alpha}_1 = \frac{L_1}{T_1}, \quad \hat{\alpha}_2 = \frac{L_2}{T_2} \quad (10)$$

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

If $n_1^{(1)} > n_1^{(2)}$, then the initial value of the intensity is equal to λ_2 . The estimation procedure is constructed as above. The sequence $\{q_j\}$ is constructed as follows

$$\begin{aligned} q_0 &= 0; \\ q_{2m+1} &= \min\{n_j^{(2)} : n_j^{(2)} > q_{2m}\}; \\ q_{2m+2} &= \min\{n_j^{(1)} : n_j^{(1)} > q_{2m+1}\}; \\ m &\geq 0. \end{aligned} \quad (11)$$

Here the odd instants q_{2m+1} are the estimators of the instants when the intensity changes from λ_2 to λ_1 , and the even instants q_{2m+2} are the estimators of the instants when the intensity changes from λ_1 to λ_2 .

The estimators for the parameters λ_1, λ_2 are defined by formula (9), where N_1 is the total number of events occurred at the intervals $[q_{2m+1}, q_{2m+2}]$, $q_{2m+2} \leq N$ and T_1 is the total length of these intervals; N_2 is the total number of events occurred at the intervals $[q_{2m}, q_{2m+1}]$, $q_{2m+1} \leq N$ and T_2 is the total length of these intervals; $m \geq 0$. The estimators for the parameters α_1, α_2 are defined by formula (10), where L_1 is the total number of the switching points $q_{2m+2} \leq N$, L_2 is the total number of the switching points $q_{2m+1} \leq N$, $m \geq 0$.

The quality of the obtained estimators depends on the results of the first stage of the algorithm, i.e., on the quality of the switching point detection procedure. All sequential change

point detection methods are characterised by the mean number of false alarms (or the mean length of the interval between false alarms) and by the mean delay of the change point detection. An algorithm has a good quality if the mean delay is small and the mean length of the interval between false alarms is big, but these demands contradict each other. For example, the increase the threshold $h^{(l)}$ results in the increase of the time necessary for the sum $S_i^{(l)}$ to reach the threshold. Consequently, both the mean length of the interval between false alarms and the mean delay increase. Hence, the parameters of the CUSUM procedure should provide the minimal mean delay if the mean length of the interval between false alarms is fixed.

In paper [12] this property is proven for the case of independent random variables with the known distribution function and for a single change point, i.e., for the simplest case. It is shown there that the relation between the mean delay and the logarithm of the mean time between false alarms tends to a constant if the sample volume tends to infinity. Sequential change point detection procedures having this property are considered to be optimal. The process studied in our paper is characterized by the presence of correlation between the observations, so, the theoretical study of its quality is hardly ever possible. Besides, several change points should be detected. In this case it is supposed commonly that there is a rather big time interval between neighboring change points, but in our problem this time is distributed exponentially; thus, it can be rather close to zero with a positive probability.

Paper [11] presents some discussion concerning the algorithm parameter choice and contains some general recommendations. In this paper we study the proposed algorithm via simulation and formulate more detailed recommendation on the parameter choice.

III. SIMULATION PROCEDURE AND RESULTS

The simulation was performed in R-project software. As our two procedures are identical, we supposed that initially the controlling chain was at the first state and investigated the first procedure.

The proposed algorithm has the following parameters: the memory depth k , the parameter Δ , which provides the change in the mean of the statistics $z_i^{(l)}$ from negative to positive after the change point, and the thresholds $h^{(l)}$.

First, the process without swithing of the controlling chain states, i.e., the simplest Poisson flow with the parameter λ was simulated. The thresholds $h^{(l)}$ were changed and the numbers of false alarms, i.e., the numbers of instants when the cumulative sums $S_i^{(l)}$ exceed the corresponding threshold, were counted. The procedure was implemented with the parameters $k = 15$ and $\Delta = 0,2$, the number of events was taken as $N = 10000$. For every set of parameters from 5 to 10 replications were simulated.

Table 1 demonstrates results of the simulation. Here λ is the intensity parameter, $h^{(l)}$ is the threshold, \bar{L} is the number

of false alarms taken in the average by all replications, L_{max} is the maximum number of false alarms, \bar{T} is the mean number of the flow events between false alarms taken in the average by all replications, T_{min} is the minimum mean length of the interval between false alarms. So, the average data and the worst case data are presented.

TABLE 1. INFLUENCE OF THE FLOW INTENSITY AND THE THRESHOLD ON THE FALSE ALARM CHARACTERISTICS

λ	$h^{(l)}$	\bar{L}	L_{max}	\bar{T}	T_{min}
0,5	20	77,6	86	129,41	116,53
0,5	30	13,2	16	823,54	517
0,5	40	1,3	2	7692,31	5000
0,7	15	59,75	67	168,63	148,8
0,7	20	13,8	19	647,19	491,94
0,7	25	1,75	3	5714,29	4196
1	10	48,4	53	204,43	190,6
1	15	5,6	10	1227,82	526,44
1	18	2,4	4	4167	1772,33
1,5	5	81,4	99	123,86	100,52
1,5	7	22	29	447	273,54
1,5	10	3	7	3333	1319,33
2	3	108,9	121	91,65	81,7
2	5	18,43	24	649,61	414,39
2	7	1,2	2	8333	5000
5	1	66	68	147	146,81
5	1,5	10,7	13	834,3	797,67
5	2	2	3	5000	3953

One can see that both the mean number and the maximum number of false alarms decrease with the increase of the threshold if the flow intensity is constant. The minimum length of the interval between false alarms differs from the average one to 60%. This scatter can be explained by a nonstationarity of the process.

To give recommendation concerning the choice of the threshold we calculated for every value of λ the minimum value of the threshold h^* when commonly no false alarms are registered. Fig. 4 shows the dependence of the value h^* on the value of $1/\lambda$.

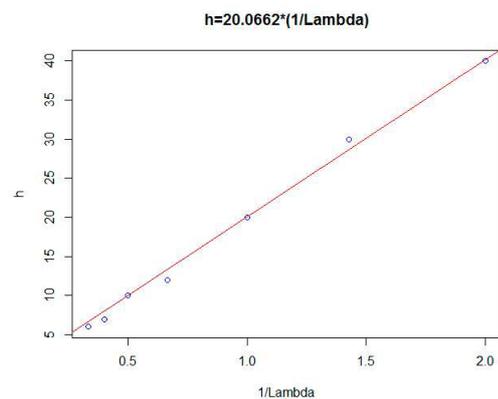


Fig.4. Dependence of the value h^* on the value of $1/\lambda$.

As it was expected, the dependence is close to linear because the mean value of the observation τ_i is equal to $1/\lambda$, and it is the mean length of the time interval between recalculating the cumulative sums. Analysis of the graph shows that false alarms are practically absent if $h^* \approx 20/\lambda$. So, there is no need to take h^* greater than $20/\lambda$ because the value $20/\lambda$ provides small false alarm probability, and the increase of the threshold value causes a big delay or even a skip of the chain state.

Table 2 illustrates the influence of the parameter Δ on the procedure characteristics. In this case $k=10$, $N = 10000$.

TABLE 2. INFLUENCE OF THE PARAMETER Δ ON THE FALSE ALARM CHARACTERISTICS

Δ	λ_1	$h^{(1)}$	\bar{L}	\bar{T}
0,2	2	5	14	744,81
0,3	2	5	5,5	1350,23
0,4	2	5	3	3828
0,5	2	5	1,5	1518
0,2	1	5	260,3	38,48
0,3	1	5	211	47,75
0,4	1	5	171,25	58,2
0,5	1	5	145	68,77
0,6	1	5	120	81,47
0,2	1	10	26,3	354,34
0,3	1	10	20	495,15
0,4	1	10	16,7	563,41
0,5	1	10	9,3	1065,1
0,6	1	10	6	1357,93

The increase in Δ causes the slower increase in the cumulative sum. Consequently, the sum hits the threshold rarely, which leads to decrease in the number of false alarms.

Then, the MMPP process with presence of swithing of the controlling chain state was simulated. The initial state of the controlling chain was the first state, i.e., the initial intensity of the flow of events was equal to λ_1 which was constant. The influence the parameter λ_2 on the delay in the change point detection at the first procedure was examined. The parameters of the controlling chain were taken as $\alpha_1=0,1$ and $\alpha_2=0,005$. For every set of parameters from 5 to 10 replications were implemented. Table 3 shows the results of the simulation.

TABLE 3. INFLUENCE OF THE THRESHOLD ON THE DELAY CHARACTERISTICS

k	$h^{(1)}$	Δ	λ_1	λ_2	δ	\bar{D}	D_{\max}
15	5	0,2	2	0,4	2	6,7	9
15	5	0,2	2	0,6	1,17	5	8
15	5	0,2	2	1	0,5	16	27
15	5	0,2	2	1,2	0,33		
10	7	0,2	2	0,4	2	6,8	8
10	7	0,2	2	0,6	1,17	13,7	30
10	7	0,2	2	0,8	0,75	16,6	23
10	7	0,2	2	1	0,5		

Here k , $h^{(1)}$ and Δ are the algorithm parameters, $\delta = 1/\lambda_2 - 1/\lambda_1$, \bar{D} is the average number of events between the change point and its detection, or the average delay, and D_{\max} is the maximum delay. If this cells are empty then there were cases when the change point was not detected.

In general, one can observe a constant increase in the average delay with the increase in the parameter $h^{(1)}$, such behavior is common for sequential change point detection procedures [13]. Besides, the average delay grows with the decrease in the parameter δ , and for $\delta \leq 0,5$ the cases when the change point is not detected start to appear.

Then we studied how the memory depth k affects the delay. Table 4 presents the results of the investigation. In addition we considered the error of the change point estimator, i.e., the difference between the instant q_1 (the first instant when the cumulative sum exceeds Δ to reach then the threshold) and the change point θ . Here \bar{E} is the average error and E_{\max} is the maximum one, $\Delta = 0,4$, $h^{(1)} = 10$.

TABLE 4. INFLUENCE OF THE MEMORY DEPTH ON THE DELAY CHARACTERISTICS

k	λ_1	λ_2	δ	\bar{D}	D_{\max}	\bar{E}	E_{\max}
15	2	0,4	2	6	10	1,1	4
15	2	0,6	1,17	8,38	13	2,1	10
20	2	0,6	1,17	10,6	14	0,7	7
25	2	0,6	1,17	7,9	12	2,8	7
15	2	0,8	0,75	11,7	17	1,7	4
20	2	0,8	0,75	19,7	29	3,7	10
25	2	0,8	0,75	20,7	54	9,3	45

For the parameters listed in Table 4 and $k < 15$ ($k = 5$ and $k = 10$), the change point in many cases was not detected, hence, the corresponding data are not presented in the table. Consequently, we considered the values of k from 15 to 25. The greater is the parameter δ the less number of steps is necessary to detect the change point; hence, the smaller values of the parameter k can be taken.

Any change point should be detected in not more than k steps. In the case of the single change point the parameter k should be taken as large as possible, because a large number of events before the change point is supposed to be observed. In the case of the multiple change points the number of observations between neighboring change points is limited, so the value of k should not exceed this number. It means that the intensity of the controlling chain state switchings significantly less then the intensity of the input flow of events, i.e., $\lambda_i > r\alpha_i$. This condition is usual for the the MMPP describing a two-regimes queueing system, because the change of the controlling chain state needs the change of dispatching rules if the system is in the same state rather long time. Usually the values $r \geq 30$ are considered. In paper [11], we recommended to choose the parameter $k \approx r/2$.

After that, the influence of the parameter Δ on the delay time with fixed value of $h^{(1)}$ was studied. According condition (1), one have to choose $\Delta < \delta$. Table 5 shows the result of the simulation for $k = 15$, $h^{(1)} = 10$.

TABLE 5. INFLUENCE OF THE PARAMETER Δ ON THE DELAY CHARACTERISTICS

Δ	λ_1	λ_2	a	δ	\bar{D}	D_{\max}
0,2	2	0,6	0,3	1,17	8,2	9
0,4	2	0,6	0,3	1,17	13,8	14
0,5	2	0,6	0,3	1,17	13,75	15
0,2	2	0,4	0,2	2	1,5	2
0,4	2	0,4	0,2	2	2,4	4
0,6	2	0,4	0,2	2	4,17	7
0,8	2	0,4	0,2	2	7,25	34

In [11] we recommended to choose the parameter $\Delta \approx \delta/2$. Table 5 shows that if $\Delta > \delta/2$ the mean and the maximum delay exceeds the value of k , thus, the risk of the change point skip increases. So, the more appropriate choice of the parameter Δ is $\Delta < \delta/2$. Under this condition, the mean delay time does not depend significantly on the parameter Δ ; hence, it is enough for the parameter to satisfy condition (1). Usually one does not know the exact value of δ , but it can be interpreted as the minimal difference of the average lengths of the interval between the input flow events in two states of the control chain that should be detected.

IV. RECOMMENDATIONS ON THE PARAMETER CHOICE

Analysing the simulation results we can give some recommendation on the parameter choice for the proposed algorithm.

The CUSUM algorithm for MMPP parameter estimation can be used in the following conditions:

- the intensity of the controlling chain state switchings is significantly less than the intensity of the input flow of events, i.e., $\lambda_i > r\alpha_i, r \geq 30$;
- the difference of the mean lengths of the interval between the input flow events in different states of the controlling chain $\delta = 1/\lambda_2 - 1/\lambda_1$ is not less than 0,5;
- the threshold parameters $h^{(i)} \approx 20/\lambda_i$;
- the memory depth parameter k belongs to the interval [15,25];
- the parameter Δ is less than $\delta/2$.

If these conditions do not hold true then we can not guarantee the correct change point detection because either false alarms or skips of the change could appear rather often.

As the algorithm is applied a posteriori to a fixed volume sample, one has a possibility to implement it several times for different sets of parameters until the conditions above are fulfilled.

Acknowledgment

We are grateful to Prof. Sergey Vorobeychikov from Tomsk state university for his support and useful comments.

References

1. M.F. Neuts. "A versatile Markovian point process." J. Appl. Probab., vol.16, pp. 764–774. 1979.
2. D.M. Lucantoni, K. S. Meier-Hellstern, and M. F. Neuts. "A single server queue with server vacations and a class of non-renewal arrival processes." Adv. Appl. Prob., vol. 22, pp. 676–705, 1990.
3. S. Asmussen. "Phase-type distributions and related point processes: Fitting and recent advances." In: Chakravarty and Alfa (eds.), Matrix-analytic methods in stochastic models, NY: Marcel Dekker. Lect. Notes Pure Appl. Math., vol. 183, pp. 137–149, 1997.
4. I. Gerhardt, B.L.Nelson. "On capturing dependence in point processes: Matching moments and other techniques." Technical report, Northwestern univ., 2009.
5. Sh. Lu. "Markov modulated Poisson process associated with state dependent marks and its application to the deep earthquakes." Ann Inst Stat Math, vol. 64, pp. 87–106, 2010.
6. S. R. Mahabhashyam, N. Gautam. "On Queues with Markov Modulated Service Rates." Queueing Systems, vol. 51 (1-2), pp. 89-113, 2005.
7. T. Yoshihara, Sh. Kasahara and Yu. Takahashi. "Practical Time-Scale Fitting of Self-Similar Traffic with Markov-Modulated Poisson Process." Telecommunication Systems, vol. 17:1-2, pp. 185–211, 2001.
8. S. L. Scott, P. Smyth. "The Markov modulated Poisson process and Markov Poisson cascade with applications to web traffic data." Bayesian Statistics, vol. 7, pp. 671–680, 2003.
9. A. Pawling, T. Schoenharl, P. Yan, and G. Madey, "WIPER: An emergency response system," in Proceedings of 5th International ISCRAM Conference, Washington, DC, 2008.
10. H. S. Abhyankar, S. C. Graves. "Creating an inventory hedge for Markov-modulated Poisson demand: An application and a model." Manufacturing Service Oper. Management, vol. 3, pp. 306–320, 2001.
11. Yu. Burkatovskaya, T. Kabanova and S. Vorobeychikov. "CUSUM Algorithms for Parameter Estimation in Queueing Systems with Jump Intensity of the Arrival Process." Communications in Computer and Information Science, vol. 564. pp. 275–288, 2015.
12. G. Lorden. "Procedures for reacting to a change in distribution". Annals. Math. Statist., No. 42, pp. 1897–1971, 1971.
13. E.S Page. "Continuous inspection schemes." Biometrika, vol. 42, No. 1, pp. 100–115, 1956.