

Филиал Кемеровского государственного университета  
в г. Анжеро-Судженске

Национальный исследовательский  
Томский государственный университет

Кемеровский государственный университет

Институт проблем управления им. В. А. Трапезникова РАН

Институт вычислительных технологий СО РАН

**ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ  
И МАТЕМАТИЧЕСКОЕ  
МОДЕЛИРОВАНИЕ  
(ИТММ–2015)**

**Материалы XIV Международной конференции  
имени А. Ф. Терпугова  
18–22 ноября 2015 г.  
Часть 1**

Издательство Томского университета

2015

УДК 519

ББК 22.17

И74

Редколлегия:

*Р. Т. Якупов*, д-р физ.-мат. наук, профессор;

*А. А. Назаров*, д-р техн. наук, профессор;

*Т. В. Любина*, канд. физ.-мат. наук

**Информационные** технологии и математическое моделирование  
И74 (ИТММ–2015): Материалы XIV Международной конференции имени  
А. Ф. Терпугова (18–22 ноября 2015 г.). – Томск : Изд-во Том. ун-та,  
2015. – Ч. 1. – 218 с.

ISBN 978-5-7511-2382-6

В часть 1 вошли материалы докладов, представленные на XIV Международной конференции имени А. Ф. Терпугова «Информационные технологии и математическое моделирование» на секциях «Вероятностные методы и модели», «Модели и методы массового обслуживания», «Оптимизационные модели и исследование операций».

Для специалистов в области информационных технологий и математического моделирования.

**УДК 519**

**ББК 22.17**

*Конференция проводится при поддержке Российского фонда  
фундаментальных исследований (проект № 15-01-20933-г)*

ISBN 978-5-7511-2382-6

© Филиал Кемеровского государственного университета  
в г. Анжеро-Судженске, 2015

# ВЕРОЯТНОСТНЫЕ И ЭКОНОМЕТРИЧЕСКИЕ МЕТОДЫ И МОДЕЛИ

## CUSUM ALGORITHMS FOR PARAMETER ESTIMATION IN QUEUEING SYSTEMS WITH JUMP INTENSITY OF THE ARRIVAL PROCESS

*Yu. Burkatovskaya, T. Kabanova, S. Vorobeychikov*

*National Research Tomsk Polytechnic University, Tomsk, Russia*

*National Research Tomsk State University, Tomsk, Russia*

### 1. Introduction

Markovian arrival processes 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 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. One of the problems connected with MAP is the estimation of intensity parameters by observing flow of events. A survey of estimation methods is given in [3]. Its emphasis is on maximum likelihood estimation and its implementation via the EM (expectation-maximization) algorithm. This approach is developed for different conditions in [4, 5], etc. The survey [6] 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 [7]. Bayesian approach based on the a posteriori probability of the controlling chain state is developed in [8]. Quality of those methods is typically examined via simulation.

In this paper we propose a different approach to MAP parameter estimation using the sequential analysis methods described in [9] and [10]. The key idea is to consider time intervals between arrivals as a stochastic process which parameters change in random instants. First we detect these points using sequential change point detection methods. Then we estimate the intensity parameters under the assumption that the intensity is constant between detected change points.

### 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  $l$ -th state is exponentially distributed with the parameter  $\alpha_l$ ,  $l=1,2$ .

The flow of events has the exponential distribution with the intensity parameter  $\lambda_1$  or  $\lambda_2$  subject to the state of the Markovian chain. The parameters of the system  $\lambda_1$ ,  $\lambda_2$  and the instants of switching between the states are supposed to be unknown. We also suppose that  $\lambda_l \ll \alpha_l$ , i.e. changes of the

controlling chain states occur more rarely than observed events. The sequence of instants of arriving events is observed. The problem is to estimate the parameters  $\lambda_1, \lambda_2, \alpha_1, \alpha_2$ .

### 3. Algorithm 1

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. 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  $\lambda_1, \lambda_2$  satisfy the condition

$$0 < \lambda_2 < \lambda_1; \quad 1/\lambda_2 - 1/\lambda_1 \geq d, \quad (1)$$

where  $d$  is a certain positive parameter. If the constant  $d$  is unknown then it can be chosen as the minimal difference between the mean lengths of the intervals  $\tau_i$  when the controlling chain is in different states that should be detected. 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  $n\alpha_l < \lambda_l$ . We suppose that  $n$  is rather large (for example,  $n \geq 10$ ), hence changes of the controlling chain states occur more rarely than observed events. 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 [5, 8]).

Choose then an integer parameter  $k > 1$  describing the memory depth. According to our analysis, a good choice of the parameter  $k$  is  $k \approx n/2$ . The idea is to compare the values  $\tau_i$  and  $\tau_{i-k}$ ,  $i > k$ . If there are no changes of the controlling chain state within the interval  $[t_{i-k-1}, t_i]$  then the values  $\tau_i$  and  $\tau_{i-k}$  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  $\tau_i$  and  $\tau_{i-k}$  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  $[t_{i-k-1}, t_i]$ .

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 choose the positive parameter  $\Delta < d$  and 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  $\lambda_l$ ,  $l=1,2$ ;

–  $H_{l,m}(t_{i-k}, t_{i-1})$  – the intensity of the arrival process on the interval  $[t_{i-k}, t_{i-1}]$  changed once from  $\lambda_l$  to  $\lambda_m$ ,  $l=1, m=2$  or  $l=2, m=1$ ;

The statistics  $z_i^{(j)}$ ,  $j \in \{1,2\}$  (2), (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 means of statistics (2), (3) change from negative to positive values when the intensity of the process changes. These properties determine the construction of the procedures. We introduce positive constants  $h_1$  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  $t_i$ . It is defined as follows

$$\begin{aligned} S_0^{(l)} &= \Delta; \quad l=1,2 \\ S_i^{(l)} &= \max\{0, S_{i-1}^{(l)} + z_i^{(l)}\}, \quad i > k; \\ S_i^{(l)} &= 0, \quad \text{if } S_i^{(l)} \geq h_l. \end{aligned} \quad (5)$$

Reaching the threshold  $h_l$  by the sum  $S_i^{(l)}$  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.

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.

$$\sigma_0^{(l)} = 0; \quad \sigma_m^{(l)} = \min\{t_j > \sigma_{m-1}^{(l)} : S_j^{(l)} \geq h_l\}. \quad (6)$$

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

$$n_0^{(l)} = 0; \quad n_m^{(l)} = \max\{t_j \leq \sigma_m^{(l)} : S_j^{(l)} > 0, S_{j-1}^{(l)} = 0\}. \quad (7)$$

Thus the instant  $n_m^{(l)}$  is the first instant when the cumulative sum becomes positive to reach then the threshold. We consider the instants  $n_i^{(1)}$  ( $n_i^{(2)}$ ) as the estimators for the instants when the mean length of the interval between the events increases (decreases).

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 equations (5). Then construct the sequences  $\{\sigma_m^{(l)}\}$ ,  $\{n_m^{(l)}\}$  defined by equations (6), (7). Let  $n_1^{(1)} < n_1^{(2)}$ , then the initial value of the intensity is equal to  $\lambda_1$ . Define the sequence

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

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

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

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  $\lambda_1$  to  $\lambda_2$ , and the even instants  $q_{2l+2}$  are the estimators of the instants when the intensity changes from  $\lambda_2$  to  $\lambda_1$ .

Define estimators for the parameters  $\lambda_1, \lambda_2$

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

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;  $l \geq 0$ .

Define estimators for the parameters  $\alpha_1, \alpha_2$

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

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  $q_{2l+2} \leq N$ ,  $l \geq 0$ .

The parameters  $\Delta$  and  $h_i$  affect the characteristics of the CUSUM procedure, i.e., the mean delay and the mean time between false alarms (see [10]). 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 choice of the parameter  $\Delta$  as  $\Delta \approx d/2$ . If the memory depth is equal to  $k$  then the sum  $S_i^{(l)}$  has to reach the threshold  $h_i$  in not more than  $k$  steps (while  $Ez_i^{(1)} > 0$ ). It can be provided by the choice of  $h_i$  as  $h_i < k\Delta \approx nd/4$ . Note that the parameter  $h_i$  should not be significantly less than its upper bound because it can increase the number of false alarms.

#### 4. Algorithm 2

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

$$\hat{\tau} = T/N. \quad (11)$$

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

$$1/\lambda_1 < E\hat{\tau} < 1/\lambda_2. \quad (12)$$

We introduce the sequence of the statistics

$$z_i^{(1)} = \tau_i - \hat{\tau}. \quad z_i^{(2)} = -\tau_i + \hat{\tau}. \quad (13)$$

Using in Algorithm 1 statistics (13) instead of (2), (3) we obtain Algorithm 2. The parameter  $h_i$  can be chosen as  $h_i < nd/4$ , i.e. as at the first algorithm.

### 5. Numerical simulation

The model for the considered flow and the suggested algorithms was implemented with varying parameters. The results of the simulation for Algorithm 2 are presented in the table below.

$T$	$\lambda_1$	$\lambda_2$	$\alpha_1$	$\alpha_2$	$h_1$	$h_2$	$\hat{\tau}$	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{\alpha}_1$	$\hat{\alpha}_2$
1000	5	1	0,3	0,2	0,5	0,5	0,3883	5,2336	1,2276	0,3220	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

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  $\hat{\lambda}_i$  because of not correct estimation of the controlling chain current state. On the other hand, increasing of  $h_i$  leads to 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  $\hat{\alpha}_i$  decrease because less switching points are detected on the first stage of the procedures. In the Table for  $h_1 = h_2 = 1$  one can see that the estimators  $\hat{\alpha}_i$  considerably less the real

values of the parameters  $\alpha_i$ . The best results are obtained for  $h_1 = h_2 = 0,8$  for all intensity parameter values. Thus, choice of the algorithm parameters is a rather difficult problem requiring further theoretical investigations.

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

### Conclusion

MAPs are used as models for real processes, particularly, for call-centers or http-server customers (see [3], [4]), healthcare systems (see [5]), 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 need the distribution function of the observing flow and, hence, can be applied to parameter estimation of other types of flows.

*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. 1979. No. 16. P. 764–774.
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. Prob., 1990. Vol. 22. P. 676–705.
3. Asmussen S. Phase-type distributions and related point processes: Fitting and recent advances // Chakravarthy and Alfa (eds.), Matrix-analytic methods in stochastic models, NY: Marcel Dekker. Lect. Notes Pure Appl. Math. 1997. Vol. 183. P. 137–149.
4. Breuer, L., Kume, A. An EM algorithm for Markovian arrival processes observed at discrete times // Measurement, Modelling, and Evaluation of Computing Systems and Dependability and Fault Tolerance Lecture Notes in Computer Science. 2010. Vol. 5987. P. 242–258.
5. Okamura H., Dohi T., Trivedi K. S. Markovian Arrival Process Parameter Estimation With Group Data // IEEE/ACM Transaction on Networking, 2009. Vol. 17, No. 4, P. 1326–1340.
6. Gerhardt I., Nelson B. L. On capturing dependence in point processes: Matching moments and other techniques // Technical report, Northwestern univ. 2009.
7. Duffie D., Glynn P. Estimation of Continuous-Time Markov Processes Sampled at Random Time Intervals // Econometrica. 2004. Vol. 72. P. 1773–1808.
8. 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 University Journal of Control and Computer Science. 2011. Vol. 4(17). P. 25–40.
9. Page E. S. Continuous inspection schemes // Biometrika. 1956. Vol. 42, No. 1. P. 100–115.
10. Lorden G. Procedures for reacting to a change in distribution // Annals. Math. Statist. 1971. No. 42. P. 1897–1971.

# STATISTICAL APPROACHES TO INTRUSION DETECTION: AN OVERVIEW OF SELECTED METHODS

*C. Callegari, S. Giordano, M. Pagano*

*Dept. of Information Engineering, University of Pisa, Pisa, Italy*

*CNIT Research Unit – University of Pisa, Pisa, Italy*

In recent years Internet has become the playground for providing sensitive services to an ever growing amount of end-users, most of them only partially aware of the risks deriving from information sharing on the net. Along with the wide proliferation of new services, the number and impact of security attacks have been continuously increasing. Indeed, the knowledge required to carry out an attack has been decreasing, since software tools for this aim are largely available on Web sites all over the world [1].

Recent advances in encryption, public key exchange, digital signatures, and the development of related standards have set a foundation for network security. However, network security goes beyond, because it must include security of computer systems and networks, at all levels, top to bottom. To this aim, the use of an Intrusion Detection System (IDS) is of primary importance to reveal ongoing intrusions in a network or in a system.

The goal of this paper is to provide an introduction to intrusion detection to non experts in the field, highlighting how advanced statistical techniques can be used to cope with known and unknown attacks.

## **Intrusion Detection System (IDS)**

In a nutshell, an IDS [2] is a software/hardware tool designed to reveal an intrusion while it is in act, or soon after it has occurred. IDSs are usually classified on the basis of different aspects [3]:

- the scope: Network IDSs vs. Host IDSs;
- the architecture: Centralized IDSs vs. Distributed IDSs;
- the approach used to process the input data: Stateless IDSs vs. Stateful IDSs;
- the detection technique: Misuse or Signature based IDSs vs. Anomaly based IDSs.

The last distinction is probably the most important one since these two categories are based on a completely different approach to the intrusion detection problem. Indeed a misuse-based IDS reveals the intrusions, looking for patterns (signatures of all the *known* attacks) of action that are known to be related to an intrusion. On the contrary, an anomaly-based IDS relies on the knowledge of a *reference model*, representing the *normal behavior* of the controlled system, and an intrusion is considered as a significant deviation from that model. This kind of systems can also detect never seen before intrusions, while a misuse-based IDS, not having the corresponding signature in the database, cannot detect any new attack.

State of the art in the field of intrusion detection is mainly represented by misuse-based IDSs [4]. Considering that most attacks are realized with known tools, downloaded from the Internet, a signature-based IDS could seem a good solution. Nevertheless the most dangerous attacks are those prepared *ad hoc* and, by definition, a misuse-based IDS is completely unable to block such a kind of attacks. Hence, the capabilities of the anomaly-based IDSs have been deeply investigated.

The rest of the paper provides a general description of anomaly-based IDSs as well as an overview of some of the best known statistical tools, mainly based on the past experiences of the authors. For sake of brevity, we will just introduce the different approaches to highlight how advanced statistical techniques may be used in the framework of intrusion detection (for a more detailed analysis see, for instance [5], and references therein).

### **Anomaly-based IDSs**

A wide range of different approaches to anomaly detection has been proposed in the literature, which can be classified according to different features:

- aggregation level: the IDS may act on single IP flows (e.g., considering the state transition of a TCP connection [6]), on network-mask based traffic aggregates or on random aggregates defined through sketches [7];

- input data: different traffic descriptors have been considered in the literature, such as the number, length and inter-arrival time of packets (at different levels of aggregation), number of active IP addresses or TCP/UDP ports, and so on;

- dimensionality of the input: “Classical” IDSs analyze the relevant traffic features separately, while “multidimensional” IDSs apply a single vectorial algorithm to all the traffic descriptors at the same time, leading to a processing time which is almost independent from the number of analyzed features [8];

- decision method: it represents the core of any IDS as it is responsible to decide if the current sample is compliant with the normal behaviour of the system or represents an anomaly. Some of the most commonly used algorithm will be described in the following section.

### **Statistical approaches to Anomaly detection**

A first classification of the algorithms for anomaly detection is based on the *domain* in which data are processed. Most of the approaches works on temporal series (sequence of packets or data collected over fixed-length time bins), but it is worth noticing that the analysis can be performed also in a transformed domain. For instance, in [9] the strong periodicity around its RTT (Round Trip Time) of a *normal* TCP flow is detected in the Fourier domain using its Power Spectral Density, which presents peaks at the integral multiples of the corresponding fundamental frequency. Wavelet transform is also widely used for detecting change-points in traffic statistics (e.g., in [10] the change points of CUSUM statistics are better detected considering the detail

coefficients) as well as for signal decomposition (e.g., in [11] the components at relevant time scales are extracted).

In the rest of this section we briefly describe two of the most widely used statistical approaches (and related variants), giving an intuitive explanation of their applicability in the framework of intrusion detection.

**Principal Component Analysis (PCA)** (or **discrete Karhunen–Loeve transform**) is the most commonly used techniques to analyze high dimensional data structures [12]. In a nutshell, PCA is a linear transformation that maps a coordinate space onto a new one whose axes, called Principal Components (PCs), have the property to point in the direction of maximum variance of the residual data (i.e., the difference between the original data and the data mapped onto the previous PCs). The empirical bases for the application of PCA to the anomaly detection field are provided in [13], where the authors perform an analysis of several traffic measurements taken over two backbone networks (Abilene and Geant) highlighting, by means of PCA, that these measurements have a small intrinsic dimension. This conclusion allows the authors to think that PCA could be suitable for anomaly detection. In more detail, the basic idea behind the several PCA-based detection methods is to project the measurement data onto a normal subspace (spanned by a limited number of PCs) and consider the residuals to evaluate how much anomalous the data are (see, for instance, [14] and references therein).

In spite of its wide popularity, PCA-based anomaly detection presents some intrinsic difficulties in tuning the system parameters: in particular, the number of dominant PCs may be very hard to be determined [15], causing the system to behave in a strongly unstable way. Moreover, it is intuitively clear that these methods are unsuitable when the data do not present a “linear structure”. A promising alternative is represented by **K-PCA**, which overcomes these limitations by applying the techniques of kernel methods [16]. In a nutshell, using a kernel, the originally linear operations of PCA are done in a reproducing kernel Hilbert space with a non-linear mapping, that can be chosen according to the intrinsic structure of data [17].

**CUSUM (or CUMulative SUM control chart)** algorithm is quite a “classical” approach to detect irregular patterns in traffic traces. Being a sequential analysis technique, typically used for change detection, the application to network anomaly detection is quite straightforward [18]. Given a sequence of observations (drawn from some stochastic process), the goal of the algorithm is to detect with the smallest possible delay a change in the distribution of the data. The assumption of the method is that the distribution before and after the change are known. Unfortunately, such constraint is usually not respected in the field of anomaly detection: indeed, we cannot suppose that the distribution after the change is known (usually neither the distribution before the change is known). This implies the need of using non parametric versions of the algorithm, such as the **multi-chart non parametric CUSUM (MNP-**

CUSUM), originally proposed in [19], which seems to be quite robust wrt. the estimation of the empirical parameters and the definition of the decision thresholds [20].

### Conclusions

In this paper we presented different approaches that can be used in revealing anomalies in traffic data. It is worth mentioning that the actual performance of an IDS depends on the considered traffic features and on the typologies of data and attacks, so the choice of the statistical tools used to characterize the normal behaviour of the users is still an open problem.

### References

1. Packet Storm Home Page [Electronic resource] // URL: <http://www.packetstormsecurity.org> (last accessed date: 07.07.2015).
2. Denning D.E. An Intrusion-Detection Model // IEEE Transactions on Software Engineering. February 1987.
3. Stallings W. Cryptography and Network Security. Principles and Practice // Pearson Education Limited, 6<sup>th</sup> edition, 2014.
4. Roesch, M., Snort – Lightweight Intrusion Detection for Network // Proc. USENIX Lisa, 1999
5. Callegari C., Coluccia A., D’Alconzo A. et al. A methodological overview on anomaly detection // Data Traffic Monitoring and Analysis. Springer Berlin Heidelberg, 2013. P. 148–183.
6. Callegari C., Vaton S., Pagano M. A New Statistical Method for Detecting Network Anomalies in TCP Traffic // European Transactions on Telecommunications (ETT). November 2010. 21 (7). P. 575–588.
7. Callegari C., Gazzarrini L., Giordano S. et al. When randomness improves the anomaly detection performance // Proceedings of 3rd International Symposium on Applied Sciences in Biomedical and Communication Technologies (ISABEL), 2010.
8. Callegari C., Casella A., Giordano S. et al. Sketch-based multidimensional IDS: A new approach for network anomaly detection. // Communications and Network Security (CNS), 2013 IEEE Conference on. October 2013. P. 350–358.
9. Cheng C.-M., Kung H.T, Tan K.-S. Use of Spectral Analysis in Defense Against DOS Attacks // IEEE Globecom. 2002.
10. Carl G., Brooks R.R., Rai S. Wavelet based Denial-of-Service detection // Computers & Security. 2006. No. 25.
11. Barford P., Kline J., Plonka D., Ron A. A Signal Analysis of Network Traffic Anomalies // ACM Internet Measurement Workshop. 2002.
12. Shyu M., Chen S., Sarinnapakorn K., Chang L. A novel anomaly detection scheme based on principal component classifier // IEEE Foundations and New Directions of Data Mining Workshop, in conjunction with ICDM03. 2003. P. 172–179.
13. Lakhina A., Papagiannaki K., Crovella M. et al. Structural analysis of network traffic flows // Proceedings of the joint international conference on Measurement and modeling of computer systems, SIGMETRICS ’04/Performance ’04. New York, USA, 2004. ACM. P. 61–72.
14. Callegari C., Gazzarrini L., Giordano S., Pagano M., and Pepe T. Improving PCA-based anomaly detection by using multiple time scale analysis and Kullback-Leibler divergence // International Journal of Communication Systems. October 2014. No. 27(10). P. 1731–1751.

15. Ringberg H., Soule A., Rexford J., Diot C. Sensitivity of PCA for traffic anomaly detection. // SIGMETRICS Perform. Eval. Rev. 2007. Vol. 35(1). P. 109–120.
16. Schoelkopf B., Smola A., Mueller K.-R. Nonlinear component analysis as a kernel eigenvalue problem // Neural Comput. July 1998. Vol. 10(5). P. 1299–1319.
17. Callegari C., Donatini L., Giordano S., Pagano M. Improving Stability of PCA-based Network Anomaly Detection by means of Kernel-PCA // Int. Journal Computational Science and Engineering (accepted for publication).
18. Tartakovsky A. G., Rozovskii B. L., Blaek R. B., Kim H. Detection of intrusions in information systems by sequential change-point methods // Statistical Methodology. 2006. No. 3(3). P. 252–293.
19. Salem O., Vaton S., Gravey A. A scalable, efficient and informative approach for anomaly-based Intrusion Detection Systems: theory and practice // International Journal of Network Management, 2010. No. 20(5). P. 271–293.
20. Callegari C., Giordano S., Pagano M. Histogram cloning and cusum: An experimental comparison between different approaches to anomaly detection // International Symposium on Performance Evaluation of Computer and Telecommunication Systems (SPECTS). July 2015.

## ON STOCHASTIC MODELS OF INTERNET TRAFFIC

*T. Czachórski<sup>1</sup>, J. Domańska<sup>1</sup>, M. Pagano<sup>2</sup>*

*<sup>1</sup>Institute of Theoretical and Applied Informatics*

*Polish Academy of Sciences, Poland*

*<sup>2</sup>Department of Information Engineering, University of Pisa, Italy*

### Introduction

The paper discusses various models of self-similar Internet traffic and techniques for estimating the intensity of Long-Range Dependence (LRD).

During the last decades, stimulated by experimental measurements in computer networks, a growing interest was paid to processes characterized by a slowly decaying correlation structure. We review some of the most relevant results in the framework of Internet traffic modeling and performance evaluation and illustrate them with our measurements and their analysis. We present different classes of stochastic models which have been used in the literature, highlighting their main features and the motivation behind their use. Then we concentrate on Markov models as the most interesting from our point of view because they may be incorporated in Markov queueing models.

#### Long Range Dependence and self-similarity

In the literature the terms long-range dependence and self-similarity are often used without distinction, although they are not equivalent concepts [1].

A continuous time process  $Y(t)$  is exactly self-similar with the Hurst parameter  $H$  if it satisfies the following condition [2]:

$$Y(t) = a^{-H} Y(at)$$

for  $t \geq 0$ ,  $a \geq 0$  and  $0 < H < 1$ . The above equality is in the sense of finite dimensional distributions and the Hurst parameter expresses the degree of the self-similarity. The process  $Y(t)$  may be nonstationary.

In the case of network traffic one usually has to deal with time series rather than a continuous process. In that context the above definition can be summarized as follows. Let  $X(t)$  be a stationary sequence representing increment process (e.g. in bytes/second). The corresponding aggregated sequence having level of aggregation  $m$ :

$$X^{(m)}(k) = \frac{1}{m} \sum_{i=1}^m X((k-1)m + i), k = 1, 2, \dots$$

is obtained by averaging  $X(t)$  over nonoverlapping blocks of length  $m$ . The following condition is satisfied for a self-similar process:

$$X = m^{1-H} X^{(m)}$$

for all integers  $m$ . A stationary sequence  $X$  is second-order self-similar if  $m^{1-H} X^{(m)}$  has the same variance and auto-correlation as  $X$  for all  $m$ . A stationary sequence  $X$  is asymptotically second-order self-similar if  $m^{1-H} X^{(m)}$  has the same variance and auto-correlation as  $X$  as  $m \rightarrow \infty$ .

Asymptotically second-order self-similar processes are also called long-range dependent processes and this is the property that network traffic exhibits [2].

### Estimation of the Hurst parameter

Starting from traffic data, one of the key problems is to verify the presence of some kind of self-similarity (at least asymptotically), which is typically done through the estimation of the Hurst parameter. Different estimation techniques have been proposed in the literature [3, 4], working directly on the aggregated time series or in a transformed domain.

*Variance time plot* is the most famous method of the first class. This method is a graphical test for LRD. The estimated value of Hurst parameter is obtained by fitting a simple least squares line through the resulting points in the plane. The methods working directly on the aggregated time series also includes other worth-mentioning approaches: *R/S statistics*, *Higuchi's method* and *Moment method*.

*Rescaled Adjusted Range (R/S) statistics* is one of the better known methods, introduced by the British hydrologist H. E. Hurst. Let  $R(n)$  be the range of the data aggregated over blocks of length  $n$  and  $S^2(n)$  be the sample variance of data aggregated at the same scale. The R/S statistics is the ratio:

$$\frac{R}{S}(n) = S^{-1}(n) \left[ \max_{0 \leq t \leq n} (X(t) - t\bar{X}(n)) - \min_{0 \leq t \leq n} (X(t) - t\bar{X}(n)) \right].$$

The parameter  $H$  can be estimated by fitting a line to the points in the  $\frac{R}{S}(n)$  vs.  $n$  in a log-log scale plot (pox plot).

*Higuchi's method* involves calculating the length of a path and, in principle, finding its fractal dimension  $D$ . Denoting by  $N$  the number of samples of the

time series, the normalized length of the corresponding curve is estimated as follows:

$$L(m) = \frac{N-1}{m^3} \sum_{i=1}^m \left[ \frac{N-i}{m} \right]^{-1} \sum_{k=0}^{\lfloor (N-i)/m \rfloor - 1} |Y_{i+km} - Y_{i+(k-1)m}|,$$

where  $m$  is essentially a block size and  $\lfloor \cdot \rfloor$  denotes the greatest integer function. A log-log plot of  $L(m)$  vs.  $m$  should produce a straight line with a slope of  $D = 2 - H$ .

*Moment method* investigates self-similarity through the behavior of absolute moments of the aggregated processes  $X^{(m)}$ :

$$\mu^{(m)}(q) = \mathbb{E} |X^{(m)}|^q.$$

According to the definition of self-similarity,  $\mu^{(m)}(q)$  is proportional to  $m^{\beta(q)}$ , where:

$$\beta(q) = q(H - 1).$$

As far as the analysis in a transformed domain is concerned, the main approaches are the following:

*Periodogram method* is based on the estimation of the spectral density through the periodogram:

$$I(\omega) = \frac{1}{2\pi N} \left| \sum_{j=0}^{N-1} X_j e^{ij\omega} \right|^2.$$

By definition, a series with LRD should have a periodogram which is proportional to  $\omega^{1-2H}$  close to the origin. Therefore, a regression of the logarithm of the periodogram on the logarithm of  $\omega$  should give a coefficient of  $1 - 2H$ , hence providing an estimation of the parameter  $H$ .

*Whittle estimator*, also based on the periodogram, involves the function:

$$Q = \int_{-\pi}^{\pi} \frac{I(\omega)}{f(\omega; \eta)} d\omega?$$

Where  $f(\omega; \eta)$  is the spectral density and  $\eta$  denotes the vector of unknown parameters. The Whittle estimator is the value of  $\eta$  which minimizes the function  $Q$ . Unlike the other estimators discussed here, the Whittle estimator is obtained through a non-graphical method. It also assumes that the parametric form of the spectral density, i.e. the function  $f(\omega; \eta)$ , is known.

*Wavelet estimator* is based on the analysis of the numerical series in the wavelet domain, taking advantages of the fact that the wavelet coefficients  $d_j$  at a given resolution level  $j$  are quasi-decorrelated; indeed, the mother wavelet  $\psi(t)$  is a band-pass signal and the power-law behavior of its Fourier transform at frequencies near 0 cancels the power-law divergence of the spectrum at the origin. Thanks to the quasi-decorrelation of coefficients  $d_j$ , a useful spectral es-

timator can be designed by performing a time average of the  $|d_{j,k}|^2$  at a given scale, that is,

$$\widehat{S}(2^{-j} \omega_0) = \frac{1}{n_j} \sum_k |d_{j,k}|^2$$

where  $n_j = 2^{-j} N$  is the available number of wavelet coefficients at octave  $j$ . In case of  $\frac{1}{|\omega|^\beta}$  processes, it is possible to design an estimator  $\widehat{H}$  for the parameter

$H$  from a simple linear regression of  $\log_2(\widehat{S}(2^{-j} \omega_0))$  on  $j$ , i.e.

$$\log_2(\widehat{S}(2^{-j} \omega_0)) = (2\widehat{H} - 1)j + c,$$

where the constant  $c$  is independent of the analyzing scale  $j$ .

Although the Hurst parameter is well defined mathematically, it is problematic to measure it properly [5]. There are several methods to estimate the Hurst parameter but they often produce conflicting results [6, 7].

We have used several methods to estimate Hurst parameter in real traces collected in our laboratory in IITiS PAN. Table 1 gives the obtained Hurst parameters for one day IITiS traces: trace 1 (6 002 874 samples), trace 2 (13 874 610 samples), trace 3 (36 135 490 samples). Our previous work [4] did not confirm the relationship between the degree of LRD and the number of transmitted packet of a given type. One can see a variance of estimators obtained with different methods.

### Markov LRD models

Markov chains and Markov-modulated processes (MMP) are well-known modeling techniques which are successful in wide variety of fields. These models are often motivated by the idea of capturing the long-range dependence (LRD) which is seen in real internet traffic and replicating the the Hurst parameter  $H$  which characterizes LRD [5].

Two-state *Markov Modulated Poisson Process* (MMPP) is also known as the Switched Poisson Process (SPP). The superposition of MMPP's is also an MMPP which is a special case of Markovian Arrival Process (MAP).

Following the model proposed in [8], a LRD process (used in our study) can be modeled as the superposition of  $d$  two-state MMPPs [9]. The  $i$ -th MMPP ( $1 \leq i \leq d$ ) can be parameterized by two square matrices:

$$\mathbf{D}_0^i = \begin{bmatrix} -(c_{1i} + \lambda_{1i}) & c_{1i} \\ c_{2i} & -(c_{2i} + \lambda_{2i}) \end{bmatrix}, \quad \mathbf{D}_1^i = \begin{bmatrix} \lambda_{1i} & 0 \\ 0 & \lambda_{2i} \end{bmatrix}.$$

Fitting algorithm [8,9] requires five input parameters. Three of them:  $\lambda^*$  (the mean rate of the process),  $\rho$  (lag 1 correlation) and the Hurst parameter  $H$  should be estimated from the real data traces. This fitting procedure was applied

to the traces of IP traffic measured at IITiS. The superposition of four MMPP's is sufficient to model asymptotic second-order self-similarity of the counting process over five time-scales.

*Hidden Markov Model* (HMM) may be seen as a probabilistic function of a (hidden) Markov chain [10, 11]. This Markov chain is composed of two variables: the hidden-state variable, whose temporal evolution follows a Markov-chain behavior ( $x_n \in \{s_1, \dots, s_N\}$  represent the (hidden) state at discrete time  $n$  with  $N$  being the number of states) and the observe variable which stochastically depends on the hidden state ( $y_n \in \{o_1, \dots, o_M\}$ ) and represents the observable at discrete time  $n$  with  $M$  being the number of observables.

Given a sequence of observable variables  $y = (y_1, y_2, \dots, y_L)$  referred to as the *training sequence*, we want to find the set of parameters such that the likelihood of the model  $L(\mathbf{y}; \lambda) = \Pr(\mathbf{y} | \lambda)$  is maximum. We solved it via the Baum-Welch algorithm, a special case of the Expectation-Maximization algorithm, that iteratively updates the parameters in order to find a local maximum point of the parameter set.

We used the well-known Bellcore trace of Internet traffic: OctExt.TL. Each line of this file contains a floating-point time stamp (representing the time in seconds since the start of a trace) and an integer length (representing the Ethernet data length in bytes). We translated the sequence of time stamps into the sequence of inter-arrival times. Then we apply a scheme using *Vector Quantization* (VQ) to translate the obtained sequence of inter-arrival times into a sequence of symbols, and training a HMM for this sequence. The quantization algorithm used is Linde-Buzo-Gray (LBG) algorithm of VQ. Vector Quantization is a clustering technique commonly used in compression, image recognition and stream encoding. It is the general approach to map a space of vector valued data to a finite set of distinct symbols, in a way to minimize distortion associated with this mapping. We consider a HMM in which the state and the observable variables are discrete.

## Conclusions

The article represents a practical approach to the problem of modeling stochastic features of Internet traffic. We summarize the main methods to determine the Hurst parameter and to include it in traffic models. In particular we show that we are able to construct efficiently Markov models of traffic with LRD, also with the use of Hidden Markov Chains. These Markov models may be a part of a computer network models aiming to evaluate its performance. Of course, the complexity of traffic models enlarges the size of the entire state space to be considered and hence the number of equations to be solved numerically. Therefore we are developing a software tool able to cope with models having hundreds of millions states.

## References

1. Lopez-Ardao J. C., Lopez-Garcia C., Suarez-Gonzalez A. et al. On the use of self-similar processes in network simulation // *ACM Transactions on Modeling and Computer Simulation*. 2000. Vol. 10, issue 2. P. 125–151.
2. Gong W.-B., Liu Y., Misra V., Towsley D. Self-similarity and long range dependence on the internet: a second look at the evidence, origins and implications // *Computer Networks*. 2005. Vol. 48. P. 377–399.
3. Beran J. *Statistics for Long-Memory Processes*. Chapman and Hall, 1994.
4. Domańska J., Domański A., Czachórski T. A Few Investigation of Long-Range Dependence in Network Traffic // *Information Science and Systems*. 2004 September. P. 137–144. Springer International Publishing.
5. Clegg R. G. A practical guide to measuring the Hurst parameter // *International Journal of Simulation*. 2006. Vol. 7, issue 2. P. 3–14.
6. Karagiannis T., Molle M., Faloutsos M. Long-Range Dependence: Ten Years of Internet Traffic Modeling // *IEEE Internet Computing*. 2004. Vol. 8, issue 5. P. 57–64.
7. Domańska J., Domański A., Czachórski T. Estimating the intensity of Long-Range Dependence in real and synthetic traffic traces // *Communications in Computer and Information Science*. 2015. Vol. 522. P. 11–22. Springer International Publishing.
8. Andersen A. T., Nielsen B. F. A Markovian Approach for Modeling Packet Traffic with Long-Range Dependence // *IEEE Journal on Selected Areas in Communications*. Vol. 16, no. 5, 1998.
9. Domańska J., Domański A., Czachórski T. Modeling packet traffic with the use of superpositions of two-state MMPPs // *Communications in Computer and Information Science*. 2014. Vol. 431. P. 24–36. Springer International Publishing.
10. Rabiner L. A tutorial on hidden Markov models and selected applications in speech recognition // *Proceedings of the IEEE*. 1989. Vol. 77, issue 2. P. 257–285.
11. Domańska J., Domański A., Czachórski T. Internet traffic source based on hidden Markov model // *Lecture Notes in Computer Science*. 2011. Vol. 6869. P. 395–404. Springer Berlin Heidelberg.

## THE ANALYSIS OF HEAT FLUXES VARIABILITY BY MOVING SEPARATION OF FINITE NORMAL MIXTURES

**S. K. Gulev<sup>1,2</sup>, V. Yu. Korolev<sup>2,3</sup>, A. K. Gorshenin<sup>3,4</sup>, K. P. Belyaev<sup>1,2</sup>**

<sup>1</sup>*P. P. Shirshov Institute of Oceanology, Moscow, Russia*

<sup>2</sup>*Lomonosov Moscow State University, Russia*

<sup>3</sup>*Federal Research Center “Computer Science and Control”  
of the Russian Academy of Sciences, Moscow, Russia*

<sup>4</sup>*Moscow State University of Information Technologies, Radioengineering and  
Electronics, Russia*

The statistical regularities in the behavior of sensible and latent turbulent heat fluxes recomputed from 6-hourly NCEP-NCAR for the period from 1948–2008 in Atlantic are analyzed. It is proposed to represent these regularities by probability distributions that are mixtures of several normal (Gaussian) laws with parameters varying in time. The method of moving separation of mixtures

is used to obtain the values of the parameters of the mixtures. This approach allows to analyze the regularities in the variation of the parameters and, hence, to capture the low-term variability which can be considered as a trend and high-term dynamics associated with diffusion or irregular variability. The weights of each of term are estimated and analyzed. Their spatial-temporal structure is mapped and investigated. The paper shows that this structure has general properties in entire Atlantic but also possesses specific features in each of Atlantic region. The analysis of the parameters identifies the areas where surface turbulent fluxes are determined by different shapes of probability density functions. The estimated extreme turbulent heat fluxes amount to 1500–2000 W/m<sup>2</sup> (for the 99<sup>th</sup> percentile) and can exceed 2000 W/m<sup>2</sup> for higher percentiles in the subpolar latitudes and western boundary current regions. The analysis of linear trends and inter-annual variability for mean and extreme fluxes shows that the strongest trends in extreme fluxes (greater than 15 W/m<sup>2</sup> per decade) in the western boundary current regions are associated with the strong diffusion parameters. In many regions the changes in extreme fluxes may be different from those for the mean fluxes at inter-annual and decadal time scales.

### **Introduction**

Surface turbulent air-sea fluxes are extremely important characteristic in geophysics. There are several data sources where it is possible to obtain the reanalyzed data of explicit and latent heat-fluxes. Surface fluxes and flux-related variables from the global reanalyzes are available also with 6-hourly resolution. The detailed assessment of heat flux characteristic, for instance, to detect and evaluate extreme fluxes requires accurate knowledge of entire distribution of turbulent heat fluxes and their variability in time.

The properties of probability distributions of turbulent fluxes are vital for assessing confident error bounds, analyzing averaged fluxes and comparing the observed fluxes with climate model simulations. Meanwhile, the knowledge of distribution of surface turbulent fluxes is necessary for the accurate analysis of cross-covariance scales in surface fluxes and for estimation of extreme turbulent fluxes which are strongly depended on space and time. The extreme turbulent fluxes reach and even exceed 1000–2000 W/m<sup>2</sup> for the mid- and subpolar latitudes particularly in the western boundary current regions.

An attempt to describe the local properties of turbulent both explicit and latent heat fluxes has been made in [1]. This paper proposed the so-called Fisher-Tippet (FT) distribution for heat-fluxes behavior; its parameters have been correctly evaluated and the goodness-of-fit has been successfully tested thus proving their adequacy. This idea allowed evaluating the spatial-temporal behavior of FT distribution parameters, linking the sought parameters with physical processes in tropical and mid-latitude air-sea interaction and numerically estimating the probabilities of extreme heat-fluxes in each of considered region.

In the present work the original idea of generalizations of time-series representation is exploited. The scheme is based on the representation of the probabil-

ity distributions of the heat-fluxes increments, i.e., one time-step differences as a mixture of several normal (Gaussian) distributions with different parameters depending on time and the corresponding weights also depending on time. The method of moving separation of mixtures is used to obtain the values of the parameters of the mixtures. This approach allows to analyze the regularities in the variation of the parameters and, hence, to capture the low-term variability which can be considered as a trend and high-term dynamics associated with diffusion or irregular variability.

### **Homogenization of the data**

Traditionally, for the statistical analysis of the stochastic regularities in the behavior of the observed time series, all the available data is used without any prior transformation. For example, in [1] a model was proposed for the probability distribution of for surface air-sea turbulent heat fluxes over the global ocean. This models successfully captures all the peculiarities of the untransformed data. However, it can hardly be used for matching the evolution of the parameters of the fluxes in time. The thing is that the sample used for the statistical analysis is not homogeneous (and possibly, not independent).

So, to avoid the negative effect of the “historical” background on the statistical regularities in the behavior of the analyzed time series, in this work we consider the time series constructed by the increments of the initially observed time series of the heat-fluxes measurements.

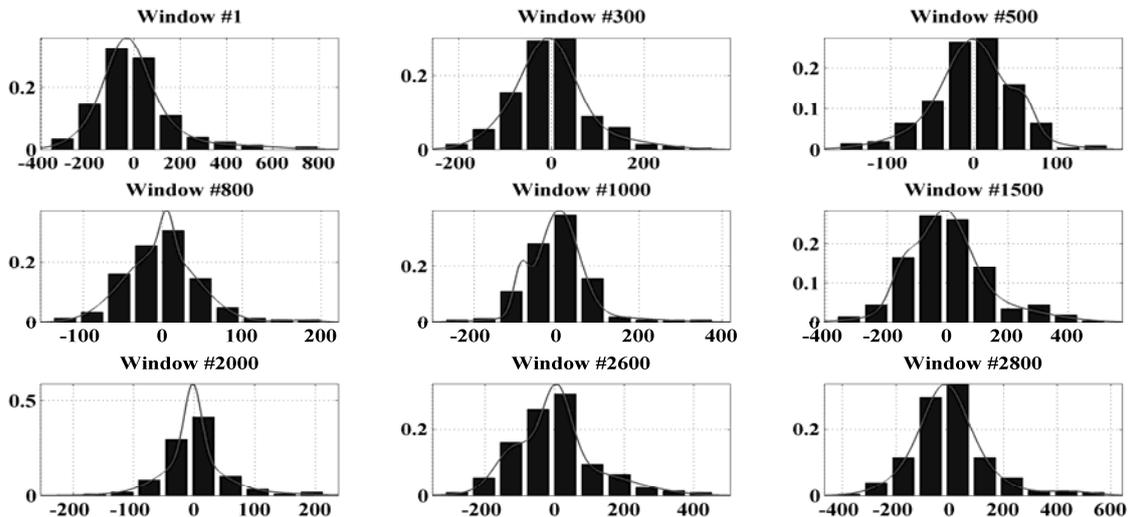
### **The method of moving separation of finite normal mixtures**

To reveal the changes of the structure of the observed stochastic processes in time, the so-called method of moving separation of mixtures (MSM method) is successfully used. This method was proposed in [2]. As examples of efficient performance of this method, the papers [3–5] should be mentioned containing the results for the financial markets, for the traffic in information systems and for the plasma turbulence correspondingly. The key point in this method is that the volatility of the process is decomposed into two components, dynamical and diffusive.

Within the framework of this method, the one-dimensional distributions of the increments of the basic process are approximated by finite location-scale mixtures of normal distributions. The theoretical background of these models is based on that finite normal mixtures are convenient approximations to general location-scale normal mixtures or normal variance-mixtures which are limit laws for the distributions of sums of a random number of independent random variables or non-homogeneous and non-stationary random walks, see details in, say, [2, 6, 7].

To analyze the dynamics of the changes in the stochastic process, the problem of statistical estimation of unknown parameters of distributions should be successively solved for a part of sample which moves in a direction of astronomical time (i.e., the initial sample is divided into sliding or moving sub-samples often called windows). Typically, the window (sub-sample) size is

fixed. Once the analyzed parameters are obtained for a current location, the window should be moved by one element of the initial sample (i.e., the method will analyze the next sub-sample). This allows to detect all possible changes in the behavior of components. As this is so, the results are smoothed since they are in some sense averaged within a window and vary from window to window. The parameters of model noticeably depend on time, as can be seen on Fig. 1 where the histograms constructed from different windows of width 200 and the densities corresponding to the finite normal mixture with the parameters estimated from the corresponding windows are presented.



*Fig. 1. The histograms constructed from different windows and the fitted finite normal mixture densities*

### Estimations of the parameters

To estimate the parameters of model, at each window the classical EM algorithm was used. The EM algorithm is an iterative numerical procedure for the maximization of the multi-parameter likelihood functions. It was suggested in [8]. Although very many modifications of this algorithm have already been proposed, the classical EM algorithm remains to be the most reliable tool for the estimation of the parameters of the finite normal mixture model.

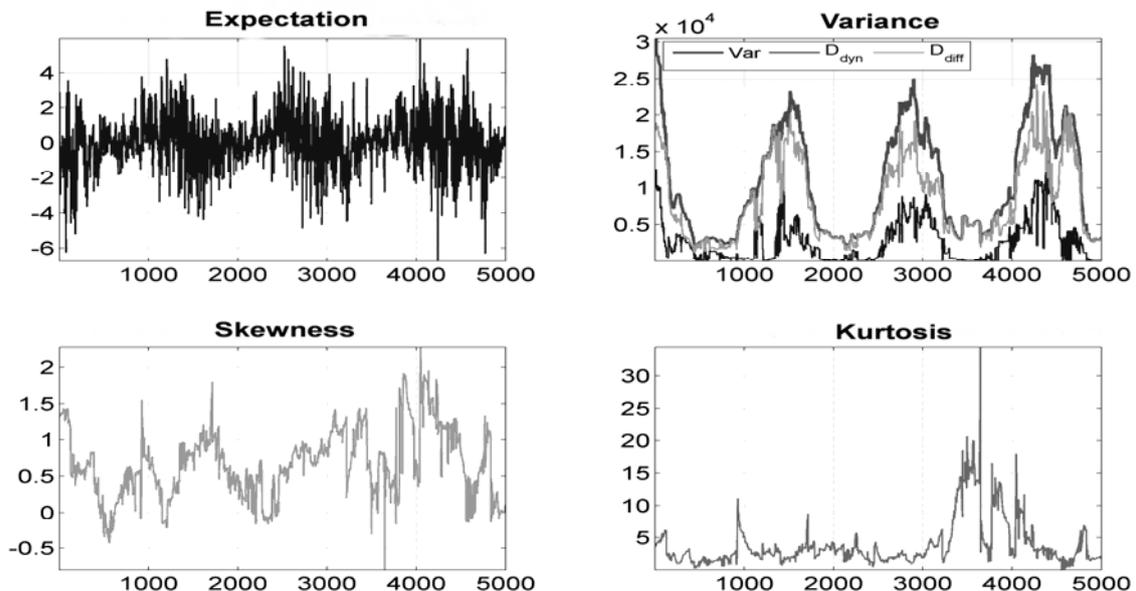
Nevertheless, in this case this algorithm has some important drawbacks. Perhaps, the main of them is its instability with respect to the initial approximation. The finite normal mixture model likelihood function has a very non-smooth surface with peaks or rills. Therefore, being a “greedy” algorithm, the EM algorithm converges to the local maximum which is closest to the starting point. The usual ways to treat this obstacle are following:

- to choose the starting point at random;
- to choose several starting points and to average the results over the number of runs;

– to choose several starting points and to use the estimate that delivers the maximum value of the target likelihood function among the runs as the result.

The third version of the EM algorithm with repeated random choice of starting points and maximization of the results over runs at each window gives most clear results.

This algorithm was applied to the analysis of temporal variation of the parameters of the distribution of increments of the heat-flux process. On Fig. 2 it is clearly seen that the expected value of the increment noticeably oscillates in time with periodically changing amplitudes. Furthermore, at each period, the amplitudes grow slower than they decrease. The periodical variation of the variance is clearly seen. It is very interesting (if not surprising) that the purely stochastic diffusive component of the variance makes greater contribution to the total variance than the dynamic component due to systematic trends. It is also interesting that the distribution of the increments is slightly asymmetric with right slope heavier than the left one. Another interesting observation is that the kurtosis of this distribution is maximum during the “calm” period.



*Fig. 2. The moment characteristics of the probability distribution of the increments of the heat-flux process*

## Conclusions

In the paper, the method of moving separation of mixtures was applied to the analysis of statistical regularities in the temporal evolution of heat-fluxes. This method was realized by a special version of the EM algorithm aimed at the maximization of the likelihood function within the class of finite normal mixture models. In the stochastic character of the evolution of heat-fluxes, one basic component with low variance can be identified accompanying by stochastically emerging and disappearing components with large variance. Some regularities in

the temporal variation of the moment characteristics of the heat-flux process increments were observed.

*The research is supported by the Russian Foundation for Basic Research (project-15-37-20851) and by the President Grant for Government Support of Young Russian Scientists (project MK-4103.2014.9).*

#### References

1. Gulev S. K., Belyaev K. P. Probability distribution characteristics for surface air-sea turbulent heat fluxes over the global ocean // *Journal of Climate*. 2012. No. 25,1(1). P. 184–206.
2. Korolev V. Yu. Probabilistic and statistical methods of decomposition of volatility of chaotic processes. 2011. Moscow: Moscow University Publishing House.
3. Gorshenin A. K., Korolev V. Yu., Tursunbayev A. M. Median modification of EM- and SEM-algorithms for separation of mixtures of probability distributions and their application to the decomposition of volatility of financial time series // *Informatics and Applications*. 2008. No. 2(4). P. 12–47.
4. Gorshenin A., Korolev V. Modeling of statistical fluctuations of information flows by mixtures of gamma distributions // *Proceedings of 27<sup>th</sup> European Conference on Modelling and Simulation*, 2013. Digitaldruck Pirrot GmbH, Dudweiler, Germany. P. 569–572.
5. Gorshenin A. K. The information technology to research the fine structure of chaotic processes in plasma by the analysis of the spectra // *Systems and Means of Informatics*. 2014. No. 24(1). P. 116–125.
6. Korolev V. Yu. Generalized hyperbolic laws as limit distributions for random sums // *Theory Probability and Its Applications*. 2014. No. 58(1). P. 63–75.
7. Korolev V. Yu., Zaks L. M. Generalized variance gamma distributions as limit laws for random sums // *Informatics and Applications*. 2013. No. 7(1). P. 105–115.
8. Dempster A., Laird N., Rubin D.: Maximum likelihood estimation from incomplete data // *Journal of the Royal Statistical Society. Series B*. 1977. No. 39(1). P. 1–38.

### **СОВМЕСТНАЯ ПЛОТНОСТЬ ВЕРОЯТНОСТЕЙ ДЛИТЕЛЬНОСТЕЙ ИНТЕРВАЛОВ МОДУЛИРОВАННОГО ОБОБЩЕННОГО ПОЛУСИНХРОННОГО ПОТОКА СОБЫТИЙ В УСЛОВИЯХ ЕГО НЕПОЛНОЙ НАБЛЮДАЕМОСТИ**

*М. А. Бахолдина, А. М. Горцев*

*Национальный исследовательский*

*Томский государственный университет, Томск, Россия*

Настоящая работа является продолжением исследований, приведенных в [1, 2], где решается задача оптимальной оценки состояний модулированного обобщенного полусинхронного потока событий, относящегося к классу дважды стохастических потоков событий (DSPPs). В настоящей работе находится явный вид совместной плотности вероятностей значений длительности интервалов между моментами наступления соседних событий потока в условиях его неполной наблюдаемости.

## 1. Постановка задачи

Интенсивность рассматриваемого потока событий представляет собой кусочно-постоянный стационарный случайный процесс  $\lambda(t)$  с двумя состояниями 1, 2:  $\lambda(t) = \lambda_1$  либо  $\lambda(t) = \lambda_2$  ( $\lambda_1 > \lambda_2 \geq 0$ ). Длительность пребывания процесса  $\lambda(t)$  (потока) в первом (во втором) состоянии распределена по экспоненциальному закону с параметром  $\beta$  ( $\alpha$ ). В течение временного интервала случайной длительности, когда  $\lambda(t) = \lambda_i$ , имеет место пуассоновский поток событий с интенсивностью  $\lambda_i$ ,  $i = 1, 2$ . Кроме того, переход из первого состояния процесса  $\lambda(t)$  во второе возможен в момент наступления события пуассоновского потока интенсивности  $\lambda_1$ ; переход осуществляется с вероятностью  $p$  ( $0 \leq p \leq 1$ ); с вероятностью  $1 - p$  процесс  $\lambda(t)$  остается в первом состоянии. Переход из второго состояния процесса  $\lambda(t)$  в первое в момент наступления события пуассоновского потока интенсивности  $\lambda_2$  невозможен. В момент окончания второго состояния процесса  $\lambda(t)$  при его переходе из второго состояния в первое инициируется с вероятностью  $\delta$  ( $0 \leq \delta \leq 1$ ) дополнительное событие в первом состоянии. В сделанных предположениях  $\lambda(t)$  – марковский процесс. Матрицы инфинитезимальных характеристик принимают вид

$$D_0 = \begin{vmatrix} -(\lambda_1 + \beta) & \beta \\ (1 - \delta)\alpha & -(\lambda_2 + \alpha) \end{vmatrix}, \quad D_1 = \begin{vmatrix} (1 - p)\lambda_1 & p\lambda_1 \\ \delta\alpha & \lambda_2 \end{vmatrix}.$$

Элементами матрицы  $D_1$  являются интенсивности переходов процесса  $\lambda(t)$  из состояния в состояние с наступлением события. Недиagonальные элементы матрицы  $D_0$  – это интенсивности переходов из состояния в состояние без наступления события. Диагональные элементы матрицы  $D_0$  – это интенсивности выхода процесса  $\lambda(t)$  из своих состояний, взятые с противоположным знаком.

После каждого зарегистрированного в момент времени  $t_k$  события наступает период мертвого времени фиксированной длительности  $T$ , в течение которого другие события потока недоступны наблюдению. По окончании периода мертвого времени первое наступившее событие снова создает период мертвого времени длительности  $T$  и т. д. (непродлевающееся мертвое время). Вариант возникающей ситуации приведен на рис. 1, где 1, 2 – состояния процесса  $\lambda(t)$ ; дополнительные события помечены буквами  $\delta$ ; периоды мертвого времени длительности  $T$  помечены штриховкой; ненаблюдаемые события отображены черными кружками, наблюдаемые  $t_1, t_2, \dots$  – белыми.

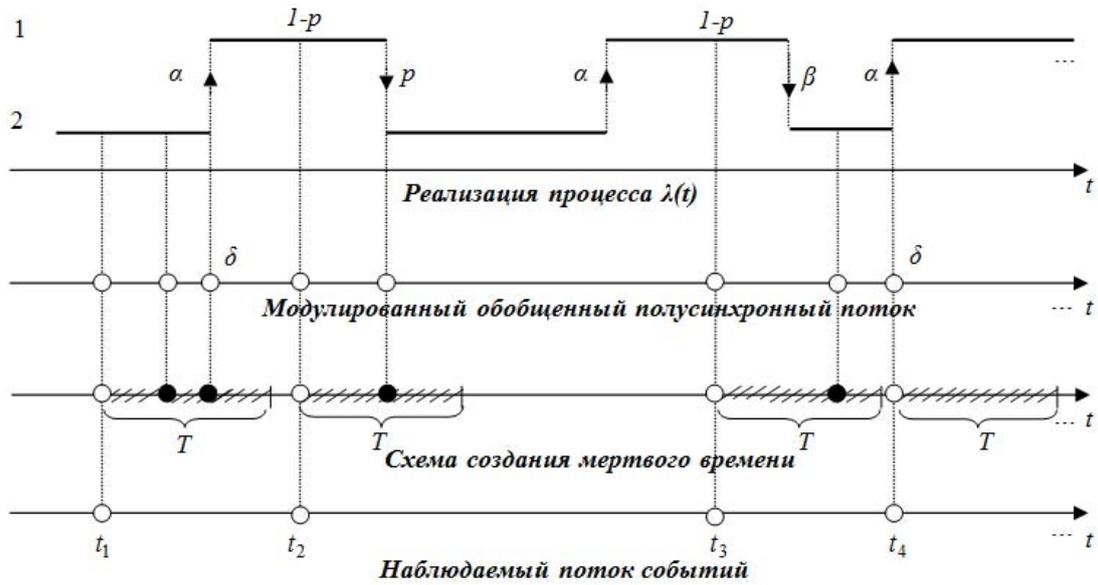


Рис. 1. Формирование наблюдаемого потока событий

Отметим, что процесс  $\lambda(t)$  является принципиально ненаблюдаемым (скрытый марковский процесс), а наблюдаемыми являются только временные моменты наступления событий наблюдаемого потока  $t_1, t_2, \dots$ . Рассматривается установившийся (стационарный) режим функционирования потока событий. В силу предпосылок последовательность моментов наступления событий  $t_1, t_2, \dots, t_k, \dots$  образует вложенную цепь Маркова  $\{\lambda(t_k)\}$ , т.е. поток обладает марковским свойством, если его эволюцию рассматривать с момента наступления события  $t_k, k = 1, 2, \dots$ .

Обозначим через  $\tau_k = t_{k+1} - t_k, k = 1, 2, \dots$ , значение длительности  $k$ -го интервала между соседними событиями наблюдаемого потока. Так как рассматривается стационарный режим, то плотность вероятностей значений длительности  $k$ -го интервала есть  $p_T(\tau_k) = p_T(\tau), \tau \geq 0$ , для любого  $k$  (индекс  $T$  подчеркивает, что плотность вероятностей зависит от длительности мертвого времени). В силу этого момент времени  $t_k$  без потери общности можно положить равным нулю, т.е. момент наступления события есть  $\tau = 0$ . Пусть теперь  $(t_k, t_{k+1}), (t_{k+1}, t_{k+2})$  – два смежных интервала с соответствующими значениями длительностей  $\tau_k = t_{k+1} - t_k, \tau_{k+1} = t_{k+2} - t_{k+1}$ . Их расположение на временной оси, в силу стационарности потока, произвольно. Тогда можно рассмотреть соседние интервалы  $(t_1, t_2), (t_2, t_3)$  с соответствующими значениями длительностей  $\tau_1 = t_2 - t_1, \tau_2 = t_3 - t_2; \tau_1 \geq 0, \tau_2 \geq 0$ . При этом  $\tau_1 = 0$  соответствует моменту  $t_1$  наступления события наблюдаемого потока;  $\tau_2 = 0$  соответствует моменту  $t_2$  наступления события наблюдаемого потока. Соответствующая совместная плотность вероятностей при этом есть  $p_T(\tau_1, \tau_2), \tau_1 \geq 0, \tau_2 \geq 0$ .

## 2. Вывод совместной плотности вероятностей $p_T(\tau_1, \tau_2)$

Рассмотрим интервал времени  $(0, \tau)$  между соседними событиями в наблюдаемом потоке. Значение длительности данного интервала есть  $\tau = T + t$ , где  $t$  – значение длительности интервала между моментом окончания периода мертвого времени и следующим событием наблюдаемого потока ( $t \geq 0$ ). Пусть  $p_{jk}(t)$  есть условная вероятность того, что на интервале  $(0, t)$  нет событий наблюдаемого потока и  $\lambda(t) = \lambda_k$  при условии, что в момент времени  $t = 0$  значение процесса  $\lambda(t)$  есть  $\lambda(0) = \lambda_j$ ,  $j, k = 1, 2$ . Соответствующую этой вероятности плотность вероятностей обозначим через  $\tilde{p}_{jk}(t)$ ,  $j, k = 1, 2$ . Введем в рассмотрение переходную вероятность  $q_{ij}(T)$  – вероятность того, что за мертвое время длительности  $T$  процесс  $\lambda(\tau)$  перейдет из состояния  $i$  (момент времени  $\tau = 0$ ) в состояние  $j$  (момент времени  $\tau = T$ ),  $i, j = 1, 2$ , и вероятность  $\pi_i(0|T)$  – условную вероятность того, что процесс  $\lambda(\tau)$  в момент времени  $\tau = 0$  находится в состоянии  $i$  ( $i = 1, 2$ ) при условии, что в этот момент времени наступило событие наблюдаемого потока, розыгрыш состояний произошел, и наступил период мертвого времени длительности  $T$ .

Пусть теперь  $\tau_1 = T + t^{(1)}$ ,  $\tau_2 = T + t^{(2)}$  – значения длительностей двух смежных интервалов между моментами наступления последовательных событий наблюдаемого потока, при этом  $\tau_1 = 0$  – момент наступления первого события,  $\tau_2 = 0$  – момент наступления второго события. В силу того, что последовательность моментов наступления событий наблюдаемого потока образует вложенную цепь Маркова, совместная плотность вероятностей  $p_T(\tau_1, \tau_2)$  принимает вид

$$p_T(\tau_1, \tau_2) = \begin{cases} 0, & 0 \leq \tau_1 < T, \quad 0 \leq \tau_2 < T, \\ \sum_{i=1}^2 \pi_i(0|T) \sum_{j=1}^2 q_{ij}(T) \sum_{k=1}^2 \tilde{p}_{jk}(\tau_1 - T) \sum_{s=1}^2 q_{ks}(T) \sum_{n=1}^2 \tilde{p}_{sn}(\tau_2 - T), & \\ \tau_1 \geq T, \quad \tau_2 \geq T, & \end{cases} \quad (1)$$

где  $\tilde{p}_{jk}(\tau_1 - T) = \tilde{p}_{jk}(t^{(1)})$ ,  $\tilde{p}_{sn}(\tau_2 - T) = \tilde{p}_{sn}(t^{(2)})$ .

Рассмотрим далее разность  $p_T(t^{(1)}, t^{(2)}) - p_T(t^{(1)})p_T(t^{(2)})$ ,  $t^{(1)} \geq 0$ ,  $t^{(2)} \geq 0$ . Отметим, что выражение для одномерной плотности вероятностей  $p_T(\tau)$ , найденное ранее [3], принимает вид

$$p_T(\tau) = \begin{cases} 0, & 0 \leq \tau < T, \\ \gamma(T)z_1 e^{-z_1(\tau-T)} + (1 - \gamma(T))z_2 e^{-z_2(\tau-T)}, & \tau \geq T, \end{cases} \quad (2)$$

$$\gamma(T) = \frac{1}{z_2 - z_1} [z_2 - \lambda_1 + (\lambda_1 - \lambda_2 - \alpha\delta)\pi_2(T)],$$

где

$$z_{1,2} = \frac{1}{2} \left( \lambda_1 + \lambda_2 + \alpha + \beta \mp \sqrt{(\lambda_1 - \lambda_2 - \alpha + \beta)^2 + 4\alpha\beta(1 - \delta)} \right), \quad 0 < z_1 < z_2, \quad (3)$$

$$\pi_1(T) = \pi_1 + [\pi_2 - \pi_2(0|T)]e^{-(p\lambda_1 + \beta + \alpha)T}, \quad (4)$$

$$\pi_2(T) = \pi_2 - [\pi_2 - \pi_2(0|T)]e^{-(p\lambda_1 + \beta + \alpha)T},$$

$$\pi_1(0|T) = \frac{\alpha[\lambda_1(1 - p + p\delta) + \delta\beta] + \lambda_1\pi_1[\lambda_2 - p(\lambda_2 + \alpha\delta)] [1 - e^{-(p\lambda_1 + \beta + \alpha)T}]}{\lambda_1\alpha + (p\lambda_1 + \beta)(\lambda_2 + \alpha\delta) + \lambda_1[\lambda_2 - p(\lambda_2 + \alpha\delta)] [1 - e^{-(p\lambda_1 + \beta + \alpha)T}]}, \quad (5)$$

$$\pi_2(0|T) = \frac{p\lambda_1(\lambda_2 + \alpha) + \lambda_2\beta + \lambda_1\pi_2[\lambda_2 - p(\lambda_2 + \alpha\delta)] [1 - e^{-(p\lambda_1 + \beta + \alpha)T}]}{\lambda_1\alpha + (p\lambda_1 + \beta)(\lambda_2 + \alpha\delta) + \lambda_1[\lambda_2 - p(\lambda_2 + \alpha\delta)] [1 - e^{-(p\lambda_1 + \beta + \alpha)T}]},$$

$$\pi_1 = \frac{\alpha}{p\lambda_1 + \beta + \alpha}, \quad \pi_2 = \frac{p\lambda_1 + \beta}{p\lambda_1 + \beta + \alpha}.$$

Принимая во внимание выражения

$$\sum_{i=1}^2 \pi_i(0|T) \sum_{j=1}^2 q_{ij}(T) = \sum_{i=1}^2 \pi_i(T) \quad \text{и} \quad p_T(\tau) = \sum_{i=1}^2 \pi_i(T) \sum_{j=1}^2 \tilde{p}_{ij}(\tau - T), \quad \tau \geq T,$$

можно показать, что

$$\begin{aligned} p_T(t^{(1)}, t^{(2)}) - p_T(t^{(1)})p_T(t^{(2)}) &= e^{-(p\lambda_1 + \beta + \alpha)T} \sum_{k=1}^2 (\tilde{p}_{1k}(t^{(2)}) - \tilde{p}_{2k}(t^{(2)})) \times \\ &\times \{ \pi_1(T)[\pi_2(0|T)\tilde{p}_{11}(t^{(1)}) - \pi_1(0|T)\tilde{p}_{12}(t^{(1)})] - \pi_2(T)[\pi_1(0|T) \times \\ &\times \tilde{p}_{22}(t^{(1)}) - \pi_2(0|T)\tilde{p}_{21}(t^{(1)})] \}. \end{aligned} \quad (6)$$

Выполняя дальнейшие преобразования, находим

$$\sum_{k=1}^2 (\tilde{p}_{1k}(t^{(2)}) - \tilde{p}_{2k}(t^{(2)})) = -\frac{\lambda_1 - \lambda_2 - \alpha\delta}{z_2 - z_1} [z_1 e^{-z_1 t^{(2)}} - z_2 e^{-z_2 t^{(2)}}], \quad (7)$$

$$\begin{aligned} \pi_1(T)[\pi_2(0|T)\tilde{p}_{11}(t^{(1)}) - \pi_1(0|T)\tilde{p}_{12}(t^{(1)})] - \pi_2(T)[\pi_1(0|T)\tilde{p}_{22}(t^{(1)}) - \\ - \pi_2(0|T)\tilde{p}_{21}(t^{(1)})] &= -\frac{1}{z_2 - z_1} [z_1 e^{-z_1 t^{(1)}} - z_2 e^{-z_2 t^{(1)}}] [\lambda_1(\pi_2(0|T) - p)\pi_1(T) + \\ &+ (\pi_2(0|T)\alpha\delta - \pi_1(0|T)\lambda_2)\pi_2(T)]. \end{aligned} \quad (8)$$

Тогда, подставляя в (6) сначала (7), (8), далее  $\tilde{p}_{jk}(t^{(1)})$ ,  $j, k = 1, 2$ ,  $\pi_i(0|T)$ ,  $i = 1, 2$ , и, наконец,  $\pi_i(T)$ ,  $i = 1, 2$ , найденные в [3], выполняя при этом достаточно трудоемкие преобразования, находим искомое выражение для совместной плотности вероятностей

$$p_T(\tau_1, \tau_2) = 0, \quad 0 \leq \tau_1 < T, \quad 0 \leq \tau_2 < T,$$

$$\begin{aligned} p_T(\tau_1, \tau_2) &= p_T(\tau_1)p_T(\tau_2) + e^{-(p\lambda_1 + \beta + \alpha)T} \gamma(T)[1 - \gamma(T)] \frac{\lambda_1[\lambda_2 - p(\lambda_2 + \alpha\delta)]}{z_1 z_2} \times \\ &\times [z_1 e^{-z_1(\tau_1 - T)} - z_2 e^{-z_2(\tau_1 - T)}] [z_1 e^{-z_1(\tau_2 - T)} - z_2 e^{-z_2(\tau_2 - T)}], \quad \tau_1 \geq T, \quad \tau_2 \geq T, \end{aligned}$$

где  $z_1 z_2 = \lambda_1 \lambda_2 + \lambda_1 \alpha + \lambda_2 \beta + \alpha \delta \beta$ ,  $z_i$ ,  $i = 1, 2$ , определены в (3),  $p_T(\tau_k)$ ,  $\gamma(T)$  определены в (2) для  $\tau = \tau_k$ ,  $k = 1, 2$ .

## Заключение

Полученные результаты делают возможным решение задачи оценивания неизвестных параметров, задающих модулированный обобщенный полусинхронный поток событий в условиях его неполной наблюдаемости. Для оценки неизвестных параметров потока можно использовать метод моментов и метод максимального правдоподобия.

### Литература

1. Бахолдина М. А., Горцев А. М. Оптимальная оценка состояний модулированного обобщенного полусинхронного потока событий при непродлеваемом мертвом времени // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2014. № 1(26). С. 13–24.
2. Bakholdina M. A., Gortsev A. M. Optimal estimation of the states of modulated semi-synchronous integrated flow of events in condition of its incomplete observability // Applied Mathematical Sciences. 2015. Vol. 9, no. 29. P. 1433–1451.
3. Бахолдина М. А., Горцев А. М. Плотность вероятностей длительности интервала между соседними событиями модулированного обобщенного полусинхронного потока событий при непродлеваемом мертвом времени // Теория вероятностей, случайные процессы, математическая статистика и приложения: матер. Междунар. науч. конф., посвящ. 80-летию проф., д-ра физ.-мат. наук Г. А. Медведева, Минск, 23–26 февр. 2015 г. / редкол.: Н. Н. Труш [и др.]. Минск: РИВШ, 2015. С. 17–22.

## СОВМЕСТНАЯ ПЛОТНОСТЬ ВЕРОЯТНОСТЕЙ ДЛИТЕЛЬНОСТИ ИНТЕРВАЛОВ МОДУЛИРОВАННОГО СИНХРОННОГО ПОТОКА СОБЫТИЙ В УСЛОВИЯХ НЕПОЛНОЙ НАБЛЮДАЕМОСТИ

*А. М. Горцев, М. Н. Сиротина*

*Национальный исследовательский*

*Томский государственный университет, Томск, Россия*

### Введение

Настоящая работа является продолжением исследования модулированного синхронного потока событий, начатое в [1, 2].

Математические модели теории массового обслуживания широко применяются при описании реальных физических, технических и других процессов и систем. В связи с бурным развитием компьютерной техники и информационных технологий появилась важная сфера приложений теории массового обслуживания – проектирование и создание информационно-вычислительных сетей, компьютерных сетей связи, спутниковых сетей и телекоммуникационных сетей [3].

На практике интенсивность входящего потока событий изменяется со временем, при этом изменения часто носят случайный характер, последнее приводит к рассмотрению дважды стохастических потоков событий [4–8], одним из которых является модулированный синхронный дважды стохастический поток [9, 10].

Одним из искажающих факторов при оценке состояний и параметров потока событий выступает мертвое время регистрирующих приборов [11], которое порождается зарегистрированным событием. Другие же события, наступившие в течение периода мертвого времени, недоступны наблюдению (теряются). Можно считать, что этот период продолжается некоторое фиксированное время (непродлевающееся мертвое время).

### 1. Постановка задачи

Рассматривается модулированный синхронный поток событий (далее – поток), интенсивность которого есть кусочно-постоянный стационарный случайный процесс  $\lambda(t)$  с двумя состояниями:  $\lambda_1, \lambda_2$  ( $\lambda_1 > \lambda_2$ ). Длительность пребывания процесса  $\lambda(t)$  (потока) в  $i$ -м состоянии распределена по экспоненциальному закону с параметром  $\alpha_i, i=1,2$ . Если процесс  $\lambda(t)$  в момент времени  $t$  находится в  $i$ -м состоянии, то на полуинтервале  $[t, t + \Delta t)$ , где  $\Delta t$  – достаточно малая величина, с вероятностью  $\alpha_i \Delta t + o(\Delta t)$  пребывание процесса  $\lambda(t)$  в  $i$ -м состоянии закончится, и процесс  $\lambda(t)$  с вероятностью, равной единице, перейдет из  $i$ -го состояния в  $j$ -е ( $i, j=1,2, i \neq j$ ). В течение временного интервала случайной длительности, когда  $\lambda(t) = \lambda_i$ , имеет место пуассоновский поток событий с интенсивностью  $\lambda_i, i=1,2$ . Кроме того, переход из первого состояния процесса  $\lambda(t)$  во второе возможен в момент наступления события пуассоновского потока интенсивности  $\lambda_1$ : переход осуществляется с вероятностью  $p$  ( $0 < p \leq 1$ ); с вероятностью  $1-p$  процесс  $\lambda(t)$  остается в первом состоянии. Переход из второго состояния процесса  $\lambda(t)$  в первое возможен также в момент наступления события пуассоновского потока интенсивности  $\lambda_2$ : переход осуществляется с вероятностью  $q$  ( $0 < q \leq 1$ ); с вероятностью  $1-q$  процесс  $\lambda(t)$  остается во втором состоянии. В сделанных предположениях  $\lambda(t)$  – марковский процесс.

После каждого зарегистрированного события наступает время фиксированной длительности  $T$  (далее – мертвое время), в течение которого другие события потока недоступны наблюдению. Рассматривается непродлевающееся мертвое время, т.е. события, наступившие в течение интервала мертвого времени, не вызывают его продления. По окончании длительности периода мертвого времени первое наступившее событие вновь генерирует период мертвого времени длительности  $T$  и т.д. Вариант возникающей ситуации приведен на рис. 1, где  $\lambda_1, \lambda_2$  – состояния процесса  $\lambda(t)$ ;  $t_1, t_2, \dots$  – моменты наступления наблюдаемых событий потока, штриховка – периоды мертвого времени длительности  $T$ , ось под номером 1 отображает исходный модулированный поток событий, под номером 2 – схему создания мертвого времени, под номером 3 – наблюдаемые события модулированного синхронного потока.

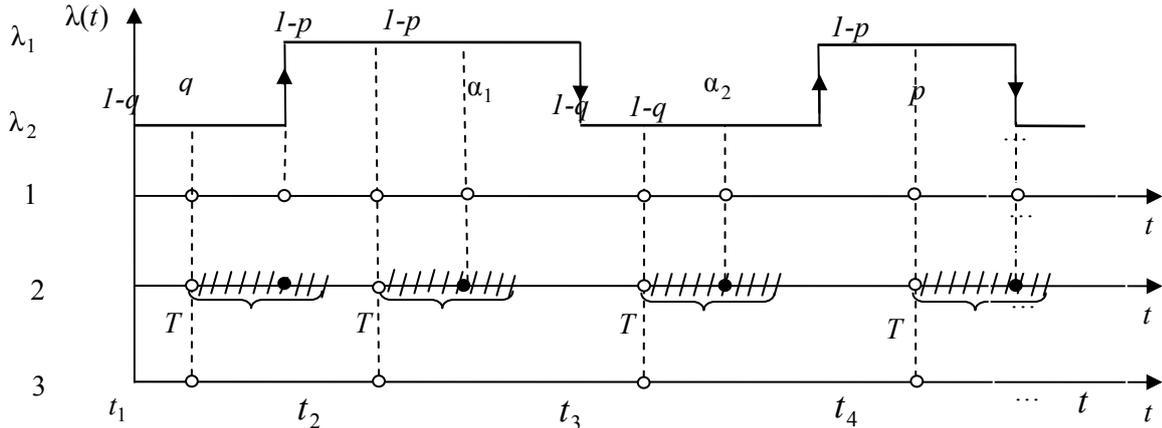


Рис. 1. Формирование наблюдаемого потока событий

Матрицы инфинитезимальных коэффициентов рассматриваемого потока имеют вид

$$D_1 = \begin{vmatrix} (1-p)\lambda_1 & p\lambda_1 \\ q\lambda_2 & (1-q)\lambda_2 \end{vmatrix}, D_0 = \begin{vmatrix} -(\lambda_1 + \alpha_1) & \alpha_1 \\ \alpha_2 & -(\lambda_2 + \alpha_2) \end{vmatrix}.$$

Элементами матрицы  $D_1$  являются интенсивности переходов процесса  $\lambda(t)$  из состояния в состояние с наступлением события. Недиагональные элементы матрицы  $D_0$  – интенсивности переходов из состояния в состояние без наступления события. Диагональные элементы матрицы  $D_0$  – интенсивности выхода процесса  $\lambda(t)$  из своих состояний, взятые с противоположным знаком. Отметим, что если  $\alpha_i = 0, i = 1, 2$ , то имеет место обычный синхронный поток событий [12].

Рассматривается стационарный режим функционирования потока. Последовательность моментов  $t_1, t_2, \dots, t_k, \dots$  наступления событий потока образует вложенную цепь Маркова, т.е. поток обладает марковским свойством, если его эволюцию рассматривать с момента  $t_k$ .

## 2. Совместная плотность вероятностей длительности интервалов модулированного синхронного потока

Пусть  $\tau_1, \tau_2$  – значения случайной величины длительности смежных интервалов между моментами наступления соседних событий потока. Тогда совместная плотность  $p_T(\tau_1, \tau_2)$  запишется в виде

$$p_T(\tau_1, \tau_2) = \begin{cases} 0, 0 \leq \tau_1 < T, 0 \leq \tau_2 < T, \\ \sum_{i=1}^2 \pi_i(0|T) \sum_{j=1}^2 q_{ij}(T) \sum_{k=1}^2 \tilde{p}_{jk}(\tau_1 - T) \sum_{s=1}^2 q_{ks}(T) \sum_{n=1}^2 \tilde{p}_{sn}(\tau_2 - T), \\ \tau_1 \geq T, \tau_2 \geq T, \end{cases} \quad (2.1)$$

где  $\tilde{p}_{jk}(\tau - T)$  ( $\tilde{p}_{sn}(\tau - T)$ ) – условная плотность вероятностей того, что на интервале  $(\tau - T, \tau)$  нет событий потока и в момент времени  $\tau$  имеет место

$\lambda(\tau) = \lambda_k(\lambda_s)$ , при условии, что в момент времени  $\tau - T$  значение процесса  $\lambda(\tau - T) = \lambda_j(\lambda_n)$  ( $j, k, s, n=1,2$ );  $q_{ij}(T)(q_{ks}(T))$  – вероятность того, что за время  $T$  процесс  $\lambda(t)$  перешел из  $i$ -го ( $k$ -го) состояния в  $j$ -е ( $s$ -е),  $i, j, k, s=1,2$ ;  $\pi_i(0|T)$  – условная стационарная вероятность того, что в момент  $\tau = 0$  процесс  $\lambda(t)$  находится в состоянии  $i$ , событие потока наступило, и наступило мертвое время длительности  $T$ . При этом

$$p_T(\tau_i) = \sum_{i=1}^2 \pi_i(0|T) \sum_{j=1}^2 q_{ij}(T) \sum_{k=1}^2 \tilde{p}_{jk}(\tau_i - T), \tau_i \geq T, i=1,2 \quad [2].$$

Обозначим  $t = \tau - T, \tau \geq T$ . Введем  $p_{jk}(t)$  – переходную вероятность того, что на интервале  $(0, t)$  нет событий потока и в момент времени  $t$  имеет место  $\lambda(t) = \lambda_k$ , при условии, что в момент времени  $t = 0$  значение процесса  $\lambda(0) = \lambda_j$  ( $j, k=1,2$ ). Тогда для введенных вероятностей  $p_{jk}(t)$  имеет место система дифференциальных уравнений:

$$\begin{aligned} p_{11}'(t) &= -(\lambda_1 + \alpha_1)p_{11}(t) + \alpha_2 p_{12}(t), & p_{12}'(t) &= -(\lambda_2 + \alpha_2)p_{12}(t) + \alpha_1 p_{11}(t), \\ p_{22}'(t) &= -(\lambda_2 + \alpha_2)p_{22}(t) + \alpha_1 p_{21}(t), & p_{21}'(t) &= -(\lambda_1 + \alpha_1)p_{21}(t) + \alpha_2 p_{22}(t), \\ p_{11}(0) &= 1, & p_{12}(0) &= 0, & p_{22}(0) &= 1, & p_{21}(0) &= 0. \end{aligned}$$

Решая полученную систему дифференциальных уравнений, находим

$$\begin{aligned} p_{11}(t) &= \frac{1}{z_2 - z_1} \left[ (\lambda_2 + \alpha_2 - z_1) e^{-z_1 t} - (\lambda_2 + \alpha_2 - z_2) e^{-z_2 t} \right], \\ p_{12}(t) &= \frac{\alpha_1}{z_2 - z_1} \left[ e^{-z_1 t} - e^{-z_2 t} \right], & p_{21}(t) &= \frac{\alpha_2}{z_2 - z_1} \left[ e^{-z_1 t} - e^{-z_2 t} \right], \\ p_{22}(t) &= \frac{1}{z_2 - z_1} \left[ (\lambda_1 + \alpha_1 - z_1) e^{-z_1 t} - (\lambda_1 + \alpha_1 - z_2) e^{-z_2 t} \right], \\ z_{1,2} &= (\alpha_1 + \lambda_1 + \alpha_2 + \lambda_2) \mp \sqrt{(\alpha_1 + \lambda_1 - \alpha_2 - \lambda_2)^2 + 4\alpha_1\alpha_2}. \end{aligned} \quad (2.2)$$

При этом плотности вероятностей  $\tilde{p}_{jk}(t) = \tilde{p}_{jk}(\tau - T)$  ( $\tilde{p}_{sn}(t) = \tilde{p}_{sn}(\tau - T)$ )  $j, k, s, n=1, 2$ , из формулы (2.1) для плотности вероятностей  $p_T(\tau_1, \tau_2)$  запишутся в виде

$$\begin{aligned} \tilde{p}_{11}(t) &= p_{11}(t)\lambda_1(1-p) + p_{12}(t)\lambda_2q, & \tilde{p}_{12}(t) &= p_{11}(t)\lambda_1p + p_{12}(t)\lambda_2(1-q), \\ \tilde{p}_{22}(t) &= p_{22}(t)\lambda_2(1-q) + p_{21}(t)\lambda_1p, & \tilde{p}_{21}(t) &= p_{22}(t)\lambda_2q + p_{21}(t)\lambda_1(1-p), \end{aligned} \quad (2.3)$$

где вероятности  $p_{jk}(t)$ ,  $j, k = 1, 2$ , определены в (2.2).

Введем  $q_{ij}(\tau)$  – вероятность того, что за время  $T$  процесс  $\lambda(t)$  перейдет из  $i$ -го состояния в  $j$ -е,  $i, j = 1, 2$ . Для данных вероятностей имеет место система дифференциальных уравнений ( $0 \leq \tau \leq T$ ):

$$\begin{aligned}
q_{11}'(\tau) &= -(\alpha_1 + p\lambda_1)q_{11}(\tau) + (\alpha_2 + q\lambda_2)q_{12}(\tau), \\
q_{12}'(\tau) &= -(\alpha_2 + q\lambda_2)q_{12}(\tau) + (\alpha_1 + p\lambda_1)q_{21}(\tau), \\
q_{21}'(\tau) &= -(\alpha_2 + q\lambda_2)q_{21}(\tau) + (\alpha_1 + p\lambda_1)q_{22}(\tau), \\
q_{22}'(\tau) &= -(\alpha_1 + p\lambda_1)q_{22}(\tau) + (\alpha_2 + q\lambda_2)q_{21}(\tau), \\
q_{11}(0) &= 1, q_{12}(0) = 0, q_{22}(0) = 1, q_{21}(0) = 0.
\end{aligned}$$

Решая данную систему дифференциальных уравнений и в полученном решении полагая  $\tau = T$ , находим значения  $q_{ij}(T)$  ( $q_{ks}(T)$ ),  $i, j, k, s=1,2$ , из формулы (2.1) для плотности вероятностей  $p_T(\tau_1, \tau_2)$ :

$$\begin{aligned}
q_{11}(T) &= \pi_2 + \pi_1 e^{-(\alpha_1 + \alpha_2 + p\lambda_1 + q\lambda_2)T}, \quad q_{12}(T) = \pi_2 - \pi_2 e^{-(\alpha_1 + \alpha_2 + p\lambda_1 + q\lambda_2)T}, \\
q_{21}(T) &= \pi_1 - \pi_1 e^{-(\alpha_1 + \alpha_2 + p\lambda_1 + q\lambda_2)T}, \quad q_{22}(T) = \pi_1 + \pi_2 e^{-(\alpha_1 + \alpha_2 + p\lambda_1 + q\lambda_2)T}, \quad (2.4) \\
\pi_1 &= \frac{\alpha_2 + q\lambda_2}{\alpha_1 + p\lambda_1 + \alpha_2 + q\lambda_2}, \quad \pi_2 = \frac{\alpha_1 + p\lambda_1}{\alpha_1 + p\lambda_1 + \alpha_2 + q\lambda_2},
\end{aligned}$$

где  $\pi_i, i=1,2$  – априорная стационарная вероятность того, что в произвольный момент времени поток находится в  $i$ -м состоянии.

Вероятности  $\pi_i(0|T), i=1,2$ , из (2.1) можно получить в виде [2]

$$\begin{aligned}
\pi_1(0|T) &= \frac{q\lambda_2\alpha_1 + (1-p)\lambda_1(\lambda_2 + \alpha_2) - \lambda_1\lambda_2(1-p-q)(\pi_2 + \pi_1 e^{-(\alpha_1 + \alpha_2 + p\lambda_1 + q\lambda_2)T})}{z_1 z_2 - \lambda_1\lambda_2(1-p-q)e^{-(\alpha_1 + \alpha_2 + p\lambda_1 + q\lambda_2)T}}, \\
\pi_2(0|T) &= \frac{(1-q)\lambda_2(\lambda_1 + \alpha_1) + p\lambda_1\alpha_2 - \lambda_1\lambda_2(1-p-q)(\pi_1 + \pi_2 e^{-(\alpha_1 + \alpha_2 + p\lambda_1 + q\lambda_2)T})}{z_1 z_2 - \lambda_1\lambda_2(1-p-q)e^{-(\alpha_1 + \alpha_2 + p\lambda_1 + q\lambda_2)T}}, \\
z_1 z_2 &= \lambda_1\lambda_2 + \lambda_1\alpha_2 + \lambda_2\alpha_1. \quad (2.5)
\end{aligned}$$

При этом вероятности  $\pi_i(T), i=1,2$ , можно получить в виде [2]:

$$\begin{aligned}
\pi_1(T) &= \pi_1 - (\pi_1 - \pi_1(0|T))e^{-(\alpha_1 + \alpha_2 + p\lambda_1 + q\lambda_2)T}, \\
\pi_2(T) &= \pi_2 - (\pi_2 - \pi_2(0|T))e^{-(\alpha_1 + \alpha_2 + p\lambda_1 + q\lambda_2)T}, \quad (2.6)
\end{aligned}$$

где  $\pi_i, i=1,2$ , определены в (2.4),  $\pi_i(0|T), i=1,2$ , определены в (2.5).

Несложно показать, что

$$\begin{aligned}
\pi_1(T) &= \pi_1(0|T)q_{11}(T) + \pi_2(0|T)q_{21}(T), \\
\pi_2(T) &= \pi_2(0|T)q_{12}(T) + \pi_1(0|T)q_{22}(T). \quad (2.7)
\end{aligned}$$

Используя формулу (2.7) и сделав замены  $t_1 = \tau_1 - T, t_2 = \tau_2 - T$ , несложно получить формулу для совместной плотности  $p_T(\tau_1, \tau_2)$  в виде

$$p_T(\tau_1, \tau_2) = \begin{cases} 0, & 0 \leq \tau_1 < T, 0 \leq \tau_2 < T, \\ \sum_{i=1}^2 \pi_i(T) \sum_{j=1}^2 \tilde{p}_{ij}(t_1) \sum_{k=1}^2 q_{jk}(T) \sum_{s=1}^2 \tilde{p}_{ks}(t_2), & t_1 \geq 0, t_2 \geq 0, \end{cases} \quad (2.8)$$

Подставляя значения  $p_{jk}(t), j, k=1,2$ , из (2.2) в формулу (2.3) для плотности вероятностей  $\tilde{p}_{jk}(t), \tilde{p}_{sn}(t), j, k, s, n=1,2$ , затем подставляя полу-

ченные значения  $\tilde{p}_{jk}(t), \tilde{p}_{sn}(t), j, k, s, n = 1, 2$  из (2.3),  $q_{jk}(T), q_{ks}(T), j, k, s = 1, 2$ , из (2.4),  $\pi_i(T), i = 1, 2$ , из (2.7) в формулу для совместной плотности вероятности  $p_T(\tau_1, \tau_2)$  (2.8) и проделывая достаточно трудоемкие преобразования, получим следующее выражение для  $p_T(\tau_1, \tau_2)$ :

$$p_T(\tau_1, \tau_2) = 0, 0 \leq \tau_1 < T, 0 \leq \tau_2 < T, \quad (2.9)$$

$$p_T(\tau_1, \tau_2) = p_T(\tau_1)p_T(\tau_2) + e^{-(\alpha_1 + \alpha_2 + p\lambda_1 + q\lambda_2)T} \frac{\lambda_1\lambda_2(1-p-q)}{z_1z_2} \gamma(T)(1-\gamma(T)) \times \\ \times (z_1e^{-z_1(\tau_1-T)} - z_2e^{-z_2(\tau_1-T)})(z_1e^{-z_1(\tau_2-T)} - z_2e^{-z_2(\tau_2-T)}), \tau_1 \geq T, \tau_2 \geq T,$$

где  $\gamma(T)(1-\gamma(T))$  имеет вид

$$\gamma(T)(1-\gamma(T)) = \quad (2.10) \\ = \frac{(\lambda_1 - \lambda_2)((\alpha_1 + p\lambda)\pi_1(0) - (\alpha_2 + q\lambda_2)\pi_2(0))((p+q)\lambda_1\lambda_2 + \lambda_1\alpha_2 + \lambda_2\alpha_1)}{(z_2 - z_1)^2(z_1z_2 - \lambda_1\lambda_2(1-p-q))e^{-(\alpha_1 + \alpha_2 + p\lambda_1 + q\lambda_2)T}} \times \\ \times \left\{ z_1z_2 + e^{-(\alpha_1 + \alpha_2 + p\lambda_1 + q\lambda_2)T} ((\alpha_1 + p\lambda_1 + \alpha_2 + q\lambda_2)(\alpha_1 + \lambda_1 + \alpha_2 + \lambda_2) - 2z_1z_2) - \right. \\ \left. - e^{-2(\alpha_1 + \alpha_2 + p\lambda_1 + q\lambda_2)T} ((\alpha_1 + p\lambda_1 + \alpha_2 + q\lambda_2)(\lambda_1(1-p) + \lambda_2(1-q)) - z_1z_2) \right\} \times \\ \times (z_1e^{-z_1(\tau_1-T)} - z_2e^{-z_2(\tau_1-T)})(z_1e^{-z_1(\tau_2-T)} - z_2e^{-z_2(\tau_2-T)})z_1z_2,$$

где  $z_{1,2}$  определены в (2.2),  $z_1z_2$  определены в (2.5),  $\pi_i(0), i = 1, 2$ , определены в (2.5) при  $T=0$ .

### 3. Условия рекуррентности модулированного синхронного потока

Рассмотрим частные случаи, когда модулированный синхронный поток событий, функционирующий в условиях мертвого времени, становится рекуррентным потоком.

1) Из формулы (2.9) для  $p_T(\tau_1, \tau_2)$  следует первое условие рекуррентности наблюдаемого потока. Если  $1-p-q=0$ , то совместная плотность (2.9) факторизуется:  $p_T(\tau_1, \tau_2) = p_T(\tau_1)p_T(\tau_2), \tau_1 \geq T, \tau_2 \geq T$ .

При этом плотность вероятностей  $p_T(\tau)$  [2] примет вид

$$p_T(\tau) = 0, 0 \leq \tau < T, \\ p_T(\tau) = \gamma(T)z_1e^{-z_1(\tau-T)} + (1-\gamma(T))z_2e^{-z_2(\tau-T)}, \tau \geq T, \\ \gamma(T) = \frac{1}{z_2 - z_1} (z_2 - \lambda_1\pi_1 - \lambda_2\pi_2 + (\lambda_1(\pi_1 - q) + \lambda_2(\pi_2 - p))e^{-(\alpha_1 + \alpha_2 + p\lambda_1 + q\lambda_2)T}), \\ 1 - \gamma(T) = \frac{1}{z_2 - z_1} (-z_1 + \lambda_1\pi_1 + \lambda_2\pi_2 - (\lambda_1(\pi_1 - q) + \lambda_2(\pi_2 - p))e^{-(\alpha_1 + \alpha_2 + p\lambda_1 + q\lambda_2)T}),$$

где  $z_{1,2}$  определены в (2.2),  $\pi_i, i = 1, 2$  определены в (2.4).

2) Из формулы (2.10) для произведения  $\gamma(T)(1-\gamma(T))$  следует второе условие рекуррентности наблюдаемого потока. Если  $(\alpha_1 + p\lambda)\pi_1(0) - (\alpha_2 + q\lambda_2)\pi_2(0) = 0$ , то совместная плотность (2.9) факторизуется:  $p_T(\tau_1, \tau_2) = p_T(\tau_1)p_T(\tau_2), \tau_1 \geq T, \tau_2 \geq T$ , при этом  $\gamma(T) = 1$ .

Тогда плотность вероятностей  $p_T(\tau)$  [2] примет вид

$$p_T(\tau) = 0, 0 \leq \tau < T,$$
$$p_T(\tau) = z_1 e^{-z_1(\tau-T)}, \tau \geq T,$$

где  $z_1$  определена в (2.2).

### Заключение

В ходе исследования получен явный вид  $p_T(\tau_1, \tau_2)$  совместной плотности вероятностей значений длительности интервалов между моментами наступления соседних событий модулированного синхронного потока событий, функционирующего в условиях мертвого времени. Приведены условия рекуррентности потока. Данные формулы позволят в дальнейшем произвести оценку параметров потока по методу максимального правдоподобия либо по методу моментов.

### Литература

1. Горцев А. М., Сиротина М. Н. Совместная плотность вероятностей длительности интервалов модулированного синхронного потока событий // ИТММ–2014: матер. XIII Междунар. науч.-практ. конф. им. А.Ф. Терпугова. Анжеро-Судженск, Россия, 20–22 ноября, 2014. Томск: Изд-во Том. ун-та, 2014. Ч. 2. С. 147–153.
2. Сиротина М. Н., Горцев А. М. Плотность вероятностей длительности интервала между соседними событиями в модулированном синхронном потоке при непродлеваемом мертвом времени // Материалы Международной научной конференции, посвященной 80-летию профессора, доктора физико-математических наук Геннадия Алексеевича Медведева. Белоруссия, Минск, 23–26 февр. 2015. Минск: Изд-во РИВШ, 2015. С. 299–304.
3. Дудин А. Н., Клименок В. Н. Системы массового обслуживания с коррелированными потоками. Минск: Изд-во БГУ, 2000. 175 с.
4. Kingman, J. F.C. On doubly stochastic Poisson process // Proceedings Cambridge Philosophical Society. 1964. Vol. 60, no. 4. P. 923–930.
5. Башарин Г. П., Кокотушкин В. А., Наумов В. А. О методе эквивалентных замен расчета фрагментов сетей связи // Изд. АН СССР. Техн. кибернетика. 1979. № 6. С. 92–99.
6. Neuts M. F. A versatile Markov point process // Journal of Applied Probability. 1979. Vol. 16. P. 764–779.
7. Lucantoni, D. M. New results on the single server queue with a batch markovian arrival process // Communication in Statistics Stochastic Models. 1991. Vol. 7. P. 1–46.
8. Card H. C. Doubly stochastic Poisson processes in artificial neural learning // Neural Networks, IEEE Transactions. 1998. Vol. 9, issue 1. P. 229–231.
9. Горцев А. М., Голофастова М. Н. Оптимальная оценка состояний модулированного синхронного дважды стохастического потока событий // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2013. № 2(23). С. 42–53.
10. Апанасович В. В., Коляда А. А., Чернявский А. Ф. Статистический анализ случайных потоков в физическом эксперименте // Минск: Университетское, 1988. 254 с.
11. Сиротина М. Н. Оптимальная оценка состояний модулированного синхронного дважды стохастического потока событий в условиях непродлеваемого мертвого

времени // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2014. № 1(26). С. 63–72.

12. Bushlanov I. V., Gortsev A. M., Nezhel'skaya L. A. Estimating parameters of the synchronous twofold-stochastic flow of events // Automation and Remote Control. 2008. Vol. 69, no. 9. P. 1517–1533.

## **ВЕРОЯТНОСТНЫЙ АНАЛИЗ ДЛЯ ИССЛЕДОВАНИЯ ЕВРОПЕЙСКОГО ЭКЗОТИЧЕСКОГО ОПЦИОНА НА ФОНДОВЫЙ ИНДЕКС**

*Е. Ю. Данилюк*

*Национальный исследовательский*

*Томский государственный университет, Томск, Россия*

На сегодняшний день финансовые инструменты торговли и хеджирования рисков [1] на рынке деривативов представлены фьючерсами, форвардами и опционами, в том числе экзотическими опционами [2–6]. Последние вызывают особый интерес у инвесторов ввиду широкого разнообразия платежных схем по опционам [7] и являются объектом исследования стохастической финансовой математики [8]. Европейский опцион продажи – это вторичная ценная бумага, или контракт, по которому покупатель опциона (держатель) обладает правом продать оговоренный в договоре базовый актив в фиксированный момент по определенной цене, а продавец дериватива за премию обязан удовлетворить платежное обязательство в момент исполнения опциона [1].

Исследование посвящено европейскому экзотическому опциону с ограничением выплат для продавца и гарантированным доходом для держателя в случае, когда базисным активом выступает фондовый индекс. Функция выплат исследуемого опциона имеет вид

$$f_T(i_T) = \min\{(K_1 - i_T)^+, K_2\}, \quad (1)$$

где  $i_T$  – значение рискованного актива в момент экспирации  $T$ ;  $K_1$  – цена исполнения (страйковая цена);  $K_2$  – оговоренная контрактом величина, ограничивающая выплаты для продавца опциона, с одной стороны, и гарантирующая доход для владельца дериватива – с другой;  $a^+ = \max(a; 0)$ . Фондовый индекс в (1) представляет собой индикатор состояния рынка ценных бумаг, используемый инвесторами и финансовыми менеджерами, чтобы описать рынок и сравнить доходы от торговли специфическими инструментами [9, 10]. Фондовый индекс рассчитывается на основе цен акций индексной корзины (как правило, взвешенных значений цен акций), текущее значение индекса в момент времени  $t$  определяется как

$$i_t = \frac{1}{b} \sum_{k=1}^n V_k S_t^k, \quad (2)$$

где  $S_t^k$  и  $V_k$  – соответственно текущее значение цены и количество акций с номером  $k$ , входящих в индексную корзину;  $b = \sum_{k=1}^n V_k$ .

Согласно (1) платежное обязательство европейского стандартного опциона продажи  $f_T(i_T) = (K_1 - i_T)^+$  лежит в основе изучаемого экзотического опциона. Если в момент времени  $T$  состояние рынка такое, что  $K_1 > i_T$ , то владелец опциона получает выплату в размере  $(K_1 - i_T)$ , если  $K_1 - i_T > K_2$ , или  $K_2$ , если  $K_1 - i_T < K_2$ . В противном случае (если  $K_1 < i_T$ ) держатель дериватива ничего не получает.

Обозначим через  $E\{\cdot\}$  математическое ожидание;  $N\{a; b\}$  – плотность нормального распределения с параметрами  $a$  и  $b$ ;  $\Phi(x) = \int_{-\infty}^x \varphi(y) dy$ ,  $\varphi(y) = (1/\sqrt{2\pi}) \exp\{-y^2/2\}$  – функция распределения и функция плотности распределения вероятностей Лапласа соответственно.

Рассмотрение задачи проводится на полном, безарбитражном и риск-нейтральном финансовом рынке двух активов, а именно: рисковом (фондовый индекс)  $i_t$  и безрисковом (банковский счет)  $B_t$  ценных бумаг. Под ценой фондового индекса будем понимать его значение, определяемое ценами акций  $S_t^k$ ,  $k = 1, 2, \dots, n$ , из индексной корзины. В свою очередь, эволюция стоимости каждого из активов задается на стохастическом базисе  $(\Omega, F, \mathbb{F} = (F_t)_{t>0}, \mathbf{P})$  [5, 7]. Текущие цены  $S_t^k$  и  $B_t$  в течение интервала времени  $t \in [0; T]$  определяются уравнениями

$$dS_t^k = S_t^k (\mu_k dt + \sigma_k dW_t), \quad S_t^k = S_0^k \exp\left\{\left(\mu_k - \frac{\sigma_k^2}{2}\right)t + \sigma_k W_t\right\}, \quad (3)$$

$$dB_t = rB_t dt, \quad B_t = B_0 \exp\{rt\}, \quad (4)$$

где  $W_t$  – стандартный винеровский процесс,  $S_0^k > 0$  – начальная стоимость акции с номером  $k$ ,  $\mu_k \in R = (-\infty, +\infty)$  – параметр «сноса»,  $\sigma > 0$  – параметр волатильности в модели геометрического броуновского движения,  $B_0 > 0$  – начальная стоимость безрискового актива,  $r > 0$  – безрисковая процентная ставка,  $k = 1, 2, \dots, n$ .

В каждый момент интервала времени  $t \in [0; T]$  инвестор формирует самофинансируемый портфель  $\pi_t = (\beta_t, \gamma_t)$ , где пара  $F_t$ -измеримых процессов  $\beta_t$  и  $\gamma_t$  – доли безрискового и рисковомого активов в портфеле соответственно, который обеспечивает капитал  $X_t = \beta_t B_t + \gamma_t i_t$ .

Задача заключается в том, чтобы для опциона с платежным обязательством (1) рассчитать рациональную цену  $P_T = X_0$ , необходимый для покрытия стоимости опциона капитал  $X_t$ , а также отвечающий капиталу портфель ценных бумаг  $\pi_t = (\beta_t, \gamma_t)$  таким образом, чтобы в момент экспи-

рации  $T$  дериватива гарантировать выполнение платежного обязательства  $X_T = f_T(i_T)$ .

Все приведенные ниже результаты получены с учетом основных финансовых соотношений (5)–(7) [3–5, 7, 8] и в предположении существования единственной риск-нейтральной меры  $\mathbf{P}^*$ , относительно которой процесс дисконтированной цены акции  $\widehat{S}_t^k = S_t^k / B_t$ , образующей базисный актив (2), является мартингалом, что обеспечивает разрешимость поставленной задачи [3–5, 7, 8]:

$$P_T = e^{-rT} E^* \{f_T(i_T)\}, \quad (5)$$

$$X_t = E^* \{e^{-r(T-t)} f_T(i_T) | i_t\}, \quad (6)$$

$$\beta_t = \frac{\partial X_t(s)}{\partial s} \Big|_{s=B_t}, \quad \gamma_t = \frac{X_t - \beta_t B_t}{i_t}, \quad (7)$$

где  $E^* \{\cdot\}$  – усреднение по риск-нейтральной мере  $\mathbf{P}^*$ .

**Теорема 1.** Для европейского экзотического опциона продажи с функцией выплат (1) текущее значение минимального инвестиционного портфеля  $X_t$  определяется уравнением (8)

$$X_t = K_1 e^{-r(T-t)} [\Phi(y_1(T-t)) - \Phi(y_0(T-t))] + K_2 e^{-r(T-t)} \Phi(y_0(T-t)) - \sum_{k=1}^n D_t^k [\Phi(y_1(T-t) - \sigma_k \sqrt{T-t}) - \Phi(y_0(T-t) - \sigma_k \sqrt{T-t})], \quad (8)$$

где

$$y_0(T-t) = \min \left\{ \frac{1}{\sigma_k \sqrt{T-t}} \left[ \ln \left( \frac{K_1 - K_2}{D_t^k} \right) - \left( r - \frac{\sigma_k^2}{2} \right) t \right] \right\}, \quad k = 1, 2, \dots, n, \quad (9)$$

$$y_1(T-t) = \min \left\{ \frac{1}{\sigma_k \sqrt{T-t}} \left[ \ln \left( \frac{K_1}{D_t^k} \right) - \left( r - \frac{\sigma_k^2}{2} \right) t \right] \right\}, \quad k = 1, 2, \dots, n, \quad (10)$$

$$D_t^k = \frac{V_k S_t^k}{b}, \quad b = \sum_{k=1}^n V_k. \quad (11)$$

**Теорема 2.** Рациональная стоимость европейского экзотического опциона продажи с ограничением выплат для продавца и гарантированным доходом для владельца опциона на фондовый индекс задается (12)

$$P_T = K_1 e^{-rT} [\Phi(y_1(T)) - \Phi(y_0(T))] + K_2 e^{-rT} \Phi(y_0(T)) - \sum_{k=1}^n D_0^k [\Phi(y_1(T) - \sigma_k \sqrt{T}) - \Phi(y_0(T) - \sigma_k \sqrt{T})], \quad (12)$$

где  $y_0(T)$ ,  $y_1(T)$ ,  $D_0^k$  определяются (9), (10), (11) при  $t = 0$  соответственно.

**Теорема 3.** Для европейского экзотического опциона продажи с платежным обязательством (1) текущее значение оптимального хеджирующего портфеля  $\pi_t = (\beta_t, \gamma_t)$  определяется как

$$\begin{aligned} \beta_t = & K_1 e^{-rT} [\Phi(y_1(T-t)) - \Phi(y_0(T-t))] + K_2 e^{-rT} \Phi(y_1(T-t)) + \\ & + e^{-rT} \sum_{k=1}^n D_t^k \left[ \frac{\varphi(y_1(T-t) - \sigma_k \sqrt{T-t})}{\sigma_m \sqrt{T-t}} - \frac{\varphi(y_0(T-t) - \sigma_k \sqrt{T-t})}{\sigma_l \sqrt{T-t}} \right] - \quad (13) \\ & - K_1 e^{-rT} \left[ \frac{\varphi(y_1(T-t))}{\sigma_m \sqrt{T-t}} - \frac{\varphi(y_0(T-t))}{\sigma_l \sqrt{T-t}} \right] - \frac{K_2 e^{-rT}}{\sigma_l \sqrt{T-t}} \varphi(y_1(T-t)), \end{aligned}$$

$$\begin{aligned} \gamma_t = & \frac{1}{i_t} \left\{ K_1 e^{-r(T-t)} \frac{\varphi(y_1(T-t))}{\sigma_m \sqrt{T-t}} - (K_1 - K_2) e^{-r(T-t)} \frac{\varphi(y_0(T-t))}{\sigma_l \sqrt{T-t}} + \right. \\ & + \sum_{k=1}^n D_t^k \left[ \frac{\varphi(y_0(T-t) - \sigma_k \sqrt{T-t})}{\sigma_l \sqrt{T-t}} - \frac{\varphi(y_1(T-t) - \sigma_k \sqrt{T-t})}{\sigma_m \sqrt{T-t}} - \right. \quad (14) \\ & \left. \left. - \Phi(y_1(T-t) - \sigma_k \sqrt{T-t}) + \Phi(y_0(T-t) - \sigma_k \sqrt{T-t}) \right] \right\}, \end{aligned}$$

где  $\sigma_l$  и  $\sigma_m$ ,  $l, m = 1, 2, \dots, n$ , – волатильности, при которых выражения

$$\begin{aligned} & \frac{1}{\sigma_k \sqrt{T-t}} \left[ \ln \left( \frac{K_1 - K_2}{D_t^k} \right) - \left( r - \frac{\sigma_k^2}{2} \right) t \right], \quad k = 1, 2, \dots, n, \\ & \frac{1}{\sigma_k \sqrt{T-t}} \left[ \ln \left( \frac{K_1}{D_t^k} \right) - \left( r - \frac{\sigma_k^2}{2} \right) t \right], \quad k = 1, 2, \dots, n, \end{aligned}$$

минимальны в каждый момент времени  $t \in [0; T]$ .

**Теорема 4.** Коэффициенты чувствительности, определяющие зависимость рациональной стоимости экзотического опциона продажи с функцией выплат (1) от страйковой цены  $P_T^{K_1} = \partial P_T / \partial K_1$ ; от величины  $K_2$ , ограничивающей выплаты по опциону,  $P_T^{K_2} = \partial P_T / \partial K_2$ ; от момента исполнения  $T$   $P_T^T = \partial P_T / \partial T$ , задаются формулами (15)–(17):

$$\begin{aligned} P_T^{K_1} = & e^{-rT} [\Phi(y_1(T)) - \Phi(y_0(T))] + \frac{K_2 e^{-rT}}{(K_1 - K_2)} \frac{\varphi(y_0(T))}{\sigma_l \sqrt{T}} + \\ & + K_1 e^{-rT} \left[ \frac{1}{K_1} \frac{\varphi(y_1(T))}{\sigma_m \sqrt{T}} - \frac{1}{(K_1 - K_2)} \frac{\varphi(y_0(T))}{\sigma_l \sqrt{T}} \right] - \quad (15) \\ & - \sum_{k=1}^n D_0^k \left[ \frac{1}{K_1} \frac{\varphi(y_1(T) - \sigma_k \sqrt{T})}{\sigma_m \sqrt{T}} - \frac{1}{(K_1 - K_2)} \frac{\varphi(y_0(T) - \sigma_k \sqrt{T})}{\sigma_l \sqrt{T}} \right], \end{aligned}$$

$$P_T^{K_2} = \frac{e^{-rT} [\varphi(y_0(T)) + \Phi(y_0(T))]}{\sigma_l \sqrt{T}} - \sum_{k=1}^n \frac{D_0^k}{(K_1 - K_2)} \frac{\varphi(y_0(T) - \sigma_k \sqrt{T})}{\sigma_l \sqrt{T}}, \quad (16)$$

$$\begin{aligned} P_T^T = & (K_1 - K_2) e^{-rT} \left[ r \Phi(y_0(T)) + \frac{y_0(T) \varphi(y_0(T))}{2T} \right] - \\ & - K_1 e^{-rT} \left[ r \Phi(y_1(T)) + \frac{y_1(T) \varphi(y_1(T))}{2T} \right] + \frac{1}{2T} \sum_{k=1}^n D_0^k [(y_1(T) - \sigma_k \sqrt{T}) \times \quad (17) \\ & \times \varphi(y_1(T) - \sigma_k \sqrt{T}) - (y_0(T) - \sigma_k \sqrt{T}) \varphi(y_0(T) - \sigma_k \sqrt{T})], \end{aligned}$$

$k, l, m = 1, 2, \dots, n$ ,  $t = 0$ .

**Теорема 5.** Коэффициент чувствительности  $P_T^{S_0^p} = \partial P_T / \partial S_0^p$ , определяющий зависимость цены европейского экзотического опциона продажи с функцией выплат (1) от начальной цены акции с номером  $p$   $S_0^p$ ,  $p = 1, 2, \dots, n$ , с учетом (2) определяется (18):

$$P_T^{S_0^p} = \begin{cases} -(V_p/b) [\Phi(y_1(T) - \sigma_p \sqrt{T}) - \Phi(y_0(T) - \sigma_p \sqrt{T})], & \text{if } p \neq l, m, \\ -(S_0^p \sigma_p \sqrt{T})^{-1} \sum_{k=1}^n D_0^k \phi(y_0(T) - \sigma_k \sqrt{T}) - (V_p/b) \times \\ \times [\Phi(y_1(T) - \sigma_p \sqrt{T}) - \Phi(y_0(T) - \sigma_p \sqrt{T})], & \text{if } p = l, p \neq m, \\ (S_0^p \sigma_p \sqrt{T})^{-1} \sum_{k=1}^n D_0^k \phi(y_0(T) - \sigma_k \sqrt{T}) + (V_p/b) \times \\ \times [\Phi(y_1(T) - \sigma_p \sqrt{T}) - \Phi(y_0(T) - \sigma_p \sqrt{T})], & \text{if } p \neq l, p = m. \end{cases} \quad (18)$$

Исследование коэффициентов чувствительности  $P_T^{K_1}$ ,  $P_T^{K_2}$ ,  $P_T^T$ ,  $P_T^{S_0^p}$  показало, что стоимость европейского экзотического опциона продажи с функцией выплат (1) является возрастающей функцией страйка  $P_T^{K_1} > 0$  и ограничивающей выплаты продавца опциона величины  $P_T^{K_2} > 0$  и убывающей функцией момента экспирации  $P_T^T < 0$ . Экономическая интерпретация данных свойств следующая: увеличение страйка  $K_1$  приводит к увеличению вероятности того, что  $K_1$  превзойдет  $i_T$ . Таким образом, вероятность больших выплат по опциону растет и увеличивает стоимость опциона. Чем больше величина гарантированного дохода по опциону, тем меньше риск для владельца дериватива, а за меньший риск следует платить больше. С увеличением времени экспирации становится сложнее прогнозировать значение индекса в момент исполнения опциона, цена опциона падает. Аналитически не удастся однозначно определить зависимость стоимости опциона от начальной цены каждой из акций, входящих в индексную корзину.

#### References

1. Hull J. Options, futures and other derivatives. Moscow: Williams, 2013. 1072 p.
2. Rubinstein M. Exotic Options // Finance working paper. 1991. № 220. P. 5–43.
3. Буренин А. Н. Рынок ценных бумаг и производных финансовых инструментов – М.: НТО, 2011. 394 с.
4. Буренин А. Н. Форварды, фьючерсы, опционы, экзотические и погодные производные. М.: НТО, 2011. 465 с.
5. Ширяев А. Н., Кабанов Ю. М., Крамков Д. О., Мельников А. В. К теории расчетов опционов европейского и американского типов. II. Непрерывное время // Теория вероятностей и ее применения. 2006. Т. 39, №. 1. С. 61–102.
6. Zang P. G. An Introduction to exotic optios // European Financial Management. 1995. Vol. 1, no. 1. P. 87–95.
7. Shiryaev A. N. Essentials of Stochastic Finance: Facts, Models, Theory. Hackensack (New Jersey): World Scientific Publishing Company, 1999. 834 p.

8. Melnikov A. V., Volkov S. N., Nechaev M. L. Mathematics of financial obligations (Translations of Mathematical Monographs). Providence, Rhode Island: American Mathematical Society, 2002. Vol. 212. 194 p.

9. Amenc N., Goltz F., Le Sourd V. Assessing the Quality of Stock Market Indices. EDHEC Publication, 2006. 124 p.

10. Wikipedia, the Free Encyclopedia [Electronic resource]. URL: [https:// en. wikipedia.org/wiki/Stock\\_market\\_index](https://en.wikipedia.org/wiki/Stock_market_index)

## **О $k$ -АДАПТИВНЫХ КОМБИНИРОВАННЫХ ОЦЕНКАХ ВЕРОЯТНОСТИ**

*Ю. Г. Дмитриев, П. Ф. Тарасенко*

*Национальный исследовательский*

*Томский государственный университет, Томск, Россия*

**1. Введение.** При исследовании математических моделей разнообразных систем массового обслуживания (СМО) [1] возникает необходимость расчета тех или иных вероятностных характеристик систем. Обычно это распределения вероятностей состояния системы, распределения времени обслуживания заявки, вероятности некоторых событий, моменты различных порядков, производящие и характеристические функции и т.д. Для практики представляют интерес статистические оценки вероятностных характеристик, построенные по данным, полученным в результате наблюдений за реальной системой. При этом результаты теоретических расчетов математической модели СМО и (или) имитационного моделирования могут использоваться как дополнительная информация при построении статистических оценок характеристик реальных систем с целью повышения их точности или сокращения числа реальных наблюдений при сохранении заданной точности. Кроме того, исследователь может, исходя из своего опыта и знаний, сделать свои предположения о значениях вероятностных характеристик (априорные догадки), которые также можно использовать при их статистическом оценивании.

В работах [2, 3] предложены комбинированные оценки вероятностных характеристик в виде линейных функционалов, учитывающих одновременно эмпирическую оценку и априорную догадку, а также исследованы их асимптотические (при увеличении объема наблюдений) свойства. Методы статистического оценивания вероятности с использованием дополнительной априорной информации рассмотрены в [4]. В данной работе предлагаются  $k$ -адаптивные комбинированные оценки вероятности и анализируются их свойства при конечном объеме наблюдений.

**2. Постановка задачи и структура оценки.** Пусть  $X_1, \dots, X_n$  – независимые результаты эксперимента объема  $n$ , а  $P = P(B)$  – вероятность некоторого события  $B$ , которое может наступить в данном эксперименте.

Пусть  $p_a$  – априорная догадка, которая выступает в качестве возможного значения вероятности  $P$ . Задача состоит в построении оценки для неизвестной вероятности, учитывающей совместно эмпирическую оценку  $\hat{P} = \hat{P}(B) = n^{-1} \sum_{i=1}^n I_B(X_i)$  и  $p_a$ , здесь  $I_B(\cdot)$  – индикаторная функция.

Следуя работам [2, 3], рассмотрим комбинированную оценку

$$\hat{P}_\lambda = (1 - \lambda)\hat{P} + \lambda p_a, \quad (1)$$

где оптимальный весовой коэффициент  $\lambda$  выбран из условия минимума среднеквадратической ошибки (СКО)  $S^2(\lambda) = M[\hat{P}_\lambda - P]^2$  и определяется выражением

$$\lambda = (1 + n\Delta^2/\sigma^2)^{-1}. \quad (2)$$

Здесь  $\sigma^2 = P(1 - P)$ ,  $\Delta = P - p_a$  – величина отклонения априорной догадки от истинного значения искомой вероятности. Весовой коэффициент  $\lambda$  изменяется в пределах  $0 < \lambda \leq 1$  и показывает, какое влияние оказывает каждое из слагаемых в комбинированной оценке (1).

Формулу (1) можно записать как

$$\hat{P}_\lambda = \hat{P} - \lambda(\hat{P} - p_a), \quad (3)$$

$$\hat{P}_\lambda = p_a + (1 - \lambda)(\hat{P} - p_a). \quad (4)$$

Формулу (3) можно трактовать следующим образом: исходной является оценка  $\hat{P}$ , а априорная догадка  $p_a$  выступает в качестве дополнительной информации, которой располагает исследователь до проведения эксперимента. В формуле (4) исходной оценкой является априорная догадка  $p_a$ , а  $\hat{P}$  выступает как дополнительная информация, содержащаяся в экспериментальных данных. В соответствии с (3) и (4) минимум СКО представим в двух формах записи

$$S^2(\lambda) = M[\hat{P} - P]^2 - \frac{[M(\hat{P} - P)]^2}{D\hat{P} + (P - p_a)^2} = D\hat{P} - \frac{(D\hat{P})^2}{D\hat{P} + \Delta^2} = (1 - \lambda)\sigma^2 / n, \quad (5)$$

где дисперсия  $D\hat{P} = P(1 - P)/n$  характеризует точность оценки  $\hat{P}$ , а коэффициент  $(1 - \lambda)$  показывает, во сколько раз уменьшается СКО комбинированной оценки (1) по сравнению с  $\hat{P}$ .

$$S^2(\lambda) = M[p_a - P]^2 - \frac{[M(p_a - P)]^2}{D\hat{P} + (P - p_a)^2} = \Delta^2 - \frac{\Delta^4}{D\hat{P} + \Delta^2} = \lambda \Delta^2. \quad (6)$$

Здесь  $\Delta^2$  характеризует точность оценивания (в смысле СКО), когда оценкой выступает  $p_a$ , коэффициент  $\lambda$  показывает, во сколько раз уменьшается СКО комбинированной оценки за счет привлечения  $\hat{P}$  по сравнению с  $p_a$ .

Величина  $(D\hat{P})^2 / (D\hat{P} + \Delta^2)$  в (5) задает выигрыш в точности оценивания за счет привлечения априорной догадки  $p_a$ , если исходной оценкой является  $\hat{P}$ . Величина  $(\Delta)^4 / (D\hat{P} + \Delta^2)$  в (6) характеризует выигрыш в точности оценивания от использования  $\hat{P}$ , если исходной оценкой является априорная догадка  $p_a$ .

**3. Адаптивные оценки.** Оптимальный коэффициент  $\lambda$ , как правило, неизвестен. Это обстоятельство затрудняет практическое использование комбинированной оценки (1). Выходом из этого положения является построение адаптивных комбинированных оценок вероятности путем замены неизвестного значения  $\lambda$  на его статистическую оценку. Однако при этом возникает вопрос, при каких условиях адаптивные комбинированные оценки являются предпочтительнее обычной оценки  $\hat{P}$  по величине СКО.

Заменяя неизвестное  $P$  на  $\hat{P}$ , в качестве первой оценки  $\lambda$  возьмем  $\hat{\lambda}_1 = (1 + n\hat{\Delta}^2 / \hat{\sigma}^2)^{-1}$ , где  $\hat{\Delta} = \hat{P} - p_a$  – оценка отклонения (смещения),  $\hat{\sigma}^2 = \hat{P}(1 - \hat{P})$  – оценка дисперсии. Подставляя  $\hat{\lambda}_1$  в (1), получим первую адаптивную оценку  $\hat{P}_1 = \hat{P} - \hat{\lambda}_1(\hat{P} - p_a)$ . Используя  $\hat{P}_1$  в оценивании смещения  $\Delta$ , получим  $\hat{\Delta}_1 = \hat{P}_1 - p_a$ , затем вторую оценку  $\hat{\lambda}_2 = (1 + n\hat{\Delta}_1^2 / \hat{\sigma}^2)^{-1}$  и  $\hat{P}_2 = \hat{P} - \hat{\lambda}_2(\hat{P} - p_a)$ . Повторяя такую процедуру последовательно  $k$  раз, приходим к следующим выражениям для оценок:

$$\begin{aligned} \hat{P}_{k,1} &= \hat{P} - \hat{\lambda}_{k,1}(\hat{P} - p_a) = p_a + (1 - \hat{\lambda}_{k,1})(\hat{P} - p_a), \\ \hat{\lambda}_{k,1} &= (1 + n\hat{\Delta}_{k-1}^2 / \hat{\sigma}^2)^{-1} = (1 + q_{k-1}^2(\sqrt{n}\hat{\Delta} / \hat{\sigma}))^{-1}, \end{aligned} \quad (7)$$

где  $\hat{\Delta}_{k-1,1} = \hat{P}_{k-1,1} - p_a$ ,  $\hat{\Delta}_0 = \hat{\Delta}$ ,  $q_k(x) = xq(q_{k-1}(x))$ ,  $k \in \{1, 2, 3, \dots\}$ ,  $q(x) = x^2 / (1 + x^2)$ ,  $q_0(x) = x$ ,  $x \in R^1$  и, как показано в [3],

$$\lim_{k \rightarrow \infty} q_k(x) = \begin{cases} (x - \sqrt{x^2 - 4}) / 2, & x \leq 2, \\ 0, & |x| < 2, \\ (x + \sqrt{x^2 - 4}) / 2, & x \geq 2. \end{cases}$$

Оценку  $\hat{P}_{k,1}$  назовем  $k$ -адаптивной. Подчеркнем, что здесь на каждом шаге при оценивании  $\Delta$  использовалась априорная догадка  $p_a$ . Данная оценка является смещенной в отличие от  $\hat{P}$ . Поскольку величина  $v = n\hat{P}$

имеет биномиальное распределение с параметрами  $(n, P)$ , то СКО оценки  $\hat{P}_{k,1}$  можно вычислить по формуле

$$S_{k,1}^2 = M[\hat{P}_{k,1} - P]^2 = \sum_{m=0}^n [\Psi_{k,1}(m, n, p_a) - P]^2 \frac{n!}{m!(n-m)!} P^m (1-P)^{n-m}, \quad (8)$$

$$\text{где } \Psi_{k,1}(m, n, p_a) = \frac{m}{n} - \frac{m - np_a}{n(1 + q_{k-1}^2 [\sqrt{n(m - np_a)} / \sqrt{m(n-m)}])}.$$

В качестве второй оценки  $\lambda$  возьмем  $\hat{\lambda}_{k,2} = \left(1 + q_{k-1}^2 (\sqrt{n\hat{\Delta}} / \sqrt{p_a(1-p_a)})\right)^{-1}$ , здесь неизвестная дисперсия  $\sigma^2 = P(1-P)$  заменена на  $\sigma_a^2 = p_a(1-p_a)$ . Подставляя в (1) вместо  $\lambda$  оценку  $\hat{\lambda}_{k,2}$ , получим вторую адаптивную комбинированную оценку  $\hat{P}_{k,2} = \hat{P} - \hat{\lambda}_{k,2}(\hat{P} - p_a)$ . Эта оценка также является смещенной, и ее СКО вычисляется по формуле

$$S_{k,2}^2 = M[\hat{P}_{k,2} - P]^2 = \sum_{m=0}^n [\Psi_{k,2}(m, n, p_a) - P]^2 \frac{n!}{m!(n-m)!} P^m (1-P)^{n-m}, \quad (9)$$

$$\text{где } \Psi_{k,2}(m, n, p_a) = \frac{m}{n} - \frac{m - np_a}{n(1 + q_{k-1}^2 [\sqrt{n(m - np_a)} / \sqrt{p_a(1-p_a)}])}.$$

В качестве третьей оценки  $\lambda$  возьмем  $\hat{\lambda}_{k,3} = \left(1 + q_{k-1}^2 (\sqrt{n\hat{\Delta}} / 0.5)\right)^{-1}$ , где неизвестная дисперсия  $\sigma^2 = P(1-P)$  заменена на свое максимальное значение, равное 0,25. Подставляя  $\hat{\lambda}_{k,3}$  в (1), получаем третью адаптивную комбинированную оценку  $\hat{P}_{k,3} = \hat{P} - \hat{\lambda}_{k,3}(\hat{P} - p_a)$ . Эта оценка является смещенной, и ее СКО вычисляется по формуле

$$S_{k,3}^2 = M[\hat{P}_{k,3} - P]^2 = \sum_{m=0}^n [\Psi_{k,3}(m, n, p_a) - P]^2 \frac{n!}{m!(n-m)!} P^m (1-P)^{n-m}, \quad (10)$$

$$\text{где } \Psi_{k,3}(m, n, p_a) = \frac{m}{n} - \frac{m - np_a}{n(1 + q_{k-1}^2 [\sqrt{n(m - np_a)} / \sqrt{0.25}])}.$$

**4. Сравнение СКО оценок.** Для сравнения СКО оценок введем отношения

$$e_0 = \frac{D\hat{P}}{D\hat{P}} = 1, e = \frac{S^2(\lambda)}{D\hat{P}}, e_{k,1} = \frac{S_{k,1}^2}{D\hat{P}}, e_{k,2} = \frac{S_{k,2}^2}{D\hat{P}}, e_{k,3} = \frac{S_{k,3}^2}{D\hat{P}}.$$

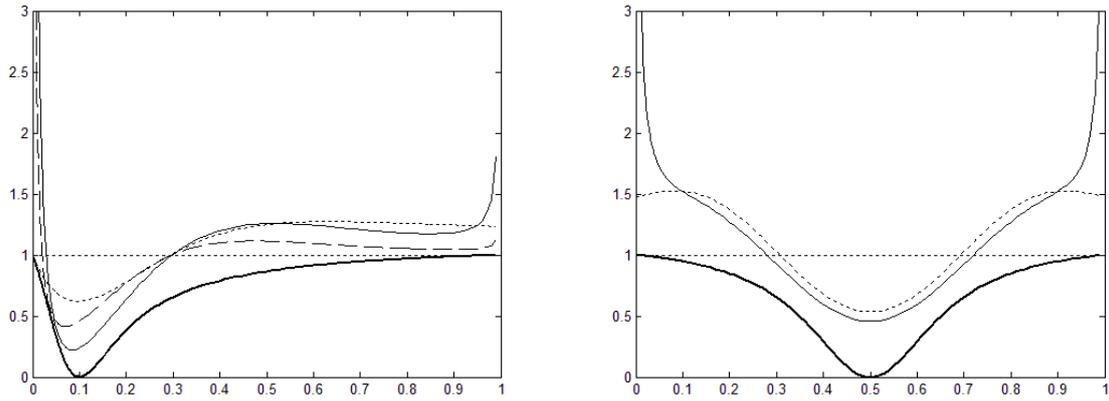


Рис. 1. Зависимость отношений  $e_{1,1}$  (точечная линия),  $e_{1,2}$  (пунктирная линия) и  $e_{1,3}$  (сплошная линия) от оцениваемой вероятности  $P$ . Жирная линия соответствует оптимальной оценке (т.е. отношению  $e$ ). Объем выборки равен 10, априорная догадка  $p_a = 0.1$  на левом графике,  $p_a = 0.5$  на правом графике. Если  $p_a = 0.5$ , то  $e_{1,2} = e_{1,3}$  и пунктирная линия совпадает со сплошной линией на правом графике

Формулы (8)–(10) позволяют вычислить значения этих отношений в зависимости от величин  $P$ ,  $k$ ,  $n$  и  $p_a$ . Отношения  $e_{k,1}$ ,  $e_{k,2}$  и  $e_{k,3}$  выделяют интервалы значений  $P$ , в которых эти отношения меньше, чем  $e_0 = 1$ . Наличие таких интервалов для  $P$  совместно с объемом наблюдений  $n$  и значениями  $p_a$  задают условия, при которых адаптивные комбинированные оценки предпочтительнее обычной оценки  $\hat{P}$  по величине СКО.

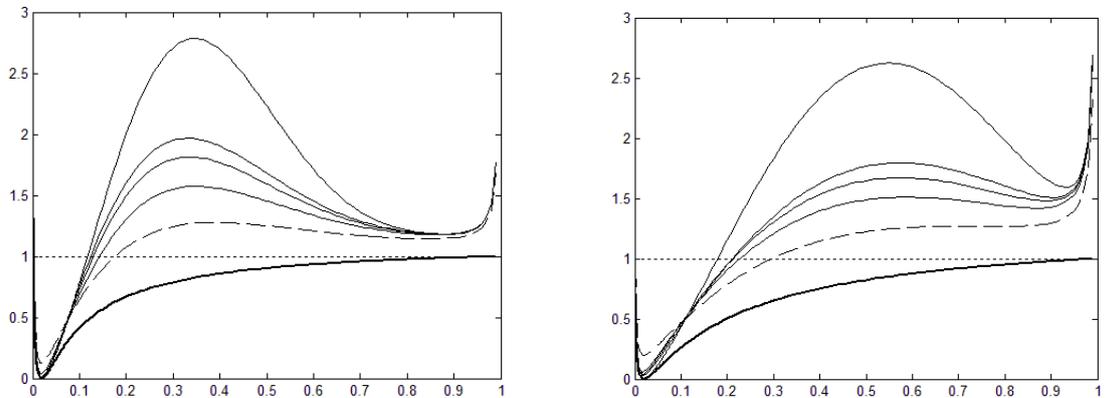


Рис. 2. Зависимость отношения  $e_{k,3}$  от оцениваемой вероятности  $P$  при различных  $k$ . Графики иллюстрируют точность оценивания малой вероятности с априорной догадкой  $p_a = 0.02$ . Жирная линия соответствует оптимальной оценке (т.е. отношению  $e$ ), пунктирная линия соответствует случаю  $k=1$ , тонкие сплошные линии отвечают случаям  $k=2$ ,  $k=3$ ,  $k=4$ ,  $k=\infty$ . Объем выборки  $n=5$  для левого графика,  $n=10$  для правого графика

Рисунки показывают, что существуют интервалы значений  $P$ , при которых комбинированные оценки точнее по СКО в сравнении с  $\hat{P}$ . С возрастанием  $k$  точность увеличивается в окрестности малых отклонений

$\Delta = P - p_a$  и уменьшается при превышении некоторых пороговых значений отклонений.

*Статья написана в рамках научного проекта (№ 8.1.98.2015), выполненного при поддержке программы «Научный фонд Томского государственного университета им. Д. И. Менделеева» в 2015 г.*

#### Литература

1. Гарайшина И. Р., Моисеева С. П., Назаров А. А. Методы исследования коррелированных потоков и специальных систем массового обслуживания. Томск: Изд-во НТЛ, 2010. 204 с.

2. Дмитриев Ю. Г., Тарасенко П. Ф. Использование априорной информации в статистической обработке экспериментальных данных // Изв. вузов. Физика. 1992. № 9. С. 136–142.

3. Dmitriev Yu., Tarassenko P., Ustinov Yu. / A. Dudin et al. (Eds.) // ITMM. 2014. CCIS 487. 2014. P. 82–90.

4. Дмитриев Ю. Г., Устинов Ю. К. Статистическое оценивание распределений вероятности с использованием дополнительной информации. Томск: Изд-во Том. ун-та, 1988. 194 с.

## **ПРИМЕНЕНИЕ МЕТОДОВ СТАТИСТИЧЕСКОГО АНАЛИЗА В ИССЛЕДОВАНИИ ВЛИЯНИЯ НЕКОТОРЫХ ФАКТОРОВ НА СЕБЕСТОИМОСТЬ СЕЛЬСКОХОЗЯЙСТВЕННОЙ ПРОДУКЦИИ В СЕЛЬСКОХОЗЯЙСТВЕННЫХ ПРЕДПРИЯТИЯХ КЕМЕРОВСКОЙ ОБЛАСТИ**

*А. И. Киселева*

*Кемеровский государственный университет, Кемерово, Россия*

Всесторонняя оценка состояния и развития животноводства важна, так как эта отрасль обеспечивает население высококачественными продуктами питания и снабжает многие отрасли промышленности (мясную, маслодельно-сыроваренную, кожевенную и др.) необходимым сырьем.

Целью работы является применение экономико-статистических методов в исследовании влияния некоторых факторов на себестоимость сельскохозяйственной продукции в сельскохозяйственных предприятиях Кемеровской области.

Исследования в данной работе проводились по 31 предприятию Кемеровской области.

В работе использованы следующие приемы и методы исследований:

- метод группировок;
- корреляционно-регрессионный анализ;
- метод наименьших квадратов (МНК).

В расчетах использовались статистические данные по Кемеровской области за период 2009–2013 гг., взятые на сайте Единой межведомствен-

ной информационно-статистической системы Федеральной службы государственной статистики [1].

Для исследования использовались различные пакеты программ статистического анализа: MS Excel, STATISTICA, SPSS, REGRESS и др.

Для характеристики объёма производства молока использовали показатель валового надоя молока. В статистической практике определяют валовой надой молока от коров, овцематок, коз, кобыл, буйволиц и некоторых других видов животных [2].

Для изучения статистических взаимосвязей применялись два метода анализа – корреляционный и регрессионный. Задачи корреляционного анализа сводятся к измерению тесноты связи между факторами, выявлению неизвестных причин связей и оценке факторов, вызывающих максимальное влияние на результат.

В качестве факторов влияния на себестоимость молока были выбраны: затраты на корма, надой и число коров.

Задача регрессионного анализа лежит в сфере установления формы зависимости, определения уравнения регрессии и его использования для оценки неизвестных значений зависимой переменной [3].

Средняя величина рассчитывалась по формуле  $x = \sum x_i / N$ .

Средняя арифметическая взвешенная, как и простая, определялась отношением общего объёма явления к объёму совокупности:  $x = \sum x_i n_i / \sum n_i$ .

Группировку сельскохозяйственных предприятий производили по численности коров (голов).

Для анализа зависимости между себестоимостью 1 ц молока, приростом и надоем и числом коров разобрана корреляционная матрица.

Аналитическая группировка сельскохозяйственных предприятий Кемеровской области по численности коров представлена в табл. 1.

Таблица 1

**Аналитическая группировка сельскохозяйственных предприятий Кемеровской области по численности коров**

№	Группы совхозов по числу коров, гол.	Число предприятий	Ср. числ., гол.	Ср. надой, кг	Ср. затраты корма, тыс. р.	Ср. с/ст. 1 ц молока, р.	Ср. полная себ-ть реализ. молока, тыс. р.
1	60–269	7	167	2359,4	825,3	447,7	1202,1
2	270–479	11	326	2528,3	1274,5	421,5	2438,4
3	480–689	4	586	2264,7	2039,5	338,5	3947
4	690–899	5	765	3090	3626,8	303,6	4106,6
5	900–1109	3	1032	3275	3951,6	253,6	5001,3
6	1110–1319	1	1314	4069	7320	278	6891
<b>Итого:</b>		<b>31</b>	<b>4190</b>	<b>17586,4</b>	<b>19037,7</b>	<b>2042,9</b>	<b>2358,4</b>
<b>Среднее значение</b>			<b>698,3</b>	<b>2931,1</b>	<b>3173,0</b>	<b>340,5</b>	<b>3931,1</b>

Из табл. 1 видно, что в тех хозяйствах, где численность коров выше, выше практически все показатели.

Модель влияния затрат на корма, надоя и числа коров на себестоимость 1 ц молока показана в табл. 2.

Таблица 2

**Модель влияния затрат на корма, надоя и числа коров на себестоимость 1 ц молока**

Параметры	$y$	$x_1$	$x_2$	$x_3$
Среднее значение	370,4	2105,3	2668,7	494,3
Медиана	345,0	1474,0	2397,0	388,0
Максимальное значение	788,0	7320,0	4219,0	1314,0
Минимальное значение	205,0	0,0	1870,0	60,0
Среднеквадратическое отклонение	132,8	1712,3	681,0	321,7
Коэффициент асимметрии	1,3	1,3	0,9	0,8
Экцесс	4,9	4,5	2,8	2,8
Jarque-Bera	13,5	11,8	4,7	3,7
Вероятность, %	0,1	0,3	9,3	15,6
Количество наблюдений	31	31	31	31

Здесь  $y$  – себестоимость 1 ц. молока, руб. (результативный признак);  $x_1$  – затраты на корма, тыс. руб. (факторный признак);  $x_2$  – надоя, кг (факторный признак);  $x_3$  – число коров, гол. (факторный признак).

Анализируя данные описательных статистик группы переменных, можно сказать, что среднее значение себестоимости 1 ц молока – 370,4 руб., средние расходы на корма – 2105,3 тыс. руб., надоя – 2668,7 кг и числа коров – 494,3. Максимальное значение себестоимости 1 ц молока – 788 руб., расходов на корма – 7320 тыс. руб., надоя – 4219 кг, числа коров – 1314. Среднеквадратическое отклонение для себестоимости 1 ц молока, расходов на корма, надоя и числа коров равны соответственно 132,8; 1712,3; 681,0 и 321,7.

Для анализа зависимости между себестоимостью 1 ц молока, приростом и надоем и числом коров разберем корреляционную матрицу (табл. 3).

Таблица 3

**Корреляционная матрица**

	$y$	$x_1$	$x_2$	$x_3$
$y$	1,000000	-0,117609	-0,470327	-0,482291
$x_1$	-0,117609	1,000000	0,497259	0,824470
$x_2$	-0,470327	0,497259	1,000000	0,539081
$x_3$	-0,482291	0,824470	0,539081	1,000000

Данные корреляционной матрицы позволяют судить о наличии зависимости между себестоимостью 1 ц молока и расходами на корма: связь обратная слабая ( $r_{xy} = -0,12$ ), между себестоимостью 1 ц молока и надоем – связь обратная умеренная ( $r_{xy} = -0,47$ ), между себестоимостью 1 ц молока и числом коров связь обратная умеренная ( $r_{xy} = -0,48$ ). Обратная связь свиде-

тельствует о том, что при увеличении одного показателя, второй будет уменьшаться. Прослеживается взаимосвязь между расходами на корма и надоем – прямая умеренная ( $r_{xy} = 0,5$ ); также существует взаимосвязь между надоем и числом голов – связь прямая заметная ( $r_{xy} = 0,54$ ) и между расходами на корма и числом коров – связь прямая очень тесная ( $r_{xy} = 0,84$ ). Это объясняется тем, что чем больше голов скота, тем выше расходы на корма. В результате анализа видно, что факторы не автокоррелированы.

Модель парной линейной регрессии представлена в табл. 4. Зависимая переменная  $y$ , метод наименьших квадратов, диапазон наблюдений 1–31, использовалось 31 наблюдение.

Уравнение регрессии имеет вид:

$$y = 0,0725x_1 - 0,0714x_2 - 0,4357x_3 + 623,6764.$$

Значения коэффициентов уравнения регрессии интерпретируются следующим образом: увеличение затрат на корма на 0,05 единицы, уменьшение надоя на 0,07 единицы и уменьшение числа голов на 0,23 единицы влечет повышение реализации мяса на 623,7 единицы.

Таблица 4

**Модель парной линейной регрессии**

Переменная	Коэффициент при соответствующей переменной	Стандартная ошибка	$t$ -значение	Вероятность отклонения гипотезы $H_0$
$x_1$	0,072500	0,017376	4,172502	0,0003
$x_2$	- 0,071391	0,029354	-2,432024	0,0219
$x_3$	- 0,435747	0,095259	-4,574359	0,0001
$C$	0,072500	0,017376	4,172502	0,0003
Коэффициент детерминации	0,571359	Среднее значение зависимости переменной		370,3871
Приспособленный коэффициент	0,523732	Стандартное отклонение зависимой переменной		132,8086
Стандартная ошибка регрессии	91,65407	Информационный критерий Акайка		11,99383
Сумма квадратов отклонений	226812.6	Критерий Шварца		12,17886
Тест на логарифм. уравнения	- 181,9044	$F$ -значение		11,99658
Статистики Дарбина–Ватсона	1,777453	Вероятность $F$ -значения		0,000036

Размер и структура переработки с/х продукции по группам представлены в табл. 5.

Таблица 5

**Размер и структура переработки с/х продукции по группам**

№	Ср. числ.		Ср. надой		Ср. затраты корма		Ср. с/ст. 1 ц молока		Ср. полная себ-ть реализ. молока	
	гол.	% к ито-гу	кг	% к ито-гу	тыс. руб.	% к ито-гу	руб.	% к ито-гу	тыс. руб.	% к ито-гу
1	167	4,0	2359,4	13,4	825,3	4,3	447,7	21,9	1202,1	5,1
2	326	7,8	2528,3	14,4	1274,5	6,7	421,5	20,6	2438,4	10,3
3	586	14,0	2264,7	12,9	2039,5	10,7	338,5	16,6	3947	16,7
4	765	18,3	3090	17,6	3626,8	19,1	303,6	14,9	4106,6	17,4
5	1032	24,6	3275	18,6	3951,6	20,8	253,6	12,4	5001,3	21,2
6	1314	31,4	4069	23,1	7320	38,5	278	13,6	6891	29,2
<b>все го</b>	<b>4190</b>	<b>100</b>	<b>17586,4</b>	<b>100</b>	<b>19037,7</b>	<b>100</b>	<b>2042,9</b>	<b>100</b>	<b>23586,4</b>	<b>100</b>

Из данных таблицы видно, что большим размером численностью голов выделяется шестая группа (31,4 %), но в этой группе всего одно предприятие; на втором месте пятая группа – 24,6 % (три предприятия) и третье место занимает четвертая группа – 18,3 % (пять предприятий).

По надоям молока шестая группа также находится на первом месте – 23,1 %; на втором и третьем местах – пятая и четвертая группы соответственно.

Средняя стоимость центнера молока наибольшая в первой группе – почти 22 %, следом идет вторая группа – 20,6 %, а затем третья группа – 16,6 %, далее вниз по порядку.

Так как затраты на корма наибольшие в шестой группе (38,5 %), то очевидно, что и себестоимость в этой группе будет наибольшая (29,2 %).

Аналогично видим из таблицы, что чем меньше затрат, тем меньше себестоимость реализованного молока.

Таким образом, с помощью экономико-статистических методов было исследовано влияние факторов на себестоимость сельскохозяйственной продукции в сельскохозяйственных предприятиях Кемеровской области. Построены модель влияния затрат на корма, надоя и числа коров на себестоимость 1 ц молока; корреляционная матрица зависимости между себестоимостью 1 ц молока и расходами на корма; модель парной линейной регрессии.

## Литература

1. Единая межведомственная информационно-статистическая система [Электронный ресурс]. URL: <http://www.fedstat.ru/> (дата обращения: 03.12.2014).
2. Афанасьев В. Н. Статистика сельского хозяйства / В. Н. Афанасьев, А. И. Маркова. М.: Финансы и статистика, 2011. 272 с.
3. Елисеева И. И. Общая теория статистики: учебник для вузов / И. И. Елисеева, М. Н. Юзбашев. 5-е изд. М.: Финансы и статистика. 2008. 480 с.

# КОМБИНИРОВАННЫЕ ОЦЕНКИ ДОЛИ

*Т. О. Кошечая, Ю. Г. Дмитриев*

*Национальный исследовательский*

*Томский государственный университет, Томск, Россия*

**1. Введение.** Разнообразие проблем в статистической обработке экспериментальных данных включает в себя оценку неизвестной доли объектов в генеральной совокупности с заданным значением признака. Интерес связан с широким практическим применением, например, в статистическом анализе точности и стабильности технологических процессов, приемочном контроле качества продукции, регулировании технологических процессов, оценке и контроле надежности. Для сокращения объема дорогостоящих экспериментальных данных или повышения точности оценивания при фиксированном объеме наблюдений, разумно привлекать дополнительную информацию, которая может быть разнообразной и иметь различные источники поступления.

В данной работе предлагаются комбинированные оценки доли, учитывающие совместно обычную эмпирическую оценку и априорную догадку в форме некоторого значения искомой доли, которое высказывается исследователем или экспертом на основании своего опыта и знаний. Анализируются среднеквадратические ошибки таких оценок при конечном объеме наблюдений. Указываются условия, при которых комбинированные оценки предпочтительнее обычной оценки.

**2. Структура оценки.** Пусть имеется конечная генеральная совокупность объема  $N$  и случайным образом производится выборка без возвращения объема  $n$  из этой генеральной совокупности. Под  $P$  будем понимать долю объектов в генеральной совокупности с заданным значением признака. Пусть  $p_a$  – априорная догадка, которая выступает в качестве возможного значения доли  $P$ . Требуется оценить долю объектов в генеральной совокупности с заданным значением признака, учитывая совместно эмпирическую оценку  $\hat{P} = n^{-1} \sum_{i=1}^n I_A(X_i)$  и  $p_a$ , здесь  $I_A(X_i)$  – индикаторная функция, равная 1, если у объекта  $X_i$  признак принимает значение  $A$ .

Следуя работам [1,2], рассмотрим комбинированную оценку

$$\begin{aligned}\hat{P}_\lambda &= (1 - \lambda)\hat{P} + \lambda p_a, \\ \hat{P}_\lambda &= \hat{P} - \lambda(\hat{P} - p_a),\end{aligned}\tag{1}$$

где оптимальный весовой коэффициент  $\lambda$  выбран из условия минимума среднеквадратической ошибки (СКО)  $S^2(\lambda) = M[\hat{P}_\lambda - P]^2$  и определяется выражением

$$\lambda = (1 + n\Delta^2/\sigma^2)^{-1}. \quad (2)$$

Здесь  $\sigma^2 = P(1-P)$ ,  $\Delta = P - p_a$  величина отклонения априорной догадки от истинного значения искомой доли. Весовой коэффициент  $\lambda$  изменяется в пределах  $0 < \lambda \leq 1$  и показывает, какое влияние оказывает каждое из слагаемых в комбинированной оценке.

В соответствии с (1) минимум СКО представим в виде

$$S^2(\lambda) = M[\hat{P} - P]^2 - \frac{[M(\hat{P} - P)]^2}{D\hat{P} + (P - p_a)^2} = D\hat{P} - \frac{(D\hat{P})^2}{D\hat{P} + \Delta^2} = (1 - \lambda)\sigma^2 / n, \quad (3)$$

где дисперсия  $D\hat{P} = P(1-P)(N-n)/(N-1)n$  характеризует точность оценки  $\hat{P}$ , а коэффициент  $(1 - \lambda)$  показывает, во сколько раз уменьшается СКО комбинированной оценки (1) по сравнению с  $\hat{P}$ .

Величина  $(D\hat{P})^2/(D\hat{P} + \Delta^2)$  в (3) задает выигрыш в точности оценивания за счет привлечения априорной догадки  $p_a$ , если исходной оценкой является  $\hat{P}$ .

**3. Адаптивные оценки.** Как правило, оптимальный коэффициент  $\lambda$  в (2) неизвестен. А это, в свою очередь, затрудняет использование комбинированной оценки (1) на практике. Выходом из этого положения является построение адаптивных комбинированных оценок доли путем замены неизвестного значения  $\lambda$  на его статистическую оценку. В связи с этим возникает вопрос, при каких условиях адаптивные комбинированные оценки являются предпочтительнее обычной оценки  $\hat{P}$  по величине СКО?

Заменим неизвестное  $P$  на  $\hat{P}$  и в качестве первой оценки  $\lambda$  возьмем

$$\hat{\lambda}_1 = (1 + n\hat{\Delta}^2/\hat{\sigma}^2)^{-1},$$

где  $\hat{\Delta} = \hat{P} - p_a$  – оценка отклонения, а оценка дисперсии имеет вид  $\hat{\sigma}^2 = \hat{P}(1 - \hat{P})(N - n)/(N - 1)$ . Подставляя  $\hat{\lambda}_1$  в (1), получаем первую адаптивную комбинированную оценку  $\hat{P}_1 = \hat{P} - \hat{\lambda}_1(\hat{P} - p_a)$ . Случайная величина  $n\hat{P}$  имеет гипергеометрическое распределение с параметрами  $(n, P, N)$ , поэтому СКО оценки  $\hat{P}_1$  можно вычислить по формуле

$$S_1^2 = M[\hat{P}_1 - P]^2 = \sum_{k=0}^n [\Psi_1(k) - P]^2 \frac{(Np)![N(1-p)]!n!(N-n)!}{k!(Np-k)!(n-k)![N(1-p)-(n-k)]!N!}, \quad (4)$$

где

$$\Psi_1(k) = \frac{k[(n-k)(N-n)p_a + (N-1)(k-np_a)^2]}{k(n-k)(N-n) + n(N-1)(k-np_a)^2}.$$

В качестве второй оценки  $\lambda$  возьмем

$$\hat{\lambda}_2 = (1 + 4n\hat{\Delta}^2(N-1)/(N-n))^{-1},$$

где неизвестную дисперсию  $\sigma^2 = P(1 - P)$  заменим её максимальным значением, равным  $0,25(N - n)/(N - 1)$ . Подставляя  $\hat{\lambda}_2$  в (1), получаем вторую адаптивную комбинированную оценку  $\hat{P}_2 = \hat{P} - \hat{\lambda}_2(\hat{P} - p_a)$ . СКО оценки вычисляется по формуле

$$S_2^2 = M[\hat{P}_2 - P]^2 = \sum_{k=0}^n [\Psi_2(k) - P]^2 \frac{(Np)! [N(1-p)]! n! (N-n)!}{k! (Np-k)! (n-k)! [N(1-p) - (n-k)]! N!}, \quad (5)$$

где

$$\Psi_2(k) = \frac{k}{n} - \frac{(N-n)(k - np_a)}{n(N-n) + 4(N-1)(k - np_a)^2}.$$

В качестве третьей оценки  $\lambda$  возьмем

$$\hat{\lambda}_3 = \left[ 1 + n\hat{\Delta}^2/\hat{\sigma}^2 \left( 1 - \frac{1}{1 + \hat{\Delta}^2/\hat{\sigma}^2} \right) \right]^{-1},$$

где оценка дисперсии имеет вид  $\hat{\sigma}^2 = \hat{P}(1 - \hat{P})(N - 1)/(N - n)$ .

Подставляя  $\hat{\lambda}_3$  в (1), получаем третью адаптивную комбинированную оценку  $\hat{P}_3 = \hat{P} - \hat{\lambda}_3(\hat{P} - p_a)$ . СКО оценки  $\hat{P}_3$  можно вычислить по формуле

$$S_3^2 = M[\hat{P}_3 - P]^2 = \sum_{k=0}^n [\Psi_3(k) - P]^2 \frac{(Np)! [N(1-p)]! n! (N-n)!}{k! (Np-k)! (n-k)! [N(1-p) - (n-k)]! N!}, \quad (6)$$

где

$$\Psi_3(k) = \frac{k^2(n-k)^2(N-n)^2 p_a + k p_a (n-k)(k - np_a)^2 (N-n)(N-1) + (k - np_a)^4 (N-1)^2}{[k^2(n-k)^2(N-n)^2 + k(n-k)(k - np_a)^2 (N-n)(N-1) + n(k - np_a)^4 (N-1)^2]}.$$

В качестве четвертой оценки  $\lambda$  возьмем

$$\hat{\lambda}_4 = \left[ 1 + 4n\hat{\Delta}^2(N-1)/(N-n) \left( 1 - \frac{1}{1 + 4\hat{\Delta}^2(N-1)/(N-n)} \right) \right]^{-1},$$

где неизвестную дисперсию, также как и для второй оценки, заменим её максимальным значением. Подставляя  $\hat{\lambda}_4$  в (1), получаем четвертую адаптивную комбинированную оценку  $\hat{P}_4 = \hat{P} - \hat{\lambda}_4(\hat{P} - p_a)$ . СКО оценки вычисляется по формуле

$$S_4^2 = M[\hat{P}_4 - P]^2 = \sum_{k=0}^n [\Psi_4(k) - P]^2 \frac{(Np)! [N(1-p)]! n! (N-n)!}{k! (Np-k)! (n-k)! [N(1-p) - (n-k)]! N!}, \quad (7)$$

где

$$\Psi_4(k) = \frac{k}{n} - \frac{(N-n)(k - np_a) [n^2(N-n) + 4(k - np_a)^2(N-1)]}{[n^3(N-n)^2 + 4n(k - np_a)^2(N-1)(N-n) + 16(k - np_a)^4(N-1)^2]}.$$

В отличие от  $\hat{P}$  оценки  $\hat{P}_1, \hat{P}_2, \hat{P}_3, \hat{P}_4$  являются смещенными (рис. 1, 2).

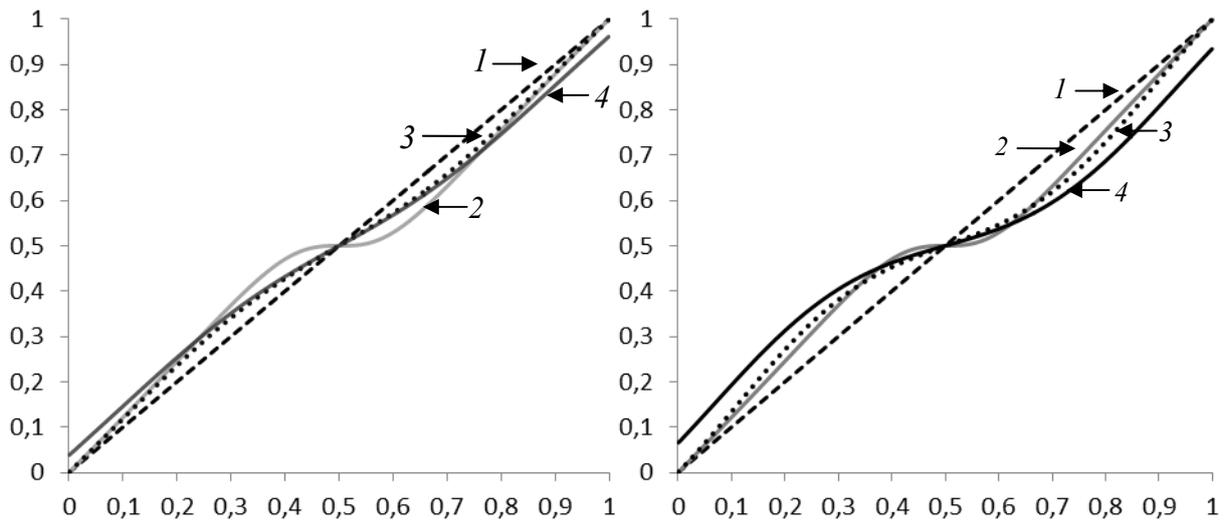


Рис. 1. Зависимость среднего значения оценки от  $P$ ,  $n = 10$ ,  $p_a = 0,5$ ,  $N = 40$   
 ( $1 - \hat{P}$ ,  $2 - \hat{P}_\lambda$ ,  $3 - \hat{P}_1$ ,  $4 - \hat{P}_2$ )

Рис. 2. Зависимость среднего значения оценки от  $P$ ,  $n = 10$ ,  $p_a = 0,5$ ,  $N = 40$   
 ( $1 - \hat{P}$ ,  $2 - \hat{P}_\lambda$ ,  $3 - \hat{P}_3$ ,  $4 - \hat{P}_4$ )

Для сравнения СКО оценок введем отношения

$$E_0 = \frac{D\hat{P}}{D\hat{P}} = 1, E = \frac{S^2}{D\hat{P}}, E_1 = \frac{S_1^2}{D\hat{P}}, E_2 = \frac{S_2^2}{D\hat{P}}, E_3 = \frac{S_3^2}{D\hat{P}}, E_4 = \frac{S_4^2}{D\hat{P}}.$$

Формулы (4)–(7) позволяют рассчитать значения этих отношений в зависимости от значений  $P$ ,  $N$ ,  $n$  и  $p_a$ . Величины  $E_1$ ,  $E_2$ ,  $E_3$  и  $E_4$  позволяют выделить интервалы значений  $P$ , при которых эти отношения меньше единицы. Наличие таких интервалов для  $P$  совместно с объемом генеральной совокупности  $N$ , объемом выборки  $n$  и значениями  $p_a$  указывают условия, при которых адаптивные комбинированные оценки предпочтительнее обычной оценки  $\hat{P}$  по величине СКО (рис. 3, 4).

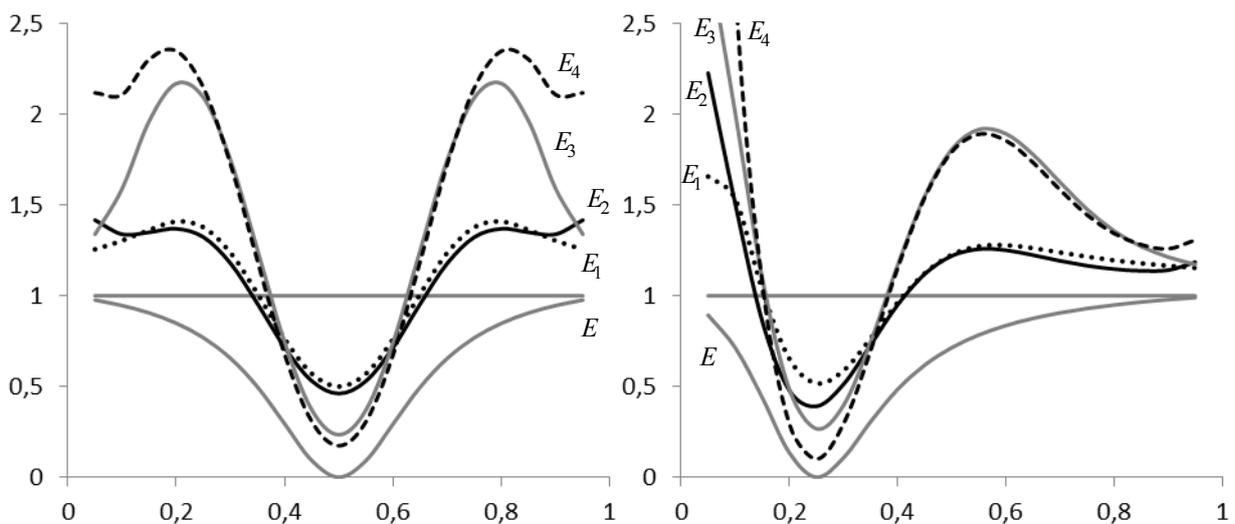


Рис. 3. Зависимость СКО оценки от  $P$ ,  
 $n = 10$ ,  $p_a = 0,5$ ,  $N = 20$

Рис. 4. Зависимость СКО оценки от  $P$ ,  
 $n = 10$ ,  $p_a = 0,25$ ,  $N = 20$

Сравнение оценок  $\hat{P}_1, \hat{P}_2$  показывает, что имеются области значений  $P$ , в которых  $\hat{P}_2$  предпочтительнее по СКО обычной оценки  $\hat{P}$  и  $\hat{P}_1$ . Так, из рис. 3 следует, что в интервале  $(0,341; 0,659)$   $E_2 < E_1$ . Также сравнение оценок  $\hat{P}_3, \hat{P}_4$  в интервале  $(0,368; 0,632)$  показывает, что  $\hat{P}_4$  предпочтительнее по СКО обычной оценки  $\hat{P}$  и  $\hat{P}_3$ , а именно,  $E_4 < E_3$ . В интервале  $(0,392; 0,608)$  оценка  $\hat{P}_4$  предпочтительнее оценок  $\hat{P}_1, \hat{P}_2, \hat{P}_3$ .

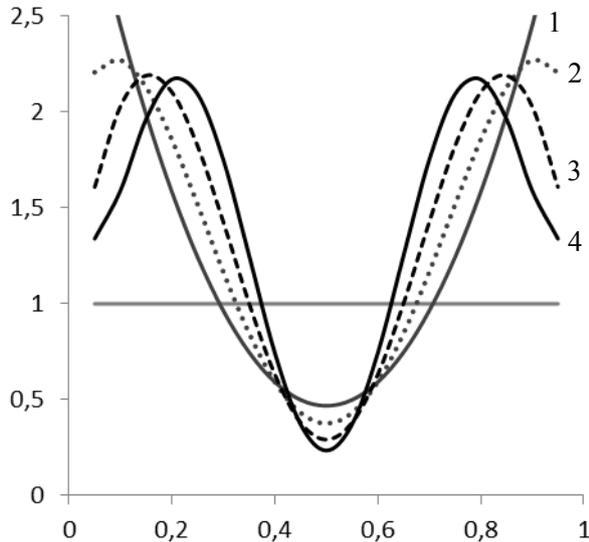


Рис. 5. Зависимость СКО оценки  $\hat{P}_3$  от  $P$   
 $p_a = 0,5, N = 20$   
 $(1 - n = 4, 2 - n = 6, 3 - n = 8, 4 - n = 10)$

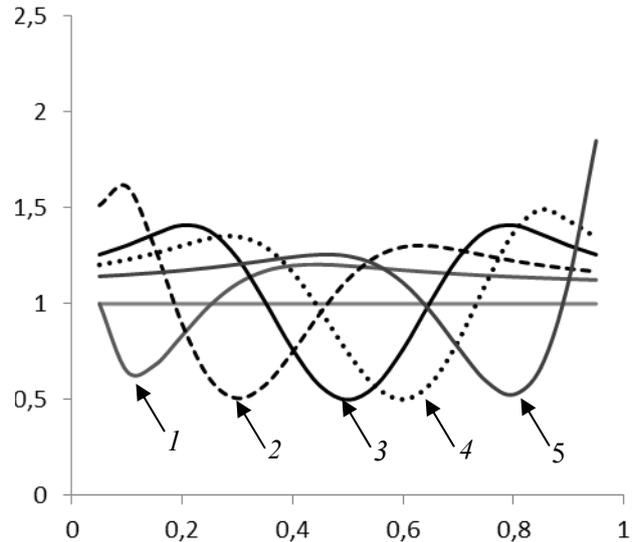


Рис. 6. Зависимость СКО оценки  $\hat{P}_1$  от  $P$ ,  
 $n = 10, N = 20$   
 $(1 - p_a = 0,1, 2 - p_a = 0,3, 3 - p_a = 0,5,$   
 $4 - p_a = 0,6, 5 - p_a = 0,8)$

Рисунок 5 показывает, что при увеличении объема наблюдений  $n$  выигрыш по СКО увеличивается, а интервал для  $P$ , где оценка  $\hat{P}_3$  предпочтительнее по СКО обычной оценки  $\hat{P}$ , уменьшается. Так, при  $n=10$  и  $P=0,5$   $E_3=0,235$ , а интервал значений для  $P$  равен  $(0,341; 0,659)$ . При  $n=4$  и  $P=0,5$   $E_3=0,521$  и интервал значений для  $P$  равен  $(0,197; 0,803)$ .

Из рис. 6 видно, что при изменении значений априорной догадки меняются выигрыш по СКО и интервалы для  $P$ , где оценка  $\hat{P}_1$  предпочтительнее по СКО обычной оценки  $\hat{P}$ . Так, при  $p_a=0,5$  максимальный выигрыш по СКО равен  $E_1=0,500$  и интервал для  $P$  равен  $(0,352; 0,648)$ . При  $p_a=0,8$   $E_1=0,527$  и интервал для  $P$  равен  $(0,638; 0,890)$ .

*Статья написана в рамках научного проекта (№ 8.1.98.2015), выполненного при поддержке программы «Научный фонд Томского государственного университета им. Д. И. Менделеева» в 2015 г.*

#### Литература

1. Дмитриев Ю. Г., Тарасенко П. Ф. Использование априорной информации в статистической обработке экспериментальных данных // Изв. вузов. Физика. 1992. № 9. С. 136–142.

## РЕЛЕЙНО-ГИСТЕРЕЗИСНОЕ УПРАВЛЕНИЕ ПОТОКОМ МОМЕНТОВ ПРОДАЖИ СКОРОПОРТЯЩЕЙСЯ ПРОДУКЦИИ

*К. И. Лившиц, Е. С. Ульянова*

*Национальный исследовательский*

*Томский государственный университет, Томск, Россия*

### 1. Математическая модель задачи

Модели управления запасами с ограниченным сроком годности интенсивно изучаются в последние годы. За это время появилось несколько обзорных статей по данной теме, например, обзоры S. K. Goyal, B. C. Giri [1], M. Bakker, J. Riezebos и R. H. Teunter [2]. Укажем еще, например, на работы V. K. Mishra и L. S. Singh [3], R. Begum, S. K. Sahu, R. R. Sahoo [4].

В настоящей работе задача поступления и сбыта скоропортящейся продукции рассматривается при следующих предположениях. Считается, что продукция поступает с постоянной скоростью  $c$ , так что за время  $t$  поступает  $ct$  единиц продукции. При хранении продукция непрерывно портится. Пусть  $S(t)$  – количество продукции в момент времени  $t$ . Тогда потери за малое время  $\Delta t$  равны  $kS(t)\Delta t$ . Будем считать далее, что величины покупок – независимые случайные величины с плотностью распределения  $\varphi(x)$ , средним значением  $M\{x\} = a$  и вторым моментом  $M\{x^2\} = a_2$ . Моменты продаж образуют пуассоновский поток, интенсивность которого  $\lambda$  зависит от цены продажи  $b$ . Считается, что интенсивность потока продаж  $\lambda$  монотонно убывает с ростом цены  $b$ .

Предполагается, что управление продажами осуществляется следующим образом. Устанавливаются два пороговых значения допустимого запаса продукции  $S_1$  и  $S_2$ , причем  $S_2 > S_1$ . В области  $S < S_1$  назначается цена продажи  $b_0$ , в области  $S > S_2$  назначается цена продажи  $b_1 < b_0$ . В области же  $S_1 \leq S \leq S_2$  назначается цена  $b = b_0$  или  $b = b_1$  в зависимости от того как процесс  $S(t)$  вошел в эту область. Если он вошел в нее через порог  $S_1$  снизу вверх, то остается  $b = b_0$ , если же он вошел в эту область через порог  $S_2$  сверху вниз, то остается  $b = b_1$ . Таким образом, значение  $b = b_1$  устанавливается при достижении запасом  $S(t)$  значения  $S_2$  и оканчивается при уменьшении запаса до значения  $S_1$ . Область  $S_1 \leq S \leq S_2$  и представляет собой область гистерезиса в управлении запасом продукции. В соответствии с этим текущая интенсивность потока моментов продаж имеет вид

$$\lambda(S) = \begin{cases} \lambda_0, & S < S_1, \\ \lambda_0 \text{ или } \lambda_1, & S_1 \leq S \leq S_2, \\ \lambda_1, & S > S_2 \end{cases} \quad (1)$$

Естественно считать, что  $C - \lambda_0 a > 0$  и  $C - \lambda_1 a < 0$ . Наконец, возможна ситуация, когда текущий спрос не может быть удовлетворен полностью. В этом случае считается, что  $S(t) < 0$ . Основная цель настоящей работы состоит в нахождении плотности распределения количества продукции для данной модели при некоторых дополнительных предположениях.

Обозначим

$$P_i(S, t) ds = \Pr\{S \leq S(t) < S + dS, \lambda(t) = \lambda_i\}, \quad i = 0, 1. \quad (2)$$

**Теорема 1.** Если  $P_i(S, t)$  дифференцируемы по  $t$ ,  $SP_i(S, t)$  дифференцируемы по  $S$ , то функции  $P_i(S, t)$  удовлетворяют системе уравнений Колмогорова:

$$\frac{\partial P_1(S, t)}{\partial t} = -\lambda_1 P_1(S, t) - \frac{\partial}{\partial S}((c - kS)P_1(S, t)) + \lambda_1 \int_0^{\infty} P_1(S + x)\varphi(x)dx, \quad S \geq S_1, \quad (3)$$

$$\begin{aligned} \frac{\partial P_0(S, t)}{\partial t} = & -\lambda_0 P_0(S, t) - \frac{\partial}{\partial S}((c - kS)P_0(S, t)) + \\ & + \lambda_0 \int_S^{S_2} P_0(x, t)\varphi(x - S)dx, \quad S_1 < S < S_2, \end{aligned} \quad (4)$$

$$\begin{aligned} \frac{\partial P_0(S, t)}{\partial t} = & -\lambda_0 P_0(S, t) - \frac{\partial}{\partial S}((c - kSI(S))P_0(S, t)) + \\ & + \lambda_0 \int_S^{S_2} P_0(x, t)\varphi(x - S)dx + \lambda_1 \int_{S_1}^{\infty} P_1(x)\varphi(x - S)dx, \quad S < S_1, \end{aligned} \quad (5)$$

где  $I(x)$  – единичная функция.

Решение системы уравнений (3)–(5) должно, очевидно, удовлетворять условию нормировки

$$\int_{S_1}^{\infty} P_1(S, t)dS + \int_{-\infty}^{S_2} P_0(S, t)dS = 1, \quad (6)$$

и функция  $P_0(S, t)$  должна быть непрерывна в точке  $S_1$ :

$$P_0(S_1 + 0, t) = P_0(S_1 - 0, t). \quad (7)$$

Безусловная плотность распределения  $P(S, t)$  количества продукции будет иметь вид

$$P(S, t) = \begin{cases} P_1(S, t), & S > S_2, \\ P_1(S, t) + P_0(S, t), & S_1 \leq S \leq S_2, \\ P_0(S, t), & S < S_1. \end{cases} \quad (8)$$

## 2. Экспоненциальное распределение величины продажи

Рассмотрим в качестве примера простейший случай, когда продажи имеют экспоненциальное распределение  $\varphi(S) = \frac{1}{a} \exp(-\frac{S}{a})$ . Обозначим

$$P_i(S) = \lim_{t \rightarrow \infty} P_i(S, t). \quad (9)$$

В стационарном режиме при  $t \rightarrow \infty$  уравнения (3) – (5) принимают вид

$$\lambda_1 P_1(S) + \frac{d}{dS} ((c - kS) P_1(S)) - \frac{\lambda_1}{a} e^{\frac{S}{a}} \int_S^{\infty} P_1(x) e^{-\frac{x}{a}} dx = 0, \quad S > S_1, \quad (10)$$

$$\lambda_0 P_0(S) + \frac{d}{dS} ((c - kS) P_0(S)) - \frac{\lambda_0}{a} e^{\frac{S}{a}} \int_S^{S_2} P_0(x) e^{-\frac{x}{a}} dx = 0, \quad S_1 \leq S \leq S_2, \quad (11)$$

$$\lambda_0 P_0(S) + \frac{d}{dS} ((c - kSI(S)) P_0(S)) - \frac{\lambda_0}{a} e^{\frac{S}{a}} \int_S^{S_2} P_0(x) e^{-\frac{x}{a}} dx - \frac{\lambda_1}{a} e^{\frac{S}{a}} \int_{S_1}^{\infty} P_1(x) e^{-\frac{x}{a}} dx = 0. \quad (12)$$

С учетом граничного условия  $P_0(-\infty) = 0$  в области  $S \leq 0$  решение уравнения (12) имеет вид

$$P_0(S) = D e^{\frac{c - \lambda_0 a S}{ca}}, \quad (13)$$

а в области  $0 < S < S_1$  решение уравнения (12)

$$P_0(S) = \left[ W_1 + W_2 \int_0^S e^{-\frac{x}{a}} (c - kx)^{\frac{\lambda_0}{k}} dx \right] e^{\frac{S}{a}} (c - kS)^{\frac{\lambda_0 - 1}{k}}. \quad (14)$$

Условие непрерывности решения в точке  $S = 0$  дает  $D = W_1 c^{\frac{\lambda_0 - 1}{k}}$ . Из уравнения (12) следует, что в точке  $S = 0$  должно выполняться условие

$$cP_0'(0 + 0) - kP_0(0 + 0) = cP_0'(0 - 0).$$

Откуда  $W_2 = 0$ . Таким образом, при  $0 < S < S_1$

$$P_0(S) = D e^{\frac{S}{a}} \left(1 - \frac{k}{c} S\right)^{\frac{\lambda_0 - 1}{k}}. \quad (15)$$

Решение уравнения (11) будем искать в виде

$$P_0(S) = \left[ W_1 + W_2 \int_{S_1}^S e^{-\frac{x}{a}} \left(1 - \frac{k}{c} x\right)^{\frac{\lambda_0}{k}} dx \right] e^{\frac{S}{a}} \left(1 - \frac{k}{c} S\right)^{\frac{\lambda_0 - 1}{k}}. \quad (16)$$

Условие непрерывности в точке  $S_1$  (7) дает  $W_1 = D$ . Далее, решение (16) должно удовлетворять исходному уравнению (11). Откуда

$$W_2 = -D \left[ a e^{-\frac{S_2}{a}} \left(1 - \frac{k}{c} S_2\right)^{\frac{\lambda_0}{k}} + \int_{S_1}^{S_2} e^{-\frac{x}{a}} \left(1 - \frac{k}{c} x\right)^{\frac{\lambda_0}{k}} dx \right]^{-1}. \quad (17)$$

Наконец, решение уравнения (10) с учетом, что в рассматриваемой модели всегда количество товара  $S \leq \frac{c}{k}$ , имеет вид

$$P_1(S) = Ae^{\frac{S}{a}} \left(1 - \frac{k}{c} S\right)^{\frac{\lambda_1-1}{k}}. \quad (18)$$

Связь между постоянными  $A$  и  $D$  находится из условия, что совокупность построенных решений должна удовлетворять уравнению (12). Откуда

$$A = -ae^{-\frac{S_1}{a}} \left(1 - \frac{k}{c} S_1\right)^{\frac{\lambda_1}{k}} W_2. \quad (19)$$

Последняя неопределенная постоянная  $D$  определяется из условия нормировки (6).

Таким образом, плотность распределения количества продукции определяется соотношениями (13), (15), (16), (18), а входящие в эти соотношения постоянные связаны между собой соотношениями (18), (19) и (6).

При  $S_2 = S_1$  – релейное управление ценой продажи плотность распределения количества продукции  $P(S)$  принимает вид

$$P(S) = \begin{cases} De^{\frac{c-\lambda_0 a}{ca} S}, & S < 0, \\ D \left(1 - \frac{k}{c} S\right)^{\frac{\lambda_0-1}{k}} e^{\frac{S}{a}}, & 0 \leq S \leq S_1, \\ D \left(1 - \frac{k}{c} S_0\right)^{\frac{\lambda_0-\lambda_1}{k}} \left(1 - \frac{k}{c} S\right)^{\frac{\lambda_1-1}{k}} e^{\frac{S}{a}}, & S_1 < S \leq \frac{c}{k}, \end{cases} \quad (20)$$

где постоянная  $D$  определяется условием нормировки.

### 3. Диффузионная аппроксимация процесса производства и сбыта при релейном управлении ценой продажи

В общем случае найти решение системы уравнений (3)–(5) не удастся даже в стационарном режиме. Поэтому представляет интерес построение приближенных решений уравнений. Рассмотрим, далее, случай релейного управления ценой продажи, когда порог  $S_2 = S_1$ . Будем предполагать, что скорость производства  $c = CN$ , интенсивности потоков покупок  $\lambda_0 = \Lambda_0 N$ ,  $\lambda_1 = \Lambda_1 N$ , порог  $S_1 = S_0 N$ , где  $N \gg 1$ . Обозначив  $\varepsilon^2 = 1/N$ , введем процесс  $\xi(t) = \varepsilon^2 S(t)$ . Пусть  $h(z, t) = \frac{\partial \Pr\{\xi(t) < z\}}{\partial z}$ . Можно показать, что

при  $\varepsilon \ll 1$  плотность распределения  $h(z, t)$  удовлетворяет уравнению

$$\frac{\partial h(z, t)}{\partial t} = -\frac{\partial}{\partial z} [(C - \Lambda_1 a - kz)h(z, t)] + \frac{\Lambda_1 a_2}{2} \varepsilon^2 \frac{\partial^2 h(z, t)}{\partial z^2} \quad (21)$$

в области  $z \geq S_0$  и уравнению

$$\frac{\partial h(z,t)}{\partial t} = -\frac{\partial}{\partial z}[(c - \Lambda_0 a - kzI(z))h(z,t)] + \frac{\Lambda_0 a_2}{2} \varepsilon^2 \frac{\partial^2 h(z,t)}{\partial z^2} \quad (22)$$

в области  $z \leq S_0$ . В стационарном режиме при  $t \rightarrow \infty$  плотность распределения  $h(z) = \lim_{t \rightarrow \infty} h(z,t)$  имеет вид

$$h(z) = \begin{cases} Ae^{\frac{(c-\Lambda_0 a)^2}{\Lambda_0 a_2 \varepsilon^2 k} z} e^{\frac{2(c-\Lambda_0 a)}{\Lambda_0 a_2 \varepsilon^2} z}, & z < 0, \\ Ae^{\frac{(kz+c-\Lambda_0 a)^2}{\Lambda_0 a_2 \varepsilon^2 k}}, & 0 \leq z \leq s_0, \\ Be^{\frac{(kz+c-\Lambda_1 a)^2}{\Lambda_1 a_2 \varepsilon^2 k}}, & z > s_0, \end{cases} \quad (23)$$

где постоянные  $A, B$  определяются условием нормировки и некоторым дополнительным условием в точке  $S_0$ .

### Заключение

В работе получены уравнения, определяющие плотность распределения количества скоропортящейся продукции при непрерывном ее поступлении и релейно-гистерезисным управлении интенсивностью потока продаж. Получены решения этих уравнений в случае экспоненциального распределения величин продаж и построена диффузионная аппроксимация процесса производства и сбыта продукции при релейном управлении интенсивностью продаж. Аналогично вышеизложенному могут быть исследованы и другие модели управления производством и сбытом скоропортящейся продукции, например модель с релейно-гистерезисным управлением скоростью производства.

### Литература

1. Goyal S. K., Giri B. C. // European Journal of Operational Research. 2001. Vol. 134 (1). P. 1–16.
2. Bakker M., J. Riezebos J., Teunter R. H. // European Journal of Operational Research. 2012. Vol. 221. P. 275–284.
3. Mishra V. K. // Journal of Industrial Engineering and Management. 2013. Vol. 6(2). P. 495–506.
4. Begum R., Sahu S. K., Sahoo R. R. // British Journal of Applied Science & Technology. 2012. Vol. 2(2). P. 112–131.

# ПЛОТНОСТЬ ВЕРОЯТНОСТЕЙ ДЛИТЕЛЬНОСТИ ИНТЕРВАЛА МОДУЛИРОВАННОГО МАР-ПОТОКА СОБЫТИЙ В УСЛОВИЯХ НЕПРОДЛЕВАЮЩЕГОСЯ МЁРТВОГО ВРЕМЕНИ

*Л. А. Нежелская*

*Национальный исследовательский*

*Томский государственный университет, Томск, Россия*

## 1. Введение

Математические модели теории массового обслуживания широко используются для описания реальных физических, технических и других процессов и систем. Интенсивное развитие компьютерной техники и информационных технологий во многом определило важную сферу приложений теории массового обслуживания – проектирование и создание информационно-вычислительных сетей, телекоммуникационных сетей и т.д. Всё это послужило стимулом к созданию адекватных математических моделей реальных информационных потоков, функционирующих в цифровых сетях интегрального обслуживания, так называемых дважды стохастических потоков событий.

Данная статья является непосредственным развитием исследований, проведённых в [1–7]. При исследовании потоков событий можно сформулировать два класса задач: 1) оценивание состояний потока событий [1–3]; 2) оценивание параметров потока [4–7].

Одним из искажающих факторов при оценке состояний и параметров потока выступает мёртвое время регистрирующих приборов [8], которое порождается зарегистрированным событием. Другие события, которые наступают в течение периода мёртвого времени, недоступны наблюдению (теряются). Этот период продолжается некоторое фиксированное время (непродлевающееся мёртвое время). В качестве примера приведём CSMA/CD – протокол случайного множественного доступа с обнаружением конфликта, широко используемый в компьютерных сетях. В момент регистрации (обнаружения) конфликта на входе некоторого узла сети рассылается сигнал «заглушки» («пробки»); в течение времени рассылки сигнала «заглушки» заявки, поступившие в данный узел сети, получают отказ в обслуживании и направляются в источник повторных вызовов. Здесь время, в течение которого узел сети закрыт для обслуживания заявок, поступающих в него после обнаружения конфликта, можно трактовать как мёртвое время прибора, регистрирующего конфликт в узле сети.

В данной статье получен явный вид плотности вероятностей длительности интервала между соседними событиями в модулированном МАР-потоке с непродлевающимся мёртвым временем. Явный вид плотности вероятностей необходим для решения задачи оценивания параметров потока и длительности мёртвого времени.

## 2. Постановка задачи

Рассматривается модулированный МАР-поток событий с интенсивностью, представляющей собой кусочно-постоянный стационарный случайный процесс  $\lambda(t)$  с двумя состояниями:  $\lambda(t)=\lambda_1$  либо  $\lambda(t)=\lambda_2$  ( $\lambda_1 > \lambda_2 \geq 0$ ). Длительность пребывания процесса  $\lambda(t)$  в  $i$ -м состоянии,  $i=1,2$ , определяется двумя случайными величинами: первая случайная величина распределена по экспоненциальному закону  $F_i^{(1)}(t)=1-e^{-\alpha_i t}$ ,  $i=1,2$ ; в момент окончания  $i$ -го состояния процесс  $\lambda(t)$  переходит с вероятностью единица из  $i$ -го состояния в  $j$ -е,  $i, j=1,2$  ( $i \neq j$ ); вторая случайная величина распределена по экспоненциальному закону  $F_i^{(2)}(t)=1-e^{-\lambda_i t}$ ,  $i=1,2$ ; в момент окончания  $i$ -го состояния процесс  $\lambda(t)$  переходит с вероятностью  $P_1(\lambda_j|\lambda_i)$  в  $j$ -е состояние ( $i \neq j$ ) с наступлением события либо с вероятностью  $P_0(\lambda_j|\lambda_i)$  переходит в  $j$ -е состояние ( $i \neq j$ ) без наступления события, либо с вероятностью  $P_1(\lambda_i|\lambda_i)$  переходит в  $i$ -е состояние с наступлением события. При этом  $P_0(\lambda_j|\lambda_i)+P_1(\lambda_j|\lambda_i)+P_1(\lambda_i|\lambda_i)=1$ ,  $i, j=1,2$ ,  $i \neq j$ . Первая и вторая случайные величины являются независимыми друг от друга. В сделанных предположениях  $\lambda(t)$  – марковский процесс. Матрицы инфинитезимальных характеристик процесса  $\lambda(t)$  при этом примут вид:

$$\mathbf{D}_0 = \begin{vmatrix} -(\alpha_1 + \lambda_1) & \alpha_1 + \lambda_1 P_0(\lambda_2|\lambda_1) \\ \alpha_2 + \lambda_2 P_0(\lambda_1|\lambda_2) & -(\alpha_2 + \lambda_2) \end{vmatrix}, \quad \mathbf{D}_1 = \begin{vmatrix} \lambda_1 P_1(\lambda_1|\lambda_1) & \lambda_1 P_1(\lambda_2|\lambda_1) \\ \lambda_2 P_1(\lambda_1|\lambda_2) & \lambda_2 P_1(\lambda_2|\lambda_2) \end{vmatrix}.$$

Элементами матрицы  $\mathbf{D}_1$  являются интенсивности перехода процесса  $\lambda(t)$  из состояния в состояние с наступлением события. Диагональные элементы матрицы  $\mathbf{D}_0$  – интенсивности выхода процесса  $\lambda(t)$  из своих состояний, взятые с противоположным знаком. Недиагональные элементы матрицы  $\mathbf{D}_0$  – интенсивности переходов процесса  $\lambda(t)$  из состояния в состояние без наступления события. Следует заметить, что если  $\alpha_i = 0$ ,  $i=1,2$ , имеет место обычный МАР-поток событий [2].

После каждого зарегистрированного в момент времени  $t_k$  события наступает время фиксированной длительности  $T$  (мёртвое время), в течение которого другие события исходного модулированного МАР-потока недоступны наблюдению. По окончании мёртвого времени первое наступившее событие снова создает период мёртвого времени длительности  $T$  и т.д. Вариант возникающей ситуации показан на рис. 1, где  $t_1, t_2, \dots$  – моменты наступления событий в наблюдаемом потоке; 1 и 2 – состояния случайного процесса  $\lambda(t)$ ; чёрными кружками обозначены события модулированного МАР-потока, не доступные наблюдению; штриховкой – длительность мёртвого времени.



Рис. 1. Формирование наблюдаемого потока событий

Процесс  $\lambda(t)$  является принципиально ненаблюдаемым, т.е. имеем скрытый марковский процесс; наблюдаемыми являются только временные моменты наступления событий потока  $t_1, t_2, \dots$ . Рассматривается стационарный режим функционирования потока. В силу предпосылок в моменты времени  $t_1, t_2, \dots, t_k$  наступления событий потока последовательность  $\{\lambda(t_k)\}$  представляет собой вложенную цепь Маркова, т.е. наблюдаемый поток обладает марковским свойством, если его эволюцию рассматривать с момента  $t_k$  – момента наступления события потока,  $k = 1, 2, \dots$ .

Обозначим  $\tau_k = t_{k+1} - t_k$ ,  $k = 1, 2, \dots$ , – значение длительности  $k$ -го интервала между соседними событиями в наблюдаемом потоке. Так как рассматривается стационарный режим, то плотность вероятностей значений длительности  $k$ -го интервала  $p(\tau_k) = p(\tau)$ ,  $\tau \geq 0$ , для любого  $k$ . В силу этого момент времени  $t_k$  наступления события без ограничения общности можно положить равным нулю или, что то же самое, момент наступления события есть  $\tau = 0$ . С другой стороны, поскольку зарегистрированное в момент времени  $\tau = 0$  событие создаёт период мёртвого времени длительности  $T$ , то полагаем  $\tau = T + t$ , где  $t$  – значение длительности интервала между моментом окончания мёртвого времени  $\tau = T$  и моментом времени, когда наступает следующее событие в наблюдаемом потоке ( $t \geq 0$ ). Здесь предполагается, что значение  $T$  точно известно.

Цель данной статьи заключается в получении явного вида функции плотности распределения вероятностей  $p_T(\tau)$  длительности интервала в модулированном МАР-потоке событий с непродлевающимся мёртвым временем (далее – наблюдаемый поток или поток).

### 3. Выражение для плотности вероятностей $p_T(\tau)$

Функция плотности вероятностей длительности интервала между соседними событиями потока определяется формулой

$$\begin{cases} p_T(\tau) = 0, & 0 \leq \tau < T, \\ p_T(\tau - T) = \sum_{i=1}^2 \pi_i(0|T) \sum_{j=1}^2 q_{ij}(T) \sum_{k=1}^2 \tilde{p}_{jk}(\tau - T), & \tau \geq T, \end{cases} \quad (1)$$

где  $\tilde{p}_{jk}(\tau - T)$  – плотность вероятностей того, что процесс  $\lambda(\tau)$  изменяет своё состояние с  $j$  на  $k$  без наступления события на интервале  $(T, T + t)$  и событие наступает в момент времени  $\tau = T + t$ ,  $j, k = 1, 2$ ;  $q_{ij}(T)$  – переходная вероятность того, что процесс  $\lambda(\tau) = \lambda_j$  в момент окончания мёртвого времени  $\tau = T$  при условии, что процесс  $\lambda(0) = \lambda_i$  в момент времени  $\tau = 0$ ,  $i, j = 1, 2$ ;  $\pi_i(0|T)$  – условная стационарная вероятность того, что процесс  $\lambda(\tau)$  находится в состоянии  $i$  в момент времени  $\tau = 0$  при условии, что в этот же момент времени  $\tau = 0$  наступило событие потока и наступил период мёртвого времени длительности  $T$  ( $\pi_1(0|T) + \pi_2(0|T) = 1$ ).

Введём в рассмотрение переходные вероятности  $p_{jk}(t)$  того, что на интервале  $(T, T + t)$  нет событий потока и в момент времени  $T + t$  значение процесса  $\lambda(T + t) = \lambda_k$  при условии, что в момент времени  $\tau = T$  значение процесса  $\lambda(T) = \lambda_j$ ,  $j, k = 1, 2$ . Припишем моменту окончания мёртвого времени значение  $t = 0$ . Рассматривая поведение процесса  $\lambda(t)$  на интервале  $(0, t + \Delta t)$  и используя  $\Delta$ -метод, приходим к следующей системе дифференциальных уравнений с начальными условиями:

$$\begin{aligned} p'_{11}(t) &= -(\alpha_1 + \lambda_1)p_{11}(t) + (\alpha_2 + \lambda_2 P_0(\lambda_1|\lambda_2))p_{12}(t), \\ p'_{12}(t) &= (\alpha_1 + \lambda_1 P_0(\lambda_2|\lambda_1))p_{11}(t) - (\alpha_2 + \lambda_2)p_{12}(t), \\ p'_{21}(t) &= -(\alpha_1 + \lambda_1)p_{21}(t) + (\alpha_2 + \lambda_2 P_0(\lambda_1|\lambda_2))p_{22}(t), \\ p'_{22}(t) &= (\alpha_1 + \lambda_1 P_0(\lambda_2|\lambda_1))p_{21}(t) - (\alpha_2 + \lambda_2)p_{22}(t), \\ p_{11}(0) &= 1, \quad p_{12}(0) = 0, \quad p_{21}(0) = 0, \quad p_{22}(0) = 0. \end{aligned}$$

Решая полученную систему, находим вероятности  $p_{jk}(t)$ ,  $j, k = 1, 2$ :

$$\begin{aligned} p_{11}(t) &= \frac{1}{z_2 - z_1} \left[ (\lambda_2 + \alpha_2 - z_1) e^{-z_1 t} - (\lambda_2 + \alpha_2 - z_2) e^{-z_2 t} \right] \\ p_{12}(t) &= \frac{\alpha_1 + \lambda_1 P_0(\lambda_2|\lambda_1)}{z_2 - z_1} \left[ e^{-z_1 t} - e^{-z_2 t} \right] \\ p_{21}(t) &= \frac{\alpha_2 + \lambda_2 P_0(\lambda_1|\lambda_2)}{z_2 - z_1} \left[ e^{-z_1 t} - e^{-z_2 t} \right] \end{aligned}$$

$$p_{22}(t) = \frac{1}{z_2 - z_1} \left[ (\lambda_1 + \alpha_1 - z_1) e^{-z_1 t} - (\lambda_1 + \alpha_1 - z_2) e^{-z_2 t} \right]$$

$$z_1, z_2 = \frac{1}{2} \left[ (\lambda_1 + \lambda_2 + \alpha_1 + \alpha_2) \mp \sqrt{(\lambda_1 - \lambda_2 + \alpha_1 - \alpha_2)^2 + 4(\alpha_1 + \lambda_1 P_0(\lambda_2 | \lambda_1))(\alpha_2 + \lambda_2 P_0(\lambda_1 | \lambda_2))} \right]$$

$$0 < z_1 < z_2. \quad (2)$$

Плотности вероятностей  $\tilde{p}_{jk}(t)$ ,  $j, k = 1, 2$ , при этом примут вид

$$\begin{aligned} \tilde{p}_{11}(t) &= \lambda_1 P_1(\lambda_1 | \lambda_1) p_{11}(t) + \lambda_2 P_1(\lambda_1 | \lambda_2) p_{12}(t), \\ \tilde{p}_{12}(t) &= \lambda_1 P_1(\lambda_2 | \lambda_1) p_{11}(t) + \lambda_2 P_1(\lambda_2 | \lambda_2) p_{12}(t), \\ \tilde{p}_{21}(t) &= \lambda_1 P_1(\lambda_1 | \lambda_1) p_{21}(t) + \lambda_2 P_1(\lambda_1 | \lambda_2) p_{22}(t), \\ \tilde{p}_{22}(t) &= \lambda_1 P_1(\lambda_2 | \lambda_1) p_{21}(t) + \lambda_2 P_1(\lambda_2 | \lambda_2) p_{22}(t), \end{aligned} \quad (3)$$

где вероятности  $p_{jk}(t)$ ,  $j, k = 1, 2$ , определены в (2).

Так как последовательность  $\{\lambda(t_k)\}$  есть вложенная цепь Маркова, то для вероятностей  $\pi_i(0|T)$  справедливы следующие уравнения:

$$\begin{aligned} \pi_1(0|T) &= \pi_1(0|T)\pi_{11} + \pi_2(0|T)\pi_{21}, \quad \pi_2(0|T) = \pi_1(0|T)\pi_{12} + \pi_2(0|T)\pi_{22}, \\ \pi_1(0|T) + \pi_2(0|T) &= 1, \end{aligned} \quad (4)$$

где  $\pi_{ij}$  – переходные вероятности того, что процесс  $\lambda(t)$  перейдёт из состояния  $i$  в состояние  $j$  ( $i, j = 1, 2$ ) за время, которое пройдёт от момента  $\tau = 0$  наступления события потока до момента времени наступления следующего события потока.

Для получения  $\pi_{ij}$  введём переходные вероятности  $q_{ij}(\tau)$  того, что в момент времени  $\tau$  значение процесса  $\lambda(\tau) = \lambda_j$  при условии, что в момент времени  $\tau = 0$  значение процесса  $\lambda(0) = \lambda_i$ ,  $i, j = 1, 2$ . Для указанных вероятностей справедлива следующая система дифференциальных уравнений с начальными условиями:

$$\begin{aligned} q'_{11}(\tau) &= [-(\alpha_1 + \lambda_1) + \lambda_1 P_1(\lambda_1 | \lambda_1)] q_{11}(\tau) + (\alpha_2 + \lambda_2 [1 - P_1(\lambda_2 | \lambda_2)]) q_{12}(\tau), \\ q'_{12}(\tau) &= [-(\alpha_2 + \lambda_2) + \lambda_2 P_1(\lambda_2 | \lambda_2)] q_{12}(\tau) + (\alpha_1 + \lambda_1 [1 - P_1(\lambda_1 | \lambda_1)]) q_{11}(\tau), \\ q'_{21}(\tau) &= [-(\alpha_1 + \lambda_1) + \lambda_1 P_1(\lambda_1 | \lambda_1)] q_{21}(\tau) + (\alpha_2 + \lambda_2 [1 - P_1(\lambda_2 | \lambda_2)]) q_{22}(\tau), \\ q'_{22}(\tau) &= [-(\alpha_2 + \lambda_2) + \lambda_2 P_1(\lambda_2 | \lambda_2)] q_{22}(\tau) + (\alpha_1 + \lambda_1 [1 - P_1(\lambda_1 | \lambda_1)]) q_{21}(\tau), \\ q_{11}(0) &= 1, \quad q_{12}(0) = 0, \quad q_{21}(0) = 0, \quad q_{22}(0) = 1. \end{aligned} \quad (5)$$

Решая систему (5), находим вероятности  $q_{ij}(\tau)$ ,  $i, j = 1, 2$ :

$$\begin{aligned} q_{11}(\tau) &= \pi_1 + \pi_2 e^{-a\tau}, \quad q_{12}(\tau) = \pi_2 - \pi_2 e^{-a\tau}, \\ q_{21}(\tau) &= \pi_1 - \pi_1 e^{-a\tau}, \quad q_{22}(\tau) = \pi_2 + \pi_1 e^{-a\tau}, \\ a &= \alpha_1 + \alpha_2 + \lambda_1 [1 - P_1(\lambda_1 | \lambda_1)] + \lambda_2 [1 - P_1(\lambda_2 | \lambda_2)] \end{aligned}$$

$$\pi_1 = \frac{\alpha_2 + \lambda_2 [1 - P_1(\lambda_2 | \lambda_2)]}{a}, \quad \pi_2 = \frac{\alpha_1 + \lambda_1 [1 - P_1(\lambda_1 | \lambda_1)]}{a}, \quad (6)$$

где априорные финальные вероятности  $\pi_i$ ,  $i=1,2$ , состояния  $i$  процесса  $\lambda(t)$  получены в [1].

В силу марковости процесса  $\lambda(t)$  для переходных вероятностей  $\pi_{ij}$ ,  $i, j=1,2$ , справедлива следующая система уравнений:

$$\begin{aligned} \pi_{11} &= q_{11}(T)p_{11} + q_{12}(T)p_{21}, & \pi_{12} &= q_{12}(T)p_{22} + q_{11}(T)p_{12}, \\ \pi_{21} &= q_{22}(T)p_{21} + q_{21}(T)p_{11}, & \pi_{22} &= q_{21}(T)p_{12} + q_{22}(T)p_{22}, \\ \pi_{11} + \pi_{12} &= 1, & \pi_{21} + \pi_{22} &= 1, \end{aligned} \quad (7)$$

где  $q_{ij}(T)$  получены по формулам (6) при  $\tau=T$ , а  $p_{ij}$  – переходные вероятности того, что процесс  $\lambda(t)$  перейдёт из состояния  $i$  в состояние  $j$  за интервал времени между моментом  $t=0$  окончания мёртвого времени и моментом наступления следующего события потока. Так как  $t$  – произвольный момент времени, то переходные вероятности  $p_{ij}$ ,  $i, j=1,2$ , запишутся в виде

$$p_{ij} = \int_0^{\infty} \tilde{p}_{ij}(t) dt. \quad (8)$$

Подставляя сначала (3) в (8), затем (2) в (8) и вычисляя соответствующие интегралы, находим

$$\begin{aligned} p_{11} &= \frac{\lambda_1 P_1(\lambda_1 | \lambda_1)(\lambda_2 + \alpha_2) + \lambda_2 P_1(\lambda_1 | \lambda_2)(\alpha_1 + \lambda_1 P_0(\lambda_2 | \lambda_1))}{(\lambda_1 + \alpha_1)(\lambda_2 + \alpha_2) - (\alpha_1 + \lambda_1 P_0(\lambda_2 | \lambda_1))(\alpha_2 + \lambda_2 P_0(\lambda_1 | \lambda_2))}, \\ p_{12} &= \frac{\lambda_1 P_1(\lambda_2 | \lambda_1)(\lambda_2 + \alpha_2) + \lambda_2 P_1(\lambda_2 | \lambda_2)(\alpha_1 + \lambda_1 P_0(\lambda_2 | \lambda_1))}{(\lambda_1 + \alpha_1)(\lambda_2 + \alpha_2) - (\alpha_1 + \lambda_1 P_0(\lambda_2 | \lambda_1))(\alpha_2 + \lambda_2 P_0(\lambda_1 | \lambda_2))}, \\ p_{21} &= \frac{\lambda_2 P_1(\lambda_1 | \lambda_2)(\lambda_1 + \alpha_1) + \lambda_1 P_1(\lambda_1 | \lambda_1)(\alpha_2 + \lambda_2 P_0(\lambda_1 | \lambda_2))}{(\lambda_1 + \alpha_1)(\lambda_2 + \alpha_2) - (\alpha_1 + \lambda_1 P_0(\lambda_2 | \lambda_1))(\alpha_2 + \lambda_2 P_0(\lambda_1 | \lambda_2))}, \\ p_{22} &= \frac{\lambda_2 P_1(\lambda_2 | \lambda_2)(\lambda_1 + \alpha_1) + \lambda_1 P_1(\lambda_2 | \lambda_1)(\alpha_2 + \lambda_2 P_0(\lambda_1 | \lambda_2))}{(\lambda_1 + \alpha_1)(\lambda_2 + \alpha_2) - (\alpha_1 + \lambda_1 P_0(\lambda_2 | \lambda_1))(\alpha_2 + \lambda_2 P_0(\lambda_1 | \lambda_2))}. \end{aligned} \quad (9)$$

С учётом явного вида  $q_{ij}(T)$ ,  $i, j=1,2$ , полученного из (6) при  $\tau=T$ , переходные вероятности  $\pi_{ij}$  запишутся в виде

$$\begin{aligned} \pi_{11} &= p_{11} - \pi_2(p_{11} - p_{21}) \left[ \frac{1 - e^{-aT}}{1 - e^{-aT}} \right], & \pi_{12} &= p_{12} + \pi_2(p_{22} - p_{12}) \left[ \frac{1 - e^{-aT}}{1 - e^{-aT}} \right], \\ \pi_{21} &= p_{21} + \pi_1(p_{11} - p_{21}) \left[ \frac{1 - e^{-aT}}{1 - e^{-aT}} \right], & \pi_{22} &= p_{22} - \pi_1(p_{22} - p_{12}) \left[ \frac{1 - e^{-aT}}{1 - e^{-aT}} \right] \end{aligned} \quad (10)$$

Подставляя (10) в (4) и учитывая, что  $\pi_1(0|T) + \pi_2(0|T) = 1$ , находим

$$\pi_1(0|T) = \frac{p_{21} + \pi_1(p_{11} - p_{21}) \left[ \frac{1 - e^{-aT}}{1 - e^{-aT}} \right]}{p_{12} + p_{21} - (1 - p_{11} - p_{22}) \left[ \frac{1 - e^{-aT}}{1 - e^{-aT}} \right]},$$

$$\pi_1(0|T) = \frac{p_{12} + \pi_2(p_{22} - p_{12})[1 - e^{-aT}]}{p_{12} + p_{21} - (1 - p_{11} - p_{22})[1 - e^{-aT}]}, \quad (11)$$

где  $a$  и  $\pi_i$  определены в (6), а  $p_{ij}$  – в (9). С учётом (7) формула (1) принимает вид

$$p_T(\tau - T) = \pi_1(T) \sum_{k=1}^2 \tilde{p}_{1k}(\tau - T) + \pi_2(T) \sum_{k=1}^2 \tilde{p}_{2k}(\tau - T), \quad \tau \geq T,$$

где

$$\pi_1(T) = \pi_1 + [\pi_2 - \pi_2(0|T)]e^{-aT}, \quad \pi_2(T) = \pi_2 - [\pi_2 - \pi_2(0|T)]e^{-aT}. \quad (12)$$

Так как  $t = \tau - T$ ,  $t \geq 0$ . имеем

$$p_T(t) = \pi_1(T) \sum_{k=1}^2 \tilde{p}_{1k}(t) + \pi_2(T) \sum_{k=1}^2 \tilde{p}_{2k}(t), \quad t \geq 0. \quad (13)$$

Подставляя сначала (3), затем (2) в (13) и производя достаточно трудоёмкие преобразования, находим явный вид  $p_T(\tau)$ :

$$p_T(\tau) = \begin{cases} 0, & 0 \leq \tau < T, \\ \gamma(T)z_1 e^{-z_1(\tau-T)} + (1 - \gamma(T))z_2 e^{-z_2(\tau-T)}, & \tau \geq T, \end{cases}$$

$$\gamma(T) = \frac{1}{z_2 - z_1} \{z_2 - \lambda_1 \pi_1(T)[1 - P_0(\lambda_2|\lambda_1)] - \lambda_2 \pi_2(T)[1 - P_0(\lambda_1|\lambda_2)]\},$$

где  $z_1$  и  $z_2$  определены в (2),  $\pi_1(T)$  и  $\pi_2(T)$  определены в (12).

#### 4. Заключение

Полученная формула для плотности вероятностей  $p_T(\tau)$  позволяет решить задачу оценивания неизвестных параметров и длительности мёртвого времени в модулированном MAP-потоке событий с непродлевающимся мёртвым временем методом моментов или методом максимального правдоподобия. В первом случае получается система уравнений моментов, а во втором строится функция правдоподобия. Задача оценивания является предметом дальнейшего исследования.

#### Литература

1. Nezhel'skaya L. A. Optimal State Estimation in Modulated Map Event Flows with Unextendable Dead Time // Communications in Computer and Information Sciences: proceeding of the 13<sup>th</sup> International Scientific Conference ITMM 2014 named after A.F. Terpu-gov "Information Technologies and Mathematical Modelling" (November 20–22, 2014). Cham Heidelberg New York Dordrecht London: Springer, 2014. P. 342–350.
2. Gortsev A. M., Nezhel'skaya L. A., Solov'ev A. A. Optimal State Estimation in MAP Event Flows with Unextendable Dead Time // Automation and Remote Control. 2012. Vol. 73(8). P. 1316–1326.
3. Gortsev A. M., Nezhel'skaya L. A., Shevchenko T. I. Estimation of the states of an MC-stream of events in the presence of measurement errors // Russian Physics Journal. 1993. Vol. 36(12). P. 1153–1167.

4. Gortsev A. M., Nezhel'skaya L. A. An asynchronous double stochastic flow with initiation of superfluous events // Discrete Mathematics and Applications. 2011. Vol. 21(3). P. 283–290.

5. Bushlanov I. V., Gortsev A. M., Nezhel'skaya L. A. Estimating parameters of the synchronous twofold-stochastic flow of events // Automation and Remote Control. 2008. Vol 69(9). P. 1517–1533.

6. Gortsev, A. M., Nezhel'skaya L. A. Estimation of the dead time period and intensities of the synchronous double stochastic event flow // Radiotekhnika 2004. No. (10). P. 8–16.

7. Gortsev A. M., Nezhel'skaya L. A. Estimation of the dead-time period and parameters of a semi-synchronous double stochastic stream of events // Measurement Technuques. 2003. Vol. 46(6). P. 536–545.

8. Apanasovich V. V., Kolyada A. A., Chernyavskii A. F. Statistical Analysis of Random Flows in a Physical Experiment. Minsk: Universitetskoe, 1988.

## АДАПТИВНАЯ ФИЛЬТРАЦИЯ ИЗОБРАЖЕНИЙ НА ОСНОВЕ ВЗВЕШЕННОГО МЕТОДА МАКСИМАЛЬНОГО ПРАВДОПОДОБИЯ

*В. А. Симахин, О. С. Черепанов*

*Курганский государственный университет, Курган, Россия*

### 1. Постановка задачи. Введение

Рассмотрим статистическую модель растрового изображения RGB [1] вида

$$c(\mathbf{z}_i) = c^*(\mathbf{z}_i) + \varepsilon_i, \quad (1.1)$$

где  $c^*(\mathbf{z}_i)$  – значение цвета пикселя исходного изображения с координатами  $\mathbf{z}_i = (x_k, y_j)^T$  ( $i = \overline{1, N}$ ),  $\varepsilon$  – независимая от  $\mathbf{z}$  случайная величина с априорной функцией распределения  $G(x)$  и плотностью  $g(x)$ ,  $c(\mathbf{z}_i)$  – цвет пикселя наблюдаемого зашумленного изображения с координатами  $\mathbf{z}_i$  с функцией распределения

$$F = (1 - p)G + pS - \text{(супермодель Тьюки)}, \quad (1.2)$$

где  $F$  – распределение  $c(\mathbf{z})$  с плотностью  $f$ ;  $S$  – распределение выбросов;  $p$  – доля выбросов.

Требуется по зашумленному изображению  $c(\mathbf{z})$  восстановить исходное изображение  $c^*(\mathbf{z}_0)$  в точке  $\mathbf{z}_0$ .

Как правило, распределения  $S$  и доля выбросов  $p$  в супермодели (1.2) неизвестны, поэтому задача (1.1) относится к классу непараметрических задач робастной статистики.

Для описания изображений исследователи используют разнообразные математические модели, которые и приводят к различным алгоритмам фильтрации изображений [1–3]. Статистическая модель типа (1.1) позволяет при синтезе алгоритмов фильтрации учитывать как наличие выбросов, так и априорную неопределенность о виде распределения шума

$G$ . Синтез алгоритмов фильтрации для непараметрической задачи давно привлекает внимание исследователей [2, 3]. Для синтеза непараметрических алгоритмов фильтрации используется метод локальной аппроксимаций [2–4], а учет присутствия выбросов приводит к созданию робастных непараметрических фильтров. Основной идеей построения робастных фильтров является нахождение минимаксных решений для выбранных критериев робастности на заданных супермоделях [5, 6]. Критерии робастности и эффективности оказались противоречивыми, поэтому робастные процедуры на супермоделях могут иметь чрезвычайно низкую эффективность для конкретных распределений шумов  $G$  [6–9]. Этот факт и стимулирует поиск непараметрических алгоритмов, которые адаптируются, как по виду распределения шумов, так и по наличию выбросов. Одним из таких подходов синтеза адаптивных решений является взвешенный метод максимального правдоподобия (ВММП) [7–9].

В данной работе на основе ВММП синтезированы полунепараметрические и непараметрические адаптивные алгоритмы фильтрации изображений.

## 2. Взвешенный метод максимального правдоподобия.

### Фильтрация изображений

#### 2.1. Взвешенный метод максимального правдоподобия

Без потери общности будем рассматривать задачу фильтрации монохромного изображения. Задача фильтрации (1.1) сводится к задаче двумерной локальной регрессии. В окрестности точки  $z_0$  функцию яркости пикселя можно представить в следующем виде:

$$\hat{c}(z) = \sum_{j=1}^k \theta_j \varphi_j(z - z_0), \quad (2.1)$$

где  $\theta = (\theta_1, \dots, \theta_k)^T$  – вектор неизвестных параметров,  $\varphi = (\varphi_1, \dots, \varphi_k)^T$  – система линейно-независимых функций.

В соответствии с локальным методом максимального правдоподобия [4] для модели (1.2) эффективные оценки для вектора параметров  $\theta$  будут определяться из системы уравнения вида

$$\sum_{i=1}^N \left( \frac{\partial}{\partial \theta_j} \ln f(\varepsilon_i) \Big|_{\theta = \theta_N} \right) K_{1h}(z_i - z_0) = 0, \quad j=1, \dots, k, \quad (2.2)$$

$$\varepsilon_i = c_i - \sum_{j=1}^k \theta_{Nj} \varphi_j(z_i - z_0), \quad (2.3)$$

где  $K(z)$  – двумерная ядерная функция.

Систему уравнений (2.2) с учетом (1.2) представим в виде уравнений взвешенного максимального правдоподобия (ВММП) [7–9]:

$$\sum_{i=1}^N \left( \frac{\partial}{\partial \theta_j} \ln g(\varepsilon_i) \Big|_{\theta=\theta_N} \right) g^l(\varepsilon_i) K_{lh}(z_i - z_0) = 0, j=1, \dots, k, \quad (2.4)$$

где  $l$  – параметр радикальности.

Представление оценок в виде (2.4) позволяет свести всю неизвестную информацию о выбросах и локальных отклонениях от идеальной модели  $g(z)$  к одному неизвестному параметру радикальности  $l$ . Физически параметр радикальности  $l$  осуществляет процесс «мягкого» усечения как по виду априорного распределения, так и по величине выбросов шума. Система уравнения (2.4) порождает класс устойчивых оценок параметров по критерию устойчивости [6]. При  $l=0$  получаем оптимальные фильтры локального метода максимального правдоподобия, при  $l=1$  – фильтры максимальной устойчивости. В случае нормального распределения при  $l=0$  получаем классический гауссовский фильтр [1].

Исследования эффективности оценок ВММП [8, 9] показывают, что их эффективность в значительной степени зависит от выбора параметра радикальности  $l$ , который, как правило, неизвестен. В работе [7] был предложен непараметрический алгоритм адаптации по  $l$  на основе бутстреп-процедур (локальная адаптация). Для этого на основе бутстреп-процедур находится оценка условной дисперсии  $V$  фильтра в точке  $z_0$  в зависимости от  $l$  и размеров окна и осуществляется поиск минимума по  $l$ .

При синтезе эффективных фильтров в непараметрической постановке задачи возникает проблема, связанная с адаптацией также и по виду неизвестного распределения шума  $G$  (глобальная адаптация). Для глобальной адаптации используются непараметрические оценки плотности  $g(z)$  [8].

## 2.2. Полунапараметрические адаптивные фильтры

Рассмотрим полунапараметрические алгоритмы фильтрации на основе ВММП. В данной постановке задачи предполагается знание вида распределения шума. Ниже, в качестве примеров, рассмотрим адаптивные алгоритмы фильтрации для шумов с нормальным распределением и распределением Лапласа. Для других видов распределения шумов алгоритмы фильтрации можно получить аналогично.

### 2.2.1. Адаптивный гауссовский фильтр

Пусть шум имеет нормальное распределение  $g(z)$ . В соответствии с (2.4) адаптивный гауссовский фильтр будет определяться системой уравнений вида

$$\sum_{i=1}^N \varphi_j(z_i - z_0) \cdot \varepsilon_i \cdot g^l(\varepsilon_i, \sigma_N) K_{lh}(z_i - z_0) = 0, j = 1..k, \quad (2.5)$$

$$\sum_{i=1}^N \left( \left( \frac{\varepsilon_i}{\sigma_N} \right)^2 - \frac{1}{l+1} \right) \cdot g^l(\varepsilon_i, \sigma_N) K_{lh}(z_i - z_0) = 0,$$

где  $\varepsilon_i$  определено в (2.3);  $\sigma_N^2(z_0)$  – оценка локальной дисперсии шума;  $g$  – плотность нормального распределения со средним 0 и дисперсией  $\sigma_N^2$ .

При  $l=0$  и  $k=1$  получаем классический гауссовский фильтр.

### 2.2.2. Адаптивный медианный фильтр

Пусть шум имеет распределение Лапласа. В соответствии с (2.4) адаптивный медианный фильтр будет определяться системой уравнений вида

$$\sum_{i=1}^N \varphi_j(z_i - z_0) \cdot \text{Sign}(\varepsilon_i) \cdot g^l(\varepsilon_i, s_N) K_{1h}(z_i - z_0) = 0, j = 1..k, \quad (2.6)$$

$$\sum_{i=1}^N \left( \frac{|\varepsilon_i|}{s_N} - \frac{1}{l+1} \right) \cdot g^l(\varepsilon_i, s_N) K_{1h}(z_i - z_0) = 0,$$

где  $\varepsilon_i$  определено в (2.3);  $s_N$  – оценка локального параметра масштаба шума;  $g$  – плотность распределения Лапласа с параметром сдвига 0 и параметром масштаба  $s_N$ .

При  $l=0$ ,  $k=1$  и квадратной функции ядра получаем классический медианный фильтр [1].

### 2.3. Непараметрические адаптивные фильтры

Рассмотрим робастные непараметрические алгоритмы фильтрации. В данном случае вид распределения шумов неизвестен. Для оценки неизвестной плотности используем симметризованную оценку плотности Розенблатта–Парзена:

$$\tilde{g}_N(\varepsilon) = \frac{1}{2Nh_N} \sum_{i=1}^N \left[ K_2\left(\frac{\varepsilon - \varepsilon_i}{h_{Nc}}\right) + K_2\left(\frac{\varepsilon + \varepsilon_i}{h_{Nc}}\right) \right], \quad (2.7)$$

где  $h_{Nc}$  – параметр размытости;  $K_2(x)$  – ядерная функция.

Подставляя (2.7) в (2.4), получим робастный непараметрический фильтр в следующем виде:

$$\begin{aligned} & \sum_{i=1}^N \tilde{g}_N^{l-1}(\varepsilon_i) K_1(z_i) \times \\ & \times \sum_{\substack{t=1 \\ t \neq i}}^N K_1(z_t) \left( K_2'\left(\frac{\varepsilon_i - \varepsilon_t}{h_{Nc}}\right) (\varphi_j(z_i) - \varphi_j(z_t)) + K_2'\left(\frac{\varepsilon_i + \varepsilon_t}{h_{Nc}}\right) (\varphi_j(z_i) + \varphi_j(z_t)) \right) = 0, \\ & \sum_{i=1}^N \tilde{g}_N^l(\varepsilon_i) K_1(z_i) \left( 1 + \frac{1}{2Nh_{Nc} \tilde{g}_N(\varepsilon_i)} \times \right. \\ & \left. \times \sum_{\substack{t=1 \\ t \neq i}}^N K_1(z_t) \left( K_2'\left(\frac{\varepsilon_i - \varepsilon_t}{h_{Nc}}\right) \left(\frac{\varepsilon_i - \varepsilon_t}{h_{Nc}}\right) + K_2'\left(\frac{\varepsilon_i + \varepsilon_t}{h_{Nc}}\right) \left(\frac{\varepsilon_i + \varepsilon_t}{h_{Nc}}\right) \right) \right) = 0, \end{aligned} \quad (2.8)$$

Алгоритм (2.8) осуществляет глобальную адаптацию по распределению шума  $G$ . Для отбраковки выбросов осуществляется локальная адаптация по параметру  $l$ .

### 3. Моделирование

Эффективность полунепараметрического фильтра (2.5) сравнивалась с эффективностью классических фильтров: скользящего среднего и медианного фильтра на 8-битном тестовом изображении «Lenna» при наличии следующих моделей шумов: нормального распределения (НР), 10 %-ного импульсного шума типа «соль» (СОЛЬ),  $\hat{c}(x, y) = \theta_0$ .

Качество фильтрации изображения оценивалось по среднеквадратическому отклонению (СКО/эффективность) по отношению к минимальному СКО. В таблице представлены результаты работы фильтров по критерию СКО/эффективность.

**Значение СКО/эффективность фильтров на изображении «Lenna»**

Фильтр	Скользящее среднее 5x5	Медианный 7x7	ВММП $l=0,4$
СКО (НР)	16,43/0,81	18,07/0,74	<b>13,35/1,00</b>
СКО (СОЛЬ)	24,71/0,42	14,00/0,75	<b>10,47/1,00</b>

### Заключение

В работе синтезированы новые адаптивные полунепараметрические (2.5), (2.6) и непараметрические (2.8) алгоритмы фильтрации растровых изображений на основе ВММП для разных уровней априорной информации о виде распределения шумов. Моделирование показывает, что полунепараметрический фильтр (2.5) на тестовом изображении «Lenna» на 25 % превосходит популярный медианный фильтр.

### Литература

1. Гонсалес Р., Вудс Р. Цифровая обработка изображений. М.: Техносфера, 2005. 1072 с.
2. Katkovic V., Egiazarian K., Astola J. Local Approximation Techniques in Signal and Image Processing. Washington: SPIE, The International Society for Optical Engineering, 2006.
3. Takeda H., Farsiu S., Milanfar P. Kernel regression for image processing and reconstruction // IEEE Transactions on Image Processing. 2007. Vol. 16, № 2. P. 349–366.
4. Fan J., Gijbels I. Local Polynomial Modelling and its Applications, ser. Monographs on Statistics and Applied Probability. London; New York: Chapman and Hall, 1996.
5. Шуленин В. П. Математическая статистика. Ч. 3: Робастная статистика. Томск: Изд-во НТЛ, 2012. 518 с.
6. Шурыгин А. М. Прикладная статистика. Робастность. Оценивание. Прогноз. М.: Финансы и статистика, 2000. 223 с.
7. Simakhin V. A., Nonparametric robust regression estimate // Proceeding SPIE. 2006. Vol. 6522. P. 130–139.
8. Симахин В. А. Робастные непараметрические оценки. – Germany: LAMBERT Academic Publishing. 2011. 292 с.
9. Simakhin V. A., Cherepanov O. S. Robust Semiparametric Regression // Communication in Computer and Information Science. 2014. Vol. 487. P. 397–405.

# МОДЕЛИ И МЕТОДЫ МАССОВОГО ОБСЛУЖИВАНИЯ

## A $M^{(2)} | M^{(2)} | 1$ RETRIAL QUEUEING SYSTEM WITH R-PERSISTENT EXCLUSION OF ALTERNATIVE CUSTOMERS

*A. Nazarov, Ya. Chernikova*

*National Research Tomsk State University, Tomsk, Russia*

### Introduction

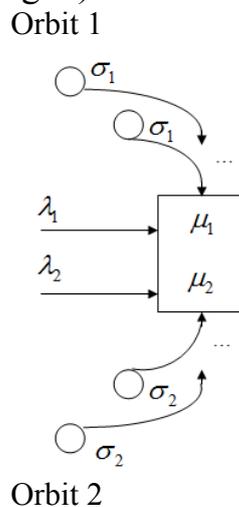
Queueing systems in which arriving customers who find all servers and waiting positions (if any) occupied may retry for service after a period of time are called Retrial queues [1–3]. A review of the main results on this topic can be found in [4]. Retrial queues have been widely used as mathematical models of different communication systems: shared bus local area networks operating under transmission protocols like CSMA/CD (Carrier Sense Multiple Access with Collision Detection), cellular mobile networks, in computer and communications networks, IP networks. Priority control is also widely used in production practice, transportation management, etc.

Several authors including B. D. Choi [6–10], N. Rengnanathan [11], G. V. Krishna Reedy, Y.J. Zhu have studied priority queues. These authors and several others have studied single server or multi-server queues with two or more priority classes under preemptive or non-preemptive priority rules. B. D. Choi analyzed a  $M|G|1$  retrial queueing systems with two types of calls and finite capacity, Moreno, P consider An  $M|G|1$  retrial queue with recurrent customers and general retrial times.

In this paper, we study the retrial queueing system  $M^{(2)} | M^{(2)} | 1$  with  $r$ -persistent exclusion of alternative customers.

### Problem statement

We consider two-dimension retrial queueing system with  $r$ -persistent collision of alternative customers (Fig. 1).



*Fig. 1. Retrial queueing system  $M^{(2)} | M^{(2)} | 1$*

We assume, that two arrival processes in the system are distributed by the stationary Poisson process with intensity  $\lambda_1$  and, respectively. Customer, which finds the free server, occupies it during a random time which distributed by exponential law with intensity and  $\mu_2$ , respectively. If at the time of arrival, customer of the first type finds the server busy with a customer of the first type, then it goes to the orbit 1 (in the orbit for customer of the first type), where it performs a random delay with duration determined by exponential distribution with intensity  $\sigma_1$ . From the orbit 1, after the random delay, the customer is trying to occupy the server again. If at the time of arrival, customer of the first type finds the server busy with a customer of the second type, then an arrived customer with probability  $\sigma_2$  replaces the customer, which was in service, and occupies the server, and with probability  $\sigma_1$  it goes to the orbit 1.

The same goes for the second type customer. If at the time of arrival, customer of the second type finds the server busy with a customer of the second type, then it goes to the orbit 2 (in the orbit for customer of the second type), where it performs a random delay with duration determined by exponential distribution with intensity  $\sigma_2$ . From the orbit 2, after the random delay, the customer is trying to occupy the server again. If at the time of arrival, customer of the second type finds the server busy with a customer of the first type, then an arrived customer with probability  $\sigma_1$  replaces the customer, which was in service, and occupies the server, and with probability  $\sigma_2$  it goes to the orbit 2.

Let  $i_1(t)$  be the number of customers in the orbit 1 and  $i_2(t)$  be the number of customers in the orbit 2, define the server state in the following way:

$$k(t) = \begin{cases} 0, & \text{if server is free,} \\ 1, & \text{if server is busy with a customer of the first type,} \\ 2, & \text{if server is busy with a customer of the second type.} \end{cases}$$

We would like to solve a problem of computation of stationary probability distribution of the number of customers in the orbit 1, orbit 2 and server state.

### **System of Kolmogorov differential equations**

We consider Markovian process  $\{k(t), i_1(t), i_2(t)\}$ .

Let us denote by  $P\{k(t) = k, i_1(t) = i_1, i_2(t) = i_2\} = P_k(i_1, i_2, t)$  a probability that, at the moment  $t$ , the server in the state  $k$  and  $i_1$  customers are in the orbit 1,  $i_2$  customers are in the orbit 2.

We write system of differential Kolmogorov's equations for the probability distribution

$$\frac{\partial P_0(i_1, i_2, t)}{\partial t} = -(\lambda_1 + \lambda_2 + i_1\sigma_1 + i_2\sigma_2)P_0(i_1, i_2, t) + \mu_1 P_1(i_1, i_2, t) + \mu_2 P_2(i_1, i_2, t),$$

$$\begin{aligned} \frac{\partial P_1(i_1, i_2, t)}{\partial t} = & -(\lambda_1 + \lambda_2 + \mu_1 + r_2 i_2 \sigma_2) P_1(i_1, i_2, t) + (1 - r_2) \lambda_2 P_1(i_1, i_2 - 1, t) + \\ & + \lambda_1 P_0(i_1, i_2, t) + (i_1 + 1) \sigma_1 P_0(i_1 + 1, i_2, t) + \lambda_1 P_1(i_1 - 1, i_2, t) + r_1 \lambda_1 P_2(i_1, i_2 - 1, t) + \\ & + r_1 (i_1 + 1) \sigma_1 P_2(i_1 + 1, i_2 - 1, t), \end{aligned} \quad (1)$$

$$\begin{aligned} \frac{\partial P_2(i_1, i_2, t)}{\partial t} = & -(\lambda_1 + \lambda_2 + \mu_2 + r_1 i_1 \sigma_1) P_2(i_1, i_2, t) + (1 - r_1) \lambda_1 P_2(i_1 - 1, i_2, t) + \\ & + \lambda_2 P_0(i_1, i_2, t) + (i_2 + 1) \sigma_2 P_0(i_1, i_2 + 1, t) + \lambda_2 P_2(i_1, i_2 - 1, t) + r_2 \lambda_2 P_1(i_1 - 1, i_2, t) + \\ & + r_2 (i_2 + 1) \sigma_2 P_1(i_1 - 1, i_2 + 1, t). \end{aligned}$$

### Equations for partial characteristic function

We write partial characteristic function in the following form:

$$H_k(u_1, u_2, t) = \sum_{i_1=0}^{\infty} \sum_{i_2=0}^{\infty} e^{ju_1 i_1} e^{ju_2 i_2} P_k(i_1, i_2, t), \quad k = 0, 1, 2$$

where  $j = \sqrt{-1}$  is imaginary unit. We rewrite the system (1) for partial characteristic function.

We can rewrite system (1) as:

$$\begin{aligned} & -(\lambda_1 + \lambda_2) H_0(u_1, u_2) + j\sigma_1 \frac{\partial H_0(u_1, u_2)}{\partial u_1} + j\sigma_2 \frac{\partial H_0(u_1, u_2)}{\partial u_2} + \\ & + \mu_1 H_1(u_1, u_2) + \mu_2 H_2(u_1, u_2) = 0, \\ & -(\lambda_1 + \lambda_2 + \mu_1) H_1(u_1, u_2) + j\sigma_2 r_2 \frac{\partial H_1(u_1, u_2)}{\partial u_2} - j\sigma_1 e^{-ju_1} \frac{\partial H_0(u_1, u_2)}{\partial u_1} + \\ & + (1 - r_2) \lambda_2 e^{ju_2} H_1(u_1, u_2) + \lambda_1 H_0(u_1, u_2) + \lambda_1 e^{ju_1} H_1(u_1, u_2) + \\ & + r_1 \lambda_1 e^{ju_2} H_2(u_1, u_2) - jr_1 \sigma_1 e^{j(u_2 - u_1)} \frac{\partial H_2(u_1, u_2)}{\partial u_1} = 0, \\ & -(\lambda_1 + \lambda_2 + \mu_2) H_2(u_1, u_2) + j\sigma_1 r_1 \frac{\partial H_2(u_1, u_2)}{\partial u_1} - j\sigma_2 e^{-ju_2} \frac{\partial H_0(u_1, u_2)}{\partial u_2} + \\ & + (1 - r_1) \lambda_1 e^{ju_1} H_2(u_1, u_2) + \lambda_2 H_0(u_1, u_2) + \lambda_2 e^{ju_2} H_2(u_1, u_2) + \\ & + r_2 \lambda_2 e^{ju_1} H_1(u_1, u_2) - jr_2 \sigma_2 e^{j(u_1 - u_2)} \frac{\partial H_1(u_1, u_2)}{\partial u_2} = 0. \end{aligned} \quad (2)$$

We will solve system (2) using the method of asymptotic analysis under condition of long delay ( $\sigma \rightarrow 0$ ).

We find partial characteristic function in following form:

$$H(u_1, u_2) = \exp \left\{ j \frac{u_1}{\sigma} x_1 + j \frac{u_2}{\sigma} x_2 + \frac{(ju_1)^2}{2\sigma} Q_{11} + \frac{(ju_2)^2}{2\sigma} Q_{22} + \frac{ju_1 ju_2}{\sigma} Q_{12} \right\}$$

where values  $x_1, x_2$  are asymptotic means of the number of customers in the orbits, where values  $Q_{11}, Q_{22}$  are variance,  $Q_{12}$  is correlation coefficient.

### Numerical realization

For example, we take the parameters of arrival processes as:

$$\lambda_1 = 3, \lambda_2 = 2.$$

If the parameters of exponential law service are fixed as follow is:

$$\mu_1 = 10, \mu_2 = 20.$$

The parameters of a random delay with duration determined by exponential distribution are fixed in following form:

$$\sigma_1 = 0.02, \sigma_2 = 0.03.$$

So as  $\sigma_1 = \gamma_1 \sigma$ ,  $\sigma_2 = \gamma_2 \sigma$ , then we will take  $\gamma_1 = 2$ ,  $\gamma_2 = 3$ .

Probability of displacement of the customer from the server by the customer of the first type  $r_1 = 1$ . Probability of displacement of the customer from the server by the customer of the second type  $r_1 = 1$ .

We have values of asymptotic means of the number of customers in the orbits with these values of parameters

$$x_1 = 1, x_2 = 0,444,$$

and variance

$$Q_{11} = 1,152, Q_{22} = 0,308$$

and correlation coefficient

$$r = 0,421.$$

### Conclusion

In the paper we study the retrial queueing system  $M^{(2)}|M^{(2)}|1$  with r-persistent exclusion of alternative customers by method of asymptotic analysis under condition of long delay. Stationary probability distribution of server states and values of asymptotic means of the number of customers in the orbits are obtained. Two-dimension marginal distribution of the number of customers in the orbit 1, in the orbit 2 is asymptotically Gaussian. We obtain the numerical realization for the considered parameters.

### References

1. Artalejo J. R. A Classified Bibliography of Research on Retrial Queues // Progress in 1990–1999. Top. 1999. Vol. 7, issue 2. P. 187–211.
2. Artalejo J. R. Accessible Bibliography on Retrial Queues // Mathematical and Computer Modeling. 1999. Vol. 30, issue 1–2. P. 1–6.
3. Artalejo J. R. Accessible Bibliography on Retrial Queues // Progress in 2000–2009. Mathematical and Computer Modeling. 2010. Vol. 51. P. 1071–1081.
4. Falin G. I. A Survey of Retrial Queues // Queuing Systems. 1990. Vol. 7. P. 127–167.
5. Falin G. I., Artalejo J. R., Martin M. On the single retrial queue with priority customers // Queueing Systems. 1993. Vol. 14, no. 3–4. P. 439–455.
6. Choi B. D., Chang, Y. Single Server Retrial Queues with Priority Calls. Mathematical and Computer Modeling. 1999. Vol. 30, no. 3–4. P. 7–32.
7. Choi B. D., Choi K. B., Lee Y. W. M/G/1 retrial queueing systems with two types of calls and finite capacity // Queueing Systems. 1995. Vol. 19. P. 215–229.

8. Choi B.D., Park, K. K. The M/G/1 retrial queue with bernoulli schedule // Queueing Systems 1990. Vol. 7, no. 2. P. 219–227.
9. Choi B. D., Shin, Y. W., Ahn W. C. Retrial queues with collision arising from un-slotted CSMA/CD protocol // Queueing Systems. 1992. Vol. 11, no. 4. P. 335–356.
10. Choi B. D., Park, K. K., Pearce C. E. M. An M/M/1 retrial queue with control policy and general retrial times // Queueing Systems. 1993. Vol. 14, no. 3–4. P. 275–292.

## **ANALYZING BLOCKING PROBABILITY IN LTE WIRELESS NETWORK VIA QUEUING SYSTEM WITH FINITE AMOUNT OF RESOURCES**

*K. E. Samouylov, E. S. Sopin, O. G. Vikhrova*

*Peoples` Friendship University of Russia, Moscow, Russia*

According to analytics, the global mobile data traffic will grow three times faster than fixed traffic by 2019. The number of user`s mobile devices is supposed to increase from 4.1 billion to 4.9 billion while the number of mobile device connections can reach even 10 billion. Mobile video traffic will be up to 72 percent of the global mobile traffic. As the number of wireless connections tends to increase significantly, it results in dramatic mobile traffic growth. Mobile service providers face the challenge to utilize limited radio resources efficiently. In this paper, we propose a mathematical model of radio resources allocation in broadband wireless networks such as LTE-Advanced in terms of queuing systems and evaluate blocking probability and average amount of occupied resources.

### **Introduction**

High popularity of various multimedia mobile services is conditioned by huge amount of mobile devices and attractiveness of mobile services. Users prefer to use their smart phones, tablets, etc. for business needs and entertaining via wireless networks. To deliver a service with expected service quality providers have focused on methods and techniques to utilize resources more efficiently and more flexible. In modern high-speed wireless networks such as LTE and LTE-Advanced, every mobile session will be allocated with radio resources based on type of service, distance between mobile device and base station and multiple access scheme.

We consider that each mobile device can transfer data to the LTE base station with specific data rate. Let`s denote maximum bitrate for customer  $i$  as  $c_i^{\max} = w \log_2 \left( 1 + \gamma_i p_{\max} \right)$ , where  $w$  is a spectral bandwidth,  $p_{\max}$  is maximum transmit power of the base station and  $\gamma_i$  is signal-to-noise ratio for the session, that depends on distance between user device and base station, possible obstacles between them, etc. New session will be accepted only if there is

enough resources to serve it, i.e.  $\sum \frac{c_i}{i c_i^{\max}} \leq 1$ , where  $c_i$  required data is rate for

session  $i$ . In order to evaluate network performance measures we design a mathematical model in terms of queuing system with limited resources to analyze blocking probability and average amount of occupied radio resources.

### Model description

In paper [2] multiple servers queuing system is considered where each customer occupies on arrival some random amount of limited resources. At the end of service time the amount of allocated resources is fully released, see fig.1. If there is not enough resources to meet customer requirements, the customer is denied service. Random variables (RV) of required resources assumed to be independent of arrival and service processes, mutually independent and identically distributed. In this type of queue, we have to remember a vector of allocated resources for each customer. Thus, it significantly complicates state space of the corresponding random process and its analysis.

To simplify the model we offer to track only total amount of occupied resources. As soon as we don't know how many resources have been allocated to each customer, we assume that amount of released resources on a departure of a customer is also random and may differ from allocated one.

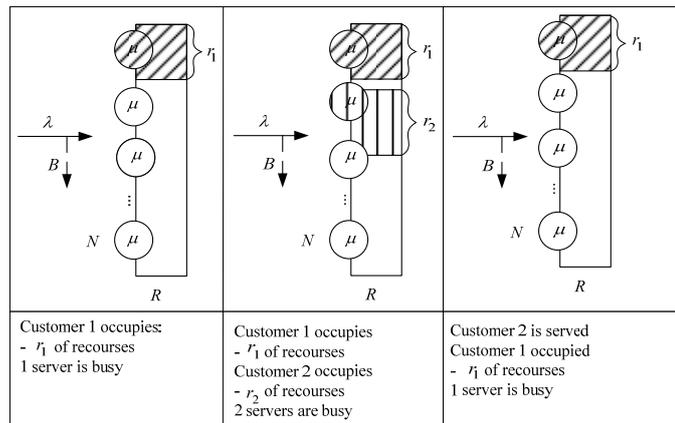


Fig. 1. Diagram of a general model

Given the total amount of allocated resources and the number of customers in system, RV of released resources are independent from past behavior of the system and its cumulative distribution function (CDF) can be obtained using the Bayes theorem.

In [2] it was shown that average amount of occupied resources of initial and simplified models are very close to each other in case of Poisson arrival process and exponential service distribution time. Later in [4], simulations showed that not only average values but also steady state distribution of allocated resources for both models are very close. Finally, in [3], it was analytically proved that steady-state distributions of total occupied resources and number of customers are equal. Note, that some generalizations discussed in [6] include a

system in which the service time and the amount of resources allocated to the customer are dependent random variables, and each customer has three random characteristics: the number of devices required for the service, resources and service time.

In this paper we analyze simplified model with  $N$  servers and limited amount  $R$  of a discrete resource (fig. 2). Customers arrive according to the Poisson process with rate  $\lambda$ . Service times are mutually independent, independent of arriving process and are exponentially distributed with rate  $\mu$ . Let us denote  $\xi(t)$  - number of customers in the system at time  $t$ , and  $\delta(t)$  - amount of total occupied resources,  $\delta(t) < R$ . Customer  $i$  requires  $r_i \geq 0$  units of discrete resources, RVs  $r_i$  are mutually independent and identically distributed with CDF  $F(x)$ , mean  $m$  and variance  $\sigma^2$ .

Customer  $i$  will be lost in case the system doesn't have enough resources ( $R - \delta(t) < r_i$ ) or if there is no free servers, i.e.  $\xi(t) = N$ . Total amount of occupied resources  $\delta(t)$  increases by  $r_i \geq 0$  immediately after arrival of a customer.

Total amount of occupied resources  $\delta(\tau_i)$  will decrease by random value  $v_i$  at time  $\tau_i$  as soon as customer  $i$  is served. Given the number of total occupied resources  $\delta(\tau_i) = y$  and number of customers in the system  $\xi(\tau_i) = k$ , RVs  $v_i$  are independent from previous system behavior and have CDF  $F_k(x|y) = P(v_i \leq x | \xi(\tau_i) = k; \delta(\tau_i) = y)$ ,  $0 \leq x \leq y$ .

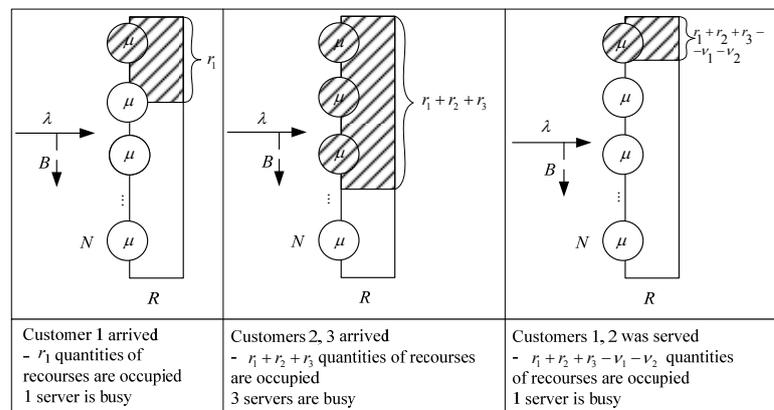


Fig. 2. Diagram of a simplified model

RVs  $r_i$  attain values  $j = \overline{0, R}$ , and  $p_j = P(r_i = j)$  is a probability that arriving customer requires  $j$  units of resource. Let's denote  $p_j^{(k)}$ ,  $j = \overline{0, R}$  a  $k$ -fold convolution of probabilities  $p_j$ , which allows to calculate the total amount of occupied resources  $y = \sum_{i=0}^k r_i$ . CDF  $F_k(x|j)$  of amount of resources released by a customer on departure is piecewise-constant function and it has a saltus at

$x=i, i=\overline{0,R}$  on a value  $p_{ij}^k = \frac{p_i p_j^{(k-1)}}{p_j^{(k)}}$ ,  $0 < i \leq j$ ,  $j=\overline{0,R}$ . In general case

convolutions  $p_j^{(k)}$ ,  $k \geq 2$  can be calculated using probabilities  $p_i$ ,  $0 \leq i \leq R$  according to the following recurrent formula:

$$p_j^{(k)} = \sum_{i=0}^j p_i p_{j-i}^{(k-1)}. \quad (1)$$

Random process  $X(t) = (\xi(t), \delta(t))$  over the set  $X = \bigcup_{k=0}^N X_k$ ,  $X_k = \{(k, i) | 0 \leq i \leq R, p_i^{(k)} > 0\}$  is a Markov chain. Figure 3 shows the state transitions diagram, where  $1 \leq i < N$ ,  $0 \leq s \leq j \leq q \leq R$ .

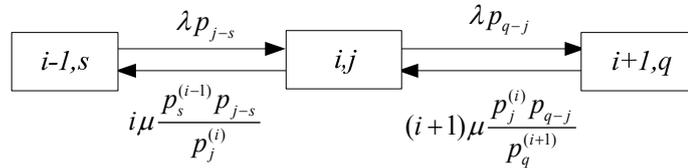


Fig. 3. State transition diagram

Infinitesimal matrix  $A = [a((i, j), (k, r))]$  has a block three-diagonal structure with main diagonal blocks  $\Psi_0, \Psi_1, \dots, \Psi_N$ , upper blocks  $\Lambda_1, \dots, \Lambda_N$  and lower blocks  $\mathbf{M}_0, \dots, \mathbf{M}_{N-1}$ .

Having infinitesimal matrix  $A$  we obtain stationary distribution for  $X(t) = (\xi(t), \delta(t))$ , where

$$q_{0,0} = \lim_{t \rightarrow \infty} P\{\xi(t) = 0, \delta(t) = 0\}, \quad (2)$$

$$q_{k,i} = \lim_{t \rightarrow \infty} P\{\xi(t) = k, \delta(t) = i\}, 1 \leq k \leq N, 0 \leq i \leq R. \quad (3)$$

Denote vectors of state probabilities  $q_0 = q_{0,0}$ ,  $q_i = (q_{i,j})_{j=0,R}$ ,  $i = 1, \dots, N$ .

Then, the set of equilibrium equations can be written in vector form:

$$\mathbf{q}_0 \Psi_0 - \mathbf{q}_1^T \mathbf{M}_0 = \mathbf{0}, \quad (4)$$

$$\mathbf{q}_i^T \Psi_i - \mathbf{q}_{i+1}^T \mathbf{M}_i - \mathbf{q}_{i-1}^T \Lambda_i = \mathbf{0}, i = 1, \dots, N-1, \quad (5)$$

$$\mathbf{q}_N^T \Psi_N - \mathbf{q}_{N-1}^T \Lambda_N = \mathbf{0}. \quad (6)$$

Matrix  $A$  is indecomposable and conservative with three diagonal blocks thus we apply UL matrix decomposition techniques [1].

In [3] analytical solution of the set of equilibrium equations was obtained for initial model. In the case of discrete numbers of allocated resources we derive more simple equations to calculate steady-state probabilities distribution:

$$q_{k,\bullet} = \lim_{t \rightarrow \infty} P\{\xi(t) = k\} = p_0 \frac{\rho^k}{k!} \sum_{i=0}^R p_i^{(k)}, 0 < k \leq N, \quad (7)$$

$$q_{k,j} = \lim_{t \rightarrow \infty} P\{\xi(t) = k; \delta(t) = j\} = p_0 \frac{\rho^k}{k!} p_j^{(k)}, \quad 0 \leq j \leq R, 0 < k \leq N, \quad (8)$$

$$p_0 = \left( 1 + \sum_{k=1}^N \frac{\rho^k}{k!} \sum_{i=0}^R p_i^{(k)} \right)^{-1}. \quad (9)$$

Thus the average amount of occupied resources  $b$  is easier to calculate:

$$b = \sum_{k=0}^N b_k q_{k,\bullet} = p_0 \sum_{k=0}^N \frac{\rho^k}{k!} \sum_{i=0}^R i p_i^{(k)}. \quad (10)$$

System blocking probability  $B$  can also be obtained as follows:

$$B = 1 - \sum_{k=0}^{N-1} \sum_{i=0}^R q_{k,i} \sum_{j=0}^{R-i} p_j = 1 - p_0 \sum_{k=0}^{N-1} \frac{\rho^k}{k!} \sum_{i=0}^R p_i^{(k+1)}. \quad (11)$$

### Numerical results

We considered different distributions for resource allocation such as binomial, shifted binomial and geometric distributions. We also assume that total amount of system resources  $R$  is equal to the number of servers  $N$ .

To evaluate systems characteristics we calculated  $k$ -fold convolutions  $p_j^{(k)}$  for each suggested distributions using formula (1).

We consider the total amount of system resources and the number of servers is equal and  $N = R = 100$ , while the arriving rate is  $\lambda = \{12, 13, \dots, 20\}$  and average service time  $\mu^{-1} = 1$ .

### Conclusion

In this paper, we analyzed simplified mathematical model in case of discrete distribution of radio resources in LTE base station. We suggested analytical and numerical methods to calculate the most interesting characteristics of the model. Besides, we investigated relation between system performance characteristics and distribution of allocated resources.

In result, we state that high variance of distribution of allocated resources leads to lower average number of customers in the system and average amount of occupied resources. In heavy load states system will accept customer's sessions with minimal requirements.

The authors wish to express their appreciation to Prof. V. Naumov for suggesting the problem and his help during the research.

*This work was partially supported by RFBR, projects No. 14-07-0090, 15-07-03608, 15-07-03051.*

### References

1. Naumov V. A. Numerical methods for Markov queues analysis // PFUR, Moscow, 1985. 36 p.
2. Naumov V. A., Samouylov K. E. On the modeling of queuing systems with multiple resources // PFUR Bulletin. Series Informatics. Mathematics. Physics. 2014. No. 3. P. 58–62.

3. Naumov V. A., Samuoylov K. E, Samuylov A. K. On total amount of resources occupied by customers // Automation and Remote Control (in print).
4. Naumov V., Samuoylov K., Sopin E., Andreev S. Two approaches to analysis of queuing systems with limited resources / Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT) // IEEE. 2014. P. 485–488.
5. Pyattaev A., Johnsson K., Surak A. Network-assisted D2D communications: Implementing a technology prototype for cellular traffic offloading / Wireless Communications and Networking Conference (WCNC) // IEEE. 2014. P. 3266–3271.
6. Tikhonenko O. M. Generalized Erlang problem for service systems with finite total capacity // Problems of Information Transmission. 7.2005. Vol. 41, issue 3. P. 243–253.

**DISCRETE-EVENT SIMULATION OF CLOSED QUEUING SYSTEMS  
BASED ON LOGICAL-ALGEBRAIC APPROACH  
(ON THE EXAMPLE OF RESOURCE USING IN  
INFOCOMMUNICATION SYSTEM)**

*N. P. Vashkevich, S. A. Zinkin, D. V. Pashchenko  
Penza State University, Penza, Russia*

Let's consider the question of building of formal network specifications describing the distribution of units of a resource  $R$  between the requests of different types in a telecommunication system. Detailed specifications should ensure effective implementation of a network of software modules that control distributed network resource, for example, replicated databases on the assumption that each copy of the database is placed on its own external storage device (ESD).

The purpose of this article is to construct a discrete-event model of queuing based on network formalism of abstract machines. The model manages the distributed abstract resource  $R$ . The Genrich-Glaesser model of resource using is the basis of our model [1–3].

The typical structure of a composite info-communication network with distributed storage resource  $R$  is shown in Fig. 1. There, the letters  $M$  and  $K$  designate global network (WAN) routers and local network (LAN) switches, respectively, node-units of resource  $R$  are indicated by black circles, on which database management systems are installed, and to which ESD for storing information indicated as cylinders are connected. Darkened circles represent nodes-sources of requests to use the distributed resource  $R$ .

Let's use the illustrative example proposed in works [1, 2] and [3] as the basis of the description of system functioning. In works [1, 2] model of resource  $R$  using for the two types of requests illustrates proposed by the author networks called "predicate-transition," and in work [3] the same model is described by the networks of abstract state machines. The disadvantage of these models is that without using of priority mechanisms between the two types of requests time of e-requests servicing increases abruptly with the growth of number of s-requests and the system becomes practically inefficient. Besides, in listed works the pro-

cedural component of the model is not formalized. Proposed in work [4, 5] formalism of abstract machine networks SeAM and its extension RSeAM and implementation of networks' data based on FS-technology described in [6, 7], take into greater account the mechanisms of directly interpretable specifications, provide a hierarchical design of distributed programs. SeAM-expressions are constructed on basis of superposition of  $\alpha$ -disjunction operators, and in expressions for advanced option of RSeAM superposition of  $\alpha$ -iterations' operators from algebra of algorithms [8–10] is additionally used, although for conventional and cyclic expressions recording another suitable notation could be chosen.

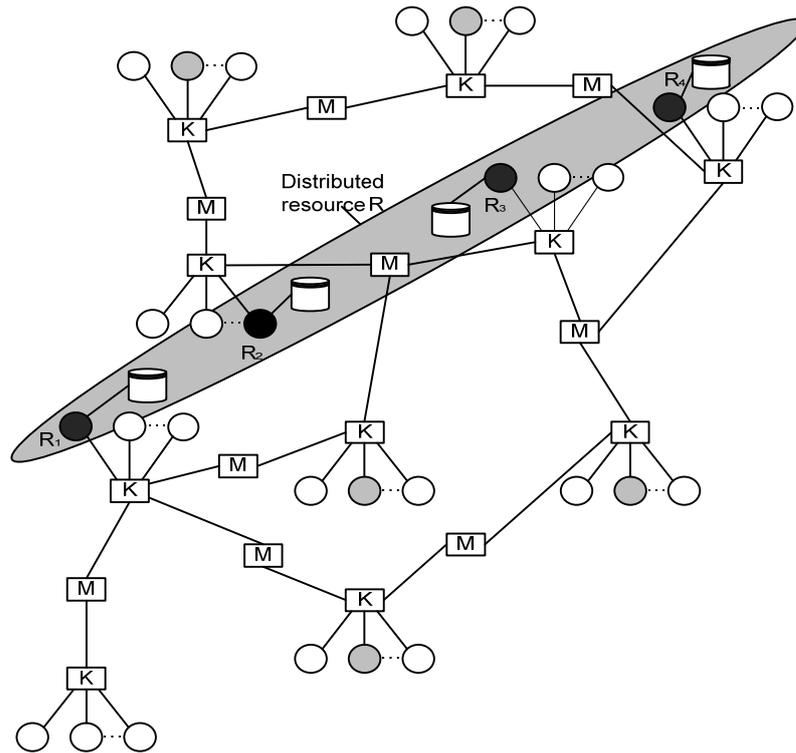


Fig. 1. The implementation of a distributed resource in computer network

Any logical formulas with predicate symbols used as logical variables are called conditions in terms for modules and SeAM and RSeAM. Expressions can use quantifiers, functional symbols with equality predicate and expressions themselves may be based on using the logic of the first and higher orders. Besides, quantified operators of choice  $\tilde{\exists}!$ ,  $\tilde{\exists}!!$ ,  $\tilde{\forall}$  and  $\tilde{\forall}!!$ , to which an operator of "underlining" is applied [4, 5], here can be used as logical variables. When the operator  $\tilde{\exists}!$  is performed, an arbitrary tuple is selected from the area of predicate's truth. When the operator  $\tilde{\exists}!!$  is performed, a single tuple is selected from the area of predicate's truth. When the operator  $\tilde{\forall}$  is performed, there is a selection of all tuples that make the area of corresponding predicate's truth. The operator  $\tilde{\forall}!!$  allows selecting of all tuples from the area of predicate's truth in case when its area of definition matches the area of its truth. In all cases, it is meant,

that predicate is described by expression from the right of the quantified operator's symbol. Each of described quantified operators which perform the sample of tuples from relations, in case of its using in the conditional part of expression for the module, is associated with elementary logical condition, which is true in case of successful execution of the operator and false otherwise. This compliance is made by the operator "underlining".

Let's consider a network of abstract machines, that is implemented in the form of causally related software modules and that is controlling the distributed resource  $R$ . We consider that  $A = \{a_1, a_2, \dots, a_{|A|}\}$  is a set of requests to ESD modules connected to the computer network and representing, in the aggregate, the distributed resource, consisting of many units of  $R$  (and containing the required copies of the same database), and  $T$  is the set of "tokens" or "tickets" which enable access to units of the resource  $R$ . The number of tokens  $n_T = |T|$  matches to the number of units  $n_R = |R|$  of resource  $R$ . Abstract structure of modeled system as a queuing system is presented in Fig. 2.

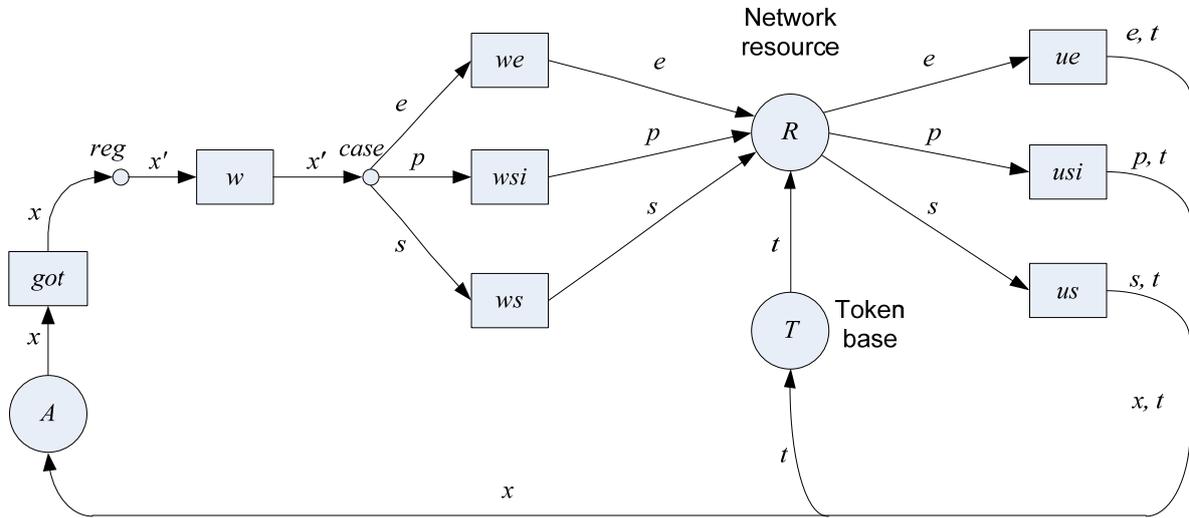


Fig. 2. The model of info-communication system in the form of a closed queuing system

In the figure rectangles represent the queues in queuing system, presented in model by unary predicates, with queue names which are used below as indices of corresponding predicate names. The circles represent sets of requests, tokens and units of the network resource. Let's pay attention to some of the modules of an abstract model.

SeAM-expression for module  $m_0$  describes the setting of the request servicing mode and its transfer from queue  $p_{got}$  to queue  $p_w$ :

$$m_0 = [(\exists!x \in A) p_{got}(x)] [(\exists!m \in M) p_{reg}(m)]$$

$$(\{p_{got}(x) \leftarrow \text{false}, f_m(x) \leftarrow m, p_w(x) \leftarrow \text{true}\} \vee R^E) \vee R^E).$$

The following predicates and one function are used in an expression for the module  $m_0$ :

$p_{got}(x)$  is a unary predicate that determines the readiness of the client of the network ESD to output the request  $x \in A$ ;

$p_{reg}(m)$  is a unary predicate that specifies one of the three modes of requests' processing,  $m \in M$ ,  $M = \{e, p, s\}$ :  $e$  is a request that requires all the resource  $R$  units, represented by a set of modules ESD on magnetic disks, for its implementation;  $p$  is a request that requires a certain unit of the resource  $R$ , i. e. the addressed ESD module;  $s$  is such a request, when any resource  $R$  unit is enough to perform it;

$f_m(x)$  is a function used to indicate the mode or type of request  $x$  processing;

$p_w(x)$  is a unary predicate, its truth area represents the total queue of all requests (irrespective to the type of request).

The module  $m_e$  performs selecting of  $e$ -type requests from the queue  $p_w$  and generates the queue  $p_{we}$  from the selected requests;  $m_{sif}$  module chooses from the queue  $p_w$  requests of  $p$ -type and puts them in a queue  $p_{wsif}$ , when the module  $m_{si}$  is performed request of  $p$ -type is passed from queue  $p_{wsif}$  to queue  $p_{wsi}$  with formed number (address) of a resource  $R$  unit; module  $m_{si}$  for requests of the  $p$ -type generates a number  $i$  of resource  $R$  unit by an operator  $(\exists! i \in I) p_{irand}(i)$ , wherein  $p_{irand}(i)$  is a unary predicate; module  $m_s$  selects from the queue  $p_w(x)$  requests of  $s$ -type and places them in a queue  $p_{ws}(x)$ ; module  $m_{ate}$  performs the necessary checking to further performing of  $e$ -type request; when the module  $m_{ati}$  is performed, request of  $p$ -type is transmitted from the queue  $p_{wsi}$  to queue  $p_{usi}$  in case of the availability of the corresponding token with the required number;  $m_{ta}$  module selects a request of  $s$ -type from the queue  $p_{ws}$  and transmits it to queue  $p_{us}$ , where it will stay until the end of the implementation of work operation;  $m_{ue}$  module captures the end of the  $e$ -type request servicing and then deletes this request from the queue  $p_{ue}$ ; module  $m_{usi}$  returns a  $p$ -type request back into input queue  $p_{got}$  after the completion of work operation by a certain module of ESD;  $m_{us}$  module is performed similar to the previous module  $m_{usi}$ : it captures the end of  $s$ -request servicing, returns it back to the input queue  $p_{got}$ , and returns the used token, as well.

Concluding, we note that questions of the using of replicated databases examined, for example, in [11–18] are among the complex and require further development. The solution of this problem shows that the networks of abstract machines have large expressive abilities in representation of asynchronous distributed systems, including the abilities to represent local management, parallel, conflict and asynchronous events. Possibility of hierarchical simulation based on them and non-interpreting of network model provides the ability to describe the queuing systems at different levels of abstraction and the system hierarchy. These networks have abilities for reconfiguration and self-modification, this fact extends their functionality significantly. Formalism used there corresponds to such a trend in networking software creating, in which the borders between network operating systems, distributed database management system and a distri-

buted application are removed, and the project "is seen" by the developer at all its stages.

#### References

1. Genrich, H. J. Predicate/transition nets // *Lecture Notes in Computer Science*. Vol. 254. Springer-Verlag, 1986. P. 207–247.
2. Genrich H. J. Equivalence transformations of PrT-Nets // *Lecture Notes in Computer Science*. Vol. 424. Springer-Verlag, 1990. P. 179–208.
3. Glaesser U. Combining abstract state machines with predicate/transition nets // *Lecture Notes in Computer Science*. Vol. 1333. Springer-Verlag, 1997. P. 108–122.
4. Zinkin S. A. Seti abstractnyh mashin vysshyyh poryadkov v proektirovanii system i setey hraneniya i obrabotki dannyh (bazovyi formalizm i ego rasshireniya) // *Izvestiya vysshyyh uchebnyh zavadeniy. Povolzhskiy region. Tehnicheskie nauki*. 2007. № 3. P. 13–22.
5. Zinkin S. A. Seti abstractnyh mashin vysshyyh poryadkov v proektirovanii system i setey hraneniya i obrabotki dannyh (mekhanizmy interpretatsii i varianty ispol'zovaniya) // *Izvestiya vysshyyh uchebnyh zavadeniy. Povolzhskiy region. Tehnicheskie nauki*. 2007. № 4. P. 37–51.
6. Zinkin S. A. Realizatsiya bar'ernoy sinhronizatsii i upravlenie protsessami v virtual'nom setevom diskovom massive // *Informatsionnye tehnologii*. 2008. № 12. P. 22–29.
7. Zinkin S. A. Elementy novoy ob'ektno-orientirovannoy tehnologii dlya modelirovaniya i realizatsii sistem i setey hraneniya i obrabotki dannyh. // *Informatsionnye tehnologii*. 2008. № 10. P. 20–27.
8. Glushkov V. M. Algebra. Yaziki. Programirovanie / V. M. Glushkov, G. E. Tseytlin, E. L. Yushchenko. Kiev: Naukova dumka, 1978. 320 p.
9. Glushkov V. M. Metody simvol'noy multiobrabotki / V. M. Glushkov, G. E. Tseytlin, E. L. Yushchenko. Kiev: Naukova dumka, 1980. 252 p.
10. Kapitonova U. V. Matematicheskaya teoriya proektirovaniya vychislitel'nyh system / U. V. Kapitonova, A. A. Letichevskiy. M.: Nauka, 1988. 296 p.
11. Khan S., Ahmad I. Replicating Data Objects in Large Distributed Database Systems: an Axiomatic Game Theoretic Mechanism Design Approach // *Distributed and Parallel Databases*. Springer Netherlands, 2010. Vol. 28, no. 2. P. 187–218.
12. Özsu M. T., Valduriez P. Principles of Distributed Database Systems. New York: Springer-Verlag, 2011. 866 p.
13. A Distributed System for Processing Declarative Reachability Queries over Partitioned Graphs / M. Sarwat, S. Elnikety, Y. He, M. F. Mokbel // *VLDB 2014: The 40th International Conference on Very Large Data Bases*. Hangzhou, China, 2014. P. 1918–1929.
14. Sourì A., Pashazadeh S., Navin A. H. Consistency of Data Replication Protocols in Database Systems: A Review // *International Journal on Information Theory (IJIT)*. 2014. Vol. 3, no. 4. P. 19–32.
15. Nukarapu D. T., Bin T., Liqiang W., Shiyong L. Data Replication in Data Intensive Scientific Applications with Performance Guarantee // *IEEE Transactions on Parallel and Distributed Systems*. 2011. Vol. 22. P. 1299–1306.
16. Choi S. Youn H., Dynamic hybrid replication effectively combining tree and grid topology // *The Journal of Supercomputing*. 2012. Vol. 59. P. 1289–1311.
17. Hassan O., Ramaswamy L., Miller J. et al. Replication in Overlay Networks: A Multi-objective Optimization Approach, in *Collaborative Computing // Networking, Applications and Worksharing*. Vol. 10. Springer Berlin Heidelberg, 2009. P. 512–528.
18. Lloret J., Garcia M., Tomas J., Rodrigues J. J. Architecture and protocol for inter-cloud communication // *Information Sciences*. 2014. Vol. 258. P. 434–451.

# ОБ ОДНОЙ МОДЕЛИ СИСТЕМЫ ОБСЛУЖИВАНИЯ-ЗАПАСАНИЯ С ОГРАНИЧЕННОЙ ОЧЕРЕДЬЮ НЕТЕРПЕЛИВЫХ ЗАЯВОК

*С. А. Багирова*

*Бакинский государственный университет, Баку, Азербайджан*

В классических моделях систем обслуживания (СО) предполагается, что они имеют неограниченные ресурсы, и потому в них потери заявок связаны лишь с ограниченным количеством каналов и/или мест для их ожидания в очереди; иными словами, обслуживание заявок не приводит к уменьшению ресурсов системы. С другой стороны, в классических моделях управления запасами (УЗ) не учитываются возможности потери и/или образования очереди расходуемых заявок ( $p$ -заявок) при наличии необходимого количества ресурсов системы; иными словами, в них предполагается, что система имеет неограниченное число каналов для отпуска требуемых ресурсов. Вместе с тем в ряде реальных систем обеспечения материальными ресурсами эти предположения не выполняются. Это объясняется тем, что, во-первых, число каналов, которые осуществляют отпуск ресурсов  $p$ -заявкам, всегда является конечным, и, во-вторых, обслуживание каждой  $p$ -заявки приводит к уменьшению уровня запасов системы, и потому, при определенном (критическом) уровне запасов следует прекратить обслуживание  $p$ -заявок. Иными словами, в рамках единой модели наряду с уровнем ресурсов системы необходимо учитывать еще и число расходуемых заявок. Такие системы называются системами обслуживания-запасания (СОЗ) [1, 2]. Они также называются системами обслуживания с встречными потоками [3, 4] или системами транспортно-складского типа [5, 6].

Анализ литературы показал, что модели СОЗ недостаточно исследованы. Исходя из этого, здесь изучается модель СОЗ при достаточно общих предположениях относительно поведения разнотипных заявок.

Изучаемая СОЗ имеет склад ограниченного объема  $S$ . В эту систему поступает пуассоновский поток расходуемых  $p$ -заявок с интенсивностью  $\lambda_p$ . Для простоты изложения предположим, что каждая  $p$ -заявка требует ресурса единичного размера. Отпуск ресурсов осуществляется через единственный канал системы. Время обслуживания  $p$ -заявок является экспоненциально распределенной случайной величиной с параметром  $\mu_p$ . После завершения обслуживания  $p$ -заявки уровень ресурсов в складе системы уменьшается на единицу.

Максимальная длина очереди  $p$ -заявок может быть равна  $N$ , т.е. если  $p$ -заявка поступила в момент, когда в очереди уже имеются  $N$  таких заявок, то независимо от уровня ресурсов в складе системы она теряется с вероятностью 1. Предполагается, что в моменты поступления  $p$ -заявки не

имеют информацию об уровне ресурсов в складе системы, и они принимаются независимо от уровня ресурсов в складе системы, если имеются свободные места в буфере системы. Вместе с тем нетерпеливость  $p$ -заявок проявляется в период ожидания в очереди: допустимые времена ожидания в очереди  $p$ -заявок, когда уровень ресурсов в складе системы равен  $i$ , являются независимыми случайными величинами, которые имеют экспоненциальные распределения со средними  $\tau^{-1}(i)$ . Иными словами, нетерпеливая  $p$ -заявка теряется из очереди, если до окончания допустимого интервала ожидания не освобождается канал обслуживания.

Обслуживание  $p$ -заявки не осуществляется, если в системе отсутствуют ресурсы, т.е. отпуск ресурсов к  $p$ -заявкам продолжается, пока склад системы не является пустым. Пополнения склада системы ресурсами выполняются с помощью снабжающих заявок ( $c$ -заявок) и осуществляются согласно схеме  $(s, S)$ . Это означает, что когда уровень ресурсов в складе системы становится меньше или равно некоторой пороговой величине  $s$ , отправляется заказ на вышестоящий склад на поставку ресурсов объема  $S - s$ . При этом требуется, чтобы после выполнения заказа уровень ресурсов в складе системы должен быть не меньше указанной пороговой величины  $s$ . Следовательно, для предотвращения случаев многократных заказов ресурса необходимо выполнение соотношения  $s < S/2$ . Сделанный заказ выполняется с некоторой задержкой  $c$ -заявок, вызванной доставкой и выгрузкой ресурсов в склад данной системы, т.е. время выполнения заказа является ненулевым. Оно распределено экспоненциально с параметром  $\nu(i)$ , который в общем случае зависит от текущего уровня  $i$  ресурсов в складе системы,  $i = 0, 1, \dots, [S/2] - 1$ , где  $[a]$  означает целую часть  $a$ .

Обслуживание  $c$ -заявок и  $p$ -заявок осуществляется на различных каналах, и эти процессы не зависят друг от друга. Иными словами, допускается отпуск ресурсов  $p$ -заявкам в период времени их выгрузки в склад системы (т.е. во время обслуживания  $c$ -заявок).

Задача состоит в определении совместного распределения уровня ресурсов в складе системы и длины очереди  $p$ -заявок. Оно позволит определить также усредненные характеристики изучаемой СОЗ – средний уровень ресурсов в складе ( $Q_{av}$ ), среднюю длину очереди  $p$ -заявок ( $L_{av}$ ) и вероятность потери  $p$ -заявок ( $PB_p$ ).

Данная СОЗ описывается двумерной цепью Маркова (ЦМ) с состояниями вида  $(m, n)$ , где  $m$  – уровень ресурсов в складе,  $n$  – число  $p$ -заявок в очереди. Фазовое пространство состояний (ФПС) этой ЦМ определяется так:

$$E = \{(m, n) : m = 0, 1, \dots, S; n = 0, 1, \dots, N\}. \quad (1)$$

Рассмотрим проблемы определения элементов производящей матрицы ( $Q$ -матрицы) данной цепи Маркова. Интенсивность перехода из состояния  $(m_1, n_1) \in E$  в другое состояние  $(m_2, n_2) \in E$  обозначим через  $q((m_1, n_1), (m_2, n_2))$ . В общем случае переходы между состояниями ФПС (1) связаны со следующими событиями: (i) с поступлением  $p$ -заявок, (ii) с завершением обслуживания  $p$ -заявок, (iii) с уходом  $p$ -заявок из очереди из-за нетерпеливости и (iv) с поступлением ресурсов из вышестоящего склада.

Исходя из этого, необходимо различить следующие случаи при определении исходного состояния  $(m_1, n_1) \in E$ : 1)  $m_1 > s$ ; 2)  $m_1 \leq s$ .

Сначала рассмотрим случай  $m_1 > s$ . В этом случае выходы из данного состояния  $(m_1, n_1)$  из-за события типа (iv) не возможны, так как в таких состояниях не осуществляется пополнение склада ресурсами. Если поступает некоторая  $p$ -заявка (события типа (i)), то она присоединяется к очереди при выполнении условия  $n_1 < N$ ; иными словами, в моменты поступления  $p$ -заявки при выполнении условия  $n_1 < N$  осуществляется переход из данного состояния в состояние  $(m_1, n_1 + 1) \in E$ . Интенсивность такого перехода равна  $\lambda_p$ . Если в исходном состоянии выполняется условие  $n_1 = N$ , то поступившая  $p$ -заявка теряется. После завершения обслуживания  $p$ -заявки (события типа (ii)) в исходном состоянии  $(m_1, n_1) \in E, n_1 > 0$  осуществляется переход в состояние  $(m_1 - 1, n_1 - 1) \in E$ . Интенсивность такого перехода равна  $\mu_p$ . Если в исходном состоянии  $(m_1, n_1) \in E, n_1 > 0$ , некоторая  $p$ -заявка уходит из очереди необслуженной (события типа (iii)), то происходит переход из данного состояния в состояние  $(m_1, n_1 - 1) \in E$ . Интенсивность такого перехода равна  $n_1 \tau(m_1)$ . Следовательно, для случаев  $m_1 > s$  указанные выше элементы  $Q$ -матрицы определяются так:

$$q((m_1, n_1), (m_2, n_2)) = \begin{cases} \lambda_p, & \text{если } m_2 = m_1, n_2 = n_1 + 1, n_1 \leq N - 1 \\ \mu_p, & \text{если } m_2 = m_1 - 1, n_2 = n_1 - 1, n_1 > 0, \\ n_1 \tau(m_1), & \text{если } m_2 = m_1, n_2 = n_1 - 1, n_1 > 0, \\ 0 & \text{в остальных случаях.} \end{cases} \quad (2)$$

Пусть теперь в исходном состоянии  $(m_1, n_1) \in E$  выполняется условие  $m_1 \leq s$ . В этих состояниях интенсивности переходов для указанных выше событий типа (i)–(iii) определяются аналогично соотношениям (2). Вместе с тем в этих состояниях в момент поступления заказа из вышестоящего склада объема  $S - s$  происходит переход в состояние  $(m_1 + S - s, n_1)$ ; ин-

тенсивность такого перехода равна  $\nu(m_1)$ . Следовательно, для случаев  $m_1 \leq s$  указанные выше элементы  $Q$ -матрицы определяются так:

$$q((m_1, n_1), (m_2, n_2)) = \begin{cases} \lambda_p, & \text{если } m_2 = m_1, n_2 = n_1 + 1, n_1 \leq N - 1, \\ \mu_p, & \text{если } m_2 = m_1 - 1, n_2 = n_1 - 1, n_1 > 0, \\ n_1 \tau(m_1), & \text{если } m_2 = m_1, n_2 = n_1 - 1, n_1 > 0, \\ \nu(m_1), & \text{если } m_1 \leq s, m_2 = m_1 + S - s, n_2 = n_1, \\ 0 & \text{в остальных случаях.} \end{cases} \quad (3)$$

С учетом соотношений (2) и (3) заключаем, что все состояния этой двумерной ЦМ являются сообщающимися, следовательно, в этой системе существует стационарный режим. Пусть  $p(m, n)$  означает стационарную вероятность состояния  $(m, n) \in E$ . Эти вероятности удовлетворяют систему уравнений равновесия (СУР), которая составляется на основе соотношений (2) и (3). Она имеет следующий вид:

случай  $m > s$ :

$$(\lambda_p I(n < N) + \mu_p I(n > 0) + n\tau(m))p(m, n) = \lambda_p p(m, n - 1)I(n > 0) + \mu_p p(m + 1, n + 1) + (n + 1)\tau(m)p(m, n + 1) + \nu(m + s - S)p(m + s - S, n); \quad (4)$$

случай  $m \leq s$ :

$$(\lambda_p I(n < N) + \mu_p I(n > 0) + n\tau(m) + \nu(m))p(m, n) = \lambda_p p(m, n - 1)I(n > 0) + \mu_p p(m + 1, n + 1) + (n + 1)\tau(m)p(m, n + 1). \quad (5)$$

К этой СУР добавляются условия нормировки:

$$\sum_{(m, n) \in E} p(m, n) = 1. \quad (6)$$

Относительно решения СУР (4)–(6), отметим, что, к сожалению, из-за сложной структуры ее матрицы не удастся найти ее аналитическое решение. Вместе с тем для моделей умеренной размерности для этой цели целесообразно использовать итеративные методы линейной алгебры, в частности, метод Гаусса–Зейделя.

После нахождения совместного распределения уровня ресурсов в складе системы и длины очереди  $p$ -заявок можно вычислить усредненные характеристик исследуемой СОЗ. Так, средняя длина очереди  $p$ -заявок ( $L_{av}$ ) и средний уровень ресурсов в складе ( $S_{av}$ ) определяются как математические ожидания соответствующих случайных величин. Иными словами, эти параметры определяются из следующих формул:

$$L_{av} = \sum_{n=1}^N n \sum_{m=0}^S p(m, n); \quad (7)$$

$$S_{av} = \sum_{m=1}^S m \sum_{n=0}^N p(m, n). \quad (8)$$

Потеря  $p$ -заявок происходит в следующих случаях:

1) в момент поступления такой заявки в очереди отсутствует свободное место;

2)  $p$ -заявка теряется из очереди, если до окончания допустимого интервала ожидания не освобождается канал обслуживания. Следовательно, для вычисления вероятности потери  $p$ -заявок ( $PB_p$ ) получим следующую формулу:

$$PB_p = \sum_{m=0}^S p(m, N) + \frac{1}{\lambda_p} \sum_{n=1}^N n \sum_{m=0}^S \tau(m) p(m, n). \quad (9)$$

Таким образом, здесь разработан точный метод для вычисления характеристик изучаемой модели СОЗ. Он является эффективным для применения в системах с умеренными значениями объема склада и максимальной длины очереди  $p$ -заявок. Вместе с тем в реальных системах указанные величины принимают достаточно большие значения, и в связи с этим размерность ФПС (1) становится чрезмерно большой. Это обстоятельство делает проблематичным точное вычисление стационарного распределения соответствующей двумерной цепи Маркова, и, таким образом, возникают вычислительные трудности при нахождении характеристик (7)–(9) изучаемой СОЗ. Поэтому приходится ограничиваться приближенными вычислениями стационарного распределения и характеристик системы. С этой целью можно использовать алгоритмы фазового укрупнения состояний двумерных цепей Маркова [7]. Эти задачи являются объектами исследования последующих публикаций.

#### Литература

1. Schwarz M., Daduna H. Queuing systems with inventory management with random lead times and with backordering // *Mathematical methods of operations research*. 2006. Vol. 64, issue 3. P. 383–414.
2. Schwarz M., Sauer C., Daduna H. et al. M/M/1 queuing systems with inventory // *Queuing systems. Theory and applications*. 2006. Vol. 54, issue 1. P. 55–78.
3. Постан М. Я. Применение Марковских процессов для моделирования систем обслуживания встречных транспортных потоков / Препринт № 88-6. Институт кибернетики НАН Украины. Киев, 1989. 14 с.
4. Меликов А. З., Фаталиева М. Р. Управление запасами систем обслуживания разнотипных встречных потоков с учетом текущей ситуации // *Электронное моделирование*. 1997. Т. 19, № 6. С. 106–115.
5. Меликов А. З., Молчанов А. А. Оптимизация процесса запасаения ресурсов в системах транспортно-складского типа // *Кибернетика и системный анализ*. 1992. № 3. С. 179–184.
6. Меликов А. З. Марковская модель процесса накопления в системах транспортно-складского типа // *Электронное моделирование*. 1996. Т. 18, № 6. С. 79–83.
7. Ponomarenko L., Kim C. S., Melikov A. Performance analysis and optimization of multi-traffic on communication networks. Springer, 2010.

# СИСТЕМА $M / GI / \infty$ В ПОЛУМАРКОВСКОЙ СЛУЧАЙНОЙ СРЕДЕ

*Г. В. Баймеева, А. А. Назаров*

*Национальный исследовательский*

*Томский государственный университет, Томск, Россия*

Рассматривается система массового обслуживания  $M / GI / \infty$ , функционирующая в полумарковской случайной среде. У этой системы имеется один простейший входящий поток с интенсивностью  $\lambda_s N$  и неограниченное число обслуживающих приборов, на каждом из которых время обслуживания распределено по закону  $B_s(x)$ , где  $s = \overline{1, K}$  – это состояние полумарковского процесса  $s(t)$ , который определяется полумарковской матрицей вида

$$Q(x) = PA(Nx).$$

Здесь матрица  $P$  – это переходная матрица вероятностей процесса  $s(t)$  и матрица  $A(Nx)$  – это диагональная матрица с функциями распределения  $N$ -кратного времени пребывания процесса  $s(t)$  в каждом из своих состояний. Мы используем большой параметр  $N$ , который в данном случае определяет условие высокой интенсивности входящего потока.

Отметим, что мы рассматриваем ситуацию, когда время обслуживания заявок, находящихся в системе, не меняется с переходом среды  $s(t)$  из одного состояния в другое. Другими словами, заявки, прибывшие в систему, обслуживаются только по одному закону до самого своего ухода.

Таким образом, процесс изменения состояния системы и среды обозначим через  $\{i(t), s(t)\}$ , где  $i(t) = i \geq 0$  – число заявок в системе в момент времени  $t$ . Этот процесс, очевидно, не является марковским. Для того чтобы марковизировать его, воспользуемся методом динамического просеивания и методом дополнительной переменной.

Суть метода динамического просеивания заключается в том, чтобы свести задачу исследования немарковского процесса к исследованию нестационарного потока событий. Допустим, что в момент времени  $t_0$  в системе нет заявок. Выберем произвольный момент времени  $T$  и на временной оси  $(t_0, T)$  будем отмечать моменты поступления заявок в систему через входящий поток. Будем говорить, что заявка «просеивается» с вероятностью

$$S_s(t) = 1 - B_s(T - t),$$

если она поступила в систему в момент времени  $t < T$  и не обслужилась полностью до момента времени  $T$ . Таким образом, можем заключить, что просеянные заявки в момент времени  $T$  будут находиться в системе. Обозначим через  $n(t)$  – число событий просеянного потока, наступивших до момента времени  $t$ . Следующее равенство всегда имеет место:

$$i(T) = n(T). \quad (1)$$

Это означает, что характеристики процесса  $i(t)$  в момент времени  $T$  совпадают с характеристиками величины  $n(T)$ .

Теперь выберем момент времени  $t_0$  таким образом, что для всех  $t < t_0$  в системе нет клиентов, т.е.

$$S_s(t) = 1 - B_s(T - t) = 0, t < t_0.$$

Достаточно очевидно, что следует положить  $t_0 = -\infty$ .

Из состояния  $n(t) = n$  в момент времени  $t$  возможны следующие переходы процесса  $n(t)$ :

$$n(t + \Delta t) = \begin{cases} n + 1, & \text{с вероятностью } \lambda_s \Delta t S_s(t) + o(\Delta t), \\ n, & \text{с вероятностью } 1 - \lambda_s \Delta t S_s(t) + o(\Delta t), \end{cases} \quad s = \overline{1, K}.$$

Далее, для марковизации процесса  $s(t)$  воспользуемся методом дополнительной переменной. Введем компоненту  $z(t)$ , имеющую смысл остаточного времени пребывания процесса  $s(t)$  в текущем состоянии. Другими словами,  $z(t)$  – это промежуток времени от момента времени  $t$  до перехода процесса в следующее состояние. Из вышесказанного следует, что марковский трехмерный процесс  $\{s(t), n(t), z(t)\}$  описывает процесс изменения состояния системы и среды с течением времени. Определим вероятности этих состояний в виде

$$P(s, n, z, t) = P\{s(t) = s, n(t) = n, z(t) < \frac{z}{N}\}. \quad (2)$$

Используя формулу полной вероятности, запишем систему дифференциальных уравнений Колмогорова, определяющую вероятности (2):

$$\begin{aligned} & \frac{1}{N} \frac{\partial P(s, n, z, t)}{\partial t} - \frac{\partial P(s, n, z, t)}{\partial z} + \frac{\partial P(s, n, 0, t)}{\partial z} = \\ & = \lambda_s S_s(t) \{P(s, n - 1, z, t) - P(s, n, z, t)\} + A_s(z) \sum_{k=1}^K p_{ks} \frac{\partial P(k, n, 0, t)}{\partial z}. \end{aligned} \quad (3)$$

Начальное условие к такой системе при выполнении условия согласованности  $z = \infty$  запишем в виде

$$P(s, n, z, t_0) = \begin{cases} 0, & \text{если } n > 0, \\ \mathbf{r}(s, z), & \text{если } n = 0, \end{cases} \quad s = \overline{1, K}. \quad (4)$$

Здесь  $\mathbf{r}(z)$  – вектор стационарного распределения вероятностей состояний процесса  $s(t)$ .

Введем частичные характеристические функции вида

$$H(s, u, z, t) = \sum_{n=0}^{\infty} e^{jun} P(s, n, z, t), \quad s = \overline{1, K}, \quad (5)$$

и перепишем систему (3) в виде системы уравнений, определяющей функции (5):

$$\begin{aligned} & \frac{1}{N} \frac{\partial H(s,u,z,t)}{\partial t} - \frac{\partial H(s,u,z,t)}{\partial z} + \frac{\partial H(s,u,0,t)}{\partial z} = \\ & = \lambda_s S_s(t) (e^{ju} - 1) H(s,u,z,t) + A_s(z) \sum_{k=1}^K p_{ks} \frac{\partial H(k,u,0,t)}{\partial z}, \quad s = \overline{1, K}. \end{aligned} \quad (6)$$

Введем векторно-матричные обозначения:

$$\begin{aligned} \mathbf{\Lambda} &= \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_K \end{bmatrix}, \quad \mathbf{S}(t) = \begin{bmatrix} S_1(t) & 0 & \dots & 0 \\ 0 & S_2(t) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & S_K(t) \end{bmatrix}, \\ \mathbf{H}(u,t) &= [H(1,u,t) \quad H(2,u,t) \quad \dots \quad H(K,u,t)], \\ \mathbf{P} &= \begin{bmatrix} p_{11} & p_{12} & \dots & p_{1K} \\ p_{21} & p_{22} & \dots & p_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ p_{K1} & p_{K2} & \dots & p_{KK} \end{bmatrix}, \quad \mathbf{A}(x) = \begin{bmatrix} A_1(x) & 0 & \dots & 0 \\ 0 & A_2(x) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & A_K(x) \end{bmatrix}. \end{aligned}$$

С их помощью перепишем систему (6) в следующем виде:

$$\begin{aligned} & \frac{1}{N} \frac{\partial \mathbf{H}(u,z,t)}{\partial t} - \frac{\partial \mathbf{H}(u,z,t)}{\partial z} + \frac{\partial \mathbf{H}(u,0,t)}{\partial z} [\mathbf{I} - \mathbf{PA}(z)] = \\ & = (e^{ju} - 1) \mathbf{H}(u,z,t) \mathbf{\Lambda S}(t). \end{aligned} \quad (7)$$

Начальное условие для решения этой системы запишется в виде

$$\mathbf{H}(u,z,t_0) = \mathbf{r}(z). \quad (8)$$

Далее найдем приближение первого порядка к рассматриваемой характеристической функции методом асимптотического анализа. Введем обозначения:

$$\varepsilon = \frac{1}{N}, \quad u = \varepsilon w, \quad \mathbf{H}(u,z,t) = \mathbf{F}_1(w,z,t,\varepsilon).$$

Перепишем с их использованием систему (7) в следующем виде

$$\begin{aligned} & \varepsilon \frac{\partial \mathbf{F}_1(w,z,t,\varepsilon)}{\partial t} - \frac{\partial \mathbf{F}_1(w,z,t,\varepsilon)}{\partial z} + \frac{\partial \mathbf{F}_1(w,0,t,\varepsilon)}{\partial z} [\mathbf{I} - \mathbf{PA}(z)] = \\ & = (e^{j\varepsilon w} - 1) \mathbf{F}_1(w,z,t,\varepsilon) \mathbf{\Lambda S}(t). \end{aligned} \quad (9)$$

**Этап 1.** При  $\varepsilon \rightarrow 0$  получим, что

$$\frac{\partial \mathbf{F}_1(w,z,t)}{\partial z} = \frac{\partial \mathbf{F}_1(w,0,t)}{\partial z} [\mathbf{I} - \mathbf{PA}(z)]. \quad (10)$$

Решение системы (9) будем искать в виде

$$\mathbf{F}_1(w,z,t) = \mathbf{r}(z) \Phi_1(w,t), \quad (11)$$

где  $\Phi_1(w,t)$  – некоторая скалярная функция, которая будет определена позже. После подстановки этой замены в (10), получаем, что

$$\mathbf{r}(z) = \mathbf{r}'(0) \int_0^z [\mathbf{I} - \mathbf{PA}(x)] dx.$$

Вектор стационарного распределения вероятностей состояний среды  $\mathbf{r}(z)$  определяется по формуле (12):

$$\mathbf{r}(z) = \frac{1}{a} \int_0^z \mathbf{r}[\mathbf{I} - \mathbf{P}\mathbf{A}(x)] dx, \quad a = \mathbf{r}\mathbf{A}\mathbf{e}, \quad (12)$$

где  $\mathbf{r}$  – вектор стационарного распределения вероятностей состояний вложенной цепи Маркова процесса  $s(t)$ ,  $\mathbf{A}$  – это диагональная матрица, содержащая на главной диагонали средние значения функций распределения

$$A_s(x), \quad s = \overline{1, K}.$$

**Этап 2.** В векторно-матричном уравнении (9) положим  $z = \infty$ , поделим обе части на  $\varepsilon$  и устремим  $\varepsilon \rightarrow 0$ . Тогда получим, что

$$\frac{\partial \mathbf{F}_1(w, t)}{\partial t} \mathbf{e} = jw \mathbf{F}_1(w, t) \mathbf{A} \mathbf{S}(t) \mathbf{e}.$$

Подставляя в последнее соотношение замену (11), получим обыкновенное дифференциальное уравнение первого порядка относительно функции  $\Phi_1(w, t)$ :

$$\frac{\partial \Phi_1(w, t)}{\partial t} = jw \Phi_1(w, t) \frac{1}{a} \mathbf{r} \mathbf{A} \mathbf{A} \cdot \mathbf{S}(t) \mathbf{e}.$$

Его решение с учетом начального условия  $\Phi_1(w, t_0) = 1$  запишется в виде

$$\Phi_1(w, t) = \exp\{jw \kappa_1(t)\}, \quad \kappa_1(t) = \frac{1}{a} \int_{-\infty}^t \mathbf{r} \mathbf{A} \mathbf{A} \cdot \mathbf{S}(\tau) \mathbf{e} d\tau, \quad (13)$$

где  $w = uN$ .

Таким образом, был найден вид первого приближения к характеристической функции числа событий потока  $n(t)$ :

$$M\{e^{jun(t)}\} = \mathbf{H}(u, t) \mathbf{e} \approx \mathbf{F}_1(w, t, \varepsilon) \mathbf{e} = \exp\{juk_1(t)N\}.$$

Используя равенство (1), мы можем получить вид первого приближения к характеристической функции числа заявок в системе в момент времени  $T$ :

$$M\{e^{juiT}\} = M\{e^{jun(T)}\} = \mathbf{H}(u, T) \mathbf{e} \approx h_1(u, T) = \exp\{juk_1(T)N\}. \quad (14)$$

Посчитаем в явном виде величину  $\kappa_1(T)$ :

$$\kappa_1(T) = \frac{1}{a} \int_{-\infty}^T \mathbf{r} \mathbf{A} \mathbf{A} \cdot \mathbf{S}(t) \mathbf{e} dt = \frac{1}{a} \int_{-\infty}^T \mathbf{r} \mathbf{A} \mathbf{A} [\mathbf{I} - \mathbf{B}(T-t)] \mathbf{e} dt.$$

Здесь  $\mathbf{B}(t)$  – диагональная матрица, содержащая на главной диагонали условные функции распределения времени обслуживания при каждом из со-

стояний  $s = \overline{1, K}$  процесса  $s(t)$ . В последнем интеграле сделаем замену  $\tau = T - t$ :

$$\frac{1}{a} \int_{-\infty}^T \mathbf{r} \Lambda \Lambda [\mathbf{I} - \mathbf{B}(T - t)] e^{at} dt = \frac{1}{a} \mathbf{r} \Lambda \Lambda \int_0^{\infty} [\mathbf{I} - \mathbf{B}(\tau)] d\tau e = \frac{1}{a} \mathbf{r} \Lambda \Lambda \cdot \mathbf{B}e.$$

Здесь  $\mathbf{B}$  – диагональная матрица, содержащая на главной диагонали условные средние значения времени обслуживания заявки. Таким образом, окончательный вид среднего значения числа заявок в системе в момент времени  $T$  определяется в виде

$$\kappa_1(T) = \frac{1}{a} \mathbf{r} \Lambda \Lambda \cdot \mathbf{B}e. \quad (15)$$

Таким образом, была рассмотрена система массового обслуживания  $M(\lambda_s) / G(B_s(x)) / \infty$ , в которой параметр  $\lambda_s$  простейшего входящего потока и распределение вероятностей длительности времени обслуживания  $B_s(x)$  меняются в зависимости от состояния внешнего полумарковского процесса  $s(t) = s = \overline{1, K}$ . С использованием методов динамического просеивания, дополнительной переменной и асимптотического анализа в условиях растущей интенсивности входящего потока и частой смены состояний среды было получено приближенное распределение вероятностей числа заявок в системе в момент времени  $T$ . В дальнейшем планируется провести асимптотический анализ второго порядка данной системы для получения более точного приближения к виду характеристической функции числа заявок в системе.

#### Литература

1. Baykal-Gursoy M., Xiao W. Stochastic Decomposition in M/M/∞ Queues with Markov Modulated Service Rates // Queueing Syst. 2004. Vol. 48. P. 75–88.
2. O'Connell C.A., Purdie P. The M/M/∞ Queue in a Random Environment // J. Appl. Prob. 1986. Vol. 23. P. 175–184.
3. Blom J., Kella O., Mandjes M., Thorsdottir H. Markov-Modulated Infinite-Server Queues with General Service Times // Queueing Syst. 2014. Vol. 76. P. 403–424.
4. D'Auria B. M/M/∞ queues in semi-Markovian random environment // Queueing Syst. 2008. Vol. 58. P. 221–237.
5. Fralix B.H., Adan I.J.B.F. An Infinite-Server Queue Influenced by a Semi-Markovian Environment // Queueing Syst. 2009. Vol. 61. P. 65–84.
6. Назаров А. А. Теория вероятностей и случайных процессов: учеб. пособие / А. А. Назаров, А. Ф. Терпугов. 2-е изд., испр. Томск: Изд-во НТЛ, 2010. 204 с.
7. Назаров А. А. Теория массового обслуживания: учеб. пособие / А. А. Назаров, А. Ф. Терпугов. Томск: Изд-во НТЛ, 2004. 228 с.
8. Фихтенгольц Г. М. Курс дифференциального и интегрального исчисления: в 3 т. / Г. М. Фихтенгольц. 8-е изд. М.: Физматлит, 2001. Т. 1–3. 851 с.

# СРАВНИТЕЛЬНЫЙ АНАЛИЗ ДИСЦИПЛИН ОБСЛУЖИВАНИЯ ДЛЯ СИСТЕМ ПОЛЛИНГА С ДВУМЯ ОЧЕРЕДЯМИ

*Ю. В. Гайдамака, Э. Р. Зарипова*

*Российский университет дружбы народов, Москва, Россия*

## **Введение**

Мультимедийная подсистема IMS (англ., IP Multimedia Subsystem) предназначена для надлежащего обслуживания мультимедийных и интерактивных телекоммуникационных услуг. Основным сигнальным протоколом подсистемы IMS является протокол установления сессий (англ. Session Initiation Protocol, SIP), отвечающий за соединение и поддержание соединения. Типичная процедура установления сессии предполагает два типа сигнальных сообщений: иницирующие и продолжающие сессию, где продолжающие сессию сообщения обладают более высоким приоритетом. Для анализа процесса обслуживания сигнальных сообщений с приоритетом авторами предложено использование несимметричных поллинговых моделей с двумя очередями и ненулевым временем переключения. Для данных, применимых к анализу сигнального трафика, оценено время ожидания начала обслуживания для шлюзовой и исчерпывающей дисциплины обслуживания очередей. Анализ симметричных систем поллинга с одинаковым временем переключения показал преимущество исчерпывающей дисциплины обслуживания.

## **Система поллинга с двумя очередями**

Рассмотрим две наиболее распространенные дисциплины обслуживания очередей: шлюзовую и исчерпывающую [1–3]. При исчерпывающей дисциплине обслуживания прибор, подключаясь к очереди, обслуживает заявки в очереди до тех пор, пока она не опустеет, затем переключается к следующей очереди. При шлюзовой дисциплине прибор при подключении к очереди фиксирует число заявок и обслуживает только их, затем переключается к следующей очереди. Необходимо отметить, что в общем случае время переключения прибора между очередями ненулевое.

Исследование сигнального трафика протокола SIP позволяет выделить в процедурах два типа сообщений: Invite и non-Invite.

Сообщения типа Invite инициируют процедуры установления сессии, авторизации и другие процедуры. Остальные сообщения (200 Ok, Ack, Bye, 100 Trying, 180 Ringing и др.) образуют non-Invite сообщения, которые обладают большим приоритетом, так как они увеличивают количество успешно завершенных процедур. После успешного обслуживания сообщения Invite серверу необходимо в последующем обработать еще несколько сообщений типа non-Invite, завершающих сессию [4]. Сигнальные сообщения имеют так называемое «время жизни», поэтому

важно оценивать время пребывания сообщения в сервере протокола обслуживания сессий, которое складывается из времени ожидания начала обслуживания и времени обслуживания. Время ожидания начала обслуживания оценивается авторами в данной работе.

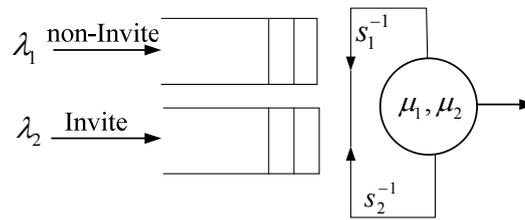


Рис. 1. Система поллинга с двумя очередями

На рис. 1 показана система поллинга с двумя входящими потоками заявок и ненулевым переключением между очередями. Разделение сигнальных сообщений на два типа предложено в [5] для контроля за перегрузками, возникающими в сетях связи.

В математической модели процесса обслуживания разнотипных сигнальных сообщений в виде системы поллинга предположим, что входящие потоки пуассоновские с интенсивностями  $\lambda_i$ , а среднее время обслуживания заявки из очереди  $Q_i$  равно  $b_i = 1 / \mu_i$ ,  $i = \overline{1, 2}$ . Пусть среднее время переключения к очереди  $Q_i$  равно  $s_i$ , а  $s = s_1 + s_2$  – полное время переключения между очередями. Определим нагрузку, образованную заявками из  $Q_i$  очереди равную  $\rho_i = \lambda_i / \mu_i$ . Тогда нагрузка от двух очередей равна  $\rho = \rho_1 + \rho_2$ . Гамильтонов цикл  $C = \rho C + s$  равен времени обслуживания прибором заявок в одном цикле и определяется в выбранных обозначениях следующим образом:  $C = \frac{s}{1 - \rho}$  [1].

### Шлюзовая дисциплина обслуживания очередей

В [6] определено преобразование Лапласа–Стилтьеса (ПЛС) функции распределения (ФР) времени ожидания начала обслуживания при шлюзовой дисциплине  $\tilde{W}_i(x) = \frac{1 - \rho}{s} \times \frac{G_i(\tilde{B}_i(x)) - G_i(1 - x / \lambda_i)}{x - \lambda_i + \lambda_i \tilde{B}_i(x)}$ , где  $\tilde{B}_i(x)$  – ПЛС ФР времени обслуживания заявок из очереди  $Q_i$  со средним значением  $b_i$  и  $k$ -м моментом  $b_i^{(k)}$ ;  $G_i(\cdot)$  – производящая функция числа заявок в системе поллинга в момент опроса очереди  $Q_i$ ,  $i = \overline{1, 2}$ .

Для несимметричной системы поллинга определено среднее время ожидания начала обслуживания при шлюзовой дисциплине, показано среднее время ожидания начала обслуживания для заявок из очереди  $Q_1$ ;  $s_i^{(2)}$  – второй момент времени переключения к очереди  $Q_i$ .

$$\omega_1 = \frac{(1 + \rho_1)}{2s(1 - \rho_1\rho_2)(2\rho_1\rho_2 + \rho + 1)} \times \left\{ (Cb_2^{(2)}\lambda_2 + s_1^{(2)})(1 + \rho_1\rho_2) + \frac{2s_1^2(\rho - \rho_1^2\rho_2^2)}{(1 - \rho)} + (Cb_1^{(2)}\lambda_1 + s_2^{(2)}) \times (1 + 2\rho_2 - 2\rho\rho_2^2 - \rho_1\rho_2) + \frac{2s_1s_2(1 + \rho + \rho_1\rho_2(1 - \rho - 2\rho_1\rho_2))}{(1 - \rho)} + \frac{2s_2^2(\rho_2^2 + \rho_2^3 - \rho_2^4 + \rho_1)}{(1 - \rho)} + \frac{2s_2^2(2\rho_1\rho_2 - 2\rho\rho_1\rho_2^2 - \rho_1^2\rho_2)}{(1 - \rho)} \right\}$$

При симметричной системе поллинга  $\rho_1 = \rho_2 = \rho_i$ ,  $s_1 = s_2 = s_i$ ,  $b_1^{(2)} = b_2^{(2)} = b_i^{(2)}$ ,  $\lambda_1 = \lambda_2 = \lambda_i$ , среднее время ожидания начала обслуживания равно

$$\omega_{gated\_symm} = \frac{b_i^{(2)}\lambda_i}{(1 - 2\rho_i)} + \frac{s_i^{(2)}}{2s_i} + \frac{s_i(4\rho_i + 1)}{2(1 - 2\rho_i)}.$$

### Исчерпывающая дисциплина обслуживания очередей

В [6] определено ПЛС ФР времени ожидания начала обслуживания при исчерпывающей дисциплине  $\tilde{W}_i(x) = \frac{1 - \rho}{s} \times \frac{1 - G_i(1 - x / \lambda_i)}{x - \lambda_i + \lambda_i \tilde{B}_i(x)}$ .

В случае несимметричной системы поллинга с исчерпывающей дисциплиной обслуживания определено среднее время ожидания, для заявок из очереди  $Q_1$  формула представлена в следующем виде:

$$\omega_1 = \frac{s_1^{(2)}(1 - \rho - \rho_2 + 2\rho\rho_2)}{2s(1 - \rho + 2\rho_1\rho_2)} + \frac{s_2^{(2)}(1 - \rho_1)}{2s(1 - \rho + 2\rho_1\rho_2)} + \frac{s_1s_2(1 - \rho_1)}{s(1 - \rho)} + \frac{\lambda_1b_1^{(2)}(1 - \rho - \rho_2 + 2\rho_2\rho)}{2(1 - \rho + 2\rho_1\rho_2)(1 - \rho)} + \frac{b_2^{(2)}\lambda_2(1 - \rho_1)}{2(1 - \rho + 2\rho_1\rho_2)(1 - \rho)} + \frac{s_2^2\rho_1\rho_2(1 - \rho_1)}{s(1 - \rho + 2\rho_1\rho_2)(1 - \rho)} + \frac{s_1^2\rho_2(1 - \rho - \rho_2 + 2\rho_1\rho_2 + \rho_2^2)}{s(1 - \rho + 2\rho_1\rho_2)(1 - \rho)}.$$

Определим аналогично среднее время ожидания начала обслуживания при исчерпывающей дисциплине в симметричной системе поллинга:

$$\omega_{ex\_symm} = \frac{b_i^{(2)}\lambda_i}{(1 - 2\rho_i)} + \frac{s_i^{(2)}}{2s_i} + \frac{s_i}{2(1 - 2\rho_i)}.$$

Заметим, что исчерпывающая дисциплина обслуживания для симметричных систем поллинга является эффективной, так как среднее время ожидания при шлюзовой дисциплине всегда больше среднего времени ожидания при исчерпывающей дисциплине:

$$\omega_{gated\_symm} - \omega_{ex\_symm} = \frac{4\rho_i s_i}{2(1 - 2\rho_i)}.$$

## Численный эксперимент

Исходные данные взяты из [4, 7]:  $s_i = 2$  мс – время переключения к очереди  $Q_i$ ,  $i = \overline{1, 2}$ .  $b_1 = 4$  мс – среднее время обслуживания 1-заявки (non-Invite);  $b_2 = 10$  мс – среднее время обслуживания 2-заявки (Invite);  $\lambda_1 = 6\lambda_2$  для простейшей процедуры установления сессии.

На рис. 2 представлен график среднего времени ожидания начала обслуживания для данных, применимых к анализу сигнального трафика при экспоненциальном времени обслуживания и времени переключения между очередями.

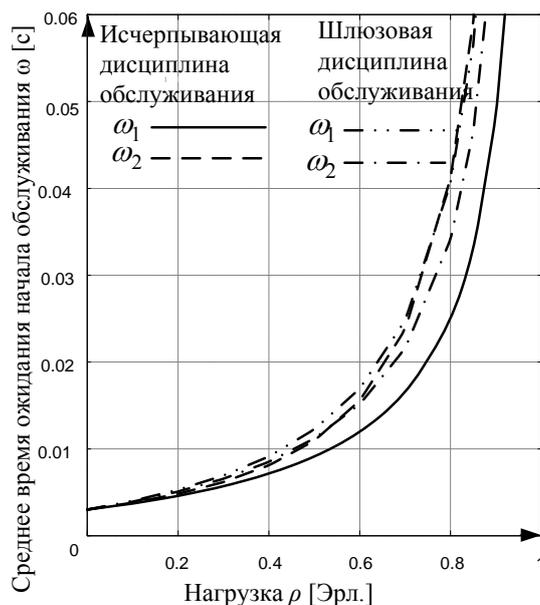


Рис. 2. Среднее время ожидания начала обслуживания

Заметим, что заявки из очереди  $Q_1$  являются приоритетными в обслуживании. Для заявок из первой очереди эффективной является исчерпывающая дисциплина, так как она имеет меньшее время ожидания начала обслуживания. Заметим, что в случае симметричных систем поллинга исчерпывающая дисциплина также всегда имеет меньшее среднее время ожидания начала обслуживания.

В [7] анализировались системы поллинга с мгновенным переключением между очередями. Исчерпывающая дисциплина обслуживания показала эффективность по следующим показателям: меньшая вероятность пребывания системы в режиме перегрузки, меньшая длина приоритетной очереди. Результаты текущей работы и [7] показывают эффективность исчерпывающей дисциплины по отношению к шлюзовой для данных, применимых к анализу сигнального трафика.

Для эффективной исчерпывающей дисциплины обслуживания очередей планируется найти формулы вторых моментов времени ожидания начала обслуживания для несимметричных систем поллинга с ненулевым временем переключения в явном виде.

*Исследование выполнено при финансовой поддержке РФФИ в рамках научных проектов 15-07-03051, 15-07-03608.*

#### Литература

1. Вишнеvский В. М., Семенова О. В. Системы поллинга: Теория и применение в широкополосных беспроводных сетях. М.: Техносфера, 2007. 312 с.
2. Гайдамака Ю. В., Зарипова Э. Р., Болотова Г. О. Разработка модели функционирования SIP-сервера в виде системы поллинга с дисциплиной шлюзового обслуживания // Т-Comm – Телекоммуникации и Транспорт. 2013. № 11. С. 69–72.
3. Гайдамака Ю. В., Зарипова Э. Р., Вихрова О. Г. Применение системы поллинга с исчерпывающей дисциплиной обслуживания к анализу SIP-сервера // Т-Comm – Телекоммуникации и транспорт. 2013. № 11. С. 73–76.
4. Gaidamaka Yu. V., Zaripova E. R. Session setup delay estimation methods for IMS-based IPTV services // Lecture Notes in Computer Science. 8638. 2014. P. 408–418.
5. Ohta M. Overload control in a SIP Signalling Network // International Journal of electrical and Electronics engineering. 2009. P. 87–92.
6. Kudoh S., Takagi H., Hashida O. Second moments of the waiting time in symmetric polling systems // Journal of the Operations Research Society of Japan, 2000. Vol. 43, no. 2. P. 306–316.
7. Shorgin S., Samouylov K., Gaidamaka Y., Etezov Sh. Polling System with Threshold Control for Modeling of SIP Server under Overload // Springer International Publishing. Advances in Intelligent Systems and Computing. 2014. Vol. 240. P. 97–107.

## **ИССЛЕДОВАНИЕ МАТЕМАТИЧЕСКОЙ МОДЕЛИ СТРАХОВОЙ КОМПАНИИ В ВИДЕ СИСТЕМЫ МАССОВОГО ОБСЛУЖИВАНИЯ С НЕОГРАНИЧЕННЫМ КОЛИЧЕСТВОМ ПРИБОРОВ С УЧЕТОМ НЕЯВНОЙ РЕКЛАМЫ**

*Д. Д. Даммер*

*Национальный исследовательский*

*Томский государственный университет, Томск, Россия*

Рассмотрим модель страховой компании с неограниченным страховым полем [1] в виде системы массового обслуживания с неограниченным числом обслуживающих приборов (рис. 1). Срок действия договора страхования соответствует длительности обслуживания заявки на приборе. Пусть в компанию поступают риски, образуя поток событий, с интенсивностью, зависящей от числа застрахованных рисков. Интенсивность этого потока будем определять двумя составляющими: параметром  $\lambda$ , который определяет поток рисков, приходящих независимо от застрахованных рисков, и параметром  $\alpha$ , который определяет поток рисков, находящихся под воздействием «неявной рекламы». Модели с таким потоком входящих рисков рассматриваются в работе [2], но методы исследования моделей носят иной характер, и поток требований на выплату страховых сумм никаким образом не учитывается.

Каждый риск, находящийся в компании, на протяжении длительности действия договора страхования независимо от других рисков генерирует с интенсивностью  $\gamma$  требование на выплату страховых сумм. И эти требования также образуют простейший поток событий. Естественно считать, что требование риска на выплату определяется наступлением страхового случая. Величину продолжительности договора страхования для каждого риска, находящегося в компании, будем считать случайной величиной, распределённой по экспоненциальному закону с параметром  $\mu$ .

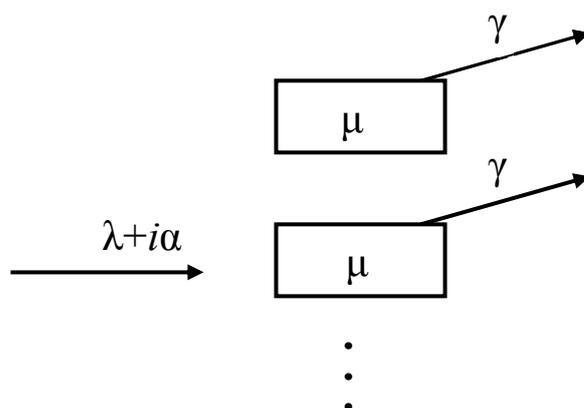


Рис. 1. Модель страховой компании в виде системы массового обслуживания с неограниченным количеством приборов

Обозначим:  $n(t)$  – число требований на выплату за интервал времени  $[0, t]$ ;  $i(t)$  – число страховых рисков, находящихся в компании в момент времени  $t$ ;  $P(i, n, t) = \mathbf{P}\{i(t) = i, n(t) = n\}$  – вероятность того, что в момент времени  $t$  в компании находится  $i$  застрахованных рисков и число требований на выплату страховых сумм к этому же моменту составило  $n$ . Задача состоит в том, чтобы найти это распределение двумерного процесса  $(i(t), n(t))$ .

Используя  $\Delta t$  метод, составим систему дифференциальных уравнений Колмогорова [3] для распределения вероятностей  $P(i, n, t)$ . Сначала запишем допредельные равенства:

$$P(i, n, t + \Delta t) = P(i, n, t)(1 - \lambda\Delta t - i\mu\Delta t - i\gamma\Delta t) + \\ + P(i - 1, n, t)(\lambda + (i - 1)\alpha)\Delta t + \\ + P(i + 1, n, t)(i + 1)\mu\Delta t + P(i, n - 1, t)i\gamma\Delta t + o(\Delta t).$$

Перейдем к пределу при  $\Delta t \rightarrow 0$ , получим систему дифференциальных уравнений:

$$\frac{\partial P(i, n, t)}{\partial t} = -P(i, n, t)(\lambda + i\mu + i\gamma + i\alpha) + P(i - 1, n, t)(\lambda + (i - 1)\alpha) + \\ + P(i + 1, n, t)(i + 1)\mu + P(i, n - 1, t)i\gamma. \quad (1)$$

Для решения системы (1) определим функцию

$$H(u, z, t) = \sum_{i, n=0}^{\infty} e^{iuj} z^n P(i, n, t),$$

характеристическую по  $u$  и производящую по  $z$ , где  $j$  – мнимая единица. Тогда из системы (1) с учетом свойств характеристических и производящих функций получим дифференциальное уравнение в частных производных первого порядка относительно функции  $H(u, z, t)$ :

$$\begin{aligned} \frac{\partial H(u, z, t)}{\partial t} + j[\alpha e^{ju} + \mu e^{-ju} + \gamma z - \mu - \gamma - \alpha] \frac{\partial H(u, z, t)}{\partial u} = \\ = \lambda(e^{ju} - 1)H(u, z, t). \end{aligned} \quad (2)$$

Решение этого дифференциального уравнения определяется решением системы обыкновенных дифференциальных уравнений для характеристических кривых [4]:

$$\frac{dt}{1} = \frac{du}{-j(\alpha + \mu + \gamma - \alpha e^{ju} - \mu e^{-ju} - \gamma z)} = \frac{dH(u, z, t)}{\lambda H(u, z, t)(e^{ju} - 1)}.$$

Найдем два первых интеграла системы. Сначала рассмотрим уравнение

$$dt = \frac{du}{j(\alpha(e^{ju} - 1) + \mu(e^{-ju} - 1) - \gamma(1 - z))}.$$

Сделаем замену переменных:  $e^{ju} - 1 = v$ . Тогда

$$dt = \frac{du}{-(\alpha v^2 + (\alpha - \mu - \gamma)(1 - z) - \gamma(1 - z))}$$

или

$$dt = \frac{dv}{-\alpha(v - v_1)(v - v_2)}, \quad (3)$$

где

$$v_{1,2} = \frac{1}{2} \left( 1 - \frac{\mu}{\alpha} - \frac{\gamma}{\alpha}(1 - z) \right) \pm \frac{\sqrt{D}}{2}, \quad (4)$$

$$D = \left( 1 - \frac{\mu}{\alpha} - \frac{\gamma}{\alpha}(1 - z) \right)^2 + 4 \frac{\gamma}{\alpha}(1 - z) > 0.$$

Таким образом, корни  $v_1$  и  $v_2$  различные и вещественные, причем при условии  $\alpha < \mu$ ,  $v_1 > 0$  и  $v_2 \leq 0$ .

Решение уравнения (3) дает выражение

$$C_1 = e^{-\alpha(v_1 - v_2)t} \left( \frac{v - v_2}{v - v_1} \right). \quad (5)$$

Другой первый интеграл найдем из уравнения

$$\frac{dH(u, z, t)}{\lambda H(u, z, t)(e^{ju} - 1)} = \frac{du}{j(\alpha(e^{ju} - 1) + \mu(e^{-ju} - 1) - \gamma(1 - z))}. \quad (6)$$

Сделаем аналогичную замену переменных  $e^{ju} - 1 = v$ . Введем функцию  $H_1(v, z, t) = H(u, z, t)$ , тогда можем записать

$$\frac{dH_1(v, z, t)}{H_1(v, z, t)} = \frac{\lambda v du}{-\alpha(v - v_1(z))(v - v_2(z))}. \quad (7)$$

Решение этого уравнения будет иметь вид

$$H_1(v, z, t) = C_2 \left[ \frac{(v - v_2(z))^{v_2(z)}}{(v - v_1(z))^{v_1(z)}} \right]^{\frac{\lambda}{\alpha(v_1(z) - v_2(z))}}. \quad (8)$$

Введем произвольную дифференцируемую функцию  $\varphi(C_1) = C_2$ . Тогда общее решение уравнения (6) с учетом (5) будет иметь вид

$$H_1(v, z, t) = \varphi \left[ e^{-\alpha(v_1(z) - v_2(z))t} \left( \frac{v - v_2(z)}{v - v_1(z)} \right) \right] \left[ \frac{(v - v_2(z))^{v_2(z)}}{(v - v_1(z))^{v_1(z)}} \right]^{\frac{\lambda}{\alpha(v_1(z) - v_2(z))}}. \quad (9)$$

Определим частное решение с помощью начальных условий. Для этого запишем значение функции  $H(u, z, t)$  при  $t = 0$ . Имеем

$$H(u, z, 0) = \sum_{i, n=0}^{\infty} e^{iuj} z^n P(i, n, 0) = \sum_{i=0}^n e^{jui} P(i),$$

так как в начальный момент (т.е. в момент, когда страховая компания только начинает свою работу) требований на выплату страховых сумм не было, и значит

$$P(i, n, 0) = \begin{cases} P(i), & n = 0; \\ 0, & n > 0. \end{cases}$$

Обозначим  $G(u) = H(u, z, 0)$ . Тогда, решение дифференциального уравнения относительно функции  $G(u)$ , с учетом условия  $G(0) = 1$ , будет иметь вид

$$G(u) = \left( \frac{1 - \frac{\alpha}{\mu}}{1 - e^{ju} \frac{\alpha}{\mu}} \right)^{\frac{\lambda}{\alpha}}. \quad (10)$$

Теперь выражение (9) при  $t = 0$  можем записать в виде

$$\left( \frac{1 - \frac{\alpha}{\mu}}{1 - (v+1) \frac{\alpha}{\mu}} \right)^{\frac{\lambda}{\alpha}} = \varphi \left[ \left( \frac{v - v_2(z)}{v - v_1(z)} \right) \right] \left[ \frac{(v - v_2(z))^{v_2(z)}}{(v - v_1(z))^{v_1(z)}} \right]^{\frac{\lambda}{\alpha(v_1(z) - v_2(z))}}.$$

Введем обозначения:  $x = \frac{v - v_1(z)}{v - v_2(z)}$ , тогда

$$\varphi(x) = \left[ \frac{\left(1 - \frac{\alpha}{\mu}\right)(v_2(z) - v_1(z))}{(1-x) - \frac{\alpha}{\mu}[v_2(z) + 1 - x(1 + v_1(z))]} \right]^{\frac{\lambda}{\alpha}} x^{\frac{\lambda v_2(z)}{\alpha(v_2(z) - v_1(z))}}.$$

Таким образом, можем записать выражение для функции  $H(u, z, t)$ :

$$\begin{aligned} H(u, z, t) = & e^{\lambda v_2(z)t} \left[ \left(1 - \frac{\alpha}{\mu}\right)(v_1(z) - v_2(z)) \right]^{\frac{\lambda}{\alpha}} \times \\ & \times \left\{ (v_1(z) - e^{ju} + 1) \left[ 1 - \frac{\alpha}{\mu}(1 + v_2(z)) \right] - \right. \\ & \left. - (v_2(z) - e^{ju} + 1) e^{\alpha(v_2(z) - v_1(z))t} \left[ 1 - \frac{\alpha}{\mu}(1 + v_2(z)) \right] \right\}^{\frac{\lambda}{\alpha}}. \end{aligned} \quad (11)$$

Найдем математические ожидания и дисперсии числа рисков и числа требований на страховые выплаты. Имеем:

$$\left. \frac{\partial H(u, z, t)}{\partial u} \right|_{\substack{u=0 \\ z=1}} = M\{i(t)\} = \frac{\lambda}{\mu - \alpha}, \quad (12)$$

$$\left. \frac{\partial H(u, z, t)}{\partial z} \right|_{\substack{u=0 \\ z=1}} = M\{n(t)\} = \frac{\lambda\gamma}{\mu - \alpha} t. \quad (13)$$

Для дисперсий получены следующие выражения:

$$D\{i(t)\} = \frac{\lambda\mu}{(\mu - \alpha)^2}, \quad (14)$$

$$D\{n(t)\} = 2 \frac{\lambda\mu\gamma^2}{(\mu - \alpha)^3} t + \frac{\lambda\gamma}{\mu - \alpha} t - 2 \frac{\lambda\mu\gamma^2}{(\mu - \alpha)^4} (1 - e^{-(\mu - \alpha)t}). \quad (15)$$

Формулы (12) и (14) совпадают с результатом, полученным в работе (2), где исследуется одномерный процесс числа застрахованных в компании рисков.

Найдем коэффициент корреляции процессов  $i(t)$  и  $n(t)$ . Имея выражение для функции  $H(u, z, t)$ , найдем смешанный момент первого порядка:

$$\frac{1}{j} \left. \frac{\partial^2 H(u, z, t)}{\partial u \partial z} \right|_{\substack{u=0 \\ z=1}} = M\{i(t)n(t)\},$$

тогда с учетом выражений (12) – (15) получим

$$r_{in}(t) = \frac{\lambda\gamma\mu(1 - e^{-(\mu - \alpha)t})}{\sqrt{\lambda\mu \left\{ 2\lambda\mu\gamma^2(\mu - \alpha)t + \lambda\gamma(\mu - \alpha)^3 t - 2\lambda\mu\gamma^2(1 - e^{-(\mu - \alpha)t}) \right\}}}.$$

Таким образом, в данной работе построена математическая модель страховой компании в виде системы массового обслуживания с неограниченным количеством обслуживающих приборов. Найдена характеристическая функция двумерного процесса числа застрахованных в компании рисков и числа требований на страховые выплаты. Показано, что полученные результаты являются обобщением частных случаев.

#### Литература

1. Гафуров Ш. Р., Гугнин В. И., Аманов С. Н. Язык бизнеса. Ташкент: Шарк, 1995. 738 с.
2. Глухова Е. В., Змеев О. А., Лившиц К. И. Математические модели страхования. Томск: Изд-во Том. ун-та, 2004. 180 с.
3. Назаров А. А., Терпугов А. Ф. Теория массового обслуживания. Томск: Изд-во НТЛ, 2005. 228 с.
4. Эльсгольц Л. Э. Дифференциальные уравнения и вариационное исчисление. М.: Наука, 1969. 424 с.

## **ИССЛЕДОВАНИЕ ПОТОКА ПОВТОРНЫХ ОБРАЩЕНИЙ В СМО С ПОВТОРНЫМ ОБСЛУЖИВАНИЕМ ЗАЯВОК МЕТОДОМ АСИМПТОТИЧЕСКОГО АНАЛИЗА**

*Л. А. Задиранова*

*Национальный исследовательский*

*Томский государственный университет, Томск, Россия*

Рассмотрим системы массового обслуживания с неограниченным числом обслуживающих устройств, в качестве моделей входящих потоков взяты пуассоновский поток (M), марковский модулированный поток (ММРР) и рекуррентный поток (GI).

Продолжительность обслуживания заявки является случайной величиной и имеет экспоненциальное распределение с параметром  $\mu$ . Поступившая заявка занимает любой из свободных приборов, завершив обслуживание на котором, с вероятностью  $1-r$  покидает систему или с вероятностью  $r$  возвращается для повторного обслуживания.

Ставится задача исследования потока заявок, обратившихся в рассматриваемые системы за время  $t$  для повторного обслуживания (поток повторных обращений).

### **Поток повторных обращений в системе $M|M|_{\infty}$ с повторным обслуживанием**

Рассмотрим систему массового обслуживания с неограниченным числом приборов, на вход которой поступает простейший поток заявок с интенсивностью  $\lambda$ . Обозначим  $i(t)$  – число занятых приборов в момент времени  $t$ ;  $n(t)$  – число заявок, обратившихся в систему за время  $t$  для повторного обслуживания, тогда двумерный поток  $\{i(t), n(t)\}$  является марковским [1].

Для распределений вероятностей  $P(i, n, t) = P\{i(t) = i, n(t) = n\}$  запишем систему дифференциальных уравнений Колмогорова вида

$$\frac{\partial P(i, n, t)}{\partial t} = -\lambda P(i, n, t) - i\mu P(i, n, t) + \lambda P(i-1, n, t) + i\mu r P(i, n-1, t) + \mu(1-r)(1+i)P(i+1, n, t), \text{ где } i = \overline{0, \infty}, n = \overline{0, \infty}. \quad (1)$$

Введем частичные характеристические функции в виде [2]

$$H(u, w, t) = \sum_i \sum_n e^{ju} e^{jwn} P(i, n, t),$$

тогда из (1) получаем следующее дифференциальное уравнение:

$$\frac{\partial H(u, w, t)}{\partial t} = \lambda(e^{ju} - 1)H(u, w, t) + j\mu(1 - re^{jw} - (1-r)e^{-ju}) \frac{\partial H(u, w, t)}{\partial u}. \quad (2)$$

Полученное уравнение позволяет определить основные вероятностные характеристики рассматриваемой системы, в том числе и для потока повторных обращений в систему.

Сформулируем вспомогательное утверждение относительно вида асимптотической характеристической функции числа занятых приборов в рассматриваемой системе.

**Лемма.** *Асимптотическое приближение характеристической функции числа занятых приборов в системе  $M|M|\infty$  в условии растущего времени обслуживания имеет вид*

$$h(u) = \exp\left\{\frac{j\lambda u}{\mu(1-r)}\right\}.$$

Проведем исследование потока повторных обращений в систему за время  $t$  с помощью метода асимптотического анализа в условии растущего времени обслуживания [3]. Для этого обозначим

$$\mu = \varepsilon, \quad u = \varepsilon y, \quad H(u, w, t) = F(y, w, t, \varepsilon)$$

и перепишем (2) в виде

$$\frac{\partial F(y, w, t, \varepsilon)}{\partial t} = \lambda(e^{j\varepsilon y} - 1)F(y, w, t, \varepsilon) + j(1 - re^{jw} - (1-r)e^{-j\varepsilon y}) \frac{\partial F(y, w, t, \varepsilon)}{\partial y}. \quad (3)$$

**Теорема 1.** *Предельное, при  $\varepsilon \rightarrow 0$ , значение функции  $F(y, w, t)$  решения  $F(y, w, t, \varepsilon)$  уравнения (3) имеет вид*

$$F(y, w, t) = \exp\left\{\frac{\lambda r t (e^{jw} - 1)}{1-r} + \frac{j\lambda y}{1-r}\right\}. \quad (4)$$

*Доказательство.*

Выполняя в уравнении (4) предельный переход при  $\varepsilon \rightarrow 0$ , получаем дифференциальное уравнение первого порядка:

$$\frac{\partial F(y, w, t)}{\partial t} = jr(1 - e^{jw}) \frac{\partial F(y, w, t)}{\partial y}. \quad (5)$$

Общее решение полученного уравнения имеет вид

$$F(y, w, t) = \varphi \left( t + \frac{jy}{r(e^{jw} - 1)} \right),$$

где  $\varphi(y)$  – некоторая функция, вид которой определим, используя начальное условие.

Рассмотрим функцию  $F(y, w, t)$  в нулевой момент времени, очевидно, что данная функция не будет зависеть от  $w$ , и начальное условие имеет вид

$$F(y, w, 0) = \Phi(y), \quad (6)$$

где  $\Phi(y)$  – асимптотическое приближение характеристической функции распределения числа занятых приборов в системе в условии растущего времени обслуживания заявок, вид которого определен выше:

$$\Phi(y) = \exp \left\{ \frac{j\lambda y}{1-r} \right\}.$$

Таким образом, решение уравнения (4), удовлетворяющее начальному условию (6), запишем в виде

$$F(y, w, t) = \exp \left\{ \frac{\lambda t r (e^{jw} - 1)}{1-r} + \frac{j\lambda y}{1-r} \right\},$$

которое совпадает с равенством (4). Теорема доказана.

Полагая в (4)  $y = 0$ , имеем асимптотическое приближение характеристической функции числа повторных обращений, поступивших в систему за время  $t$ , в условии растущего времени обслуживания

$$h(w, t) = M \{ e^{jwn(t)} \} = H(0, w, t) = F(0, w, t, \varepsilon) \approx F(0, w, t) = \exp \left\{ \frac{\lambda r t (e^{jw} - 1)}{1-r} \right\}.$$

### **Поток повторных обращений в системе ММРР|М| $\infty$ с повторным обслуживанием**

Рассмотрим систему массового обслуживания с неограниченным числом приборов, на вход которой поступает марковский модулированный поток (ММРР), управляемый цепью Маркова  $k(t)$  с конечным числом состояний,  $k(t) = 1, 2, \dots, K$ , заданной матрицей инфинитезимальных характеристик  $\mathbf{Q} = \|q_{ij}\|$ ,  $i, j = 1, 2, \dots, K$ , и матрицей условных интенсивностей  $\mathbf{\Lambda}$  [4].

Обозначим  $i(t)$  – число занятых приборов в момент времени  $t$ ;  $n(t)$  – число повторных заявок, обратившихся за время  $t$ ,  $k(t)$  – состояние управляющей цепи Маркова. Трехмерный процесс  $\{k(t), i(t), n(t)\}$  является марковским.

Для распределения вероятностей  $P(k, i, n, t) = P\{k(t) = k, i(t) = i, n(t) = n\}$  можно записать систему дифференциальных уравнений Колмогорова вида

$$\frac{\partial P(k, i, n, t)}{\partial t} = -\lambda_k P(k, i, n, t) - i\mu P(k, i, n, t) + \lambda_k P(k, i-1, n, t) +$$

$$+\mu(1-r)(1+i)P(k, i+1, n, t) + \mu irP(k, i, n-1, t) + \sum P(v, i, n, t)q_{vk}, \quad (7)$$

$$k, v=1, 2, \dots, K, \quad i, n=1, 2, 3, \dots$$

Введя частичные характеристические функции, запишем систему (7) в виде дифференциального матричного уравнения

$$\frac{\partial \mathbf{H}(u, w, t)}{\partial t} + j\mu(re^{jw} - 1 + (1-r)e^{-ju}) \frac{\partial \mathbf{H}(u, w, t)}{\partial u} = \mathbf{H}(u, w, t)[(e^{ju} - 1)\mathbf{\Lambda} + \mathbf{Q}], \quad (8)$$

где

$$\mathbf{H}(u, w, t) = [H(1, u, w, t), H(2, u, w, t), \dots, H(K, u, w, t)],$$

$$\mathbf{\Lambda} = \begin{bmatrix} \lambda_1 & 0 & \cdot & 0 \\ 0 & \lambda_2 & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & \lambda_K \end{bmatrix}, \quad \mathbf{Q} = \begin{bmatrix} q_{11} & q_{12} & \cdot & q_{1K} \\ q_{21} & q_{22} & \cdot & q_{2K} \\ \cdot & \cdot & \cdot & \cdot \\ q_{K1} & q_{K2} & \cdot & q_{KK} \end{bmatrix}.$$

Найдем асимптотическую характеристическую функцию числа повторных обращений в системе ММРР|M| $\infty$  за время  $t$  в условии растущего времени, обозначим

$$\mu = \varepsilon, \quad u = \varepsilon y, \quad \mathbf{H}(u, w, t) = \mathbf{F}(y, w, t, \varepsilon)$$

и перепишем уравнение (8) с учетом введенных обозначений:

$$\frac{\partial \mathbf{F}(y, w, t, \varepsilon)}{\partial t} + j(re^{jw} - 1 + (1-r)e^{-j\varepsilon y}) \frac{\partial \mathbf{F}(y, w, t, \varepsilon)}{\partial y} =$$

$$= \mathbf{F}(y, w, t, \varepsilon) \left[ (e^{j\varepsilon y} - 1)\mathbf{\Lambda} + \mathbf{Q} \right]. \quad (9)$$

Сформулируем теорему, доказательство которой можно провести аналогично доказательству теоремы 1.

**Теорема 2.** Сумма компонентов предельного, при  $\varepsilon \rightarrow 0$ , значения вектор-функции  $\mathbf{F}(y, w, t)$  решения  $\mathbf{F}(y, w, t, \varepsilon)$  уравнения (9) имеет вид

$$\mathbf{F}(y, w, t)\mathbf{E} = \exp \left\{ \frac{rkt(e^{jw} - 1)}{1-r} + \frac{jyk}{1-r} \right\}, \quad (10)$$

где величина  $k$  определяется выражением  $k = \mathbf{R}\mathbf{\Lambda}\mathbf{E}$ .

Полагая в (10)  $y = 0$ , имеем асимптотическое приближение характеристической функции числа повторных обращений, поступивших в систему за время  $t$ , в условии растущего времени обслуживания:

$$h(w, t) = M \{ e^{jwn(t)} \} = \mathbf{H}(0, w, t)\mathbf{E} =$$

$$= \mathbf{F}(0, w, t, \varepsilon)\mathbf{E} \approx \mathbf{F}(0, w, t)\mathbf{E} = \exp \left\{ \frac{\lambda rt(e^{jw} - 1)}{1-r} \right\}.$$

### Поток повторных обращений в системе GI|M| $\infty$ с повторным обслуживанием

Рассмотрим систему массового обслуживания с неограниченным числом обслуживающих приборов, на вход которой поступает рекуррентный

поток заявок с функцией распределения длин интервалов между моментами поступления заявок  $A(x)$ .

Обозначим  $i(t)$  – число занятых приборов в системе в момент времени  $t$ ;  $n(t)$  – число заявок, обратившихся в систему за время  $t$  для повторного обслуживания. Так как полученный случайный процесс  $\{i(t), n(t)\}$  немарковский, то марковизируем его, введя дополнительную переменную  $z(t)$ , равную длине интервала от момента  $t$  до момента поступления следующей заявки. Тогда трехмерный процесс  $\{z(t), i(t), n(t)\}$  будет марковским.

Обозначим распределение вероятностей значений полученного марковского процесса  $P(z, i, n, t) = P\{z(t) < z, i(t) = i, n(t) = n\}$ , тогда получаем систему дифференциальных уравнений Колмогорова:

$$\begin{aligned} \frac{\partial P(z, i, n, t)}{\partial t} = & \frac{\partial P(z, i, n, t)}{\partial z} - \frac{\partial P(0, i, n, t)}{\partial z} + \\ & + i\mu r P(z, i, n-1, t) + \mu(1+i)(1-r)P(z, i+1, n, t) + \\ & + A(z) \frac{\partial P(0, i-1, n, t)}{\partial z} - i\mu P(z, i, n, t), \text{ где } i = \overline{0, \infty}, n = \overline{0, \infty}. \end{aligned} \quad (11)$$

Введя частичные характеристические функции, запишем систему (11) в виде

$$\begin{aligned} \frac{\partial H(z, u, w, t)}{\partial t} = & \frac{\partial H(z, u, w, t)}{\partial z} - j\mu \left[ (1-r)e^{-ju} + re^{jw} - 1 \right] \frac{\partial H(z, u, w, t)}{\partial u} + \\ & + (e^{ju} A(z) - 1) \frac{\partial H(0, u, w, t)}{\partial z}. \end{aligned} \quad (12)$$

Определим характеристики потока повторных обращений в рассматриваемой системе методом асимптотического анализа. Рассмотрим систему GI|M| $\infty$  с повторным обслуживанием в условии растущего времени обслуживания, для этого обозначим

$$\mu = \varepsilon, u = \varepsilon y, H(z, u, w, t) = F(z, y, w, t, \varepsilon)$$

и перепишем (12) с учетом введенных обозначений:

$$\begin{aligned} \frac{\partial F(z, y, w, t, \varepsilon)}{\partial t} = & \frac{\partial F(z, y, w, t, \varepsilon)}{\partial z} + (e^{j\varepsilon y} A(z) - 1) \frac{\partial F(0, y, w, t, \varepsilon)}{\partial z} - \\ & - j((1-r)e^{-j\varepsilon y} + re^{jy} - 1) \frac{\partial F(z, y, w, t, \varepsilon)}{\partial y} = 0. \end{aligned} \quad (13)$$

Сформулируем теорему, доказательство которой можно провести аналогично доказательству теоремы 1.

**Теорема 3.** *Предельное, при  $\varepsilon \rightarrow 0$ , значение функции  $F(y, w, t)$  решения  $F(y, w, t, \varepsilon)$  уравнения (13) имеет вид*

$$F(y, w, t) = \exp \left\{ \frac{r\lambda t}{1-r} (e^{jw} - 1) + \frac{j\lambda y}{1-r} \right\}. \quad (14)$$

Полагая в (14)  $y=0$ , имеем асимптотическое приближение характеристической функции числа повторных обращений, поступивших в систему за время  $t$ , в условии растущего времени обслуживания

$$h(w,t) = M\{e^{jwn(t)}\} = H(0,w,t) = F(0,w,t,\varepsilon) \approx F(0,w,t) = \exp\left\{\frac{\lambda rt(e^{jw} - 1)}{1-r}\right\}.$$

Таким образом, в настоящей работе рассмотрены математические модели систем  $M|M|\infty$ ,  $MMPP|M|\infty$  и  $GI|M|\infty$  с повторными обращениями, получено асимптотическое приближение характеристических функций потока суммарных обращений в условии растущего времени обслуживания для каждой системы.

#### Литература

1. Морозова А. С. Исследование математических моделей стимулирования сбыта продукции: дис. ... канд. физ.-мат. наук / Филиал Кемеровского гос. ун-та в г. Анжеро-Судженске. Анжеро-Судженск, 2007. 115 с.
2. Назаров А. А. Теория массового обслуживания: учеб. пособие / А. А. Назаров, А. Ф. Терпугов. Томск: Изд-во НТЛ, 2004. 228 с.
3. Назаров А. А. Методы асимптотического анализа в теории массового обслуживания / А. А. Назаров, С. П. Моисеева. Томск: Изд-во НТЛ, 2006. 112 с.
4. Жидкова Л. А. Исследование числа занятых приборов в системе  $MMPP|M|\infty$  с повторными обращениями / Л. А. Жидкова, С. П. Моисеева // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2014. № 1 (26). С. 53–62.

## ИССЛЕДОВАНИЕ ПОТОКА СУММАРНЫХ ОБРАЩЕНИЙ В СИСТЕМЕ $MMPP|M|\infty$ С ПОВТОРНЫМ ОБЛУЖИВАНИЕМ ТРЕБОВАНИЙ

*Л. А. Задиранова, С. П. Моисеева*

*Национальный исследовательский*

*Томский государственный университет, Томск, Россия*

Рассмотрим бесконечнолинейную систему массового обслуживания с входящим  $MMPP$ -поток, управляемым цепью Маркова  $k(t)$  с конечным числом состояний,  $k(t) = 1, 2, \dots, K$ , заданной матрицей инфинитезимальных характеристик  $Q = \|q_{ij}\|$ ,  $i, j = 1, 2, \dots, K$ , и матрицей условных интенсивностей  $\Lambda$ . Продолжительность обслуживания заявки имеет экспоненциальное распределение с параметром  $\mu$ . Поступившая заявка занимает любой из свободных приборов, завершив обслуживание на котором, с вероятностью  $1-r$  покидает систему или с вероятностью  $r$  возвращается в неё для повторного обслуживания [1].

Ставится задача исследования потока суммарных обращений в систему  $MMPP|M|\infty$  с повторным обслуживанием заявок.

Обозначим  $i(t)$  – число занятых приборов в момент времени  $t$ ;  $m(t)$  – суммарное число заявок, обратившихся в систему за время  $t$  для повторно-го и первичного обслуживания;  $k(t)$  – состояние управляющей цепи Маркова.

Очевидно, что процесс  $\{i(t), m(t)\}$  не является марковским, так как интенсивность поступления заявок в рассматриваемую систему зависит от состояния управляющей цепи Маркова  $k(t)$ , поэтому будем рассматривать трехмерную цепь Маркова  $\{k(t), i(t), m(t)\}$ .

Для распределения вероятностей

$$P(k, i, m, t) = P\{k(t) = k, i(t) = i, m(t) = m\}$$

запишем систему дифференциальных уравнений Колмогорова вида

$$\begin{aligned} \frac{\partial P(k, i, m, t)}{\partial t} = & -\lambda_k P(k, i, m, t) - i\mu P(k, i, m, t) + \lambda_k P(k, i-1, m-1, t) + \\ & + \mu ir P(k, i, m-1, t) + \mu(1-r)(1+i)P(k, i+1, m, t) + \sum P(v, i, m, t)q_{vk}, \end{aligned}$$

где  $i = \overline{0, \infty}, m = \overline{0, \infty}$ .

Введем частичные характеристические функций [2] в виде

$$H(k, u, v, t) = \sum_{i=0}^{\infty} \sum_{m=0}^{\infty} e^{ju} e^{jvm} P(k, i, m, t),$$

тогда имеем

$$\begin{aligned} \frac{\partial H(k, u, v, t)}{\partial t} = & j\mu [1 - re^{jv} - (1-r)e^{-ju}] \frac{\partial H(k, u, v, t)}{\partial u} + \\ & + H(k, u, v, t) [\lambda_k (e^{j(u+v)} - 1)] + \sum H(v, u, v, t) q_{vk}. \end{aligned}$$

Запишем данную систему в виде дифференциального матричного уравнения

$$\frac{\partial \mathbf{H}(u, v, t)}{\partial t} + j\mu (re^{jv} - 1 + (1-r)e^{-ju}) \frac{\partial \mathbf{H}(u, v, t)}{\partial u} = \mathbf{H}(u, v, t) [(e^{j(u+v)} - 1)\mathbf{\Lambda} + \mathbf{Q}], \quad (1)$$

где

$$\begin{aligned} \mathbf{H}(u, v, t) = & [H(1, u, v, t), H(2, u, v, t), \dots, H(K, u, v, t)], \\ \mathbf{\Lambda} = & \begin{bmatrix} \lambda_1 & 0 & \cdot & 0 \\ 0 & \lambda_2 & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & \lambda_K \end{bmatrix}, \quad \mathbf{Q} = \begin{bmatrix} q_{11} & q_{12} & \cdot & q_{1K} \\ q_{21} & q_{22} & \cdot & q_{2K} \\ \cdot & \cdot & \cdot & \cdot \\ q_{K1} & q_{K2} & \cdot & q_{KK} \end{bmatrix}. \end{aligned}$$

Отметим, что получить аналитическое решение этого уравнения не удастся. В данной работе предлагается решать это уравнение методом асимптотического анализа в условиях растущего времени обслуживания, и предельно частых изменений состояний входящего потока.

Зафиксируем некоторую матрицу инфинитезимальных характеристик  $\mathbf{Q}^{(1)}$ . Положим, что  $S$  – некоторая положительная величина, при этом

$$\mathbf{Q} = S \cdot \mathbf{Q}^{(1)}.$$

Очевидно, что стационарные распределения вероятностей состояний управляющей цепи  $k(t)$ , заданной матрицами инфинитезимальных характеристик  $\mathbf{Q}^{(1)}$  и  $\mathbf{Q}=S \mathbf{Q}^{(1)}$ , совпадают (не зависят от  $S$ ), но при увеличении значений параметра  $S$  интенсивности перехода цепи Маркова  $k(t)$  из одного состояния в другое возрастают, что соответствует условию предельно частых изменений состояния потока.

Обозначим

$$\mu = \varepsilon, u = \varepsilon y, \frac{1}{S} = \varepsilon, H(u, v, t) = F(y, v, t, \varepsilon),$$

тогда перепишем уравнение (1) в виде

$$\begin{aligned} \frac{\partial \mathbf{F}(y, v, t, \varepsilon)}{\partial t} = j\varepsilon \left( 1 - (1-r)e^{-jy\varepsilon} - re^{jv} \right) \frac{\partial \mathbf{F}(y, v, t, \varepsilon)}{\varepsilon \cdot \partial y} + \\ + \mathbf{F}(y, v, t, \varepsilon) \left( \Lambda \left( e^{j(\varepsilon y + v)} - 1 \right) + S\mathbf{Q}^{(1)} \right). \end{aligned} \quad (2)$$

**Теорема.** Сумма компонентов предельного, при  $\varepsilon \rightarrow 0$ , значения вектор-функции  $F(y, v, t)$  решения  $F(y, v, t, \varepsilon)$  уравнения (2) имеет вид

$$\mathbf{F}(y, v, t)\mathbf{E} = \exp \left\{ \frac{\kappa t (e^{jv} - 1)}{1-r} + \frac{j\kappa y}{1-r} \right\}, \quad (3)$$

где  $\kappa = \mathbf{R}\Lambda\mathbf{E}$ , а  $\mathbf{R} = \mathbf{H}(0)$  – вектор стационарного распределения вероятностей состояний цепи Маркова  $k(t)$ , определяемый системой уравнений  $\mathbf{R}\mathbf{Q} = 0$  и удовлетворяющий условию нормировки  $\mathbf{R}\mathbf{E} = 1$ .

*Доказательство.*

Поделим левую и правую части уравнения (2) на  $S$

$$\begin{aligned} \varepsilon \frac{\partial \mathbf{F}(y, v, t, \varepsilon)}{\partial t} = j\varepsilon \left( 1 - (1-r)e^{-jy\varepsilon} - re^{jv} \right) \frac{\partial \mathbf{F}(y, v, t, \varepsilon)}{\partial y} + \\ + \mathbf{F}(y, v, t, \varepsilon) \left( \varepsilon \cdot \Lambda \cdot \left( e^{j(\varepsilon y + v)} - 1 \right) + \mathbf{Q}^{(1)} \right), \end{aligned}$$

и, устремив  $\varepsilon \rightarrow 0$ , получим систему

$$\mathbf{F}(y, v, t)\mathbf{Q}^{(1)} = 0,$$

поэтому ее решение имеет вид

$$\mathbf{F}(y, v, t) = \mathbf{R} \cdot \Phi(y, v, t), \quad (4)$$

где  $\mathbf{R}$  – вектор стационарного распределения состояний управляющей цепи Маркова  $k(t)$ , а  $\Phi(y, v, t)$  – некоторая скалярная функция [3]. Для определения вида этой функции умножим справа уравнение (2) на единичный вектор-столбец  $\mathbf{E}$  соответствующей размерности

$$\begin{aligned} \frac{\partial \mathbf{F}(y, v, t, \varepsilon)}{\partial t} \mathbf{E} = j \left( 1 - (1-r)e^{-jy\varepsilon} - re^{jv} \right) \frac{\partial \mathbf{F}(y, v, t, \varepsilon)}{\partial y} \mathbf{E} + \\ + \mathbf{F}(y, v, t, \varepsilon) \left( \Lambda \left( e^{j(\varepsilon y + v)} - 1 \right) + S\mathbf{Q}^{(1)} \right) \mathbf{E}. \end{aligned}$$

Выполняя в полученном уравнении предельный переход при  $\varepsilon \rightarrow 0$  и учитывая вид решения (3), имеем уравнение в частных производных [4] вида

$$\frac{\partial \Phi(y, v, t, \varepsilon)}{\partial t} = jr(1 - e^{jv}) \frac{\partial \Phi(y, v, t, \varepsilon)}{\partial y} + \Phi(y, v, t, \varepsilon) \lambda (e^{jv} - 1).$$

Запишем соответствующую систему дифференциальных уравнений:

$$\frac{dt}{1} = \frac{dy}{jr(1 - e^{jv})} = \frac{d\Phi(y, v, t)}{\lambda (e^{jv} - 1) \Phi(y, v, t)},$$

решая полученную систему дифференциальных уравнений с учетом на-

чального условия  $\Phi(y, v, 0) = \Phi(y) = \exp\left\{\frac{jky}{1-r}\right\}$  [5], имеем

$$\Phi(y, v, t) = \exp\left\{\frac{\kappa t (e^{jv} - 1)}{1-r} + \frac{jky}{1-r}\right\},$$

следовательно, решение уравнения (2) имеет вид

$$\mathbf{F}(y, v, t) = \mathbf{R} \exp\left\{\frac{\kappa t (e^{jv} - 1)}{1-r} + \frac{jky}{1-r}\right\}.$$

Теорема доказана.

Полагая в (3)  $y=0$ , имеем асимптотическое приближение характеристической функции суммарного числа обращений, поступивших в систему за время  $t$  для повторного и первичного обслуживания

$$h(v, t) = M\{e^{jvm(t)}\} = \mathbf{H}(0, v, t)\mathbf{E} = \mathbf{F}(0, v, t)\mathbf{E} = \exp\left\{\frac{\kappa t (e^{jv} - 1)}{1-r}\right\}.$$

Полученные результаты позволяют сделать вывод о том, асимптотическое приближение характеристических функций суммарного числа обращений, поступивших в рассматриваемые системы за время  $t$  для повторного и первичного обслуживания, имеют вид распределения Пуассона.

#### Литература

1. Задиранова Л. А. Асимптотический анализ потока повторных обращений в системе ММРР|M| $\infty$  с повторными обслуживанием / Л. А. Задиранова, С. П. Моисеева // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2015. № 2 (31). С. 26–54.
2. Назаров А. А. Теория массового обслуживания: учеб. пособие. / А. А. Назаров, А. Ф. Терпугов. Томск: Изд-во НТЛ, 2004. 228 с.
3. Назаров А. А. Методы асимптотического анализа в теории массового обслуживания / А. А. Назаров, С. П. Моисеева. Томск: Изд-во НТЛ, 2006. 112 с.
4. Камке Э. Справочник по дифференциальным уравнениям в частных производных первого порядка / Э. Камке. М.: Наука, 1966. 260 с.
5. Жидкова, Л. А. Исследование числа занятых приборов в системе ММРР|M| $\infty$  с повторными обращениями / Л. А. Жидкова, С. П. Моисеева // Вестник Томского госу-

## ОСОБЕННОСТИ ИМИТАЦИОННОГО МОДЕЛИРОВАНИЯ ФРАКТАЛЬНЫХ СИСТЕМ С ОЧЕРЕДЯМИ

*В. Н. Задорожный*

*Омский государственный технический университет, Омск, Россия*

### Введение

Открытие фрактальных свойств у трафика информационных сетей [1] стало ключом к пониманию причины целого ряда крупных неудач, постигших проекты сетевых устройств, направленные на обеспечение качества информационного обмена и основанные на классической теории очередей. Это открытие привело к радикальной корректировке математических моделей трафика и методов его обслуживания. При описании и анализе фрактального трафика стали широко применяться такие математические понятия, как самоподобный случайный процесс, долговременная зависимость, распределения с тяжелыми хвостами (РТХ) [1–3].

Фрактальными системами (FS) будем называть такие системы класса  $GI|GI|n|m$ , которые обладают следующими свойствами. Независимое время  $\tau_i$  между моментами поступления  $i$ -й и  $(i+1)$ -й заявок в FS имеет одну и ту же при любом  $i$  функцию распределения (ф.р.)  $A(t)$  с математическим ожиданием (м.о.)  $M(\tau_i) = \bar{\tau} < \infty$ . Независимое время  $x_i$  обслуживания (трудоемкость) любой  $i$ -й заявки имеет ф.р.  $B(t)$  с м.о.  $M(x_i) = \bar{x} < \infty$ . Хотя бы одна из ф.р.  $A(t)$ ,  $B(t)$  является асимптотически степенной, т.е. задает фрактальную случайную величину (с.в.) [4], и имеет бесконечную дисперсию. Коэффициент  $\rho$  загрузки рассматриваемых FS не превосходит единицы:  $\rho = \bar{x}/(n\bar{\tau}) \leq 1$ . Если размер  $m$  буфера конечный, то требуется при данном  $m$  определить вероятность  $P$  потери заявки (прямая задача) или найти наименьший размер  $m$ , гарантирующий, что вероятность потери будет не больше  $P$  (обратная задача). При  $m = \infty$  интерес представляет средняя длина очереди  $L$  или среднее время ожидания  $W$ .

Исследование FS аналитическими методами весьма затруднено, поэтому для их расчета широко используется имитационное моделирование (ИМ) [5]. Большое количество публикаций посвящено разработке эффективных имитационных оценок для весьма ограниченного набора характеристик узких классов FS. Однако мы не нашли работ, в которых точность реализации РТХ в ИМ FS исследовалась бы не с помощью выборочных оценок, а точными методами. В предлагаемой статье этот пробел восполняется. Точными методами выявляется проблема существенного искажения РТХ при ИМ FS. Разрабатываются методы решения этой проблемы.

## 1. Проблема искажения РТХ при ИМ FS

Типичными представителями FS являются системы Pa|M|n|m, M|Pa|n|m и Pa|Pa|n|m. Здесь символ Pa соответствует распределению Парето (РП), описываемому ф. р.

$$F(t) = 1 - \left(\frac{K}{t}\right)^\alpha, \quad \alpha > 0, \quad K > 0, \quad t \geq K \quad (1)$$

(где  $\alpha$  – параметр формы,  $K$  – минимальное значение с.в.) и кратко обозначаемому как Pa( $K; \alpha$ ). С. в.  $\xi \in \text{Pa}(K; \alpha)$  имеет при  $\alpha > 1$  конечное м.о.  $M(\xi) = \alpha K / (\alpha - 1)$ , при  $\alpha > 2$  – конечную дисперсию  $D(\xi) = \sigma_\xi^2 = \alpha K^2 / [(\alpha - 1)^2 (\alpha - 2)]$ . Если мы рассматриваем РП с бесконечной дисперсией и конечным м.о., то параметр  $\alpha$  находится в промежутке  $1 < \alpha \leq 2$ . Этот диапазон значений  $\alpha$  наиболее актуален при моделировании сетевых устройств, обслуживающих фрактальный трафик.

Если обе ф. р.  $A(t)$  и  $B(t)$  являются РП с бесконечной дисперсией, то FS адекватно учитывает все основные особенности обслуживаемого фрактального трафика [1].

Обращением ф. р. (1) получаем следующую формулу для генерации с.в.  $\xi \in \text{Pa}(K; \alpha)$ :  $\xi = K(1-u)^{-1/\alpha}$ , где  $u$  – базовая с.в. (БСВ), равномерно распределенная в промежутке от 0 до 1. В результате правомерной замены  $(1-u) \rightarrow u$ , формула принимает вид  $\xi = Ku^{-1/\alpha}$ . Эту формулу можно получить и путем обращения хвоста:  $\bar{F}(t) = 1 - F(t) = (K/t)^\alpha$ .

Причиной смещения моментов при реализации с. в. с РТХ является дискретность программных генераторов БСВ. Программные генераторы реализуют дискретную БСВ (ДБСВ)  $u'$  с множеством равновероятных значений  $\{0, \varepsilon, 2\varepsilon, 3\varepsilon, \dots, 1-\varepsilon, 1\}$  (как правило, нуль или единица из этого множества исключаются), образующих решетку с шагом  $\varepsilon$ . Так, в GPSS [6] генератор Uniform(1,0,1) реализует ДБСВ, принимающую значения  $\{0,000000, 0,000001, \dots, 0,999999\}$  с шагом дискретизации  $\varepsilon = 10^{-6}$ . В других языках программирования шаг  $\varepsilon$  обычно составляет около  $10^{-15}$ .

Итак, в формуле, реализующей с. в. с РП, вместо БСВ  $u$  участвует ДБСВ  $u'$  и, следовательно, вместо непрерывной с. в.  $\xi \in \text{Pa}(K; \alpha)$  реализуется дискретная с. в. (д. с. в.)  $\xi' = K(u')^{-1/\alpha}$  с дискретным РП DPa( $K; \alpha; \varepsilon$ ).

При ИМ FS выясняется, что в случае реализации РП и других РТХ, в отличие от реализации классических ф.р. с легкими хвостами, шаг  $\varepsilon = 10^{-15}$  и тем более шаг  $\varepsilon = 10^{-6}$  оказывается слишком большим и приводит к значительным отличиям свойств генерируемой д.с.в.  $\xi'$  от свойств с.в.  $\xi \in \text{Pa}(K; \alpha)$ . Найдем, например, начальные моменты  $\xi'^{(k)}$  д.с.в.  $\xi'$ , учитывая конечное число  $N = 1/\varepsilon$  ее возможных значений:

$$\begin{aligned}\xi'^{(k)} &= M(\xi'^k) = \sum_{i=1}^N P(\xi'_i) \xi_i^k = \sum_{i=1}^N \frac{1}{N} K^k u_i^{-\frac{k}{\alpha}} = K^k N^{-1} \sum_{i=1}^N (iN^{-1})^{-\frac{k}{\alpha}} = \\ &= K^k N^{-1} \sum_{i=1}^N i^{-\frac{k}{\alpha}} N^{\frac{k}{\alpha}} = K^k N^{\frac{k}{\alpha}-1} \sum_{i=1}^N i^{-\frac{k}{\alpha}}, \quad (\alpha > k, \quad N = 1/\varepsilon); \quad (2)\end{aligned}$$

затем вычислим при разных  $\varepsilon$  м. о.  $M(\xi')$ , дисперсию  $\sigma_{\xi'}^2$  и коэффициент вариации (к. в.)  $C_{\xi'}$  д. с. в.  $\xi'$  по вытекающим из (2) формулам:

$$M(\xi') = KN^{\frac{1}{\alpha}-1} \sum_{i=1}^N i^{-\frac{1}{\alpha}}, \quad (3)$$

$$\sigma_{\xi'}^2 = \xi'^{(2)} - E^2(\xi') = K^2 N^{\frac{2}{\alpha}-1} \left( \sum_{i=1}^N i^{-\frac{2}{\alpha}} \right) - M^2(\xi'), \quad C_{\xi'} = \frac{\sigma_{\xi'}}{M(\xi')} \quad (4)$$

и сравним с соответствующими числовыми характеристиками с. в.  $\xi \in Pa(K; \alpha)$ . Все приведенные в таблице значения точные или округлены до 5 значащих цифр.

**Числовые характеристики д. с. в.  $\xi' \in DPa(K; \alpha; \varepsilon)$   
и с. в.  $\xi \in Pa(K; \alpha)$  при  $K = 1$**

$\alpha$	$M(\xi')$			$M(\xi)$	$C_{\xi'}$			$C_{\xi}$
	$\varepsilon = 10^{-6}$	$\varepsilon = 10^{-12}$	$\varepsilon = 10^{-15}$		$\varepsilon = 10^{-6}$	$\varepsilon = 10^{-12}$	$\varepsilon = 10^{-15}$	
<b>1,01</b>	13,415	24,612	29,662	101,00	83,856	$3,9866 \cdot 10^4$	$9,7690 \cdot 10^5$	$\infty$
<b>1,1</b>	8,0297	10,154	10,549	11,000	48,305	$1,0882 \cdot 10^4$	$1,7677 \cdot 10^5$	
<b>1,2</b>	5,4565	5,9457	5,9828	6,0000	26,687	2450,9	$2,4357 \cdot 10^4$	
<b>1,9</b>	2,1089	2,1111	2,1111	2,1111	1,9590	3,6806	4,6685	
<b>2</b>	1,9985	2,0000	2,0000	2,0000	1,6136	2,4601	2,7891	

Приведенные в таблице результаты расчетов по точным формулам (3), (4) демонстрируют существенные отличия свойств генерируемой с. в.  $\xi'$  от свойств подлежащей реализации с. в.  $\xi \in Pa(K; \alpha)$ . Найденные значения  $M(\xi')$ ,  $C_{\xi'}$  – это те значения, к которым сходятся выборочные оценки м. о. и к.в. при увеличении объема реализуемых в ИМ выборок до бесконечности. Из таблицы видно, что чем тяжелее хвост РП, т.е. чем меньше  $\alpha$ , тем сильнее отличается м. о. реализуемой с.в.  $\xi'$  от м. о. подлежащей реализации с. в.  $\xi$ . Концы промежутка  $1 < \alpha \leq 2$  являются критическими зонами, точка  $\alpha = 1$  – для реализуемого м. о., точка  $\alpha = 2$  – для реализуемого к. в. Причиной искажения РП является то, что при значениях ДБСВ  $u'$ , близких при использовании формулы  $\xi = Ku^{-1/\alpha}$  к нулю, обычный шаг  $\varepsilon$  приводит к тому, что слишком много больших значений реализуемой с. в., существенных для формирования репрезентативной выборки, просто не реализуется.

**2. Влияние искажений РТХ на результаты ИМ FS**

В системе  $M|Pa|1|\infty$  время обслуживания  $x$  принадлежит распределе-

нию  $\text{Pa}(K; \alpha)$ , и если  $\alpha \in (1, 2]$ , то стационарная средняя длина  $L(\rho)$  очереди заявок бесконечна при любом  $\rho \in (0, 1)$ . Действительно, согласно формуле Полячека–Хинчина [7] здесь при любом  $\rho \in (0, 1]$  средняя длина очереди бесконечна:  $L(\rho) = \rho^2 [1 + C_\xi^2] / [2(1 - \rho)] = \infty$ , так как коэффициент вариации  $C_\xi$  при  $\alpha \in (1, 2]$  бесконечен. Но если эту FS исследовать методом ИМ, то вместо времени обслуживания  $x \in \text{Pa}(K; \alpha)$  реализуется время  $x' \in \text{DPa}(K; \alpha; \varepsilon)$  и, например, при  $K = 1$ ,  $\alpha = 2$ ,  $\varepsilon = 10^{-15}$  мы при достаточно большой длине прогона модели получаем следующий определяемый величиной  $C_\xi = 2,7891$  (см. таблицу) результат:  $L(\rho) \approx 4,39\rho^2 / (1 - \rho)$ . Этот вывод подтверждается и результатами ИМ. В рассмотренном случае имитационное решение принципиально отличается от точного.

В системе  $\text{Pa}|\text{M}|1|\infty$  среднее время ожидания может быть найдено точно [7] по формуле  $W = \sigma / [\mu(1 - \sigma)]$ , где  $\mu = 1/\bar{x}$  – средняя интенсивность обслуживания;  $\sigma$  – единственный в области  $0 \leq \sigma < 1$  корень уравнения  $\sigma = A^*(\mu - \mu\sigma)$ ;  $A^*(s) = \int_0^\infty e^{-st} dA(t)$  – преобразование Лапласа (ПЛ) ф.р.  $A(t)$ . В данном случае  $A(t)$  есть ф.р. Парето (1). В [8] получаемое уравнение для  $\sigma$  решается численным методом при  $\alpha = 1,1$ ,  $K = 1$  (т. е.  $\bar{\tau} = 11$ ) для различных  $\rho$  и определяются соответствующие точные значения среднего времени ожидания  $W$ . Например, при  $\rho = 0,4$  имеем  $\sigma = 0,964199$  и  $W = 118,5$ . Но при шаге дискретизации  $\varepsilon = 10^{-6}$  с помощью ПЛ ф.р. реализуемой д.с.в.  $\tau' \in \text{DPa}(1; 1,1; 10^{-6})$  при  $\rho = 0,4$  точно вычисляется решение  $\sigma = 0,85953$  и  $W' = 19,653$ , отличающееся от значения  $W$  более чем в пять раз. Соответствующая имитационная оценка  $W_{\text{ИМ}} \approx 19,657$ , полученная на GPSS при длине прогона около 125 млн заявок, подтверждает правильность выполненного расчета.

Очевидно, что ИМ FS  $\text{Pa}|\text{Pa}|\text{n}|\text{m}$  в общем случае также отягощено аналогичными детерминированными погрешностями результатов, и этот вывод распространяется на случаи применения в FS асимптотически-степенных и субэкспоненциальных РТХ.

### 3. Эффективные методы неискаженной реализации РТХ

Базовый каскадный алгоритм реализации РП устраняет проблему детерминированных искажений РП без перехода к низкопроизводительной и трудно реализуемой в ИМ «длинной арифметике».

Пусть имеется генератор случайных чисел  $\text{RAND1}()$  с разрядностью  $r \geq 12$  десятичных знаков и  $\text{RAND2}()$ , имеющий хотя бы 6 разрядов. Разобьем область  $t \geq K$  значений с. в.  $x \in \text{Pa}(K, \alpha)$  на интервалы точками  $K_1 = K$ ,  $K_2 = K_1 \cdot 10^{6/\alpha}$ ,  $K_3 = K_2 \cdot 10^{6/\alpha}$ , .... С учетом (1) вероятности интервалов  $(K_i, K_{i+1})$  составят  $p_i = P(K_i \leq x < K_{i+1}) = 0,999999 \cdot 10^{-6(i-1)}$ ,  $i = 1, 2, \dots$ . Представим хвост  $\bar{F}(t) = (K/t)^\alpha$  линейной комбинацией соответствующих условных

хвостов:  $\bar{F}(t) = \sum_{i=1}^{\infty} p_i \bar{F}_i(t)$ , где  $\bar{F}_i(t) = [\bar{F}(t) - \bar{F}(K_{i+1})] / p_i$ . После этого с. в.  $x$  с хвостом  $\bar{F}(t) = (K/t)^\alpha$  можно генерировать путем организуемого в виде цикла выбора с вероятностями  $p_i$  интервалов  $(K_i, K_{i+1})$  и реализации условных хвостов  $\bar{F}_i(t)$ . Проблема искажения РТХ устраняется за счет того, что все реализуемые с. в. определены на конечных интервалах.

Универсальный каскадный алгоритм (рис. 1) является обобщением базового каскадного метода, не требующим обращения условных хвостов и даже не обязывающим использовать безусловные ф. р. или хвост. Это позволяет применять его для несмещенной реализации таких РТХ, ф. р. которых не выражаются в элементарных функциях в замкнутом виде (например, для реализации логнормального распределения).

```

Input:  $K, \alpha$ 
output:  $x$ 
Step1:  $i \leftarrow 1$ 
Step2: If  $\text{Rand2}() < 10^{-6}$ 
then  $i \leftarrow i + 1$ 
        goto Step2
         $a \leftarrow 10^{-6i}$ 
         $b \leftarrow 10^{-6(i-1)}$ 
         $u \leftarrow (b - a)\text{Rand1}() + a$ 
Step3:  $x \leftarrow Ku^{-1/\alpha}$ 
return:  $x$ 

```

Рис. 1. Универсальный каскадный алгоритм, настроенный на генерацию с. в. с РП

На рис. 1 алгоритм настроен на реализацию РП. Настройка на любое другое распределение сводится к замене формулы на шаге **Step3** формулой (или последовательностью формул), реализующей требуемое безусловное распределение через БСВ  $u$ . Проблема искажения РТХ решается приведенным алгоритмом в том случае, если малым  $u$  соответствуют большие  $x$ . Как правило, это обеспечить совсем нетрудно.

### Заключение

В практике ИМ фрактальных систем редко требуются выборки такого объема, при котором отличие предложенного метода реализации РТХ от обычных методов проявлялось бы регулярно и очевидным образом. Однако предложенный каскадный метод полезен уже тем, что снимает с исследователя-экспериментатора заботу об оценке иногда весьма значимых детерминированных погрешностей ИМ, которые связаны с искажениями РТХ, обусловленными дискретностью генераторов БСВ.

### Литература

1. Leland W. E., Taqqu M. S., Willinger W., Wilson, D. V. On the Self-Similar Nature of Ethernet Traffic // IEEE/ACM Transactions on networking, Vol. 2, no. 1. February 1994. P. 1–15.

2. Crovella M. E., Taqqu M., Bestavros A. Heavy Tailed-Probability distributions in the World Wide Web. December 1997. Vol. 5(6). P. 835–846.
3. Vishnevskiy V. M. Theoretical bases of designing computer networks. Moscow: Technosphere, 2003. 512 p.
4. Mandelbrot B. The fractal geometry of nature. New York: W. H. Freeman and Co., 1982. 480 p.
5. Zwart A. P. Queueing Systems with Heavy Tails. Eindhoven University of Technology, 2001. 227 p.
6. GPSS World reference manual. Minuteman Software, Fifth Edition, Holly Springs, NC, U.S.A., 2009.
7. Kleinrock L. Queueing Systems. Vol. II Computer Applications. New York: Wiley Interscience, 1976. 576 p.
8. Zadorozhnyi V. N. Simulation modeling of fractal queues, in Dynamics of Systems // Mechanisms and Machines (Dynamics). December 2014. P. 1–4.

## О ПОТОКЕ ПОВТОРНЫХ ТРЕБОВАНИЙ В ТАНДЕМЕ ЦИКЛИЧЕСКИХ СИСТЕМ ОБСЛУЖИВАНИЯ В СТАЦИОНАРНОМ РЕЖИМЕ

*А. В. Зорин, В. А. Зорин*

*Национальный исследовательский Нижегородский государственный университет им. Н. И. Лобачевского, Нижний Новгород, Россия*

Рассматриваются две системы обслуживания. В первую систему поступают конфликтные входные потоки требований  $\Pi_1, \Pi_2$ , а во вторую систему – конфликтные входные потоки  $\Pi_3, \Pi_4$ . Входные потоки  $\Pi_1, \Pi_2, \Pi_4$  формируются во внешней случайной среде с конечным числом  $d$  состояний  $e^{(1)}, e^{(2)}, \dots, e^{(d)}$ . Смена состояния случайной среды может произойти только в моменты смены состояния любого из обслуживающих устройств. Вероятность смены состояния  $e^{(k)}$  на  $e^{(l)}$  равна  $a_{k,l}$ . При состоянии среды  $e^{(k)}$ ,  $k = 1, 2, \dots, d$ , требования по потоку  $\Pi_j, j = 1, 2, 4$ , поступают группами, так что поток групп простейший с параметром  $\lambda_j^{(k)}$ , а размер группы равен  $b$  с вероятностью  $f(b; j, k)$ . Входной поток  $\Pi_3$  образуется из повторных требований потоков  $\Pi_1, \Pi_2$ . А именно, требования потока  $\Pi_1$  после обслуживания направляются на обслуживание во вторую систему; совокупность этих требований составляет входной поток  $\Pi_2$  требований, обслуживание которых будет заключаться в перемещении из первой во вторую систему обслуживания. Каждое требование потока  $\Pi_2$  после обслуживания и независимо от других требований либо мгновенно с вероятностью  $\alpha$  начинает перемещение во вторую систему обслуживания и добавляется к потоку  $\Pi_3$  либо с вероятностью  $1 - \alpha$  покидает сообщающиеся системы обслуживания и добавляется к соответствующему выходному потоку. Предполагается, что необходимо время для того, чтобы заявка из первой системы достигла второй системы. Перемещение заявки из первой во вторую систему

осуществляется со случайной скоростью, распределение которой неизвестно. Более того, скорости перемещения разных заявок различны и имеют разное распределение. В силу этого предполагаем, что за акт обслуживания или переналадки каждая перемещающаяся заявка либо достигает второй системы обслуживания с известной вероятностью, либо не достигает ее. Выходной поток заявок, обслуживание которых заключалось в перемещении из первой системы во вторую систему, и является, наконец, входным потоком  $\Pi_3$  второй системы. Таким образом, из двух конфликтных потоков, поступающих во вторую систему обслуживания, один образуется из потока повторных вызовов из первой системы. Требования потока  $\Pi_j, j = 1, 2, \dots, 5$ , поступают в очередь  $O_j$  неограниченного объема.

Обслуживание конфликтных потоков в каждой системе осуществляется в классе циклических алгоритмов с фиксированным ритмом переключения. Для разрешения конфликтности потоков после каждого акта обслуживания требуется акт переналадки. Удобно считать, что обслуживание требований в обеих системах осуществляется единым прибором, при этом конечное число  $n$  циклически сменяемых состояний  $\Gamma^{(1)}, \Gamma^{(2)}, \dots, \Gamma^{(n)}$  и фиксированные времена  $T_r, r = 1, 2, \dots, n$  пребывания в соответствующих состояниях  $\Gamma^{(r)}$ . Наряду с традиционными функциями по обслуживанию конфликтных потоков  $\Pi_1, \Pi_2, \Pi_3, \Pi_4$  новое обслуживающее устройство также совершает обслуживание требований потока  $\Pi_5$ .

Состояние  $\Gamma^{(r)}$  может принадлежать одному из девяти классов  $\Gamma^I, \Gamma^{II}, \dots, \Gamma^{IX}$ . При  $\Gamma^{(r)} \in \Gamma^I$  обслуживаются только требования из очереди  $O_5$ ; при  $\Gamma^{(r)} \in \Gamma^{II}$  – из очередей  $O_1$  и  $O_5$ ; при  $\Gamma^{(r)} \in \Gamma^{III}$  – из очередей  $O_2$  и  $O_5$ ; при  $\Gamma^{(r)} \in \Gamma^{IV}$  – из очередей  $O_3$  и  $O_5$ ; при  $\Gamma^{(r)} \in \Gamma^V$  – из очередей  $O_4$  и  $O_5$ ; при  $\Gamma^{(r)} \in \Gamma^{VI}$  – из очередей  $O_1, O_3$  и  $O_5$ ; при  $\Gamma^{(r)} \in \Gamma^{VII}$  – из очередей  $O_2, O_3$  и  $O_5$ ; при  $\Gamma^{(r)} \in \Gamma^{VIII}$  – из очередей  $O_1, O_4$  и  $O_5$ ; наконец, при  $\Gamma^{(r)} \in \Gamma^{IX}$  обслуживаются только требования из очередей  $O_2, O_4$  и  $O_5$ . При состоянии  $\Gamma^{(r)}$  обслуживающего устройства за время  $T_r$  каждое требование в очереди  $O_5$  с вероятностью  $p_r$  завершает обслуживание, покидает очередь  $O_5$  и поступает в поток  $\Pi_3$ , а с вероятностью  $1 - p_r$  остается в очереди до следующего такта. Для задания процесса обслуживания требований первых двух потоков будем использовать потоки насыщения  $\Pi_j^{нас}, j = 1, 2$ . Введем при  $j = 1$  множество  ${}^j\Gamma = \Gamma^{II} \cup \Gamma^{VI} \cup \Gamma^{VIII}$ , при  $j = 2$  множество  ${}^j\Gamma = \Gamma^{III} \cup \Gamma^{VII} \cup \Gamma^{IX}$ . Поток насыщения  $\Pi_j^{нас}$  содержит неслучайное число  $\ell_{r,j} \geq 1$  требований при  $\Gamma^{(r)} \in {}^j\Gamma$  и 0 требований при  $\Gamma^{(r)} \notin {}^j\Gamma$ .

Данный тандем систем отражает существенные особенности участков реальных транспортных сетей. Формальное построение математической модели и анализ условий существования стационарного режима для этой модели были проведены в работах [1–3]. Напомним существенные обозначения из упомянутых публикаций. Моменты времени  $\tau_0 = 0, \tau_1, \tau_2, \dots$  –

суть моменты смены состояния обслуживающего устройства. Случайный элемент  $\Gamma_i(\omega) \in \Gamma = \{\Gamma^{(1)}, \Gamma^{(2)}, \dots, \Gamma^{(n)}\}$  есть состояние обслуживающего устройства на промежутке  $(\tau_{i-1}, \tau_i]$ ,  $\chi_i(\omega) \in \{e^{(1)}, e^{(2)}, \dots, e^{(d)}\}$  – состояние внешней среды на промежутке  $(\tau_i, \tau_{i+1}]$ ,  $\kappa_{j,i}(\omega)$  – число требований в очереди  $O_j$  в момент  $\tau_i$ ,  $\eta_{5,i}(\omega)$  – число вторичных требований в потоке  $\Pi_5$  на промежутке  $(\tau_i, \tau_{i+1}]$ ,  $\bar{\xi}_{j,i}(\omega)$  – число обслуженных требований из очереди  $O_j$  на промежутке  $(\tau_i, \tau_{i+1}]$ . В настоящей работе изучается вопрос о распределении величины  $\eta_{5,i}$  в стационарном режиме функционирования тандема систем обслуживания. Для нелокального описания потока повторных требований введем метку  $v_i = (\Gamma_i, \chi_i, \kappa_{1,i}, \kappa_{2,i}, \bar{\xi}_{1,i-1}, \bar{\xi}_{2,i-1})$  повторных требований на промежутке  $(\tau_i, \tau_{i+1}]$  и будем рассматривать его нелокальное описание в виде маркированного точечного процесса  $\{(\tau_i(\omega), \eta_{5,i-1}(\omega), v_i(\omega)); i = 0, 1, \dots\}$ . Главную роль при изучении стационарного распределения величины  $\eta_{5,i}$  играет следующее соотношение:

$$\begin{aligned} & \mathbf{P}(\{\omega : \eta_{5,i}(\omega) = b\}) = \\ &= \sum_{l=1}^d \sum_{\Gamma^{(r)} \in {}^1\Gamma} \mathbf{P}(\{\omega : \Gamma_{i+1}(\omega) = \Gamma^{(r)}, \chi_{i+1}(\omega) = e^{(l)}, \bar{\xi}_{1,i}(\omega) = b\}) + \\ &+ \sum_{l=1}^d \sum_{\Gamma^{(r)} \in {}^2\Gamma} \sum_{x=b}^{\ell_{r,2}} \mathbf{P}(\{\omega : \Gamma_{i+1}(\omega) = \Gamma^{(r)}, \chi_{i+1}(\omega) = e^{(l)}, \bar{\xi}_{2,i}(\omega) = x\}) \psi(b; x, \alpha) + \\ &+ \sum_{l=1}^d \sum_{\Gamma^{(r)} \in \Gamma \setminus ({}^1\Gamma \cup {}^2\Gamma)} \delta_{0,b} \mathbf{P}(\{\omega : \Gamma_{i+1}(\omega) = \Gamma^{(r)}, \chi_{i+1}(\omega) = e^{(l)}\}), \end{aligned}$$

где  $\delta_{0,b}$  принимает значение 1 при  $b = 0$  и значение 0 в противном случае, а  $\psi(b; x, \alpha) = C_x^b \alpha^b (1 - \alpha)^{x-b}$ .

В работе доказывается, что стохастические последовательности

$$\{(\Gamma_i(\omega), \kappa_{j,i}(\omega), \chi_i(\omega), \bar{\xi}_{j,i-1}(\omega)); i = 0, 1, \dots\}, \quad j = 1, 2; \quad (1)$$

$$\{(\Gamma_i(\omega), \kappa_{1,i}(\omega), \kappa_{2,i}(\omega), \chi_i(\omega), \bar{\xi}_{1,i-1}(\omega), \bar{\xi}_{2,i-1}(\omega)); i = 0, 1, \dots\} \quad (2)$$

являются однородными марковскими цепями, и производится классификация состояний. Приводятся необходимые и достаточные условия существования стационарного распределения цепей (1), (2). Далее эти условия предполагаются выполненными. Показывается, что при некотором упорядочивании счетного множества существенных состояний каждой из цепей вида (1) матрица её переходных вероятностей обладает блочной структурой. При этом строки блоков, начиная с некоторой, получают сдвиг вправо на одну позицию предыдущей строки. Таким образом, оказывается возможным использовать технику цензурирования марковской цепи [4] для построения устойчивого рекуррентного численного алгоритма нахождения соответствующего стационарного распределения. Получающиеся

формулы приведены в расширенной версии статьи в силу их большого размера. Для расчёта по найденным формулам была реализована программа в среде Octave [5].

*Работа выполнена в рамках фундаментальной НИР «Математическое моделирование и анализ стохастических эволюционных систем и процессов принятия решений» (№ государственной регистрации 01201456585).*

#### Литература

1. Зорин А. В. Стохастическая модель сообщающихся систем массового обслуживания с повторными вызовами и циклическим управлением в случайной среде // Кибернетика и системный анализ. 2013. Т. 49, № 6. С. 100–109.
2. Зорин А. В. Об условиях существования стационарного режима в тандеме систем массового обслуживания с циклическим управлением в случайной среде // Автоматика и вычислительная техника. 2013. № 4. С. 28–39.
3. Зорин А. В., Кузнецов Н. Ю., Кузнецов И. Н. Анализ стохастической модели сообщающихся систем массового обслуживания с повторными вызовами и циклическим алгоритмом управления в случайной среде // Вестник Нижегородского государственного университета им. Н. И. Лобачевского. 2013. № 5. С. 217–223.
4. Grassmann W. K., Heyman D. P. Equilibrium distribution of block-structured Markov chains with repeating rows // J. of Appl. Prob. 1990. Vol. 27. P. 557–576.
5. Eaton J. W. The GNU Octave Manual. Network Theory, Ltd, 2008. 568 p.

## **ИНВАРИАНТНОСТЬ СТАЦИОНАРНОГО РАСПРЕДЕЛЕНИЯ ВЕРОЯТНОСТЕЙ СОСТОЯНИЙ ЗАМКНУТОЙ СЕТИ С НЕАКТИВНЫМИ ЗАЯВКАМИ И МНОГОРЕЖИМНЫМ ОБСЛУЖИВАНИЕМ**

*Ю. С. Крук, Ю. Е. Дудовская*

*Белорусский национальный технический университет,  
Гомельский государственный университет им. Ф. Скорины,  
Гомель, Беларусь*

В работе исследуется замкнутая сеть массового обслуживания с временно неактивными заявками, многорежимными стратегиями обслуживания и информационными сигналами. Будем считать, что в сети циркулируют заявки двух типов. Первые могут обслуживаться узлами, а вторые являются временно неактивными и не обслуживаются, скапливаясь в очередях узлов. Поступающие в сеть потоки информационных сигналов позволяют заявкам менять свое состояние: из неактивного состояния переходить в состояние, когда они могут получать обслуживание, и наоборот.

Приборы в узлах сети могут функционировать в нескольких режимах, отвечающих различной степени работоспособности узлов. При переходе узла в режим с большим номером производительность узла уменьшается, ухудшается процесс обслуживания. При переходе узла в режим с меньшим номером происходит восстановление показателей процесса обслуживания,

улучшается качество обслуживания. Переходы возможны только в соседние режимы. При переключении режимов число заявок в узле не меняется.

Предполагается, что количество работы, требующееся для обслуживания заявок в узлах, распределено по произвольному закону. Для рассмотренной сети устанавливается инвариантность стационарного распределения вероятностей состояний сети по отношению к функциональной форме распределений количества работы при фиксированных первых моментах.

Исследуется замкнутая сеть массового обслуживания с множеством узлов  $J = \{1, \dots, N\}$ . В сети циркулируют  $M$  заявок. Все заявки, находящиеся в сети, подразделяются на обыкновенные, которые могут получать обслуживание, и неактивные.

В узлы сети извне поступают независимые пуассоновские потоки информационных сигналов с интенсивностями  $\nu_i$  и  $\varphi_i$ ,  $i \in J$ . Информационный сигнал, поступивший в  $i$ -й узел с интенсивностью  $\nu_i$ , уменьшает количество обыкновенных заявок на единицу и увеличивает на единицу количество неактивных заявок; в случае отсутствия в  $i$ -м узле обыкновенных заявок сигнал покидает сеть. Информационный сигнал, поступивший в  $i$ -й узел с интенсивностью  $\varphi_i$ , уменьшает на единицу количество неактивных заявок, увеличивая на единицу число обыкновенных заявок; в случае отсутствия в  $i$ -м узле неактивных заявок сигнал покидает сеть. Описанные информационные сигналы не требуют обслуживания.

Предполагается, что  $i$ -й узел может находиться в одном из  $n_i''$  режимов работы  $(n_i'' = \overline{0, r_i}, i = \overline{1, N})$ .

Состояние сети в момент времени  $t$  характеризуется вектором  $z(t) = ((n_i(t), n_i'(t), n_i''(t)), i \in J)$ , где  $(n_i(t), n_i'(t), n_i''(t))$  – состояние  $i$ -го узла в момент времени  $t$ . Здесь  $n_i(t)$  и  $n_i'(t)$  – число обыкновенных и соответственно неактивных заявок в  $i$ -м узле в момент времени  $t$ ;  $n_i(t) + n_i'(t)$  – общее число заявок в  $i$ -м узле;  $n_i''(t)$  – режим функционирования  $i$ -го узла. Случайный процесс  $z(t)$  имеет конечное фазовое пространство

$$Z = \left\{ ((n_1, n_1', n_1''), \dots, (n_N, n_N', n_N'')) : n_i, n_i' \geq 0, \sum_{i=1}^N (n_i + n_i') = M, n_i'' = \overline{0, r_i}, i \in J \right\}.$$

Нумерация обыкновенных заявок в очереди каждого узла осуществляется от «хвоста» очереди к прибору, т. е. если в  $i$ -м узле находится  $n_i$  обыкновенных заявок, то заявка, которая обслуживается, имеет номер  $n_i$ , а последняя заявка в очереди имеет номер 1. Временно неактивные заявки в очереди  $i$ -го узла нумеруются следующим образом: заявка, последняя ставшая неактивной, имеет номер  $n_i'$ . Поступающий в узел  $i$  сигнал  $\nu_i$  воздействует на обыкновенную заявку, имеющую номер 1, которая становится неактивной заявкой под номером  $n_i' + 1$ . Сигнал  $\varphi_i$  воздействует на

неактивную заявку, имеющую номер  $n'_i$ , которая становится обыкновенной заявкой под номером 1.

Назовем нулевой режим основным режимом работы. Время работы узла, находящегося в состоянии  $z_i = (n_i, n'_i, n''_i)$ , в режиме  $n''_i$  ( $n''_i = \overline{0, r_i}, i = \overline{1, N}$ ) имеет показательное распределение, при этом с интенсивностью  $\sigma_i(n_i + n'_i, n''_i)$  ( $\sigma_i(n_i + n'_i, n''_i) > 0$ )  $i$ -й узел переходит в  $(n''_i + 1)$ -й режим ( $n''_i = \overline{0, r_i - 1}$ ), а с интенсивностью  $\rho_i(n_i + n'_i, n''_i)$  ( $\rho_i(n_i + n'_i, n''_i) > 0$ ) – в  $(n''_i - 1)$ -й режим ( $n''_i = \overline{1, r_i}$ ). Переключение прибора с одного режима в другой сохраняет общее число заявок в узле.

Дисциплина обслуживания – *LCFS-PR*. Поступающая в  $i$ -й узел заявка начинает сразу обслуживаться и получает номер  $n_i + 1$ , а вытесненная заявка сохраняет номер  $n_i$  и становится первой в очереди на дообслуживание. Предполагается, что в начальный момент временно неактивные заявки в сети отсутствуют.

Если в момент времени  $t$  состояние  $i$ -го узла есть вектор  $(n_i(t), n'_i(t), n''_i(t))$  и сразу после указанного момента в этот узел поступает заявка, которая, как отмечалось выше, начинает немедленно обслуживаться, то количество работы по ее обслуживанию является случайной величиной  $\eta_i(n_i + n'_i + 1)$  с функцией распределения  $B_i(n_i + n'_i + 1, z)$  и математическим ожиданием  $\tau_i(n_i + n'_i + 1) < \infty$ . Предполагается, что  $B_i(n_i + n'_i + 1, 0) = 0$ ,  $i \in J$ .

Если в момент времени  $t$  состояние  $i$ -го узла есть вектор  $(n_i(t), n'_i(t), n''_i(t))$ , то обслуживание ведется со скоростью  $\alpha_i(n_i + n'_i)$ , т. е. зависит от состояния узла,  $i \in J$ .

Обслуживание в сети имеет не «временную», а так называемую «энергетическую» трактовку, т.е. каждая операция обслуживания характеризуется случайной величиной работы, которую необходимо выполнить.

Заявка, получившая обслуживание в  $i$ -м узле, мгновенно с вероятностью  $p_{i,j}$  переходит в узел  $j$  ( $\sum_{j \in J} p_{i,j} = 1, i \in J$ ). Не ограничивая общности рассуждений, будем считать  $p_{i,i} = 0, i \in J$ .

Для замкнутых сетей показано, что в случае неприводимости матрицы маршрутизации  $(p_{i,j})$  система уравнений трафика

$$\varepsilon_j = \sum_{i \in J} \varepsilon_i p_{i,j}, \quad j \in J, \quad (1)$$

имеет единственное с точностью до постоянного множителя положительное решение  $\varepsilon_j$  [1].

Случайный процесс  $z(t)$  не является марковским. Пусть  $\psi_{i,k}(t)$  – количество работы, которое осталось выполнить с момента  $t$  для завершения обслуживания заявки, стоящей в момент времени  $t$  на  $k$ -й позиции в  $i$ -м узле:

$$\Psi_i(t) = (\psi_{i,1}(t), \dots, \psi_{i,n_i+n'_i}(t)), \quad i \in J.$$

В силу сказанного выше

$$\frac{d\psi_{i,n_i+n'_i}(t)}{dt} = -\alpha_i(n_i + n'_i), \quad i \in J,$$

если  $(n_i, n'_i, n''_i)$  – состояние  $i$ -го узла.

Рассмотрим кусочно-линейный марковский процесс  $\zeta(t) = (z(t), \psi(t))$ , добавляя к  $z(t)$  непрерывную компоненту  $\psi(t) = (\psi_1(t), \dots, \psi_N(t))$ .

Введем следующие обозначения:

$$\begin{aligned} F(z, x) &= F(z, x_{1,1}, \dots, x_{1,n_1+n'_1}; x_{2,1}, \dots, x_{2,n_2+n'_2}; \dots; x_{N,1}, \dots, x_{N,n_N+n'_N}) = \\ &= \lim_{t \rightarrow \infty} P \left\{ z(t) = z, \psi_{i,1}(t) < x_{i,1}, \dots, \psi_{i,n_i+n'_i}(t) < x_{i,n_i+n'_i}, i \in J \right\}, \\ &z \in Z, \quad x_{k,l} \in R, \quad k = \overline{1, N}, \quad l = \overline{1, n_k + n'_k}. \end{aligned}$$

Функции  $F(z, x)$  будем называть стационарными функциями распределения вероятностей состояний кусочно-линейного процесса  $\zeta(t)$ , поскольку при каждом фиксированном  $z$  функция  $F(z, x)$  в части непрерывных компонент представляет собой функцию распределения.

Модель замкнутой сети с неактивными заявками рассмотрена в работе [2]. В [2] исследован случай, когда количество работы по обслуживанию заявки является произвольно распределенной случайной величиной. Доказана следующая теорема.

**Теорема 1.** *Марковский процесс  $\zeta(t)$  эргодичен, а стационарные функции распределения вероятностей состояний  $F(z, x)$  определяются по формулам*

$$\begin{aligned} F(z, x) &= G^{-1}(M, N) p_1(n_1, n'_1) p_2(n_2, n'_2) \dots p_N(n_N, n'_N) \times \\ &\times \prod_{i=1}^N \prod_{s=1}^{n_i+n'_i} \frac{1}{\tau_i(s)} \int_0^{x_{i,s}} (1 - B_i(s, u)) du, \quad z \in Z, \end{aligned}$$

где

$$p_i(n_i, n'_i) = \varepsilon_i^{n_i} \left( \frac{\varepsilon_i \nu_i}{\phi_i} \right)^{n'_i} \prod_{s=1}^{n_i+n'_i} \frac{\tau_i(s)}{\alpha_i(s)},$$

$\varepsilon_i$  – решения уравнений трафика,  $G(M, N)$  – нормирующая постоянная.

Для замкнутой сети с неактивными заявками, многорежимным обслуживанием и информационными сигналами справедлива следующая теорема.

**Теорема 2.** Марковский процесс  $\zeta(t)$  эргодичен. При выполнении соотношений

$$\begin{aligned} & \sigma_i(n_i + n'_i, n''_i - 1) \alpha_i(n_i + n'_i, n''_i) \rho_i(n_i + n'_i - 1, n''_i) = \\ & = \sigma_i(n_i + n'_i - 1, n''_i - 1) \alpha_i(n_i + n'_i, n''_i - 1) \rho_i(n_i + n'_i, n''_i), \end{aligned} \quad (2)$$

$$1 \leq n_i + n'_i \leq M, 1 \leq n''_i \leq r_i, 1 \leq i \leq N,$$

стационарные функции распределения вероятностей состояний процесса  $\zeta(t)$  определяются по формулам

$$\begin{aligned} F(z, x) = & G^{-1}(M, N) p_1(n_1, n'_1, n''_1) p_2(n_2, n'_2, n''_2) \dots p_N(n_N, n'_N, n''_N) \times \\ & \times \prod_{i=1}^N \prod_{s=1}^{n_i+n'_i} \frac{1}{\tau_i(s)} \int_0^{x_{i,s}} (1 - B_i(s, u)) du, \quad z \in Z, \end{aligned}$$

где

$$p_i(n_i, n'_i, n''_i) = \varepsilon_i^{n_i} \left( \frac{\varepsilon_i v_i}{\phi_i} \right)^{n'_i} \prod_{s=1}^{n_i+n'_i} \frac{\tau_i(s)}{\alpha_i(s, n''_i)} \prod_{k=1}^{n''_i} \frac{\sigma_i(0, k-1)}{\rho_i(0, k)}, \quad (3)$$

$\varepsilon_i$  – решения уравнений трафика (1);  $G(M, N)$  – постоянная, которая может быть найдена из условия нормировки

$$\sum_{((n_1, n'_1, n''_1), \dots, (n_N, n'_N, n''_N)) \in Z} p((n_1, n'_1, n''_1), \dots, (n_N, n'_N, n''_N)) = 1.$$

Пусть  $\{p(z), z \in Z\}$  – стационарное распределение вероятностей состояний процесса  $z(t)$ . Из теоремы 2 с учетом равенства  $p(z) = F(z, +\infty)$  вытекает следующее следствие.

**Следствие.** Процесс  $z(t)$  эргодичен. При выполнении условий (2) стационарное распределение вероятностей состояний

$$p(z) = G^{-1}(M, N) p_1(n_1, n'_1, n''_1) p_2(n_2, n'_2, n''_2) \dots p_N(n_N, n'_N, n''_N), \quad z \in Z,$$

не зависит от функционального вида распределений  $B_i(s, x)$ ,  $i \in J$ . Вероятности  $p_i(n_i, n'_i, n''_i)$  определяются по формуле (3).

Таким образом, для случая произвольного распределения величины работы, требующейся для обслуживания заявок, установлена инвариантность стационарного распределения вероятностей состояний сети по отношению к функциональной форме распределений величины работы при фиксированных первых моментах.

#### Литература

1. Gordon W. J. Closed queueing networks with exponential servers // Oper. Res. 1967. No. 15. P. 252–267.
2. Bojarovich J. Stationary distribution insensitivity of a closed queueing network with non-active customers / J. Bojarovich, Y. Dudovskaya // Information Technologies and Mathematical Modelling. 13<sup>th</sup> International Scientific Conference, ITMM 2014, named after A. F. Terpugov, Anzhero-Sudzhensk, Russia, November 20–22, 2014. Proceedings. Switzerland: Springer International Publishing, 2014. Vol. 487. P. 50–58.

# ИССЛЕДОВАНИЕ ВРЕМЕНИ ОЖИДАНИЯ ОБСЛУЖИВАНИЯ КЛИЕНТОВ CALL-ЦЕНТРА

*Е. Ю. Лисовская, С. П. Моисеева*

*Национальный исследовательский*

*Томский государственный университет, Томск, Россия*

Известно достаточно большое количество работ по моделированию работы центров обработки вызовов (call-центров) – это услуга сети, в которой агенты предоставляют телефонные услуги. Как правило, число операторов, работающих в таких компаниях, может быть достаточно велико. Обслуживание каждого клиента начинается незамедлительно (т.е. системы без отказов) [5]. В таких центрах соединение происходит с автоответчиком, после чего при необходимости происходит соединение с оператором. В настоящей работе предлагается математическая модель работы call-центра в виде двухфазной системы массового обслуживания. Первая фаза представляет собой систему массового обслуживания с неограниченным числом приборов, на вход которой поступает простейший поток заявок с параметром  $\bar{\lambda}$ , время обслуживания имеет произвольный закон распределения. Известно [6], что выходящий поток также будет пуассоновским с тем же параметром. После обслуживания на первой фазе с вероятностью  $r$  происходит соединение с операторами, число которых ограничено, или с вероятностью  $1-r$  покидает систему. Возможно возникновение очереди, поэтому вторая фаза представляет СМО вида  $M/GI/N/\infty$ . Ставится задача определения оптимального числа операторов, обеспечивающих работу центра, при которой время ожидания не превышает заданных значений.

## **Математическая модель**

На вход системы поступает простейший поток с параметром  $\lambda$  и после обслуживания на первой фазе с вероятностью  $r$  переходит на вторую фазу. Таким образом, на вход второй фазы поступает простейший поток с параметром  $\lambda = \bar{\lambda}r$ . Поступающая заявка занимает любой из свободных приборов или становится в очередь в случае, когда все приборы заняты. Система имеет  $N$  обслуживающих приборов, время обслуживания каждой заявки является случайной величиной с произвольной функцией распределения  $A(x)$ , одинаковой для всех приборов.

Известно, что для системы  $M|M|N|K$  получены формулы Эрланга [6]. Вместе с тем для произвольного времени обслуживания и бесконечной очереди полученная задача теоретически не решена, и представить аналитическое решение невозможно, поэтому ставится задача построения аппроксимации стационарного распределения вероятностей  $P(i)$ ,  $0 \leq i < \infty$ , числа заявок в рассматриваемой системе  $M|GI|N|\infty$ .

## Аппроксимация распределения вероятностей числа заявок в системе M|GI|N|∞

Обозначим  $\pi_i$  – аппроксимация распределения вероятностей  $P(i)$ , которые будем определять в виде составного распределения [7]:

$$\pi_i = \begin{cases} C_1 P_1(i), & 0 \leq i \leq N, \\ C_2 P_2(i - N + 1), & i \geq N, \end{cases} \quad (1)$$

Здесь  $P_1(i)$ ,  $0 \leq i \leq N$  определяются как вероятности числа занятых приборов в  $N$ -линейной СМО (M|GI|N|0) с потерями заявок и задаются формулами Эрланга [6]:

$$P_1(i) = \frac{(\lambda a)^i}{i!} \Big/ \left( \sum_{i=0}^N \frac{(\lambda a)^i}{i!} \right), \quad (2)$$

где  $a = \int_0^{\infty} (1 - A(x)) dx$  – среднее время обслуживания.

Вероятности  $P_2(i)$  определяются как вероятности числа заявок в од-нолинейной системе M|GI|1|∞ с ожиданием для случая, когда все приборы заняты. В этом случае блок занятых приборов можно считать как один прибор, который обслуживает случайное время с функцией распределения  $B(x)$ . Поэтому вероятности  $P_2(i)$ ,  $i = 0, 1, \dots$  находим по формуле Поллаче-ка–Хинчина [6] для производящей функции:

$$G(x) = \sum_{n=0}^{\infty} x^n P_2(n) = (1 - \lambda b) \frac{(1 - x) B^*(\lambda - \lambda x)}{B^*(\lambda - \lambda x) - x}. \quad (3)$$

В работе [7] показано, что функция распределения  $B(x)$  имеет вид

$$B(x) = 1 - (1 - A(x)) \left( 1 - \frac{1}{a} \int_0^x (1 - A(z)) dz \right)^{N-1}, \quad (4)$$

а значения констант  $C_1$  и  $C_2$  определяются выражениями

$$C_1 = \frac{P_2(1)}{P_2(1) + P_1(N)(1 - (P_2(0) + P_2(N)))},$$

$$C_2 = \frac{P_1(N)}{P_2(1) + P_1(N)(1 - (P_2(0) + P_2(N)))}.$$

### Вероятность немедленного обслуживания

Важной характеристикой для исследования является время ожидания заявки в очереди. Если заявка поступает в систему в тот момент времени, когда в системе есть хотя бы один свободный прибор, то она немедленно встает на обслуживание. Нетрудно показать, что вероятность немедленного обслуживания, учитывая (1), можно записать в виде

$$P_0 = \frac{(1-\lambda b) \frac{1-B^*(\lambda)}{B^*(\lambda)} \left( 1 - \frac{(\lambda a)^N}{N!} \sum_{i=0}^N \frac{(\lambda a)^i}{i!} \right)}{(1-\lambda b) \frac{1-B^*(\lambda)}{B^*(\lambda)} + \frac{(\lambda a)^N}{N!} \sum_{i=0}^N \frac{(\lambda a)^i}{i!} \left[ 1 - (1-\lambda b) \left( 1 + \frac{1-B^*(\lambda)}{B^*(\lambda)} \right) \right]}. \quad (5)$$

### Распределение вероятностей положительного времени ожидания

Если заявка поступает в систему в тот момент времени, когда все приборы заняты, тогда ее время ожидания  $\tau > 0$ , и эту величину будем называть положительным временем ожидания  $\tau^+$ .

Найдем условное распределение вероятностей  $P_{ож}(m)$ ,  $m \geq 0$ , того, что в очереди  $m$  заявок, при условии, что все приборы системы заняты.

Применяя равенство (1), запишем

$$\begin{aligned} P_{ож}(m) &= \pi(N+m) / \sum_{i=0}^{\infty} \pi(N+i) = C_2 P_2(1+m) / C_2 \sum_{i=0}^{\infty} P_2(1+i) = \\ &= P_2(1+m) / (1 - P_2(0)) = \frac{1}{\lambda b} P_2(1+m). \end{aligned} \quad (6)$$

То есть условное распределение вероятностей  $P_{ож}(m)$  того, что в очереди  $m$  заявок, при условии, что все приборы системы заняты, имеет вид

$$P_{ож}(m) = \frac{1}{\lambda b} P_2(m+1). \quad (7)$$

Найдем производящую функцию  $G_{ож}(x)$  этого распределения:

$$\begin{aligned} G_{ож}(x) &= \sum_{m=0}^{\infty} x^m P_{ож}(m) = \sum_{m=0}^{\infty} x^m \frac{1}{\lambda b} P_2(m+1) = \frac{1}{\lambda b x} \sum_{v=1}^{\infty} x^v P_{ож}(v) = \frac{1}{\lambda b x} [G(x) - P_0] = \\ &= \frac{1}{\lambda b x} [G(x) - (1 - \lambda b)] = \frac{1}{\lambda b x} \left[ (1 - \lambda b) \frac{(1-x)B^*(\lambda - \lambda x)}{B^*(\lambda - \lambda x) - x} - (1 - \lambda b) \right] = \\ &= \frac{1 - \lambda b}{\lambda b x} \frac{(1-x)B^*(\lambda - \lambda x) - B^*(\lambda - \lambda x) + x}{B^*(\lambda - \lambda x) - x} = \frac{1 - \lambda b}{\lambda b} \frac{1 - B^*(\lambda - \lambda x)}{B^*(\lambda - \lambda x) - x}, \end{aligned}$$

т.е.

$$G_{ож}(x) = \frac{1 - \lambda b}{\lambda b} \frac{1 - B^*(\lambda - \lambda x)}{B^*(\lambda - \lambda x) - x}. \quad (8)$$

Эта производящая функция получена на периоде занятости рассматриваемой системы, то есть когда заняты все ее приборы. В этом условии N-линейный блок обслуживания, определенный функцией распределения

$A(x)$ , допустимо заменить однолинейным с функцией распределения времени обслуживания  $B(x)$  вида (4).

Заявка, поступившая в систему на периоде ее занятости, с вероятностью  $P_{ож}(m)$  обнаруживает в очереди  $m$  заявок, поэтому ее время ожидания  $\tau^+$  складывается из суммарного времени обслуживания  $m$  заявок, каждое из которых имеет функцию распределения  $B(x)$  из (9) и времени до обслуживания одной заявки, имеющего функцию распределения

$$B_0(x) = \frac{1}{b} \int_0^x (1 - B(z)) dz.$$

Обозначив  $\xi_0$  – остаточное время обслуживания, а  $\xi_1, \xi_2, \dots, \xi_m$  – времена обслуживания первой, второй и  $m$ -й заявок в очереди, время ожидания можно определить как

$$\tau^+ = \xi_0 + \xi_1 + \dots + \xi_m.$$

Найдем характеристическую функцию  $h(u)$  положительного времени ожидания заявок в системе  $M|GI|N|\infty$ . Применяя формулу полной вероятности для математических ожиданий, можно записать

$$\begin{aligned} h(u) &= M \left\{ e^{ju\tau^+} \right\} = \sum_{m=0}^{\infty} M \left\{ \exp \{ ju(\xi_0 + \xi_1 + \dots + \xi_m) \} \mid m(t) = m \right\} P_{ож}(m) = \\ &= \sum_{m=0}^{\infty} M \left\{ e^{ju\xi_0} \right\} \left( M \left\{ e^{ju\xi_m} \right\} \right)^m P_{ож}(m) = \varphi_0(u) \sum_{m=0}^{\infty} \varphi(u)^m P_{ож}(m), \end{aligned} \quad (9)$$

где  $\varphi_0(u)$  и  $\varphi(u)$  – характеристические функции остаточного и полного времени обслуживания одной заявки, здесь

$$\varphi(u) = \int_0^{\infty} e^{jux} dB(x). \quad (10)$$

Последнее равенство для  $h(u)$  перепишем в виде

$$h(u) = \varphi_0(u) G_0(\varphi(u)) = \varphi_0(u) \frac{1 - \lambda b}{\lambda b} \frac{1 - B^*(\lambda - \lambda\varphi(u))}{B^*(\lambda - \lambda\varphi(u)) - \varphi(u)}. \quad (11)$$

Так как

$$B_0(x) = \frac{1}{b} \int_0^x (1 - B(z)) dz,$$

то

$$\varphi_0(u) = \int_0^{\infty} e^{jux} dB_0(x) = \frac{1}{b} \int_0^x e^{jux} (1 - B(x)) dx = \frac{1}{jub} (\varphi(u) - 1),$$

поэтому характеристическую функцию  $h(u)$  из (11) можно окончательно записать в виде

$$h(u) = \frac{1}{jub} \frac{(1-\lambda b)}{\lambda b} (\varphi(u) - 1) \frac{1 - B^*(\lambda - \lambda\varphi(u))}{B^*(\lambda - \lambda\varphi(u)) - \varphi(u)}. \quad (12)$$

Здесь  $\varphi(u)$  имеет вид (10).

Формулы (5) и (12) полностью характеризуют время ожидания заявки в очереди N-линейной системы M|GI|N| $\infty$ . Формула (5) для  $P_0$  определяет вероятность того, что время ожидания будет нулевым, а  $h(u)$  из (12) является характеристической функцией распределения вероятностей положительного времени ожидания  $\tau^+$  заявки в очереди.

Применяя характеристическую функцию  $h(u)$  из (12), среднее значение  $\bar{\tau}^+$  положительного времени ожидания заявки в очереди запишем в виде

$$\bar{\tau}^+ = \frac{1}{j} h'(u) \Big|_{u=0} = \frac{b_2}{2b(1-\lambda b)}, \quad (13)$$

где  $b$  и  $b_2$  – математическое ожидание и второй начальный момент случайной величины, определяемой функцией распределения  $B(x)$  из (4).

### Заключение

В работе построена математическая модель call-центра, получена формула для нахождения вероятности немедленного обслуживания, получена характеристическая функция распределения вероятностей положительного времени ожидания заявок в очереди. Полученные результаты позволяют определить оптимальное число приборов на второй фазе, что то же самое число работающих операторов в call-центре.

### Литература

1. Mandelbaum A., Pats G. State-dependent queues: approximations and applications // In Stochastic Networks, IMA Volumes in Mathematics / F. P. Kelly and R. J. Williams, eds. Springer, 1995. P. 239–282.
2. Borst S., Mandelbaum A., Reiman, M. Dimensioning Large Call Centers // Operations Research. 2004. Vol. 52. P. 17–34; downloadable at <http://iew3.technion.ac.il/serveng/References/references.html>
3. Liu L, Kashyap B. R. K., Templeton J. G. C. The service system M/MR/ $\infty$  with impatient customers // Queuing Systems. 1987. Vol. 2, no. 4. P. 363–372.
4. Tsoukato K. P., Makowski A. M. Heavy Traffic Analysis for A Multiplexer Driven by M/GI/infinity Input Processes [Электронный ресурс]. URL: <http://handle.dtic.mil/100.2/ADA4555833>. P. 780–789.
5. Brown L., Gans N., Mandelbaum A. et al. Statistical Analysis of a Telephone Call Center // A Queueing-Science Perspective. Journal of the American Statistical Association. 2005. Vol. 100. P. 36–50.
6. Назаров А. А. Теория массового обслуживания : учеб. пособие / А. А. Назаров, А. Ф. Терпугов. Томск: Изд-во НТЛ, 2010. 228 с.
7. Лисовская Е. Ю. Исследование процесса числа заявок в системе M|GI|N| $\infty$  / Е. Ю. Лисовская, С. П. Моисеева // Теория вероятностей, случайные процессы, математическая статистика и приложения: матер. Междунар. науч. конф., посвящ.

## ИССЛЕДОВАНИЕ ДИНАМИЧЕСКОЙ RQ-СИСТЕМЫ С ВХОДЯЩИМ MАР-ПОТОКОМ ЗАЯВОК

*Т. В. Любина, В. О. Азева*

*Филиал Кемеровского государственного университета  
в г. Анжеро-Судженске, Россия*

Рассмотрим систему массового обслуживания с источником повторных вызовов и входящим MАР-потокм заявок, управляемую динамическим протоколом доступа, которую будем называть динамической RQ-системой с входящим MАР-потокм заявок [1].

На вход RQ-системы поступает MАР-поток заявок из внешнего источника, определяемый матрицей  $Q$  инфинитезимальных характеристик  $q_{vn}$  цепи Маркова  $n(t)$ , управляющей MАР-потокм, также задан набор неотрицательных чисел  $\rho\lambda_n$  и вероятностей  $d_{mn} = 0$ , которые целесообразно определять матрицей  $D = [d_{vn}]$  и диагональной матрицей  $\rho\Lambda$  условных интенсивностей  $\rho\lambda_n$  на главной диагонали. Заявка, заставшая прибор свободным, занимает его для обслуживания в течение случайного времени, распределенного по экспоненциальному закону с параметром  $\mu$ . По завершении обслуживания заявка покидает прибор. Если во время обслуживания заявки поступает другая, то поступившая заявка переходит в источник повторных вызовов. Из источника повторных вызовов после случайной задержки заявка с динамической (зависящей от состояния источника повторных вызовов) интенсивностью  $\gamma/i$  вновь обращается к прибору с повторной попыткой его захвата;  $i$  – число заявок в источник повторных вызовов. Если прибор свободен, то заявка становится на обслуживание, если же он занят, то возвращается в источник повторных вызовов [2].

Состояние системы в момент времени  $t$  определяется трехмерной цепью Маркова  $\{k(t), n(t), i(t)\}$ , где  $i(t)$  – число заявок в ИПВ;  $n(t)$  – значения цепи Маркова, управляющей MМРР-потокм, а  $k(t)$  определяет состояние прибора следующим образом:  $k(t) = 0$ , если прибор свободен, и  $k(t) = 1$ , если прибор занят.

Обозначим  $P\{k(t) = k, n(t) = n, i(t) = i\} = P(k, n, i, t)$  – вероятность того, что в момент времени  $t$  прибор находится в состоянии  $z$ , цепь Маркова в состоянии  $n$  и в ИПВ  $i$  заявок. Таким образом, распределение вероятностей  $P(k, n, i, t)$  удовлетворяет следующей системе дифференциальных уравнений Колмогорова для распределения вероятностей  $P(k, n, i, t)$ :

$$\left\{ \begin{aligned} \frac{\partial P(0,n,i,t)}{\partial t} &= -(\rho\lambda_n + \gamma)P(0,n,i,t) + \mu P(1,n,i,t) + \\ &+ \sum_v \{P(0,v,i,t)(1-d_{vn})\} q_{vn}, \\ \frac{\partial P(1,n,i,t)}{\partial t} &= -(\rho\lambda_n + \mu)P(1,n,i,t) + \gamma P(0,n,i+1,t) + \\ &+ \rho\lambda_n P(0,n,i,t) + \rho\lambda_n P(0,n,i-1,t) + \\ &+ \sum_v \{P(1,v,i,t)(1-d_{vn}) + P(0,v,i,t)d_{vn} + P(1,v,i-1,t)d_{vn}\} q_{vn}. \end{aligned} \right. \quad (1)$$

Решение системы уравнений Колмогорова (1) достаточно полно определяет функционирование динамической RQ-системы с входящим MAP-потокм заявок. Допредельное исследование проведем методом производящих функций.

### Исследование методом производящих функций

Будем полагать, что система функционирует в стационарном режиме, т.е.  $P(k,n,i,t) \equiv P(k,n,i)$ . Запишем систему (1) для стационарного распределения в матричном виде. Обозначив вектор-строки

$$\mathbf{P}(k,i) = \{P(k,1,i), P(k,2,i), \dots, P(k,N,i)\},$$

получим

$$\begin{aligned} \mathbf{P}(0,0)(\mathbf{Q} - \rho\mathbf{\Lambda}) + \mathbf{P}(1,0)\mu - \mathbf{P}(0,0)\mathbf{A} &= 0, \quad i = 0, \\ \mathbf{P}(1,0)(\mathbf{Q} - \rho\mathbf{\Lambda} - \mu\mathbf{I}) + \mathbf{P}(0,0)\rho\mathbf{\Lambda} - \mathbf{P}(1,0)\mathbf{A} + \mathbf{P}(0,0)\mathbf{A} + \mathbf{P}(0,1)\gamma &= 0, \quad i = 0, \\ \mathbf{P}(0,i)(\mathbf{Q} - \rho\mathbf{\Lambda} - \gamma\mathbf{I}) + \mathbf{P}(1,i)\mu - \mathbf{P}(0,i)\mathbf{A} &= 0, \quad i \geq 1, \\ \mathbf{P}(1,i)(\mathbf{Q} - \rho\mathbf{\Lambda} - \mu\mathbf{I}) + \mathbf{P}(0,i)\rho\mathbf{\Lambda} + \mathbf{P}(1,i-1)\rho\mathbf{\Lambda} + \\ + \mathbf{P}(0,i+1)\gamma - \mathbf{P}(1,i)\mathbf{A} + \mathbf{P}(0,i)\mathbf{A} + \mathbf{P}(1,i-1)\mathbf{A} &= 0, \quad i \geq 1, \end{aligned} \quad (2)$$

где матрица  $\mathbf{A} = \mathbf{D} \cdot \mathbf{Q}$ , произведение Адамара двух матриц  $\mathbf{D}$  и  $\mathbf{Q}$ .

Чтобы решить систему (2), определим векторные производящие функции

$$\mathbf{G}(k,x) = \sum_{i=0}^{\infty} x^i \mathbf{P}(k,i), \quad k = 0, 1. \quad (3)$$

Из системы (2) с учетом равенства (3) получаем следующую систему для функций  $\mathbf{G}(k,x)$ :

$$\left\{ \begin{aligned} \mathbf{G}(0,x)(\mathbf{Q} - \rho\mathbf{\Lambda} - \gamma\mathbf{I} - \mathbf{A}) + \mathbf{G}(1,x)\mu &= -\gamma\mathbf{P}(0,0), \\ \mathbf{G}(0,x)((\rho\mathbf{\Lambda} + \mathbf{A})x + \gamma\mathbf{I}) + \mathbf{G}(1,x)(\mathbf{Q} + (x-1)(\rho\mathbf{\Lambda} + \mathbf{A}) - \mu\mathbf{I})x &= \gamma\mathbf{P}(0,0). \end{aligned} \right. \quad (4)$$

Из системы (4) получаем выражения для  $\mathbf{G}(0,x)$  и  $\mathbf{G}(1,x)$ :

$$\mathbf{G}(0,x) = \mathbf{P}(0,0) \left\{ \gamma\mathbf{I} + \frac{\gamma}{\mu} x (\mathbf{Q} + (x-1)(\rho\mathbf{\Lambda} + \mathbf{A}) - \mu\mathbf{I}) \right\} \left\{ (1-x)\gamma\mathbf{I} + x\mathbf{Q} - \right.$$

$$-\frac{1}{\mu}(\mathbf{Q} - \rho\mathbf{\Lambda} - \mathbf{A} - \gamma\mathbf{I})(\mathbf{Q} + (x-1)(\rho\mathbf{\Lambda} + \mathbf{A}))x \Big\}^{-1}, \quad (5)$$

$$\mathbf{G}(1, x) = -\frac{1}{\mu}[\gamma\mathbf{P}(0, 0) + \mathbf{G}(0, x)(\mathbf{Q} - \rho\mathbf{\Lambda} - \mathbf{A} - \gamma\mathbf{I})].$$

Обозначим матрицы

$$\mathbf{A}(x) = \gamma\mathbf{I} + \frac{\gamma}{\mu}x(\mathbf{Q} + (x-1)(\rho\mathbf{\Lambda} + \mathbf{A}) - \mu\mathbf{I}),$$

$$\mathbf{B}(x) = (1-x)\gamma\mathbf{I} + x\mathbf{Q} - \frac{x}{\mu}(\mathbf{Q} - \rho\mathbf{\Lambda} - \mathbf{A} - \gamma\mathbf{I})(x-1)(\mathbf{Q} + \rho\mathbf{\Lambda} + \mathbf{A}),$$

тогда равенство (5) переписется в следующем виде:

$$\mathbf{G}(0, x) = \mathbf{P}(0, 0)\mathbf{A}(x)\mathbf{B}^{-1}(x).$$

Производящая функция  $\mathbf{G}(0, x)$  определена для всех значений  $x \in [0, 1]$ , но матрица  $\mathbf{B}(x)$  вырождена при  $x = x_v$ , где  $x_v$  – корни уравнения  $|\mathbf{B}(x)| = 0$ , принадлежащие рассматриваемому интервалу  $[0, 1]$ .

Обратную матрицу  $\mathbf{B}^{-1}(x)$  запишем в виде  $\mathbf{B}^{-1}(x) = \frac{1}{|\mathbf{B}(x)|}\mathbf{D}^T(x)$ , где элементами  $D(x)_{n_1 n_2}$  матрицы  $\mathbf{D}(x)$  являются алгебраические дополнения к элементам  $B(x)_{n_1 n_2}$  матрицы  $\mathbf{B}(x)$ .

Из равенства нулю определителя  $|\mathbf{B}(x_v)| = 0$  следует, что компоненты вектора  $\mathbf{P}(0, 0)$  удовлетворяют однородной системе линейных алгебраических уравнений  $\mathbf{P}(0, 0)\mathbf{A}(x_v)\mathbf{B}^T(x_v) = 0$ .

Эта система определяет значения компонент вектора  $\mathbf{P}(0, 0)$  с точностью до мультипликативной постоянной, значение которой определяется условием нормировки. Таким образом, удалось найти выражения (5) для производящих функций  $\mathbf{G}(k, x)$ .

### Исследование методом асимптотического анализа

Нахождение явного выражения (5) для производящей функции в математических моделях RQ-систем является исключительной ситуацией, поэтому требуется разработка других методов анализа таких моделей. Наиболее плодотворным в этом направлении является метод асимптотического анализа [3], изложение которого выполним для данной модели. Это позволит выявить эффективность разработанного метода путем сравнения асимптотических результатов с допредельными, а также позволит сравнить их с асимптотическими результатами, полученными для адаптивной RQ-системы.

Систему (4) модифицируем следующим образом:

$$\begin{cases} \mathbf{H}(0,u)(\mathbf{Q} - \rho\mathbf{A} - \mathbf{A} - \gamma\mathbf{I}) + \mathbf{H}(1,u)\mu = -\gamma\mathbf{P}(0,0), \\ \mathbf{H}(0,u)(\rho\mathbf{A} + \mathbf{A} + e^{-ju}\gamma\mathbf{I}) + \mathbf{H}(1,u)(\mathbf{Q} + (e^{ju} - 1)(\rho\mathbf{A} + \mathbf{A}) - \mu\mathbf{I}) = e^{-ju}\gamma\mathbf{P}(0,0), \end{cases} \quad (6)$$

где  $\rho$  – параметр, который используется для получения предельного условия большой загрузки RQ-системы, а функция  $\mathbf{H}(k,u) = \mathbf{G}(k, e^{ju}) = \mathbf{G}(k, x)$ .

Систему (6) будем решать методом асимптотического анализа в условии большой загрузки  $\rho \uparrow S$  [4], где  $S_1$  – пропускная способность динамической RQ-системы. Обозначив  $\varepsilon = S - \rho$ , будем полагать также  $\varepsilon \rightarrow 0$  и систему (6) будем решать при выполнении этого условия. В системе (6) выполним замены

$$\rho = S - \varepsilon, \quad u = \varepsilon w, \quad \mathbf{H}_k(u) = \mathbf{F}_k(w, \varepsilon), \quad \mathbf{P}(0,0) = \varepsilon\mathbf{\Pi}(\varepsilon)$$

и перепишем в виде

$$\begin{cases} \mathbf{F}_0(w, \varepsilon)(\mathbf{Q} - (S - \varepsilon)\mathbf{A} - \mathbf{A} - \gamma\mathbf{I}) + \mathbf{F}_1(w, \varepsilon)\mu = -\gamma\varepsilon\mathbf{\Pi}(\varepsilon), \\ \mathbf{F}_0(w, \varepsilon)((S - \varepsilon)\mathbf{A} + e^{-j\varepsilon w}\gamma\mathbf{I} + \mathbf{A}) + \\ + \mathbf{F}_1(w, \varepsilon)(\mathbf{Q} + (e^{j\varepsilon w} - 1)(S - \varepsilon)\mathbf{A} - \mu\mathbf{I} - \mathbf{A}(1 - e^{j\varepsilon w})) = e^{-j\varepsilon w}\gamma\varepsilon\mathbf{\Pi}(\varepsilon). \end{cases} \quad (7)$$

**Теорема.** Значение  $S$  пропускной способности динамической RQ-системы с входящим MAP-потокм заявок равно значению корня уравнения

$$\gamma R_0 \mathbf{E} - R_1 (S\mathbf{A} + \mathbf{A}) \mathbf{E} = 0, \quad (8)$$

где вектор-строка  $\mathbf{R}_k$  – совместное распределение вероятностей состояний прибора и значений цепи Маркова, управляющей входящим MAP-потокм, которое определяется равенствами

$$\begin{aligned} \mathbf{R}_0(S) &= \mu \mathbf{R} \{ (\mu + \gamma) \mathbf{I} + S\mathbf{A} + \mathbf{A} - \mathbf{Q} \}^{-1}, \\ \mathbf{R}_1(S) &= \mathbf{R} \{ \mathbf{I} - \mu [ (\mu + \gamma) \mathbf{I} + S\mathbf{A} + \mathbf{A} - \mathbf{Q} ]^{-1} \}. \end{aligned} \quad (9)$$

*Доказательство.* Доказательство этой теоремы выполним в два этапа.

**Этап 1.** Обозначим  $\lim_{\varepsilon \rightarrow 0} \mathbf{F}_k(w, \varepsilon) = \mathbf{F}_k(w)$ , выполнив в (7) указанный предельный переход, получим систему

$$\begin{cases} \mathbf{F}_0(w)(\mathbf{Q} - S\mathbf{A} - \mathbf{A} - \gamma\mathbf{I}) + \mathbf{F}_1(w)\mu = 0, \\ \mathbf{F}_0(w)(S\mathbf{A} + \mathbf{A} + \gamma\mathbf{I}) + \mathbf{F}_1(w)(\mathbf{Q} - \mu\mathbf{I}) = 0. \end{cases} \quad (10)$$

Решение  $\mathbf{F}_k(w)$  системы (10) будем искать в виде

$$\mathbf{F}_k(w) = \mathbf{R}_k \Phi(w), \quad (11)$$

где функция  $\Phi(w)$  на бесконечности равна нулю, а  $\mathbf{R}_k$  – распределение вероятностей состояний прибора, определяемое системой

$$\begin{cases} \mathbf{R}_0(\mathbf{Q} - S\mathbf{A} - \mathbf{A} - \gamma\mathbf{I}) + \mathbf{R}_1\mu = 0, \\ \mathbf{R}_0(S\mathbf{A} + \mathbf{A} + \gamma\mathbf{I}) + \mathbf{R}_1(\mathbf{Q} - \mu\mathbf{I}) = 0. \end{cases}$$

Нетрудно показать, что  $(\mathbf{R}_0 + \mathbf{R}_1)\mathbf{Q} = 0$ , т.е.  $\mathbf{RQ} = 0$ , где  $\mathbf{R} = \mathbf{R}_0 + \mathbf{R}_1$  и удовлетворяет условию нормировки  $\mathbf{RE} = 1$ , тогда  $\mathbf{R}_0$  и  $\mathbf{R}_1$ , зависящие от  $S$ , определяются равенствами

$$\begin{aligned}\mathbf{R}_0(S) &= \mu \mathbf{R} \{(\mu + \gamma)\mathbf{I} + S\mathbf{\Lambda} + \mathbf{A} - \mathbf{Q}\}^{-1}, \\ \mathbf{R}_1(S) &= \mathbf{R} \left\{ \mathbf{I} - \mu [(\mu + \gamma)\mathbf{I} + S\mathbf{\Lambda} + \mathbf{A} - \mathbf{Q}]^{-1} \right\},\end{aligned}$$

совпадающими с (9).

**Этап 2.** Переписав (7) в виде

$$\begin{cases} \mathbf{F}_0(w, \varepsilon)(\mathbf{Q} - S\mathbf{\Lambda} - \mathbf{A} - \gamma\mathbf{I} + \varepsilon\mathbf{\Lambda}) + \mathbf{F}_1(w, \varepsilon)\mu = -\gamma\varepsilon\Pi(\varepsilon) + O(\varepsilon^2), \\ \mathbf{F}_0(w, \varepsilon)(S\mathbf{\Lambda} + \gamma\mathbf{I} - \varepsilon(\mathbf{\Lambda} + jw\gamma\mathbf{I})) + \mathbf{A} + \\ + \mathbf{F}_1(w, \varepsilon)(\mathbf{Q} - \mu\mathbf{I} + j\varepsilon w(S\mathbf{\Lambda} + \mathbf{A})) = \gamma\varepsilon\Pi(\varepsilon) + O(\varepsilon^2), \end{cases}$$

просуммируем по  $k$  и  $n$  все уравнения полученной системы, получим уравнение

$$\mathbf{F}_0(w, \varepsilon)\varepsilon jw\gamma\mathbf{E} - \mathbf{F}_1(w, \varepsilon)j\varepsilon w(S\mathbf{\Lambda} + \mathbf{A})\mathbf{E} = 0,$$

из которого имеем

$$\mathbf{F}_0(w, \varepsilon)\gamma\mathbf{E} - \mathbf{F}_1(w, \varepsilon)(S\mathbf{\Lambda} + \mathbf{A})\mathbf{E} = 0.$$

Используя (11), получим равенство

$$\mathbf{R}_0\Phi(w)\gamma\mathbf{E} - \mathbf{R}_1\Phi(w)(S\mathbf{\Lambda} + \mathbf{A})\mathbf{E} = 0,$$

откуда получаем выражение

$$\gamma\mathbf{R}_0\mathbf{E} - \mathbf{R}_1(S\mathbf{\Lambda} + \mathbf{A})\mathbf{E} = 0,$$

которое совпадает с (8) и определяет значение пропускной способности  $S$  динамической RQ-системы.

В результате исследования динамической RQ-системы с входящим МАР-потокм получено уравнение (8) для нахождения пропускной способности  $S$ .

#### Литература

1. Любина Т. В., Назаров А. А. Исследование марковской динамической RQ-системы с конфликтами заявок // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2010. № 3 (12). С. 73–84.
2. Любина Т. В., Назаров А. А. Исследование немарковской динамической RQ-системы с конфликтами заявок // Вестник Кемеровского государственного университета. 2012. № 1 (49). С. 38–44.
3. Назаров А. А., Кузнецов Д. Ю. Адаптивные сети случайного доступа. Томск: ТПУ, 2002. 256 с.
4. Любина Т. В., Назаров А. А. Исследование динамической и адаптивной RQ-систем с входящим ММРР-потокм заявок // Вестник Томского государственного университета. 2013. № 3 (24). С. 104–112.

# СТАЦИОНАРНОЕ ФУНКЦИОНИРОВАНИЕ СЕТЕЙ МАССОВОГО ОБСЛУЖИВАНИЯ С ГРУППОВЫМ ПОСТУПЛЕНИЕМ ОТРИЦАТЕЛЬНЫХ ЗАЯВОК

*Ю. В. Малинковский*

*Гомельский государственный университет им. Ф. Скорины,  
Гомель, Беларусь*

*Национальный исследовательский  
Томский государственный университет, Томск, Россия*

**1. Изолированный узел.** В систему массового обслуживания с единственным экспоненциальным прибором с интенсивностью обслуживания  $\mu$  поступает  $T+1$  независимых пуассоновских потоков заявок: поток обычных, требующих обслуживания положительных заявок с интенсивностью  $\lambda^+$ , и  $T$  потоков отрицательных заявок с интенсивностями  $\lambda_1^-, \dots, \lambda_T^-$ . Поступающая отрицательная заявка  $l$ -го потока мгновенно вычеркивает, уничтожает ровно  $l$  положительных заявок, если такое количество положительных заявок в системе имеется, и уничтожает все заявки в системе, если их число меньше  $l, l = \overline{1, T}$ . После этого она мгновенно пропадает вместе с вычеркнутыми положительными заявками, не оказывая в дальнейшем на функционирование системы никакого влияния. Состоянием системы в момент  $t$  назовем количество положительных заявок  $n(t)$  в этот момент времени. Очевидно,  $n(t)$  – цепь Маркова с непрерывным временем и пространством состояний  $Z_+$ . Ее стационарное распределение  $\{p(n), n = 0, 1, \dots\}$ , если оно существует, удовлетворяет системе уравнений равновесия для так называемых вертикальных сечений графа переходов цепи

$$\begin{aligned} \lambda^+ p(n) = & (\mu + \lambda_1^- + \dots + \lambda_T^-) p(n+1) + (\lambda_2^- + \dots + \lambda_T^-) p(n+2) + \\ & + (\lambda_3^- + \dots + \lambda_T^-) p(n+3) + \dots + \lambda_T^- p(n+T), \quad n = 0, 1, \dots \end{aligned} \quad (1)$$

Это однородное линейное разностное уравнение порядка  $T$ . Частное решение (1) ищем в виде  $p(n) = z^n$ . Подставляя его в (1), получим характеристическое уравнение

$$g(z) = \sum_{l=1}^T z^l \sum_{s=1}^T \lambda_s^- + \mu z - \lambda^+ = \sum_{s=1}^T \lambda_s^- \sum_{l=1}^s z^l + \mu z - \lambda^+ = 0. \quad (2)$$

Докажем достаточность условия

$$\rho = \frac{\lambda^+}{\mu + \sum_{t=1}^T t \lambda_t^-} < 1 \quad (3)$$

для эргодичности процесса  $n(t)$ . Сначала используем теорему Декарта [1]. В (2) ровно одна переменная знака при переходе от  $\mu$  к  $\lambda^+$ . Значит, (2) имеет

ровно один положительный корень. При этом  $g(0) = -\lambda^+ < 0$  и, в силу (3),  $g(1) = \sum_{t=1}^T t\lambda_t^- + \mu - \lambda^+ > 0$ . Поэтому этот корень  $z_0 \in (0,1)$ . Поэтому уравнение равновесия (1) имеет решение  $p(n) = z_0^n$ , причем из условия нормировки  $C = 1 - z_0$ , т. е. совпадает с геометрическим распределением

$$p(n) = (1 - z_0)z_0^n, \quad n = 0, 1, \dots \quad (4)$$

Применим эргодическую теорему Фостера [2]. Для того чтобы неприводимая консервативная регулярная цепь Маркова с непрерывным временем была эргодична, необходимо и достаточно, чтобы система уравнений равновесия имела ненулевое решение такое, что  $\sum_{n=0}^{\infty} |p(n)| < \infty$ . При выполнении условия (3) уравнение (2), как мы видели, имеет корень  $z_0 \in (0,1)$ , причем (4) – частное решение системы уравнений равновесия (1). А ряд  $\sum_{n=0}^{\infty} |p(n)|$  сходится как сумма членов геометрической прогрессии со знаменателем, меньшим единицы. Неприводимость и консервативность цепи  $n(t)$  очевидны, а регулярность следует из того, что интенсивность выхода  $q(n)$  процесса  $n(t)$  из состояния  $n$  ограничена [3]. Значит, условие (3) достаточно для эргодичности  $n(t)$ , а при его выполнении эргодическое распределение имеет форму (4).

**Лемма 1.** 1. Если  $\rho > 1$ , то все корни характеристического уравнения (2) по модулю строго больше единицы.

2. Если  $\rho = 1$ , то на окружности  $|z| = 1$  характеристическое уравнение (2) имеет единственный корень  $|z| = 1$ , причем простой, а остальные корни (2) по модулю строго больше единицы.

Доказательство пункта 1 основано на применении теоремы Руше, пункта 2 – на применении модификации теоремы Руше, предложенной В. И. Клименок [4] и весьма полезной для исследования условий эргодичности процессов теории массового обслуживания.

**Лемма 2.** Если  $Q_j(n)$  – некоторые многочлены от переменной  $n$ ,  $0 \leq \phi_j < 2\pi$ , причем  $\phi_j \neq \phi_m$  при  $j \neq m$ ,  $j, m = 1, \dots, k$ , и  $\sum_{j=1}^k Q_j(n)e^{in\phi_j} \rightarrow 0$  при  $n \rightarrow \infty$ , то  $Q_j(n) \equiv 0$  ( $j = 1, \dots, k$ ).

*Доказательство.* Обозначим  $A = \max(\deg Q_1(n), \dots, \deg Q_k(n))$ , тогда

$n^A \sum_{j=1}^k \frac{Q_j(n)}{n^A} e^{in\varphi_j} \rightarrow 0$  при  $n \rightarrow \infty$ . Для  $j = 1, \dots, k$  найдется постоянная  $C_j$  та-

кая, что  $\frac{Q_j(n)}{n^A} \rightarrow C_j$ . Тогда  $n^A (\sum_j C_j e^{in\varphi_j} + o(1)) \rightarrow 0$ , а так как  $A \geq 0$ , то

тем более  $\sum_j C_j e^{in\varphi_j} + o(1) \rightarrow 0$ , откуда  $\varepsilon_n = \sum_{j=1}^k C_j e^{in\varphi_j} \rightarrow 0$  при  $n \rightarrow \infty$ . Сле-

довательно, для любого  $\varepsilon > 0$  найдется номер  $N = N_\varepsilon$  такой, что для всех  $n \geq N$  будет  $|\varepsilon_n| < \varepsilon$ . По введенному выше обозначению для  $\varepsilon_n$  имеем

$$\begin{aligned} \varepsilon_N &= \sum_{j=1}^k C_j e^{iN\varphi_j}, \\ \varepsilon_{N+1} &= \sum_{j=1}^k C_j e^{i(N+1)\varphi_j}, \\ &\dots\dots\dots \\ \varepsilon_{N+k-1} &= \sum_{j=1}^k C_j e^{i(N+k-1)\varphi_j}. \end{aligned}$$

Модуль определителя  $\Delta$  этой системы линейных относительно  $C_1, C_2, \dots, C_k$  уравнений совпадает с модулем определителя Вандермонда, т.е.  $|\Delta| = \prod_{j>k} |e^{i\varphi_j} - e^{i\varphi_k}| \neq 0$ . Раскладывая определитель  $\Delta_m$ , отличающийся

от  $\Delta$  заменой его  $m$ -го столбца столбцом свободных членов системы, по этому столбцу получим  $|\Delta_m| < C\varepsilon$ , где  $C$  – некоторая константа, не зависящая от  $\varepsilon$ . По правилу Крамера

$$|C_m| = \frac{|\Delta_m|}{|\Delta|} < \frac{C}{|\Delta|} \varepsilon,$$

откуда, в силу произвольности  $\varepsilon > 0$ , следует, что  $C_m = 0, m = 1, \dots, k$ . Следовательно,  $Q_j(n) \rightarrow 0$ , а так как  $Q_j(n)$  – многочлен, то  $Q_j(n) \equiv 0, j = 1, \dots, k$ .

**Теорема 1.** *Цепь Маркова  $n(t)$  регулярна. Для ее эргодичности необходимо и достаточно, чтобы выполнялось неравенство (3). В этом случае стационарное распределение цепи имеет форму геометрического распределения (4).*

Доказательство основывается на виде общего решения линейного однородного разностного уравнения (1) и леммах 1 и 2.

**2. Сеть массового обслуживания.** Рассматривается сеть массового обслуживания, состоящая из  $N$  однолинейных экспоненциальных узлов с интенсивностью обслуживания прибором  $i$ -го узла  $\mu_i, i = \overline{1, N}$ , в которую извне поступает  $(T+1)N$  независимых пуассоновских потоков заявок. При

этом в  $i$ -й узел поступает поток обычных (положительных) заявок с параметром  $\Lambda_i$  и  $T$  потоков отрицательных заявок с параметром  $\lambda_{il}$  ( $i = \overline{1, N}, l = \overline{1, T}$ ). Отрицательные заявки не требуют обслуживания, а отрицательная заявка  $l$ -го потока при поступлении в узел мгновенно вычеркивает, уничтожает ровно  $l$  положительных заявок, если такое количество положительных заявок в нем имеется, и уничтожает все заявки в узле, если их число меньше  $l$  ( $i = \overline{1, N}, l = \overline{1, T}$ ). Сразу после этого она мгновенно пропадает вместе с вычеркнутыми положительными заявками, не оказывая в дальнейшем на функционирование сети никакого влияния. Положительная заявка, обслуженная в  $i$ -м узле, мгновенно и независимо от других заявок переходит в  $j$ -й узел как положительная с вероятностью  $p_{ij}^+$ , как отрицательная заявка  $l$ -го потока отрицательных заявок с вероятностью  $p_{ijl}^-$  и покидает сеть с вероятностью  $p_{i0}$  ( $i, j = \overline{1, N}, l = \overline{1, T}$ ). При этом 
$$\sum_{j=1}^N \left( p_{ij}^+ + \sum_{l=1}^T p_{ijl}^- \right) + p_{i0} = 1 \quad (i = \overline{1, N}).$$
 Число мест для ожидания положительных заявок в узлах сети не ограничено. Для определенности считаем, что заявки обслуживаются в узлах в порядке поступления.

Состояние сети будем описывать случайным вектором  $\vec{n}(t) = (n_1(t), \dots, n_N(t))$ , где  $n_i(t)$  – число положительных заявок в  $i$ -м узле в момент времени  $t$ . В силу простейших предпосылок относительно входящих потоков и распределения длительностей обслуживания  $\vec{n}(t)$  – многомерная цепь Маркова с непрерывным временем и пространством состояний  $X = Z_+^N$ , где  $Z_+ = \{0, 1, \dots\}$ . Наша цель – установить условия эргодичности процесса  $\vec{n}(t)$  и найти его стационарное распределение.

Рассмотрим изолированный от сети узел, считая, что в него поступают пуассоновские потоки заявок с теми же интенсивностями, с которыми в этот узел поступают заявки соответствующих потоков в сети (которые не являются пуассоновскими). Ко всем обозначениям для изолированного узла добавим в качестве первого индекса индекс  $i$ , соответствующий номеру узла. Характеристическое уравнение (2) с подставленным в него корнем  $z_{i0}$  принимает форму тождества

$$\sum_{l=1}^T \sum_{s=1}^T \lambda_{is}^- z_{i0}^l + \mu_i z_{i0} - \lambda_i^+ = 0. \quad (5)$$

Как следует из результатов раздела 1, при выполнении условия эргодичности

$$\rho_i = \frac{\lambda_i^+}{\mu_i + \sum_{t=1}^T t \lambda_{it}^-} < 1, \quad i = \overline{1, N}, \quad (6)$$

стационарное распределение изолированного узла имеет форму

$$p_i(n_i) = (1 - z_{i0})z_{i0}^{n_i}, \quad n_i = 0, 1, \dots \quad (7)$$

Отсюда вероятность занятости прибора  $i$ -го узла в стационарном режиме равна  $z_{i0}$ . Значит, интенсивности потоков положительных и отрицательных заявок в рассматриваемой сети удовлетворяют следующей системе уравнений трафика:

$$\lambda_i^+ = \Lambda_i + \sum_{j=1}^N \mu_j z_{j0} p_{ji}^+, \quad i = \overline{1, N}, \quad (8)$$

$$\lambda_{il}^- = \lambda_{il} + \sum_{j=1}^N \mu_j z_{j0} p_{jil}^-, \quad i = \overline{1, N}, \quad l = \overline{1, T}. \quad (9)$$

С помощью теоремы непрерывности неявной функции и теоремы Брауэра о неподвижной точке можно доказать, что существует строго положительное решение системы уравнений трафика (8), (9).

Если стационарное распределение процесса  $p(\vec{n})$  существует, то оно удовлетворяет уравнениям глобального равновесия:

$$\begin{aligned} p(\vec{n}) \sum_{i=1}^N \left[ \Lambda_i + (\mu_i + \lambda_{i1} + \dots + \lambda_{iT}) I_{\{n_i \neq 0\}} \right] &= \sum_{i=1}^N \{ p(\vec{n} - \vec{e}_i) \Lambda_i I_{\{n_i \neq 0\}} + p(\vec{n} + \vec{e}_i) [\mu_i p_{i0} + \\ &+ \lambda_{i1} + (\lambda_{i2} + \dots + \lambda_{iT}) I_{\{n_i = 0\}}] + \sum_{l=2}^T p(\vec{n} - l\vec{e}_i) [\lambda_{il} + (\lambda_{il+1} + \dots + \lambda_{iT}) I_{\{n_i = 0\}}] + \\ &+ \sum_{j=1}^N [p(\vec{n} + \vec{e}_j - \vec{e}_i) \mu_j p_{ji}^+ I_{\{n_i \neq 0\}} + \sum_{l=1}^T p(\vec{n} + \vec{e}_j + l\vec{e}_i) \mu_j (p_{jil}^- + (p_{jil+1}^- + \dots + \\ &+ p_{jiT}^-) I_{\{n_i = 0\}}) + p(\vec{n} + \vec{e}_j) \mu_j (p_{jil}^- + p_{jil+1}^- + \dots + p_{jiT}^-) I_{\{n_i = 0\}}] \}, \quad \vec{n} \in Z_+^N. \end{aligned}$$

Здесь  $\vec{e}_i$  – единичный вектор  $i$ -го направления,  $I_A$  – индикатор события  $A$ , равный единице, если событие  $A$  происходит, и нулю, если происходит противоположное событие.

Основной результат имеет следующую форму.

**Теорема 2.** *Марковский процесс  $\vec{n}(t)$  регулярен, а если выполнены неравенства (6), то он эргодичен, а его стационарное распределение  $\{p(\vec{n}), \vec{n} \in Z_+^N\}$  задается как*

$$p(\vec{n}) = \prod_{i=1}^N p_i(n_i), \quad \vec{n} \in Z_+^N,$$

где  $p_i(n_i)$  и  $\rho_i$  определены равенствами (7) и (6) соответственно, а  $z_{i0}, i = \overline{1, N}$ , – корни уравнений (5), принадлежащие отрезку  $[0, 1]$ .

Регулярность очевидна. Достаточность условий (6) для эргодичности марковского процесса доказывается с помощью эргодической теоремы Фостера. Тот факт, что стационарное распределение, задаваемое в теореме 2, удовлетворяет уравнениям глобального равновесия, доказывается с по-

мощью обычной техники локального баланса. Рассмотренная модель сети обобщает классическую модель Геленбе [5].

#### Литература

1. Курош А. Г. Курс высшей алгебры. М.: Наука, 1971. 432 с.
2. Бочаров П. П., Печинкин А. В. Теория массового обслуживания. М.: РУДН, 1995. 529 с.
3. Гихман И. И., Скороход А. В. Введение в теорию случайных процессов. М.: Наука, 1977. 568 с.
4. Klimenok V. On the Modification of Rouche's Theorem for the Queueing Theory Problems // Queueing Systems. 2001. Vol. 38. P. 431–434.
5. Gelenbe E. Product-form Queueing Networks with Negative and Positive Customers // J. Appl. Prob. 1991. Vol. 28. P. 656–663.

## АНАЛИЗ МНОГОФАЗНОЙ СИСТЕМЫ ОБСЛУЖИВАНИЯ $SM/(GI/\infty)^K$ В УСЛОВИИ ВЫСОКОЙ ИНТЕНСИВНОСТИ ВХОДЯЩЕГО ПОТОКА

*А. Н. Мусеев, А. А. Назаров*

*Национальный исследовательский*

*Томский государственный университет, Томск, Россия*

**Постановка задачи.** Рассматривается многофазная система массового обслуживания с неограниченным числом приборов на каждой фазе и полумарковским входящим потоком. Система состоит из  $K$  фаз обслуживания. Время обслуживания на  $k$ -й фазе является случайной величиной с функцией распределения  $B_k(x)$ ,  $k = \overline{1, K}$ . Заявка входящего потока поступает на первую фазу. Затем, по окончании обслуживания на первой фазе, она переходит на вторую фазу и т.д. до  $K$ -й фазы, по окончании обслуживания на которой заявка покидает систему.

Входящий поток является высокоинтенсивным полумарковским, который задается полумарковской матрицей  $A(x)$  порядка  $M$ . Элементы  $A_{lm}$  этой матрицы определяются следующим образом [1]:

$$A_{lm}(x) = P \left\{ \xi_{n+1} = m, \tau_{n+1} < \frac{x}{N} \mid \xi_n = l \right\}. \quad (1)$$

Здесь  $\{\xi_n, \tau_n\}$  – стационарный двумерный марковский процесс с дискретным временем [2];  $N$  – большой по величине скалярный параметр, определяющий высокую интенсивность потока (в теоретических исследованиях будем полагать  $N \rightarrow \infty$ ). Величины  $\tau_n$  определяют длину интервалов между моментами  $t_{n-1}$  и  $t_n$  наступления последовательных событий в потоке.

Рассмотрим полумарковский процесс  $l(t)$ , определяемый следующим образом [3]:

$$l(t) = \xi_{n+1}, \quad t_n \leq t < t_{n+1}.$$

Обозначим  $z(t)$  – длину интервала времени от момента  $t$  до момента наступления следующего события в полумарковском потоке, тогда процесс  $\{l(t), z(t)\}$  является марковским. Для его совместного стационарного распределения вероятностей  $r_l(z) = P\{l(t) = l, z(t) < z/N\}$ ,  $l = \overline{1, M}$ , записанного в векторном виде  $\mathbf{r}(z) = \{r_1(z), \dots, r_M(z)\}$ , имеет место [1] дифференциальное уравнение

$$\mathbf{r}'(z) = \mathbf{r}'(0)[\mathbf{I} - \mathbf{A}(z)].$$

Здесь  $\mathbf{I}$  – единичная матрица.

В работе [1] показано, что  $\mathbf{r}'(0) = \lambda \mathbf{r}$ , где вектор-строка  $\mathbf{r}$  есть стационарное распределение состояний вложенной цепи Маркова  $\xi_n$ , который является решением уравнения Колмогорова  $\mathbf{r} = \mathbf{r} \cdot \mathbf{P}$ . Здесь  $\mathbf{P} = \lim_{z \rightarrow \infty} \mathbf{A}(z)$  есть стохастическая матрица, определяющая вероятности переходов вложенной цепи Маркова  $\xi_n$ . Величина  $\lambda$  определяется выражением

$$\lambda = \frac{1}{\mathbf{r} \mathbf{A} \mathbf{e}},$$

где  $\mathbf{A} = \int_0^{\infty} [\mathbf{P} - \mathbf{A}(x)] dx$ ;  $\mathbf{e}$  – единичный вектор-столбец.

Интенсивность полумарковского потока, заданного элементами (1), составляет  $N\lambda$  [1]. Таким образом, при  $N \rightarrow \infty$  она неограниченно возрастает. Обозначим  $i_k(t)$  – число заявок, находящихся на обслуживании на  $k$ -й фазе описанной системы ( $k = \overline{1, K}$ ) в момент времени  $t$ . Для вектора  $\mathbf{i}(t) = \{i_1(t), \dots, i_K(t)\}$  ставится задача нахождения многомерного стационарного распределения вероятностей числа заявок на фазах системы.

**Многофазное динамическое просеивание.** Прямое исследование случайного процесса  $\mathbf{i}(t)$  достаточно затруднено. Поэтому воспользуемся методом многофазного динамического просеивания, подробно описанного в [4]. Кратко опишем его.

Зафиксируем некоторый момент времени  $T$ . Обозначим через  $S_k(t)$  вероятность того, что заявка входящего потока, поступившая в момент времени  $t < T$ , в момент  $T$  будет обслуживаться на  $k$ -й фазе системы ( $k = \overline{1, K}$ ).

Через  $S_0(t) = 1 - \sum_{k=1}^K S_k(t)$  обозначим вероятность того, что указанная заявка покинет систему до момента времени  $T$ .

Определим  $K$  так называемых просеянных потоков событий. Будем считать, что заявка входящего потока, поступающая в систему в момент времени  $t$ , с вероятностью  $S_k(t)$  генерирует событие в  $k$ -м ( $k = \overline{1, K}$ ) просеянном потоке. С вероятностью  $S_0(t)$  указанная заявка не генерирует события ни в одном из потоков.

Пусть в начальный момент времени  $t_0 < T$  система пуста. Обозначим  $n_k(t)$  – число событий, наступивших в  $k$ -м просеянном потоке до момента времени  $t$ . Тогда для вектора  $\mathbf{n}(t) = \{n_1(t), \dots, n_K(t)\}$  в момент времени  $t = T$  имеем равенства

$$P\{\mathbf{i}(T) = \mathbf{i}\} = P\{\mathbf{n}(T) = \mathbf{i}\}$$

для любых значений  $\mathbf{i}$ , т.е. распределения вероятностей значений случайных процессов  $\mathbf{i}(t)$  и  $\mathbf{n}(t)$  в этот момент времени совпадают. Таким образом, получив выражение для распределения вероятностей многомерного процесса  $\mathbf{n}(t)$  и подставив  $t = T$ , получим распределение вероятностей значений исследуемого процесса  $\mathbf{i}(t)$  в момент времени  $T$ , который, вообще говоря, выбран произвольно.

В работе [4] получены следующие выражения для вероятностей  $S_k(t)$ :

$$S_k(t) = B_{k-1}^*(T-t) - B_k^*(T-t),$$

где  $B_k^*(x) = (B_1 * \dots * B_k)(x)$  есть свертка функций  $B_1(x), \dots, B_k(x)$ , если считать, что  $B_0^*(x) = 1$  и  $B_1^*(x) = B_1(x)$ .

**Уравнения Колмогорова.** Очевидно, что процесс  $\{\mathbf{n}(t), l(t), z(t)\}$  является марковским, и для его распределения вероятностей  $P(\mathbf{n}, l, z, t) = P\{\mathbf{n}(t) = \mathbf{n}, l(t) = l, z(t) < z/N\}$  можно записать систему дифференциальных уравнений Колмогорова:

$$\frac{1}{N} \frac{\partial P(\mathbf{n}, l, z, t)}{\partial t} = \frac{\partial P(\mathbf{n}, l, z, t)}{\partial z} - \frac{\partial P(\mathbf{n}, l, 0, t)}{\partial z} + \sum_{k=1}^K \sum_{m=1}^M \frac{\partial P(\mathbf{n} - \mathbf{e}_k, m, 0, t)}{\partial z} A_{ml}(z) S_k(t) + \sum_{m=1}^M \frac{\partial P(\mathbf{n}, m, 0, t)}{\partial z} A_{ml}(z) S_0(t),$$

для всех значений  $\mathbf{n}$ ,  $l = \overline{1, M}$  и  $z > 0$ . Здесь  $\mathbf{e}_k$  – вектор, все компоненты которого равны нулю за исключением  $k$ -й, которая равна единице.

Для частичных характеристических функций

$$H(\mathbf{u}, l, z, t) = \sum_{n_1=0}^{\infty} \dots \sum_{n_K=0}^{\infty} e^{ju_1 n_1 + \dots + ju_K n_K} P(\mathbf{n}, l, z, t), \quad l = \overline{1, M},$$

в векторном виде это уравнение можно записать следующим образом:

$$\frac{1}{N} \frac{\partial \mathbf{H}(\mathbf{u}, z, t)}{\partial t} = \frac{\partial \mathbf{H}(\mathbf{u}, z, t)}{\partial z} + \frac{\partial \mathbf{H}(\mathbf{u}, 0, t)}{\partial z} \left\{ \mathbf{A}(z) \left[ 1 + \sum_{k=1}^K (e^{ju_k} - 1) S_k(t) \right] - \mathbf{I} \right\} \quad (2)$$

при начальном условии

$$\mathbf{H}(\mathbf{u}, z, t_0) = \mathbf{r}(z). \quad (3)$$

Здесь  $j = \sqrt{-1}$ ,  $\mathbf{H}(\mathbf{u}, z, t) = \{H(\mathbf{u}, 1, z, t), \dots, H(\mathbf{u}, M, z, t)\}$ .

Задачу (2)–(3) будем решать асимптотически при  $N \rightarrow \infty$ , т.е. в условии высокой интенсивности входящего потока.

**Асимптотика первого порядка.** В (2)–(3) выполним следующие замены:

$$\frac{1}{N} = \varepsilon, \quad \mathbf{u} = \varepsilon \mathbf{w}, \quad \mathbf{H}(\mathbf{u}, z, t) = \mathbf{F}_1(\mathbf{w}, z, t, \varepsilon),$$

получим задачу

$$\begin{cases} \varepsilon \frac{\partial \mathbf{F}_1(\mathbf{w}, z, t, \varepsilon)}{\partial t} = \frac{\partial \mathbf{F}_1(\mathbf{w}, z, t, \varepsilon)}{\partial z} + \frac{\partial \mathbf{F}_1(\mathbf{w}, 0, t, \varepsilon)}{\partial z} \left\{ \mathbf{A}(z) \left[ 1 + \sum_{k=1}^K (e^{j\varepsilon w_k} - 1) S_k(t) \right] - \mathbf{I} \right\}, \\ \mathbf{F}_1(\mathbf{w}, z, t_0, \varepsilon) = \mathbf{r}(z). \end{cases} \quad (4)$$

Относительно асимптотического решения  $\mathbf{F}_1(\mathbf{w}, z, t) = \lim_{\varepsilon \rightarrow 0} \mathbf{F}_1(\mathbf{w}, z, t, \varepsilon)$  этой задачи имеет место следующее утверждение, которое здесь приводим без доказательства.

**Теорема 1.** Асимптотическое решение  $\mathbf{F}_1(\mathbf{w}, z, t)$  задачи (4) имеет вид

$$\mathbf{F}_1(\mathbf{w}, z, t) = \mathbf{r}(z) \exp \left\{ \lambda \sum_{k=1}^K j w_k \int_{t_0}^t S_k(x) dx \right\}.$$

**Асимптотический анализ второго порядка.** Введем функцию  $\mathbf{H}_2(\mathbf{u}, z, t)$ , определяемую выражением

$$\mathbf{H}(\mathbf{u}, z, t) = \mathbf{H}_2(\mathbf{u}, z, t) \exp \left\{ N \lambda \sum_{k=1}^K j u_k \int_{t_0}^t S_k(x) dx \right\}. \quad (5)$$

Подставим это выражение в (2)–(3), получим

$$\begin{cases} \frac{1}{N} \frac{\partial \mathbf{H}_2(\mathbf{u}, z, t)}{\partial t} + \mathbf{H}_2(\mathbf{u}, z, t) \lambda \sum_{k=1}^K j u_k S_k(t) = \frac{\partial \mathbf{H}_2(\mathbf{u}, z, t)}{\partial z} + \\ + \frac{\partial \mathbf{H}_2(\mathbf{u}, 0, t)}{\partial z} \left\{ \mathbf{A}(z) \left[ 1 + \sum_{k=1}^K (e^{j u_k} - 1) S_k(t) \right] - \mathbf{I} \right\}, \\ \mathbf{H}_2(\mathbf{u}, z, t_0) = \mathbf{r}(z). \end{cases}$$

Выполнив здесь замены

$$\frac{1}{N} = \varepsilon^2, \quad \mathbf{u} = \varepsilon \mathbf{w}, \quad \mathbf{H}_2(\mathbf{u}, z, t) = \mathbf{F}_2(\mathbf{w}, z, t, \varepsilon), \quad (6)$$

получим задачу

$$\begin{cases} \varepsilon^2 \frac{\partial \mathbf{F}_2(\mathbf{w}, z, t, \varepsilon)}{\partial t} + \mathbf{F}_2(\mathbf{w}, z, t, \varepsilon) \lambda \sum_{k=1}^K j \varepsilon w_k S_k(t) = \frac{\partial \mathbf{F}_2(\mathbf{w}, z, t, \varepsilon)}{\partial z} + \\ + \frac{\partial \mathbf{F}_2(\mathbf{w}, 0, t, \varepsilon)}{\partial z} \left\{ \mathbf{A}(z) \left[ 1 + \sum_{k=1}^K (e^{j \varepsilon w_k} - 1) S_k(t) \right] - \mathbf{I} \right\}, \\ \mathbf{F}_2(\mathbf{w}, z, t_0, \varepsilon) = \mathbf{r}(z). \end{cases} \quad (7)$$

Относительно асимптотического решения  $\mathbf{F}_2(\mathbf{w}, z, t) = \lim_{\varepsilon \rightarrow 0} \mathbf{F}_2(\mathbf{w}, z, t, \varepsilon)$  этой задачи имеет место следующее утверждение, доказательство которого в силу ограниченности объема опустим.

**Теорема 2.** Асимптотическое решение  $\mathbf{F}_2(\mathbf{w}, z, t)$  задачи (7) имеет вид

$$\mathbf{F}_2(\mathbf{w}, z, t) = \mathbf{r}(z) \exp \left\{ \lambda \sum_{k=1}^K \frac{(jw_k)^2}{2} \int_{t_0}^t S_k(x) dx + \kappa \sum_{k=1}^K \sum_{v=1}^K \frac{jw_k jw_v}{2} \int_{t_0}^t S_k(x) S_v(x) dx \right\}, \quad (8)$$

где

$$\kappa = 2\mathbf{f}'_0 \mathbf{e},$$

вектор-строка  $\mathbf{f}'_0$  удовлетворяет системе линейных алгебраических уравнений

$$\begin{cases} \mathbf{f}'_0 [\mathbf{I} - \mathbf{P}] = \lambda [\mathbf{rP} - \mathbf{r}(\infty)], \\ \mathbf{f}'_0 \mathbf{A} \mathbf{e} = \frac{\lambda^2 a_2}{2} - 1, \end{cases}$$

$$a_2 = \mathbf{r} \mathbf{A}_2 \mathbf{e}, \quad \mathbf{A}_2 = \int_0^{\infty} x^2 d\mathbf{A}(x).$$

**Стационарное распределение вероятностей числа заявок в системе SM/(GI/ $\infty$ )<sup>K</sup>.** Выполняя в (8) замены, обратные к (6), и подставляя полученное выражение в (5), получаем следующее выражение для частичной характеристической функции  $\mathbf{H}(\mathbf{u}, z, t)$ :

$$\begin{aligned} \mathbf{H}(\mathbf{u}, z, t) = \mathbf{r}(z) \exp \left\{ N\lambda \sum_{k=1}^K ju_k \int_{t_0}^t S_k(x) dx + \right. \\ \left. + N\lambda \sum_{k=1}^K \frac{(ju_k)^2}{2} \int_{t_0}^t S_k(x) dx + N\kappa \sum_{k=1}^K \sum_{v=1}^K \frac{ju_k ju_v}{2} \int_{t_0}^t S_k(x) S_v(x) dx \right\}. \end{aligned}$$

Полагая здесь  $z \rightarrow \infty$  и выполняя переход к характеристической функции  $h(\mathbf{u})$  стационарного распределения вероятностей значений исследуемого многомерного процесса  $\mathbf{i}(t)$  числа заявок на фазах системы, при  $t = T$  и полагая  $t_0 \rightarrow -\infty$ , а  $T = 0$ , получаем следующее выражение:

$$h(\mathbf{u}) = \exp \left\{ N\lambda \mathbf{juSe} + \frac{1}{2} \mathbf{ju} [N\lambda \mathbf{S} + N\kappa \mathbf{V}] \mathbf{ju}^T \right\}, \quad (9)$$

где  $\mathbf{S}$  – диагональная матрица с элементами  $S_k = \int_{-\infty}^0 S_k(t) dt$  на главной диа-

гонали,  $\mathbf{V}$  – матрица, состоящая из элементов  $V_{kv} = \int_{-\infty}^0 S_k(t) S_v(t) dt$ ,

$k, v = \overline{1, K}$ .

**Заключение.** Итак, в работе получено выражение (9) для характеристической функции стационарного распределения числа заявок на фазах многофазной системы массового обслуживания с входящим полумарковским потоком в условии его высокой интенсивности. Полученное асим-

птотическое распределение является многомерным нормальным с вектором математических ожиданий  $N\lambda Se$  и матрицей ковариаций  $M[\lambda S + \kappa V]$ . На основе проведенных численных экспериментов была определена область применимости аппроксимации (9) – установлено, что данная гауссовская аппроксимация обеспечивает расстояние Колмогорова менее 0,05 при значениях  $N \geq 10$ .

*Результаты получены в рамках выполнения госзадания Минобрнауки России № 1.511.2014/К.*

#### Литература

1. Моисеев А. Н., Назаров А. А. Асимптотический анализ высокоинтенсивного полумарковского потока событий // Доклады ТУСУРа. 2013. № 3 (29). С. 109–115.
2. Королюк В. С. Стохастические модели систем. Киев: Наук. думка, 1989. 208 с.
3. Moiseev A., Nazarov A. Asymptotic Analysis of the Infinite-Server Queueing System with High-Rate Semi-Markov Arrivals // Proc. of the IEEE International Congress on Ultra Modern Telecommunications and Control Systems (ICUMT 2014), Oct. 6–8, 2014. St. Petersburg: IEEE, 2014. P. 507–513.
4. Моисеев А. Н., Назаров А. А. Асимптотический анализ многофазной системы массового обслуживания с высокоинтенсивным рекуррентным входящим потоком // Автометрия. 2014. Т. 50, № 2. С. 67–76.

## ИССЛЕДОВАНИЕ G-СЕТЕЙ С МНОГОЛИНЕЙНЫМИ СИСТЕМАМИ И ДОХОДАМИ

*В. В. Науменко, М. А. Матальцкий*

*Гродненский государственный университет им. Я. Купалы,  
Гродно, Беларусь*

**1. Введение.** При построении и исследования математических моделей информационно-телекоммуникационных сетей (ИТС) необходимо, чтобы такие модели учитывали возможное влияние различных дестабилизирующих факторов: внезапные сбои, попадание вирусов, потеря передаваемых или обрабатываемых данных. Для учета подобных факторов Э. Геленбе были предложены G-сети [1], в которых рассматриваются дополнительные пуассоновские потоки отрицательных заявок. При поступлении в систему сети отрицательная заявка уничтожает одну положительную заявку, если таковая имеется в наличии в данной системе. Заметим, что при попадании компьютерных вирусов в ИТС из-за потери или искажения информации она несет некоторые расходы (убытки). Их учет возможно осуществить, применив в качестве модели сеть массового обслуживания (МО) с доходами (НМ-сеть) [2] с положительными и отрицательными заявками. В работе [3] исследовалась такая сеть в переходном режиме. В ней при переходе положительной заявки из одной системы обслуживания (СМО) в другую последняя получает некоторый случайный доход, а доход первой СМО уменьшается соответственно на эту величину. Отрицательная заявка,

поступающая в некоторую систему сети, в которой имеется по крайней мере одна положительная заявка, мгновенно уничтожает одну из них и наносит убыток этой СМО. После этого отрицательная заявка сразу же покидает сеть или уничтожается, если в данной СМО не было заявок. Доходами от переходов в ИТС могут, к примеру, являться онлайн-платежи и комиссия с них, денежные переводы и т.п. пользователей систем интернет-банкинга и других схожих систем. Заявками при этом являются запросы интернет-пользователей, которые пользуются услугами онлайн-платежей, а отрицательными заявками могут быть вирусы (или программы, нарушающие функционирование системы интернет-платежей) в таких ИТС или, к примеру, в случае DDoS-атаки на ИТС такого рода. Полученные результаты могут быть использованы при прогнозировании доходов в ИТС с учетом попадания в них вирусов. В данной работе проведен нестационарный анализ G-сети с доходами, но с многолинейными СМО. Рассмотрен случай, когда интенсивности входящих потоков положительных и отрицательных заявок и их обслуживания, вероятности переходов заявок между СМО зависят от времени. Предполагается, что доходы от переходов между состояниями сети являются случайными величинами (СВ) с заданными средними значениями.

**2. Нахождение ожидаемых доходов.** Рассмотрим открытую G-сеть МО с  $n$  многолинейными СМО. В СМО  $S_i$  извне поступает пуассоновский поток положительных заявок с интенсивностью  $\lambda_{0i}^+(t)$  и пуассоновский поток отрицательных заявок с интенсивностью  $\lambda_{0i}^-(t)$ ,  $i = \overline{1, n}$ . Все поступающие в сеть потоки заявок независимые. Пусть система  $S_i$  содержит  $m_i$  идентичных линий обслуживания, в каждой из которых длительности обслуживания положительных заявок в СМО  $S_i$  в момент времени  $t$  распределены по экспоненциальному закону с параметром  $\mu_i(t)$ ,  $i = \overline{1, n}$ . Положительная заявка, обслуженная в СМО  $S_i$  в момент времени  $t$ , с вероятностью  $p_{ij}^+(t)$  направляется в СМО  $S_j$  как положительная заявка, с вероятностью  $p_{ij}^-(t)$  — как отрицательная заявка, и с вероятностью  $p_{i0}(t) = 1 - \sum_{j=1}^n (p_{ij}^+(t) + p_{ij}^-(t))$  заявка уходит из сети во внешнюю среду,  $i, j = \overline{1, n}$ . Под состоянием сети в момент времени  $t$  будем понимать вектор  $k(t) = (k, t) = (k_1(t), k_2(t), \dots, k_n(t))$ , где  $k_i(t)$  — число заявок в момент времени  $t$  в системе  $S_i$ ,  $i = \overline{1, n}$ .

Рассмотрим динамику изменения доходов некоторой системы  $S_i$  сети. Обозначим через  $V_i(t)$  ее доход в момент времени  $t$ . Пусть в начальный момент времени доход системы равен  $V_i(0) = v_{i0}$ . Доход этой СМО в момент времени  $t + \Delta t$  можно представить в виде

$$V_i(t + \Delta t) = V_i(t) + \Delta V_i(t, \Delta t), \quad (1)$$

где  $\Delta V_i(t, \Delta t)$  – изменение дохода системы  $S_i$  на интервале времени  $[t, t + \Delta t)$ ,  $i = \overline{1, n}$ . Для нахождения этой величины выпишем условные вероятности событий, которые могут произойти за время  $\Delta t$ , и изменения доходов этой СМО, связанные с данными событиями.

1. С вероятностью  $\lambda_{0i}^+(t)\Delta t + o(\Delta t)$  в момент времени  $t$  в систему  $S_i$  из внешней среды поступит положительная заявка, которая принесет ей доход в размере  $r_{0i}$ , где  $r_{0i}$  – СВ, математическое ожидание (м. о.) которой равно  $M\{r_{0i}\} = a_{0i}$ ,  $i = \overline{1, n}$ .

2. С вероятностью  $\lambda_{0i}^-(t)\Delta t + o(\Delta t)$  в систему  $S_i$  в момент времени  $t$  из внешней среды поступит отрицательная заявка, которая принесет ей доход (убыток) в размере  $-\bar{r}_{0i}$ , где  $\bar{r}_{0i}$  – СВ с м.о.  $M\{\bar{r}_{0i}\} = \bar{a}_{0i}$ ,  $i = \overline{1, n}$ .

3. Если в момент времени  $t$  в системе  $S_i$  находится  $k_i(t)$  положительных заявок, то с вероятностью  $\mu_i(t)\min(k_i(t), m_i)p_{i0}(t)\Delta t + o(\Delta t)$  положительная заявка уйдет из сети во внешнюю среду, при этом общий доход СМО  $S_i$  уменьшится на величину, равную  $-R_{i0}$ , где  $M\{R_{i0}\} = b_{i0}$ ,  $i = \overline{1, n}$ .

4. Если в момент времени  $t$  в системе  $S_i$  находится положительная заявка, то после окончания ее обслуживания в СМО  $S_i$  она направляется в СМО  $S_j$  снова как положительная заявка с вероятностью  $\mu_i(t)\min(k_i(t), m_i)p_{ij}^+(t)\Delta t + o(\Delta t)$ ,  $i, j = \overline{1, n}$ ,  $i \neq j$ . При таком переходе доход системы  $S_i$  уменьшится на величину  $r_{ij}$ , а доход системы  $S_j$  увеличится на эту величину, где  $M\{r_{ij}\} = a_{ij}$ ,  $i, j = \overline{1, n}$ ,  $i \neq j$ .

5. С вероятностью  $\mu_i(t)\min(k_i(t), m_i)p_{ij}^-(t)\Delta t + o(\Delta t)$  положительная заявка, обслуженная в СМО  $S_i$ , в момент времени  $t$  направляется в СМО  $S_j$  как отрицательная заявка  $i, j = \overline{1, n}$ ,  $i \neq j$ . При таком переходе доход системы  $S_i$  уменьшится на величину  $\bar{r}_{ij}$ , а доход системы  $S_j$  не изменится, где  $M\{\bar{r}_{ij}\} = \bar{a}_{ij}$ ,  $i, j = \overline{1, n}$ ,  $i \neq j$ .

6. С вероятностью  $1 - \sum_{j=1}^n [(\lambda_{0j}^+(t) + \lambda_{0j}^-(t) + \mu_j(t))\min(k_j(t), m_j)]\Delta t + o(\Delta t)$  на отрезке времени  $[t, t + \Delta t)$  изменения состояния системы  $S_i$  не произойдет, в данном случае суммарный доход системы  $S_i$  может увеличиться (уменьшиться) на величину  $r_i\Delta t$ , где  $M\{r_i\} = c_i$ ,  $i = \overline{1, n}$ .

Аналогично, как в [3], для математического ожидания можно получить следующую приближенную формулу:

$$\begin{aligned}
M\{\Delta V_i(t, \Delta t)\} = & \left[ a_{0i} \lambda_{0i}^+(t) - \bar{a}_{0i} \lambda_{0i}^-(t) + c_i - \right. \\
& - \mu_i(t) \min(N_i(t), m_i) \left( b_{i0} p_{i0}(t) + \sum_{j=1}^n a_{ij} p_{ij}^+(t) \right) + \sum_{\substack{j=1 \\ j \neq i}}^n \mu_j(t) \min(N_j(t), m_j) a_{ji} p_{ji}^+(t) - \\
& \left. - \mu_i(t) \min(N_i(t), m_i) \sum_{j=1}^n \bar{a}_{ij} p_{ij}^-(t) \right] \Delta t + o(\Delta t). \quad (2)
\end{aligned}$$

Рассмотрим частные случаи функционирования описанной выше сети и ситуации, в которых они могут быть использованы при моделировании.

В случае сетевой атаки (типа DDoS) [4], основанной на небезграничности ресурсов атакуемой службы на компьютеры (сервера) ИТС, на которые организуется масса запросов, с которыми они заведомо не могут справиться и будут вынуждены отказать в обслуживании либо заставить ждать неприемлемо долго, могут возникнуть ситуации, что на некоторых промежутках времени в ряде СМО сети выполняется условие  $k_i(t) > m_i > 0$ , т.е.  $\min(N_i(t), m_i) = m_i$ ,  $i \in X$ , где  $X$  – множество номеров СМО, для которых выполняется данное условие. Пусть также  $X_i$  – множество номеров СМО, связанных со СМО с номером  $i$ ,  $i \in X$ . В этой ситуации из (2) получим

$$\begin{aligned}
M\{\Delta V_i(t, \Delta t)\} = & \left[ a_{0i} \lambda_{0i}^+(t) - \bar{a}_{0i} \lambda_{0i}^-(t) + c_i - \mu_i(t) m_i \left( b_{i0} p_{i0}(t) + \sum_{j \in X_i} a_{ij} p_{ij}^+(t) \right) + \right. \\
& \left. + \sum_{\substack{j \in X_i \\ j \neq i}} a_{ji} \mu_j(t) m_j p_{ji}^+(t) - \mu_i(t) m_i \sum_{j \in X_i} \bar{a}_{ij} p_{ij}^-(t) \right] \Delta t + o(\Delta t), \quad i = \overline{1, X}. \quad (3)
\end{aligned}$$

В другой ситуации при функционировании компьютеров, зараженных вредоносной троянской программой, она по команде активируется через некоторое случайное время и уничтожает большое число запросов и пакетов. В этом случае при моделировании можно предположить, что на некоторых промежутках времени в некоторых системах сети в среднем не наблюдается очередей и выполняются условия  $\min(N_i(t), m_i) = N_i(t)$ ,  $i = \overline{1, n}$ ,  $i \in Y$ , где  $Y$  – множество номеров СМО, для которых выполняется условие, описанное выше. Пусть также  $Y_i$  – множество номеров систем, связанных со СМО с номером  $i$ ,  $i \in Y$ . Тогда из (2) следует, что

$$\begin{aligned}
M\{\Delta V_i(t, \Delta t)\} = & \left[ a_{0i} \lambda_{0i}^+(t) - \bar{a}_{0i} \lambda_{0i}^-(t) + c_i - \mu_i(t) N_i(t) \left( b_{i0} p_{i0}(t) + \sum_{j \in Y_i} a_{ij} p_{ij}^+(t) \right) + \right. \\
& \left. + \sum_{\substack{j \in Y_i \\ j \neq i}} a_{ji} \mu_j(t) N_j(t) p_{ji}^+(t) - \mu_i(t) N_i(t) \sum_{j \in Y_i} \bar{a}_{ij} p_{ij}^-(t) \right] \Delta t + o(\Delta t), \quad i = \overline{1, Y}. \quad (4)
\end{aligned}$$

Введем обозначение  $v_i(t) = M\{V_i(t)\}$ ,  $i = \overline{1, n}$ . Из (1) получаем

$$v_i(t + \Delta t) = v_i(t) + M\{\Delta V_i(t, \Delta t)\} = v_i(t) + \left[ a_{0i} \lambda_{0i}^+(t) - \bar{a}_{0i} \lambda_{0i}^-(t) + c_i - \right.$$

$$\begin{aligned}
& - \left( b_{i0} p_{i0}(t) + \sum_{\substack{j=1 \\ j \neq i}}^n a_{ij} p_{ij}^+(t) \right) \mu_i(t) \min(N_i(t), m_i) + \sum_{\substack{j=1 \\ j \neq i}}^n a_{ji} p_{ji}^+(t) \mu_j(t) \min(N_j(t), m_j) - \\
& \quad - \mu_i(t) \min(N_i(t), m_i) \sum_{j=1}^n \bar{a}_{ij} p_{ij}^-(t) \Big] \Delta t + o(\Delta t), \quad i = \overline{1, n}.
\end{aligned}$$

Далее, переходя к пределу при  $\Delta t \rightarrow 0$ , получим неоднородные линейные ОДУ первого порядка для доходов систем сети:

$$\begin{aligned}
\frac{dv_i(t)}{dt} &= -\mu_i(t) \min(N_i(t), m_i) \left( b_{i0} p_{i0}(t) + \sum_{j=1}^n (a_{ij} p_{ij}^+(t) + \bar{a}_{ij} p_{ij}^-(t)) \right) + \\
&+ \sum_{\substack{j=1 \\ j \neq i}}^n a_{ji} p_{ji}^+(t) \mu_j(t) \min(N_j(t), m_j) + a_{0i} \lambda_{0i}^+(t) - \bar{a}_{0i} \lambda_{0i}^-(t) + c_i, \quad i = \overline{1, n}.
\end{aligned}$$

Задав начальные условия  $v_i(0) = v_{i0}$ ,  $i = \overline{1, n}$ , можно найти ожидаемые доходы систем сети. Таким образом,  $v_i(t) = v_{i0}(0) + \int_0^t f_i(\tau) d\tau$ , где

$$\begin{aligned}
f(\tau) &= -\mu_i(\tau) \min(N_i(\tau), m_i) \left( b_{i0} p_{i0}(\tau) + \sum_{j=1}^n (a_{ij} p_{ij}^+(\tau) + \bar{a}_{ij} p_{ij}^-(\tau)) \right) + \\
&+ \sum_{\substack{j=1 \\ j \neq i}}^n a_{ji} p_{ji}^+(\tau) \mu_j(\tau) \min(N_j(\tau), m_j) + a_{0i} \lambda_{0i}^+(\tau) - \bar{a}_{0i} \lambda_{0i}^-(\tau) + c_i, \quad i = \overline{1, n}.
\end{aligned}$$

Для двух частных случаев, когда справедливы формулы (3) и (4), получаем соответственно, что

$$\begin{aligned}
v_i(t) &= v_{i0}(0) + \int_0^t \left[ -m_i \mu_i(\tau) \left( b_{i0} p_{i0}(\tau) + \sum_{j \in X_i} (a_{ij} p_{ij}^+(\tau) + \bar{a}_{ij} p_{ij}^-(\tau)) \right) + \right. \\
&\quad \left. + \sum_{\substack{j \in X_i \\ j \neq i}} a_{ji} \mu_j(\tau) m_j p_{ji}^+(\tau) + a_{0i} \lambda_{0i}^+(\tau) - \bar{a}_{0i} \lambda_{0i}^-(\tau) \right] d\tau + c_i t, \quad i = \overline{1, X}, \\
v_i(t) &= v_{i0}(0) + \int_0^t \left[ -\mu_i(\tau) \left( b_{i0} p_{i0}(\tau) + \sum_{j \in Y_i} (a_{ij} p_{ij}^+(\tau) + \bar{a}_{ij} p_{ij}^-(\tau)) \right) N_i(\tau) + \right. \\
&\quad \left. + \sum_{\substack{j \in Y_i \\ j \neq i}} a_{ji} \mu_j(\tau) N_j(\tau) p_{ji}^+(\tau) + a_{0i} \lambda_{0i}^+(\tau) - \bar{a}_{0i} \lambda_{0i}^-(\tau) \right] d\tau + c_i t, \quad i = \overline{1, Y}.
\end{aligned}$$

Поскольку в сеть поступают пуассоновские потоки положительных заявок с интенсивностью  $\lambda_{0i}^+(t)$  и отрицательных заявок с интенсивностью  $\lambda_{0i}^-(t)$ , то можно показать, что среднее число заявок в СМО сети удовлетворяет системе дифференциальных уравнений:

$$\begin{aligned}
\frac{dN_i(t)}{dt} &= \sum_{\substack{j=1 \\ j \neq i}}^n \mu_j(t) (p_{ji}^+(t) + p_{ji}^-(t)) \min(N_j(t), m_j) - \\
&- \mu_i(t) \min(N_i(t), m_i) + \lambda_{0i}^+(t) + \lambda_{0i}^-(t), \quad i = \overline{1, n}.
\end{aligned} \tag{5}$$

Это система линейных неоднородных дифференциальных уравнений с разрывными правыми частями. Решать ее нужно путем разбиения фазового пространства на ряд областей и нахождения решения в каждой из них. Систему (5) можно решить, например, используя средства системы компьютерной математики Maple или Mathematica [5].

#### Литература

1. Gelenbe E. Product form queueing networks with negative and positive customers// Journal of Applied Probability. 1991. Vol. 28. P. 656–663.
2. Матальцкий М. А. О некоторых результатах анализа и оптимизации марковских сетей с доходами и их применении. Автоматика и телемеханика. 2009. № 10. С. 97–113.
3. Науменко В. В., Матальцкий М. А. Анализ марковских сетей с доходами, положительными и отрицательными заявками // Информатика. 2014. № 1. С. 5–14.
4. Pervasive Technology Labs at Indiana University Types of DDoS Attacks. Distributed Denial of Service Attacks (DDoS) – Resources, Pervasive Technology Labs at Indiana University. Advanced Networking Management Lab (ANML), December 3, 2009, Archived from the original on 2010-09-14, retrieved December 11, 2013.
5. Матросов А. В. Maple 6. Решение задач высшей математики и механики. СПб.: БХВ-Петербург, 2001. 526 с.

## ИССЛЕДОВАНИЕ СИСТЕМЫ МАССОВОГО ОБСЛУЖИВАНИЯ $GI|GI|_{\infty}$ С ДВУМЯ ТИПАМИ ЗАЯВОК

***Е. В. Панкратова***

*Национальный исследовательский  
Томский государственный университет, Томск, Россия*

Рассмотрим систему массового обслуживания (СМО) с неограниченным числом обслуживающих приборов. На вход системы поступает рекуррентный поток заявок двух типов, который определяется функцией распределения  $A(x)$  – длин интервалов между моментами наступления событий потока. Пришедшая заявка с вероятностью  $p_1$  определяется как заявка первого типа, а с вероятностью  $p_2$  – второго. Обслуживание заявок в зависимости от их типа осуществляется на приборах с функциями распределения времени обслуживания произвольного вида  $B_1(x)$  и  $B_2(x)$ .

Поставим задачу исследования двумерного случайного процесса  $\{l_1(t), l_2(t)\}$ , характеризующего число занятых приборов в системе в момент времени  $t$ . Так как входящий поток не является пуассоновским, то исследуемый процесс немарковский. Определим нестационарный трехмерный марковский процесс  $\{z(t), r_1(t), r_2(t)\}$ . Здесь  $z(t)$  – случайная величина, равная длине интервала от момента  $t$  до момента наступления следующего события в рекуррентном потоке;  $\{r_1(t), r_2(t)\}$  – число событий двумерного просеянного потока, наступивших до момента времени  $t$ .

$S_i(t) = P\{\tau_k^{(i)} > T - t\} = 1 - B_i(T - t)$  – вероятность того, что заявка  $i$ -го типа ( $i = 1, 2$ ) к моменту времени  $T$  не закончила обслуживание и сформировалась заявка просеянного потока [1].

Для распределения вероятностей

$$P(z, r_1, r_2, t) = P\{z(t) < z, r_1(t) = r_1, r_2(t) = r_2\}$$

запишем систему дифференциальных уравнений Колмогорова:

$$\begin{aligned} \frac{\partial P(z, r_1, r_2, t)}{\partial t} = & \frac{\partial P(z, r_1, r_2, t)}{\partial z} + \frac{\partial P(0, r_1, r_2, t)}{\partial z} (A(z) - 1) + \\ & + \frac{\partial P(0, r_1 - 1, r_2, t)}{\partial z} p_1 S_1(t) A(z) + \frac{\partial P(0, r_1, r_2 - 1, t)}{\partial z} p_2 S_2(t) A(z) - \\ & - \frac{\partial P(0, r_1, r_2, t)}{\partial z} A(z) (p_1 S_1(t) + p_2 S_2(t)). \end{aligned}$$

Начальные условия для решения этой системы в момент времени  $t_0$  определим равенством

$$P(z, r_1, r_2, t_0) = \begin{cases} R(z), r_1 = r_2 = 0, \\ 0, r_1 > r_2 > 0, \end{cases}$$

$R(z)$  – стационарное распределение вероятностей процесса  $z(t)$ .

Обозначив

$$H(z, u_1, u_2, t) = \sum_{r_1=0}^{\infty} \sum_{r_2=0}^{\infty} e^{ju_1 r_1} e^{ju_2 r_2} P(z, r_1, r_2, t),$$

получим следующую задачу Коши:

$$\begin{aligned} \frac{\partial H(z, u_1, u_2, t)}{\partial t} = & \frac{\partial H(z, u_1, u_2, t)}{\partial z} + \frac{\partial H(0, u_1, u_2, t)}{\partial z} (A(z) - 1) + \\ & + \frac{\partial H(0, u_1, u_2, t)}{\partial z} A(z) (p_1 S_1(t) (e^{ju_1} - 1) + p_2 S_2(t) (e^{ju_2} - 1)), \end{aligned} \quad (1)$$

$$H(z, u_1, u_2, t_0) = R(z).$$

### Асимптотика первого порядка

Будем искать решение системы (1) в условии эквивалентного роста времени обслуживания на приборах [2], т.е. при  $b_1 \rightarrow \infty$ ,  $b_2 \rightarrow \infty$ , где

$$b_i = \int_0^{\infty} (1 - B_i(x)) dx, i = 1, 2.$$

Обозначим  $b_1 = \frac{1}{\varepsilon}$ ,  $b_2 = \frac{1}{q\varepsilon}$  и в уравнении (1) выполним замены

$$\begin{aligned} t\varepsilon = \tau, t_0\varepsilon = \tau_0, S_1(t) = \tilde{S}_1(\tau), S_2(t) = \tilde{S}_2(\tau), \\ u_1 = \varepsilon x_1, u_2 = \varepsilon x_2, H(z, u_1, u_2, t) = F_1(z, x_1, x_2, \tau, \varepsilon). \end{aligned} \quad (2)$$

Для  $F_1(z, x_1, x_2, \tau, \varepsilon)$  получим уравнение

$$\begin{aligned} \frac{\partial F_1(z, x_1, x_2, \tau, \varepsilon)}{\partial \tau} &= \frac{\partial F_1(z, x_1, x_2, \tau, \varepsilon)}{\partial z} + \frac{\partial F_1(0, x_1, x_2, \tau, \varepsilon)}{\partial z} (A(z) - 1) + \\ &+ \frac{\partial F_1(0, x_1, x_2, \tau, \varepsilon)}{\partial z} A(z) \left( p_1 \tilde{S}_1(\tau) (e^{j\varepsilon x_1} - 1) + p_2 \tilde{S}_2(\tau) (e^{j\varepsilon x_2} - 1) \right). \end{aligned} \quad (3)$$

**Теорема 1.** *Предельное (при  $\varepsilon \rightarrow 0$ ) значение  $F_1(z, x_1, x_2, \tau)$  решения  $F_1(z, x_1, x_2, \tau, \varepsilon)$  уравнения (3) имеет вид*

$$F_1(z, x_1, x_2, \tau) = R(z) \exp \left\{ j\lambda \left[ p_1 x_1 \int_{\tau_0}^{\tau} \tilde{S}_1(w) dw + p_2 x_2 \int_{\tau_0}^{\tau} \tilde{S}_2(w) dw \right] \right\}, \quad (4)$$

где параметр  $\lambda$  определяется выражением

$$\lambda = \frac{\partial R(0)}{\partial z}. \quad (5)$$

*Доказательство.*

В уравнении (3) выполним предельный переход при  $\varepsilon \rightarrow 0$

$$\frac{\partial F_1(z, x_1, x_2, \tau)}{\partial z} - \frac{\partial F_1(0, x_1, x_2, \tau)}{\partial z} (A(z) - 1) = 0.$$

Следовательно,  $F_1(z, x_1, x_2, \tau)$  можно определить в виде

$$F_1(z, x_1, x_2, \tau) = R(z) \Phi_1(x_1, x_2, \tau). \quad (6)$$

Разложим (3) в ряд по  $\varepsilon$ , разделим на  $\varepsilon$  и устремим  $\varepsilon \rightarrow 0$ . Тогда с учетом (6) при  $z \rightarrow \infty$  получаем дифференциальное уравнение для нахождения функции  $\Phi_1(x_1, x_2, \tau)$ :

$$\frac{\partial \Phi_1(x_1, x_2, \tau)}{\partial \tau} = j\lambda \Phi_1(x_1, x_2, \tau) [p_1 x_1 \tilde{S}_1(\tau) + p_2 x_2 \tilde{S}_2(\tau)].$$

С учетом (6) и начальных условий  $\Phi_1(x_1, x_2, \tau) = 1$  получим вид  $\Phi_1(x_1, x_2, \tau)$ :

$$\Phi_1(x_1, x_2, \tau) = \exp \left\{ j\lambda \left[ p_1 x_1 \int_{\tau_0}^{\tau} \tilde{S}_1(w) dw + p_2 x_2 \int_{\tau_0}^{\tau} \tilde{S}_2(w) dw \right] \right\},$$

подставив которое в (6), получим равенство (4).

В силу замен (2) и равенства (4) можно записать асимптотическое приближённое равенство

$$\begin{aligned} H(z, u_1, u_2, t) &= F_1(z, x_1, x_2, \tau, \varepsilon) \approx F_1(z, x_1, x_2, \tau) = \\ &= R(z) \exp \left\{ j\lambda \left[ p_1 u_1 \int_{t_0}^t S_1(w) dw + p_2 u_2 \int_{t_0}^t S_2(w) dw \right] \right\}. \end{aligned}$$

При  $t = T = 0$ ,  $t_0 \rightarrow -\infty$  определим функцию

$$\begin{aligned}
h_1(u_1, u_2) &= H(\infty, u_1, u_2, 0) = \\
&= \exp \left\{ j\lambda \left[ p_1 u_1 \int_{-\infty}^0 (1 - B_1(w)) dw + p_2 u_2 \int_{-\infty}^0 (1 - B_2(w)) dw \right] \right\} = \\
&= \exp \{ j\lambda [p_1 u_1 b_1 + p_2 u_2 b_2] \},
\end{aligned}$$

которую будем называть асимптотикой первого порядка для характеристической функции процесса  $\{l_1(t), l_2(t)\}$ .

### Асимптотика второго порядка

Решение  $H(z, u_1, u_2, t)$  уравнения (1) запишем в виде произведения

$$H(z, u_1, u_2, t) = H_2(z, u_1, u_2, t) \exp \left\{ j\lambda \left[ p_1 u_1 \int_{t_0}^t S_1(w) dw + p_2 u_2 \int_{t_0}^t S_2(w) dw \right] \right\}, \quad (7)$$

подставляя которое в (1), получим уравнение для  $H_2(z, u_1, u_2, t)$  в виде

$$\begin{aligned}
&\frac{\partial H_2(z, u_1, u_2, t)}{\partial t} + H_2(z, u_1, u_2, t) j\lambda (p_1 S_1(t) + p_2 S_2(t)) = \\
&= \frac{\partial H_2(z, u_1, u_2, t)}{\partial z} + \frac{\partial H_2(0, u_1, u_2, t)}{\partial z} (A(z) - 1) + \\
&+ \frac{\partial H_2(0, u_1, u_2, t)}{\partial z} A(z) [p_1 S_1(t)(e^{ju_1} - 1) + p_2 S_2(t)(e^{ju_2} - 1)], \\
&H_2(z, u_1, u_2, t_0) = R(z).
\end{aligned} \quad (8)$$

$$\text{Введем следующие замены: } b_1 = \frac{1}{\varepsilon^2}, \quad b_2 = \frac{1}{q\varepsilon^2}, \quad t\varepsilon^2 = \tau, \quad t_0\varepsilon^2 = \tau_0, \quad (9)$$

$$\begin{aligned}
S_1(t) &= \tilde{S}_1(\tau), \quad S_2(t) = \tilde{S}_2(\tau), \quad u_1 = \varepsilon x_1, \\
u_2 &= \varepsilon x_2, \quad H_2(z, u_1, u_2, t) = F_2(z, x_1, x_2, \tau, \varepsilon).
\end{aligned}$$

**Теорема 2.** *Предельное (при  $\varepsilon \rightarrow 0$ ) значение  $F_2(z, x_1, x_2, \tau)$  функции  $F_2(z, x_1, x_2, \tau, \varepsilon)$  имеет вид*

$$\begin{aligned}
F_2(z, x_1, x_2, \tau) &= R(z) \exp \left\{ j^2 \lambda \left( p_1 \int_{\tau_0}^{\tau} \frac{x_1^2}{2} \tilde{S}_1(w) dw + p_2 \int_{\tau_0}^{\tau} \frac{x_2^2}{2} \tilde{S}_2(w) dw \right) + \right. \\
&+ p_1 p_2 x_1 x_2 \int_{\tau_0}^{\tau} \tilde{S}_1(w) \tilde{S}_2(w) dw \left( \frac{\partial f_1(0)}{\partial z} + \frac{\partial f_2(0)}{\partial z} \right) + \\
&\left. + p_1^2 x_1^2 \frac{\partial f_1(0)}{\partial z} \int_{\tau_0}^{\tau} \tilde{S}_1^2(w) dw + p_2^2 x_2^2 \frac{\partial f_2(0)}{\partial z} \int_{\tau_0}^{\tau} \tilde{S}_2^2(w) dw \right\}, \quad (10)
\end{aligned}$$

где величина  $\lambda$  определяется равенством  $\lambda = \frac{\partial R(0)}{\partial z}$ , а функция  $f_i(z)$  удовлетворяет условию  $f_i(\infty) = 0$  и является решением уравнения

$$\lambda R(z) = \frac{\partial f_i(z)}{\partial z} + \frac{\partial f_i(0)}{\partial z} (A(z) - 1) + \frac{\partial R(0)}{\partial z} A(z), i = 1, 2. \quad (11)$$

*Доказательство.*

Представим функцию  $F_2(z, x_1, x_2, \tau, \varepsilon)$  в виде разложения

$$F_2(z, x_1, x_2, \tau, \varepsilon) = \Phi_2(z, x_1, x_2, \tau) \{R(z) + \\ + j\varepsilon(p_1 x_1 \tilde{S}_1(\tau) f_1(z) + p_2 x_2 \tilde{S}_2(\tau) f_2(z))\} + O(\varepsilon^2). \quad (12)$$

Можно показать, что полученные ниже результаты не зависят от выбора значения величин  $f_1(\infty)$  и  $f_2(\infty)$ , поэтому будем полагать их равными нулю.

Подставив (12) в (8) и учитывая, что  $\frac{\partial R(z)}{\partial z} + \frac{\partial R(0)}{\partial z} (A(z) - 1) = 0$ , устремим  $\varepsilon \rightarrow 0$  и получим дифференциальное уравнение

$$\lambda R(z) = \frac{\partial f_i(z)}{\partial z} + \frac{\partial f_i(0)}{\partial z} (A(z) - 1) + \frac{\partial R(0)}{\partial z} A(z),$$

совпадающее с уравнением (11).

Для нахождения функции  $\Phi_2(x_1, x_2, \tau)$  сделаем в уравнении (8) замены (9), разложим экспоненты в ряд Тейлора до  $O(\varepsilon^2)$ , разделим обе части полученного выражения на  $\varepsilon^2$  и выполним предельный переход при  $\varepsilon \rightarrow 0$  и  $z \rightarrow \infty$ . Тогда получим следующее равенство:

$$\frac{\partial \Phi_2(x_1, x_2, \tau)}{\partial \tau} = \Phi_2(x_1, x_2, \tau) \left\{ j^2 \lambda \left( p_1 \frac{x_1^2}{2} \tilde{S}_1(\tau) + p_2 \frac{x_2^2}{2} \tilde{S}_2(\tau) \right) + \right. \\ \left. + p_1 p_2 x_1 x_2 \tilde{S}_1(\tau) \tilde{S}_2(\tau) \left( \frac{\partial f_1(0)}{\partial z} + \frac{\partial f_2(0)}{\partial z} \right) + p_1^2 x_1^2 \frac{\partial f_1(0)}{\partial z} \tilde{S}_1^2(\tau) + p_2^2 x_2^2 \frac{\partial f_2(0)}{\partial z} \tilde{S}_2^2(\tau) \right\}.$$

Учитывая, что  $\Phi_2(x_1, x_2, \tau_0) = 1$ , получим выражение

$$\Phi_2(z, x_1, x_2, \tau) = \exp \left\{ j^2 \lambda \left( p_1 \int_{\tau_0}^{\tau} \frac{x_1^2}{2} \tilde{S}_1(w) dw + p_2 \int_{\tau_0}^{\tau} \frac{x_2^2}{2} \tilde{S}_2(w) dw \right) + \right. \\ \left. + p_1 p_2 x_1 x_2 \int_{\tau_0}^{\tau} \tilde{S}_1(w) \tilde{S}_2(w) dw \left( \frac{\partial f_1(0)}{\partial z} + \frac{\partial f_2(0)}{\partial z} \right) + \right. \\ \left. + p_1^2 x_1^2 \frac{\partial f_1(0)}{\partial z} \int_{\tau_0}^{\tau} \tilde{S}_1^2(w) dw + p_2^2 x_2^2 \frac{\partial f_2(0)}{\partial z} \int_{\tau_0}^{\tau} \tilde{S}_2^2(w) dw \right\},$$

подставляя которое в (12), получим выражение (10).

В силу замены (9), а также равенства (10) для функции  $H_2(z, u_1, u_2, t)$  можно записать приближённое (асимптотическое) равенство

$$H_2(z, u_1, u_2, t) = F_2(z, x_1, x_2, \tau, \varepsilon) \approx F_2(z, x_1, x_2, \tau) =$$

$$\begin{aligned}
&= R(z) \exp \left\{ j^2 \lambda \left( p_1 \frac{u_1^2}{2} \int_{\tau_0}^{\tau} S_1(w) dw + p_2 \frac{u_2^2}{2} \int_{\tau_0}^{\tau} S_2(w) dw \right) + \right. \\
&\quad + p_1 p_2 u_1 u_2 \int_{\tau_0}^{\tau} S_1(w) S_2(w) dw \left( \frac{\partial f_1(0)}{\partial z} + \frac{\partial f_2(0)}{\partial z} \right) + \\
&\quad \left. + p_1^2 u_1^2 \frac{\partial f_1(0)}{\partial z} \int_{\tau_0}^{\tau} S_1^2(w) dw + p_2^2 u_2^2 \frac{\partial f_2(0)}{\partial z} \int_{\tau_0}^{\tau} S_2^2(w) dw \right\}.
\end{aligned}$$

Обозначим

$$\begin{aligned}
\int_{-\infty}^0 S_1^2(w) dw &= \int_0^{\infty} (1 - B_1(z))^2 dz = \beta_1, \\
\int_{-\infty}^0 S_1(w) S_2(w) dw &= \int_0^{\infty} (1 - B_1(z))(1 - B_2(z)) dz = \beta_{12}.
\end{aligned}$$

Тогда при  $t = T = 0$  и  $t_0 \rightarrow -\infty$  для характеристической функции процесса  $\{r_1(T), r_2(T)\}$  можно записать равенство

$$\begin{aligned}
&Me^{j(u_1 r_1(T) + u_2 r_2(T))} = H(\infty, u_1, u_2, T) = h_2(u_1, u_2) = \\
&= \exp \left\{ j^2 \lambda \left( p_1 \frac{u_1^2}{2} b_1 + p_2 \frac{u_2^2}{2} b_2 \right) + p_1 p_2 u_1 u_2 \left( \frac{\partial f_1(0)}{\partial z} + \frac{\partial f_2(0)}{\partial z} \right) \beta_{12} + \right. \\
&\quad \left. + p_1^2 u_1^2 \frac{\partial f_1(0)}{\partial z} \beta_1 + p_2^2 u_2^2 \frac{\partial f_2(0)}{\partial z} \beta_2 \right\},
\end{aligned}$$

которое будем называть асимптотикой второго порядка характеристических функций числа занятых приборов первого и второго типа.

#### Литература

1. Pankratova E., Moiseeva S. Queueing System with Renewal Arrival Process and Two Types of Customers // Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT). IEEE. 2014. P. 514–517.
2. Pankratova E., Moiseeva S. Queueing System  $MAP|M|\infty$  with  $n$  Types of Customers // Proc. of the 13th Intern.Sci.Conf. ITMM 2014 named after A. F. Terpigov. Anzhero-Sudzhensk, 2014. P. 356–366.

## ВЫХОДЯЩИЙ ПОТОК ЗАЯВОК В СИСТЕМЕ С ПРОГУЛКАМИ

**С. В. Пауль**

*Национальный исследовательский*

*Томский государственный университет, Томск, Россия*

Телекоммуникационные системы, математическими моделями которых могут являться однолинейные системы массового обслуживания с

«прогулками» обслуживающего прибора, достаточно часто встречаются на практике [1]. В реальных системах под «прогулкой» понимается временное прекращение обслуживания для выполнения либо других работ прибором, либо его поломкой, либо его ремонтом [2].

Рассмотрим систему массового обслуживания с одним обслуживающим прибором и очередью с неограниченным числом мест для ожидания. В систему поступает простейший поток заявок с интенсивностью  $\lambda$ . Режим работы прибора состоит из двух последовательных интервалов. В течение первого интервала заявки потока обслуживаются на приборе случайное время, распределенное по экспоненциальному закону с параметром  $\mu$ . Если заявок в очереди нет к моменту начала этого интервала или прибор обслужил все заявки, которые на этом интервале находились в очереди, то прибор все равно остается в этом режиме, ожидая прихода заявок. От момента окончания этого интервала прибор уходит на «прогулку». Во время «прогулки», пришедшие в систему, заявки накапливаются в очереди и ждут, когда прибор вернется на обслуживание.

Продолжительности этих интервалов случайные и определяются функциями распределения  $T_1(x)$  и  $T_2(x)$  соответственно. Будем рассматривать системы с дообслуживанием заявок. Целью исследования является выходящий поток заявок, обслуженных системой.

Обозначим

$m(t)$  – число заявок, обслуженных в системе за время  $t$ ;

$i(t)$  – число заявок, находящихся в системе в момент времени  $t$ ;

$k(t)$  – состояние прибора: 1 – прибор в рабочем состоянии, 2 – прибор на «прогулке»;

$z(t)$  – остаточное время пребывания прибора в соответствующем состоянии.

Рассмотрим марковский процесс  $\{m(t), i(t), k(t), z(t)\}$  и для распределения вероятностей

$$P_k(m, i, z, t) = P\{m(t) = m, i(t) = i, k(t) = k, z(t) < z\}.$$

Составим прямую систему дифференциальных уравнений Колмогорова:

$$\left\{ \begin{array}{l} \frac{\partial P_1(m, i, z, t)}{\partial t} = \frac{\partial P_1(m, i, z, t)}{\partial z} - \frac{\partial P_1(m, i, 0, t)}{\partial z} - P_1(m, i, z, t)(\lambda + \mu) + \\ \quad + P_1(m, i - 1, z, t)\lambda + P_1(m - 1, i + 1, z, t)\mu + \frac{\partial P_2(m, i, 0, t)}{\partial z} T_1(z), \\ \frac{\partial P_2(m, i, z, t)}{\partial t} = \frac{\partial P_2(m, i, z, t)}{\partial z} - \frac{\partial P_2(m, i, 0, t)}{\partial z} - \\ \quad - P_2(m, i, z, t)\lambda + P_2(m, i - 1, z, t)\lambda + \frac{\partial P_1(m, i, 0, t)}{\partial z} T_2(z). \end{array} \right.$$

Введем частичные характеристические функции

$$H_k(u, x, z, t) = \sum_{i=0}^{\infty} \sum_{m=0}^{\infty} e^{jxi} e^{jum} P_k(m, i, z, t), \text{ для которых систему дифференциальных уравнений Колмогорова перепишем в виде}$$

$$\left\{ \begin{aligned} \frac{\partial H_1(u, x, z, t)}{\partial t} &= \\ &= \frac{\partial H_1(u, x, z, t)}{\partial z} - \frac{\partial H_1(u, x, 0, t)}{\partial z} + (\lambda(e^{jx} - 1) + \mu(e^{ju} e^{-jx} - 1))H_1(u, x, z, t) + \\ &+ \frac{\partial H_2(u, x, 0, t)}{\partial z} T_1(z) + \mu(1 - e^{-jx} e^{ju}) \sum_{m=0}^{\infty} P_1(m, 0, z, t), \\ \frac{\partial H_2(u, x, z, t)}{\partial z} &= \\ &= \frac{\partial H_2(u, x, z, t)}{\partial z} - \frac{\partial H_2(u, x, 0, t)}{\partial z} + \lambda H_2(u, x, z, t)(e^{jx} - 1) + \frac{\partial H_1(u, x, 0, t)}{\partial z} T_2(z). \end{aligned} \right. \quad (1)$$

$\sum_{m=0}^{\infty} P_1(m, 0, z, t)$  – вероятность того, что прибор в режиме обслуживания,

но в системе нет заявок.

Систему (1) будем решать методом асимптотического анализа в условиях растущего времени наблюдения за входящим потоком и большой загрузки системы. Введем параметр  $N \rightarrow \infty$  и обозначим

$$\varepsilon = \frac{1}{N}, \quad \tau = \varepsilon t, \quad u = \varepsilon w, \quad x = \varepsilon y, \quad \lambda = (1 - \varepsilon)S\mu,$$

$$H_k(u, x, z, t) = F_k(w, y, z, \tau, \varepsilon), \quad \sum_{m=0}^{\infty} P_1(m, 0, z, t) = \varepsilon \pi(z, \tau, \varepsilon), \quad (2)$$

получим

$$\left\{ \begin{aligned} \varepsilon \frac{\partial F_1(w, y, z, \tau, \varepsilon)}{\partial \tau} &= \frac{\partial F_1(w, y, z, \tau, \varepsilon)}{\partial z} - \frac{\partial F_1(w, y, 0, \tau, \varepsilon)}{\partial z} + \\ &+ ((1 - \varepsilon)S\mu(e^{j\varepsilon y} - 1) + \mu(e^{j\varepsilon w} e^{-j\varepsilon y} - 1))F_1(w, y, z, \tau, \varepsilon) + \\ &+ \frac{\partial F_2(w, y, 0, \tau, \varepsilon)}{\partial z} T_1(z) + \mu(1 - e^{-j\varepsilon y} e^{j\varepsilon w}) \varepsilon \pi_1(z, \tau, \varepsilon), \\ \frac{\partial F_2(w, y, z, \tau, \varepsilon)}{\partial z} &= \frac{\partial F_2(w, y, z, \tau, \varepsilon)}{\partial z} - \frac{\partial F_2(w, y, 0, \tau, \varepsilon)}{\partial z} + \\ &+ \lambda F_2(w, y, z, \tau, \varepsilon)(e^{j\varepsilon y} - 1) + \frac{\partial F_1(w, y, 0, \tau, \varepsilon)}{\partial z} T_2(z). \end{aligned} \right. \quad (3)$$

Устремим  $\varepsilon \rightarrow 0$ , получим

$$\begin{cases} \frac{\partial F_1(u, x, z, \tau)}{\partial z} - \frac{\partial F_1(u, x, 0, \tau)}{\partial z} + \frac{\partial F_2(u, x, 0, \tau)}{\partial z} T_1(z) = 0, \\ \frac{\partial F_2(u, x, z, \tau)}{\partial z} - \frac{\partial F_2(u, x, 0, \tau)}{\partial z} + \frac{\partial F_1(u, x, 0, \tau, \delta)}{\partial z} T_2(z) = 0. \end{cases}$$

Решение данной системы будем искать в виде

$$F_k(w, y, z, \tau) = R_k(z)\Phi(w, y, \tau).$$

Тогда

$$\begin{cases} R_1(z) = \int_0^z (R'_1(0) - R'_2(0)T_1(x))dx, \\ R_2(z) = \int_0^z (R'_2(0) - R'_1(0)T_2(x))dx. \end{cases}$$

Здесь  $\{R_1(z), R_2(z)\}$  – двумерное распределение состояния прибора и величины остаточного времени пребывания прибора в нем. Устремим  $z \rightarrow \infty$ , учитывая, что  $T(\infty) = 1$ , получим

$$\begin{cases} R_1(\infty) = \int_0^\infty (R'_1(0) - R'_2(0)T_1(x))dx, \\ R_2(\infty) = \int_0^\infty (R'_2(0) - R'_1(0)T_2(x))dx. \end{cases}$$

Для сходимости несобственного интеграла необходимо выполнение следующего условия:

$$R'_1(0) - R'_2(0)T_1(\infty) = 0,$$

откуда получаем

$$R'_1(0) = R'_2(0) = R'(0),$$

$$R_k(z) = R'(0) \int_0^z (1 - T_k(x))dx.$$

Обозначим  $\int_0^\infty (1 - T_k(x))dx = T_k$  – среднее время пребывания в соответствующем состоянии прибора, тогда, с одной стороны,

$$R_1(\infty) + R_2(\infty) = 1,$$

с другой –

$$R_1(\infty) + R_2(\infty) = R'(0) \int_0^\infty (1 - T_1(x))dx + \int_0^\infty (1 - T_2(x))dx = R'(0)(T_1 + T_2).$$

Получим

$$R'(0) = \frac{1}{(T_1 + T_2)}.$$

Тогда

$$\begin{cases} R_1(\infty) = \frac{1}{T_1 + T_2} \int_0^{\infty} (1 - T_1(x)) dx = \frac{T_1}{T_1 + T_2}, \\ R_2(\infty) = \int_0^{\infty} (R_2'(0) - R_1'(0)T_2(x)) dx = \frac{T_2}{T_1 + T_2}. \end{cases}$$

В систему (3) подставим следующее разложение:

$$F_k(w, y, z, \tau, \varepsilon) = \Phi(w, y, \tau)R_k(z) + O(\varepsilon), \quad (4)$$

$$\begin{cases} \varepsilon \frac{\partial \Phi(w, y, \tau)}{\partial \tau} R_1(z) = \Phi(w, y, \tau) \{ R_1'(z) - R'(0) + (j\varepsilon\mu((S-1)y + w))R_1(z) \} + \\ + \Phi(w, y, \tau)R'(0)T_1(z) + O(\varepsilon), \\ \varepsilon \frac{\partial \Phi(w, y, \tau)}{\partial \tau} R_2(z) = \Phi(w, y, \tau) \{ R_2'(z) - R'(0) + j\varepsilon y S \mu R_2(z) \} + \\ + \Phi(w, y, \tau)R'(0)T_2(z) + O(\varepsilon). \end{cases}$$

В последней системе устремим  $z \rightarrow \infty$

$$\begin{cases} \varepsilon \frac{\partial \Phi(w, y, \tau)}{\partial \tau} R_1(\infty) = \Phi(w, y, \tau) \{ -R'(0) + (j\varepsilon\mu((S-1)y + w))R_1(\infty) \} + \\ + \Phi(w, y, \tau)R'(0) + O(\varepsilon), \\ \varepsilon \frac{\partial \Phi(w, y, \tau)}{\partial \tau} R_2(\infty) = \Phi(w, y, \tau) \{ -R'(0) + j\varepsilon y S \mu R_2(\infty) \} + \\ + \Phi(w, y, \tau)R'(0) + O(\varepsilon). \end{cases} \quad (5)$$

Просуммируем уравнения системы (5) и разделим полученное равенство на  $\varepsilon$  и устремим  $\varepsilon \rightarrow 0$

$$\frac{\partial \Phi(w, y, \tau)}{\partial \tau} = \Phi(w, y, \tau) (j\mu((S-1)y + w))S + jyS\mu(1-S),$$

$$\frac{\partial \Phi(w, y, \tau)}{\partial \tau} = j\mu S w \Phi(w, y, \tau).$$

Из последнего уравнения получим

$$\Phi(w, y, \tau) = C \exp\{j\mu S w \tau\}.$$

Будем предполагать, что в начальный момент времени наблюдения за системой она будет функционировать в стационарном режиме. Таким образом, полагая в функции  $\Phi(w, y, \tau)$   $\tau=0$ , получим, что она является асимптотической характеристической функцией числа заявок в системе:

$$\Phi(0, y, 0) = \frac{S}{S - jy\{S + \Delta\}},$$

$$\Delta = -R_1 R_2 \mu \frac{T_1 T_2}{T_1 + T_2} + R_1^2 R_2 \mu \frac{T_2^{(2)}}{2T_2} + R_2^2 R_1 \mu \frac{T_1^{(2)}}{2T_1},$$

$T_k^{(2)}$  – второй начальный момент времени пребывания прибора в состоянии  $k$ . Следовательно,

$$C = \frac{S}{S - jy\{S + \Delta\}}.$$

Тогда

$$\Phi(w, y, \tau) = \frac{S}{S - jy\{S + \Delta\}} \exp\{j\mu S w \tau\}.$$

В последнем равенстве положим  $y=0$ , получим очевидное равенство

$$\Phi(w, 0, \tau) = \exp\{jw S \mu \tau\}.$$

Имеем

$$\begin{aligned} M\{e^{jum(t)}\} &= H_1(u, 0, t) \approx \Phi(w, 0, \tau) + O(\varepsilon) = \exp\{jw S \mu t\} + O(\varepsilon) = \\ &= \exp\left\{j \frac{u}{\varepsilon} \frac{\lambda}{(1-\varepsilon)S} S \varepsilon t\right\} + O(\varepsilon) \approx \exp\{ju \lambda t\}. \end{aligned}$$

Таким образом, получили первое приближение характеристической функции числа заявок, закончивших обслуживание в системе с «прогулками» в условии растущего времени наблюдения за потоком и большой загрузки системы, которое определяет среднее количество событий выходящего потока за определенный период. Проводя аналогичные вычисления, получено асимптотическое распределение вероятностей числа обслуженных заявок в системе за время  $t$ .

#### Литература

1. Печинкин А. В. Система массового обслуживания с ненадежным прибором в дискретном времени / А. В. Печинкин, И. А. Соколов // Информ. и её примен. 2011. Т. 5, вып. 4. С. 6–17.
2. Саксонов Е. А. Метод вычисления вероятностей состояний для однолинейной системы массового обслуживания с “прогулками” обслуживающего прибора // Автомат. и телемех. 1995. Вып. 1. С. 101–106.

## АСИМПТОТИЧЕСКИЙ АНАЛИЗ ЗАМКНУТОЙ СТРУКТУРЫ МАССОВОГО ОБСЛУЖИВАНИЯ С ЗАЯВКАМИ ДВУХ КЛАССОВ

*Т. В. Русилко*

*Гродненский государственный университет им. Янки Купалы,*

*Гродно, Беларусь*

Исследуем замкнутую структуру массового обслуживания, состоящую из  $n+1$  узлов  $S_0, S_1, \dots, S_n$ . Будем считать, что в ней циркулируют однотипные заявки, которые подразделяются на заявки класса 1, обладающие абсолютным приоритетом, и класса 2 – неприоритетные. Пусть число линий обслуживания в системе массового обслуживания  $S_i$  составляет  $m_i$ ,  $i = \overline{0, n}$ . Время обслуживания заявок класса  $c$  каждой из линий системы  $S_i$  распределено по показательному закону со средним значением  $\mu_{ic}^{-1}$ ,

$i = \overline{0, n}$ ,  $c = 1, 2$ . Заявки каждого класса выбираются на обслуживание в соответствии с дисциплиной FIFO. Заявка класса  $c$ , обслуживание которой в системе  $S_i$  закончено, с вероятностью  $p_{ij}^{(c)}$  переходит в очередь системы  $S_j$ ,  $i, j = \overline{0, n}$ ,  $c = 1, 2$ . Матрицы  $P^{(c)}$ ,  $c = 1, 2$ , являются стохастическими матрицами, соответствующими неприводимой марковской цепи:

$$0 \leq p_{ij}^{(c)} \leq 1, \sum_{j=0}^n p_{ij}^{(c)} = 1, c = 1, 2.$$

Кроме того, предположим, что при переходе между системами сети заявки могут изменять класс. Будем считать, что заявка класса  $c$  после обслуживания в системе  $S_i$  станет заявкой класса  $s$  с вероятностью  $p_i^{(cs)}$ . Таким образом, заявка класса  $c$ , обслуживание которой в системе  $S_i$  закончено, переходит в очередь системы  $S_j$  как заявка класса  $s$  с вероятностью  $p_{ij}^{(cs)} = p_{ij}^{(c)} \cdot p_i^{(cs)}$ ,  $i, j = \overline{0, n}$ ,  $c, s = 1, 2$ .

Будем считать, что число заявок в структуре не постоянно и определяется процессом гибели и размножения, изменяющим число заявок в системе  $S_0$ : генерирующим новые заявки класса  $c$  с интенсивностью  $\lambda_{0c}^+$  и уничтожающим существующие с интенсивностью  $\lambda_{0c}^-$ ,  $c = 1, 2$ . Предположим, что общее число заявок в структуре ограничено константой  $K$ . Состояние структуры определяется вектором

$$k(t) = (k_{01}(t), k_{02}(t), k_{11}(t), k_{12}(t), \dots, k_{n1}(t), k_{n2}(t)),$$

где  $k_{ic}(t)$  – число заявок класса  $c$  в системе  $S_i$  в момент времени  $t$ ,  $i = \overline{0, n}$ ,  $c = 1, 2$ . Вектор  $k(t)$  в силу описанного выше является  $(2n + 2)$ -мерным марковским случайным процессом с непрерывным временем и конечным числом состояний. Очевидно, что общее число обслуживаемых заявок в структуре в момент времени  $t$  случайно и составляет  $\sum_{i=0}^n \sum_{c=1}^2 k_{ic}(t) = K(t)$ .

Проведем асимптотический анализ марковского процесса  $k(t)$  при большом числе заявок [1, 2]. Предположим, что структура функционирует в условиях большой загрузки заявками, т.е. значение  $K(t)$  достаточно велико, но ограничено:  $0 \ll K(t) \leq K$ .

Введем в рассмотрение величину

$$\varepsilon_{i1}(k_{i1}, t) = \min(k_{i1}(t), m_i), \quad (1)$$

которая означает число линий обслуживания, занятых в системе  $S_i$  заявками класса 1 в момент времени  $t$ , когда в системе  $S_i$  находится  $k_{i1}$  заявок класса 1,  $i = \overline{1, n}$ . Величина

$$\varepsilon_{i_2}(k_{i_1}, k_{i_2}, t) = \begin{cases} k_{i_2}(t), k_{i_1}(t) + k_{i_2}(t) < m_i, \\ m_i - k_{i_1}(t), k_{i_1}(t) < m_i, k_{i_1}(t) + k_{i_2}(t) \geq m_i, \\ 0, k_{i_1}(t) \geq m_i; \end{cases} \quad (2)$$

означает число линий обслуживания, занятых в системе  $S_i$  заявками класса 2 в момент времени  $t$ , когда в системе  $S_i$  находится  $k_{i_1}$  заявок класса 1 и  $k_{i_2}$  заявок класса 2,  $i = \overline{1, n}$ .

Пусть  $I_{ic} - (2n + 2)$ -вектор, все компоненты которого равны нулю, за исключением  $ic$ -й, которая равна 1,  $i = \overline{0, n}$ ,  $c = 1, 2$ . Рассмотрим все возможные переходы в состояние  $k(t + \Delta t) = (k, t + \Delta t)$  процесса  $k(t)$  за время  $\Delta t$ . Формула полной вероятности позволяет записать систему разностных уравнений для вероятностей состояний  $P(k, t)$ . Затем, используя предельный переход при  $\Delta t \rightarrow 0$ , получаем систему разностно-дифференциальных уравнений Колмогорова для вероятностей состояний структуры массового обслуживания:

$$\begin{aligned} \frac{dP(k, t)}{dt} = & \sum_{c=1}^2 \lambda_{0c}^- k_{0c}(t) (P(k + I_{0c}, t) - P(k, t)) + \sum_{c=1}^2 \lambda_{0c}^- P(k + I_{0c}, t) + \\ & + \sum_{c=1}^2 \lambda_{0c}^+ \left( K - \sum_{i=0}^n \sum_{c=1}^2 k_{ic}(t) \right) (P(k - I_{0c}, t) - P(k, t)) + \sum_{c=1}^2 \lambda_{0c}^+ P(k - I_{0c}, t) \Delta t + \\ & + \sum_{i,j=0}^n \sum_{s=1}^2 \mu_{i_1} \varepsilon_{i_1}(k_{i_1}, t) p_{ij}^{(1s)} (P(k + I_{i_1} - I_{j_s}, t) - P(k, t)) + \\ & + \sum_{i,j=0}^n \sum_{s=1}^2 \mu_{i_1} (\varepsilon_{i_1}(k_{i_1} + 1, t) - \varepsilon_{i_1}(k_{i_1}, t)) p_{ij}^{(1s)} P(k + I_{i_1} - I_{j_s}, t) + \\ & + \sum_{i,j=0}^n \sum_{s=1}^2 \mu_{i_2} \varepsilon_{i_2}(k_{i_1}, k_{i_2}, t) p_{ij}^{(1s)} (P(k + I_{i_2} - I_{j_s}, t) - P(k, t)) + \\ & + \sum_{i,j=0}^n \sum_{s=1}^2 (\mu_{i_2} (\varepsilon_{i_2}(k_{i_1}, k_{i_2} + 1, t) - \varepsilon_{i_2}(k_{i_1}, k_{i_2}, t))) p_{ij}^{(1s)} P(k + I_{i_2} - I_{j_s}, t). \end{aligned} \quad (3)$$

Решение системы (3) в аналитическом виде является трудной задачей. Поэтому далее будем рассматривать асимптотический случай большого числа заявок в структуре массового обслуживания, т.е. положим, что  $K \gg 1$ . Чтобы найти распределение вероятностей случайного вектора  $k(t)$ , перейдем к относительным переменным и будем исследовать вектор  $\xi(t) = \left( \frac{k(t)}{K} \right)$ . Возможные значения этого вектора при фиксированном  $t$

принадлежат ограниченному замкнутому множеству

$$G = \left\{ x = (x_{ic}, i = \overline{0, n}, c = 1, 2) : \sum_{i=0}^n \sum_{c=1}^2 x_{ic} \leq 1 \right\},$$

в котором они располагаются в узлах  $(2n + 2)$ -мерной решетки на расстоянии  $\varepsilon = \frac{1}{K}$  друг от друга. При увеличении  $K$  «плотность заполнения» множества  $G$  возможными компонентами вектора  $\xi(t)$  увеличивается, и становится возможным считать, что он имеет непрерывное распределение с плотностью распределения вероятностей  $p(x, t)$ , которая удовлетворяет асимптотическому соотношению  $K^{2n+2}P(k, t) \xrightarrow{K \rightarrow \infty} p(x, t)$ . Воспользуемся следующей аппроксимацией функции  $P(k, t)$ :  $K^{2n+2}P(k, t) = K^{2n+2}P(xK, t) = p(x, t)$ ,  $x \in G$ .

Обозначим  $e_{ic} = \varepsilon l_{ic}$ ,  $l_i = \frac{m_i}{K}$ ,  $i = \overline{0, n}$ ,  $c = 1, 2$ . Переписывая систему уравнений (3) для плотности  $p(x, t)$ , получим

$$\begin{aligned} \frac{\partial p(x, t)}{\partial t} = & \sum_{c=1}^2 K \lambda_{0c}^- x_{0c}(t) (p(x + e_{0c}, t) - p(x, t)) + \sum_{c=1}^2 \lambda_{0c}^- p(x + e_{0c}, t) + \\ & + \sum_{c=1}^2 K \lambda_{0c}^+ \left( 1 - \sum_{i=0}^n \sum_{c=1}^2 x_{ic}(t) \right) (p(x - e_{0c}, t) - p(x, t)) + \sum_{c=1}^2 \lambda_{0c}^+ p(x - e_{0c}, t) + \\ & + \sum_{i,j=0}^n \sum_{s=1}^2 K \mu_{i1} \varepsilon_{i1}(x_{i1}, t) p_{ij}^{(1s)} (p(x + e_{i1} - e_{js}, t) - p(x, t)) + \\ & + \sum_{i,j=0}^n \sum_{s=1}^2 \mu_{i1} \frac{\partial \varepsilon_{i1}(x_{i1}, t)}{\partial x_{i1}} p_{ij}^{(1s)} p(x + e_{i1} - e_{js}, t) + \\ & + \sum_{i,j=0}^n \sum_{s=1}^2 K \mu_{i2} \varepsilon_{i2}(x_{i1}, x_{i2}, t) p_{ij}^{(1s)} (p(x + e_{i2} - e_{js}, t) - p(k, t)) + \\ & + \sum_{i,j=0}^n \sum_{s=1}^2 \mu_{i2} \frac{\partial \varepsilon_{i2}(x_{i1}, x_{i2}, t)}{\partial x_{i2}} p_{ij}^{(2s)} p(x + e_{i2} - e_{js}, t). \end{aligned} \quad (4)$$

Будем предполагать, что  $p(x, t)$  дифференцируема по  $t$  и дважды непрерывно дифференцируема по  $x_{ic}$ ,  $i = \overline{0, n}$ ,  $c = 1, 2$ . Тогда, заменяя в (4) функции  $p(x \pm e_{ic}, t)$ ,  $p(x + e_{ic} - e_{js}, t)$ ,  $i = \overline{0, n}$ ,  $c, s = 1, 2$ , их разложениями в ряд Тейлора и учитывая, что  $\varepsilon K = 1$ , с точностью  $O(\varepsilon^2)$  получаем, что плотность распределения вероятностей  $p(x, t)$  определяется уравнением

$$\frac{\partial p(x, t)}{\partial t} = - \sum_{i=0}^n \sum_{s=1}^2 \frac{\partial}{\partial x_{ic}} (A_{is}(x, t) p(x, t)) + \frac{\varepsilon}{2} \sum_{i,j=0}^n \sum_{c,s=1}^2 \frac{\partial^2}{\partial x_{ic} \partial x_{is}} (B_{icjs}(x, t) p(x, t)), \quad (5)$$

где

$$A_{01}(x, t) = \sum_{j=0}^n \mu_{j1} \varepsilon_{j1}(x_{j1}, t) (p_{j0}^{(11)} - \delta_{j0}) - \lambda_{01}^- x_{01}(t) + \sum_{j=0}^n \mu_{j2} \varepsilon_{j2}(x_{j1}, x_{j2}, t) p_{j0}^{(21)} + \lambda_{01}^+ \left( 1 - \sum_{i=0}^n \sum_{c=1}^2 x_{ic}(t) \right), \quad (6)$$

$$A_{i1}(x, t) = \sum_{j=0}^n \mu_{j1} \varepsilon_{j1}(x_{j1}, t) (p_{ji}^{(11)} - \delta_{ji}) + \sum_{j=0}^n \mu_{j2} \varepsilon_{j2}(x_{j1}, x_{j2}, t) p_{ji}^{(21)}, \quad i = \overline{1, n}, \quad (7)$$

$$A_{02}(x, t) = \sum_{j=0}^n \mu_{j2} \varepsilon_{j2}(x_{j1}, x_{j2}, t) (p_{j0}^{(22)} - \delta_{j0}) - \lambda_{02}^- x_{02}(t) + \sum_{j=0}^n \mu_{j1} \varepsilon_{j1}(x_{j1}, t) p_{j0}^{(12)} + \lambda_{02}^+ \left( 1 - \sum_{i=0}^n \sum_{c=1}^2 x_{ic}(t) \right), \quad (8)$$

$$A_{i2}(x, t) = \sum_{j=0}^n \mu_{j2} \varepsilon_{j2}(x_{j1}, x_{j2}, t) (p_{ji}^{(22)} - \delta_{ji}) + \sum_{j=0}^n \mu_{j1} \varepsilon_{j1}(x_{j1}, t) p_{ji}^{(12)}, \quad i = \overline{1, n}, \quad (9)$$

$$B_{0101}(x, t) = \sum_{j=0}^n \mu_{j1} \varepsilon_{j1}(x_{j1}, t) (p_{j0}^{(11)} + \delta_{j0}) + \lambda_{01}^- x_{01}(t) + \sum_{j=0}^n \mu_{j2} \varepsilon_{j2}(x_{j1}, x_{j2}, t) p_{j0}^{(21)} + \lambda_{01}^+ \left( 1 - \sum_{i=0}^n \sum_{c=1}^2 x_{ic}(t) \right),$$

$$B_{i1i1}(x, t) = \sum_{j=0}^n \mu_{j1} \varepsilon_{j1}(x_{j1}, t) (p_{ji}^{(11)} + \delta_{ji}) + \sum_{j=0}^n \mu_{j2} \varepsilon_{j2}(x_{j1}, x_{j2}, t) p_{ji}^{(21)}, \quad i = \overline{1, n},$$

$$B_{0202}(x, t) = \sum_{j=0}^n \mu_{j2} \varepsilon_{j2}(x_{j1}, x_{j2}, t) (p_{j0}^{(22)} + \delta_{j0}) + \lambda_{02}^- x_{02}(t) + \sum_{j=0}^n \mu_{j1} \varepsilon_{j1}(x_{j1}, t) p_{j0}^{(12)} + \lambda_{02}^+ \left( 1 - \sum_{i=0}^n \sum_{c=1}^2 x_{ic}(t) \right),$$

$$B_{i2i2}(x, t) = \sum_{j=0}^n \mu_{j2} \varepsilon_{j2}(x_{j1}, x_{j2}, t) (p_{ji}^{(22)} + \delta_{ji}) + \sum_{j=0}^n \mu_{j1} \varepsilon_{j1}(x_{j1}, t) p_{ji}^{(12)}, \quad i = \overline{1, n},$$

$$B_{i1j1}(x, t) = -\mu_{i1} \varepsilon_{i1}(x_{i1}, t) p_{ij}^{(1s)}, \quad i \neq j,$$

$$B_{i2j2}(x, t) = -\mu_{i2} \varepsilon_{i2}(x_{i1}, x_{i2}, t) p_{ij}^{(2s)}, \quad i \neq j,$$

$\delta_{ij}$  – символ Кронекера.

Дифференциальное уравнение в частных производных второго порядка (5) является уравнением Колмогорова–Фоккера–Планка для плотности распределения вероятностей марковского процесса  $\xi(t)$ . Отсюда следует [3], что средние значения процесса  $\xi(t)$  с точностью  $O(\varepsilon^2)$  удовлетворяют системе обыкновенных дифференциальных уравнений:

$$\frac{dn_{ic}(t)}{dt} = A_{ic}(n_{ic}, t), \quad i = \overline{0, n}, \quad c = 1, 2, \quad (10)$$

где  $n_{ic}(t) = M(\xi_{ic}(t)) = M\left(\frac{k_{ic}(t)}{K}\right)$ ,  $i = \overline{0, n}$ ,  $c = 1, 2$ , – компоненты вектора  $n(t)$  среднего относительного числа заявок каждого из двух типов в системах структуры.

Анализируя представления  $A_{ic}(x, t)$ , заданные выражениями (6)–(9), а также формулы (1), (2), делаем вывод, что правая часть (10) содержит кусочно-линейные слагаемые. Поэтому при решении системы (10) будем использовать разбиение фазового пространства и решать (10) в областях линейности ее правой части. В каждой из областей разбиения система (10) превращается в систему обыкновенных линейных дифференциальных уравнений.

Решение системы (10) при определенных начальных условиях позволяет найти среднее относительное число  $n_{ic}(t)$ , а значит, и среднее число  $Kn_{ic}(t)$  заявок в каждой из систем массового обслуживания,  $i = \overline{0, n}$ ,  $c = 1, 2$ .

Приведенный метод расчета среднего относительного числа заявок в системах массового обслуживания является асимптотическим и справедлив только при большой загрузке структуры массового обслуживания заявками. Точность метода возрастает с увеличением общего числа обслуживаемых заявок. Данная методика позволяет определять среднюю загрузженность систем массового обслуживания заявками с течением времени, т.е. в переходном режиме. Это является его преимуществом, так как большинство методов анализа характеристик систем и сетей массового обслуживания дают результаты лишь в стационарном режиме.

Рассмотренная структура может быть использована в качестве математической модели процессов обработки заявок клиентов различных организаций, например страховых, туристических или логистических компаний, банков [4]. Полученные результаты могут быть полезны при прогнозировании и оптимизации процессов обслуживания клиентов.

#### Литература

1. Медведев Г. А. Об оптимизации замкнутой системы массового обслуживания // Изв. АН СССР. Техническая кибернетика. 1975. № 6. С. 65–73.
2. Медведев Г. А. Замкнутые системы массового обслуживания и их оптимизация // Изв. АН СССР. Техническая кибернетика. 1978. №6. С. 199–203.
3. Параев Ю. И. Введение в статистическую динамику процессов управления и фильтрации. М.: Советское радио, 1976. 185 с.
4. Матальцкий М. А., Русилко Т. В. Математический анализ стохастических моделей обработки исков в страховых компаниях. Гродно: ГрГУ, 2007. 335 с.

# ИССЛЕДОВАНИЕ СИСТЕМЫ МАССОВОГО ОБСЛУЖИВАНИЯ MR/M/∞ С РАЗНОТИПНЫМ ОБСЛУЖИВАНИЕМ

*Е. Г. Убонова*

*Национальный исследовательский  
Томский государственный университет, Томск, Россия*

В настоящее время все больше внимания уделяется исследованиям систем массового обслуживания. Множество практических задач может быть представлено в виде математических моделей тех или иных систем массового обслуживания. В частности, в качестве математической модели call-центра может выступать система MR/M/∞ с разнотипным обслуживанием. Данная статья является обобщением работы Е. В. Панкратовой [1].

## Постановка задачи

Рассмотрим СМО, на вход которой поступает поток марковского восстановления разнотипных заявок, заданный набором функций распределения длин интервалов  $A_1(x), A_2(x), \dots, A_k(x)$  и матрицей  $\mathbf{P}$ , вложенной по моментам наступления событий цепи Маркова  $k(t)$  с конечным числом состояний  $K$  [2].

Дисциплина обслуживания заключается в том, что заявка из MR/M/∞ входящего потока с вероятностью  $p_1$  является заявкой первого типа, с вероятностью  $p_2$  является заявкой второго типа, и занимает любой свободный прибор.

Поставим задачу исследования двумерного случайного процесса  $\{i_1(t), i_2(t)\}$ , характеризующего число занятых приборов каждого типа в момент времени  $t$ . Так как входящий поток не является пуассоновским, то рассматриваемый двумерный процесс немарковский. Рассмотрим четырехмерный марковский случайный процесс  $\{k(t), i_1(t), i_2(t), z(t)\}$ , где  $z(t)$  – длина интервала от момента времени  $t$  до момента наступления очередного события;  $k(t)$  – вложенная по моментам наступления событий цепь Маркова.

## Система дифференциальных уравнений Колмогорова

Используя  $\Delta t$ -метод [3] для распределения вероятностей

$$P(k, i_1, i_2, z, t) = P\{k(t) = k, i_1(t) = i_1, i_2(t) = i_2, z(t) < z\},$$

по теореме полной вероятности [2] можно записать следующие равенства:

$$\begin{aligned} P(k, i_1, i_2, z - \Delta t, t + \Delta t) = & P(k, i_1, i_2, z, t) - P(k, i_1, i_2, \Delta t, t) - \\ & - (i_1\mu_1 + i_2\mu_2)(P(k, i_1, i_2, z, t) - P(k, i_1, i_2, \Delta t, t)) + \\ & + \sum_{\nu} P(\nu, i_1 - 1, i_2, \Delta t, t) p_1 P_{\nu k} A_k(z) + \sum_{\nu} P(\nu, i_1, i_2 - 1, \Delta t, t) p_2 P_{\nu k} A_k(z) + \\ & + (i_1 + 1)\mu_1 \Delta t P(k, i_1 + 1, i_2, z, t) + (i_2 + 1)\mu_2 \Delta t P(k, i_1, i_2 + 1, z, t) + o(\Delta t). \end{aligned} \quad (1)$$

где  $k, \nu = 1..K$ .

Из (1) получим систему дифференциальных уравнений Колмогорова для  $k = 1..K$ :

$$\begin{aligned} \frac{\partial P(k, i_1, i_2, z, t)}{\partial t} = & \frac{\partial P(k, i_1, i_2, z, t)}{\partial z} - \frac{\partial P(k, i_1, i_2, 0, t)}{\partial z} - (i_1 \mu_1 + i_2 \mu_2) P(k, i_1, i_2, z, t) + \\ & + (i_1 + 1) \mu_1 P(k, i_1 + 1, i_2, z, t) + (i_2 + 1) \mu_2 P(k, i_1, i_2 + 1, z, t) + \\ & + \sum_v \frac{P(v, i_1 - 1, i_2, 0, t)}{\partial z} p_1 P_{vk} A_k(z) + \sum_v \frac{P(v, i_1, i_2 - 1, 0, t)}{\partial z} p_2 P_{vk} A_k(z). \end{aligned}$$

Для стационарного распределения вероятностей запишем эту систему в виде

$$\begin{aligned} \frac{\partial \Pi(k, i_1, i_2, z)}{\partial z} - \frac{\partial \Pi(k, i_1, i_2, 0)}{\partial z} - (i_1 \mu_1 + i_2 \mu_2) \Pi(k, i_1, i_2, z) + \\ + (i_1 + 1) \mu_1 \Pi(k, i_1 + 1, i_2, z) + (i_2 + 1) \mu_2 \Pi(k, i_1, i_2 + 1, z) + \\ + \sum_v \frac{\Pi(v, i_1 - 1, i_2, 0)}{\partial z} p_1 P_{vk} A_k(z) + \sum_v \frac{\Pi(v, i_1, i_2 - 1, 0)}{\partial z} p_2 P_{vk} A_k(z) = 0. \quad (2) \end{aligned}$$

### Метод характеристических функций

Для решения системы (2) введем функцию, которую будем называть частичной характеристической функцией [4]:

$$H(k, u, w, z) = \sum_{i_1=0}^{\infty} \sum_{i_2=0}^{\infty} e^{ju i_1} e^{jw i_2} \Pi(k, i_1, i_2, z),$$

где  $j = \sqrt{-1}$ .

Тогда из системы (2), учитывая, что

$$\frac{\partial H(k, u, w, z)}{\partial u} = j \sum_{i_1=0}^{\infty} \sum_{i_2=0}^{\infty} i_1 e^{ju i_1} e^{jw i_2} \Pi(k, i_1, i_2, z),$$

$$\frac{\partial H(k, u, w, z)}{\partial w} = j \sum_{i_1=0}^{\infty} \sum_{i_2=0}^{\infty} i_2 e^{ju i_1} e^{jw i_2} \Pi(k, i_1, i_2, z),$$

получаем следующую систему уравнений:

$$\begin{aligned} \frac{\partial H(k, u, w, z)}{\partial z} - \frac{\partial H(k, u, w, 0)}{\partial z} - \mu_1 j (e^{-ju} - 1) \frac{\partial H(k, u, w, z)}{\partial u} - \\ - \mu_2 j (e^{-jw} - 1) \frac{\partial H(k, u, w, z)}{\partial w} + \sum_v p_1 e^{ju} P_{vk} A_k(z) \frac{\partial H(k, u, w, 0)}{\partial z} + \\ + \sum_v p_2 e^{jw} P_{vk} A_k(z) \frac{\partial H(k, u, w, 0)}{\partial z} = 0. \end{aligned}$$

Запишем данную систему в виде дифференциального матричного уравнения

$$\begin{aligned} \frac{\partial \mathbf{H}(u, w, z)}{\partial z} + \frac{\partial \mathbf{H}(u, w, 0)}{\partial z} [p_1 e^{ju} \mathbf{PD}(z) + p_2 e^{jw} \mathbf{PD}(z) - \mathbf{I}] - \\ - \mu_1 j (e^{-ju} - 1) \frac{\partial \mathbf{H}(u, w, z)}{\partial u} - \mu_2 j (e^{-jw} - 1) \frac{\partial \mathbf{H}(u, w, z)}{\partial w} = 0, \quad (3) \end{aligned}$$

где  $\mathbf{H}(u, w, z) = [H(1, u, w, z), H(2, u, w, z), \dots, H(K, u, w, z)]$ ,

$$\mathbf{D}(z) = \begin{bmatrix} A_1(z) & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & A_K(z) \end{bmatrix}, \mathbf{P} = \begin{bmatrix} p_{11} & \dots & p_{1K} \\ \vdots & \ddots & \vdots \\ p_{K1} & \dots & p_{KK} \end{bmatrix}, \mathbf{I} = \begin{bmatrix} 1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 1 \end{bmatrix}.$$

Начальным условием решения уравнения (3) является  $\mathbf{H}(0, 0, z) = \mathbf{R}(z)$ , где  $\mathbf{R}(z)$  – вектор-функция стационарного распределения вероятностей значений двумерного марковского процесса  $\{k(t), z(t)\}$ .

### Моменты первого порядка

**Теорема 1.** *Моменты первого порядка числа заявок первого и второго типа в системе MR/M/∞ имеют следующий вид:*

$$\overline{m}_1^{(1)}(\infty) = \frac{p_1}{\mu_1} \lambda,$$

$$\overline{m}_2^{(1)}(\infty) = \frac{p_2}{\mu_2} \lambda,$$

где  $\lambda = \frac{\mathbf{R}(0)}{\partial z}$ .

*Доказательство.* Из свойств характеристической функции имеем

$$\left. \frac{\partial H(k, u, w, z)}{\partial u} \right|_{\substack{u=0 \\ w=0}} = j m_1^{(1)}(k, z),$$

$$\left. \frac{\partial H(k, u, w, z)}{\partial w} \right|_{\substack{u=0 \\ w=0}} = j m_2^{(1)}(k, z).$$

Для векторных функций имеем

$$\left. \frac{\partial \mathbf{H}(u, w, z)}{\partial u} \right|_{\substack{u=0 \\ w=0}} = j \mathbf{m}_1^{(1)}(z) = j [m_1^{(1)}(1, z), m_1^{(1)}(2, z), \dots, m_1^{(1)}(K, z)],$$

$$\left. \frac{\partial \mathbf{H}(u, w, z)}{\partial w} \right|_{\substack{u=0 \\ w=0}} = j \mathbf{m}_2^{(1)}(z) = j [m_2^{(1)}(1, z), m_2^{(1)}(2, z), \dots, m_2^{(1)}(K, z)].$$

Продифференцировав (3) по  $u$ , получаем

$$\begin{aligned} & \frac{\partial^2 \mathbf{H}(u, w, z)}{\partial z \partial u} + \frac{\partial^2 \mathbf{H}(u, w, 0)}{\partial z \partial u} [p_1 e^{ju} \mathbf{PD}(z) + p_2 e^{jw} \mathbf{PD}(z) - \mathbf{I}] + \\ & + \frac{\partial \mathbf{H}(u, w, 0)}{\partial z} j p_1 e^{ju} \mathbf{PD}(z) - j(-j) \mu_1 e^{-ju} \frac{\partial \mathbf{H}(u, w, z)}{\partial u} - \\ & - \mu_1 j (e^{-ju} - 1) \frac{\partial^2 \mathbf{H}(u, w, z)}{\partial u^2} - \mu_2 j (e^{-jw} - 1) \frac{\partial^2 \mathbf{H}(u, w, z)}{\partial w^2} = 0. \end{aligned} \quad (4)$$

Положив  $u$  и  $w$  равными нулю, получаем

$$\frac{\partial \mathbf{m}_1^{(1)}(z)}{\partial z} + \frac{\partial \mathbf{m}_1^{(1)}(0)}{\partial z} [\mathbf{PD}(z) - \mathbf{I}] + p_1 \frac{\partial \mathbf{R}(0)}{\partial z} \mathbf{PD}(z) - \mu_1 \mathbf{m}_1^{(1)}(z) = 0. \quad (5)$$

Эту систему дифференциальных уравнений будем решать при помощи преобразования Лапласа–Стилтьеса [2]

$$\varphi_1(\alpha) = \int_0^{\infty} e^{-\alpha z} d\mathbf{m}_1^{(1)}(z), D^*(\alpha) = \int_0^{\infty} e^{-\alpha z} d\mathbf{D}(z).$$

Выполнив в (5) преобразование Лапласа–Стилтьеса, получим

$$\varphi_1(\alpha)(\mu_1 - \alpha) = \frac{\partial \mathbf{m}_1^{(1)}(0)}{\partial z} [\mathbf{PD}^*(\alpha) - \mathbf{I}] + p_1 \frac{\partial \mathbf{R}(0)}{\partial z} \mathbf{PD}^*(\alpha). \quad (6)$$

Положив в (6)  $\mu_1 = \alpha$ , имеем

$$\frac{\partial \mathbf{m}_1^{(1)}(0)}{\partial z} = p_1 \frac{\partial \mathbf{R}(0)}{\partial z} \mathbf{PD}^*(\mu_1) [\mathbf{I} - \mathbf{PD}^*(\mu_1)]^{-1}.$$

Тогда

$$\begin{aligned} \mathbf{m}_1^{(1)}(\infty) = \varphi_1(0) &= \frac{1}{\mu_1} p_1 \frac{\partial \mathbf{R}(0)}{\partial z} \mathbf{PD}^*(\mu_1) [\mathbf{I} - \mathbf{PD}^*(\mu_1)]^{-1} [\mathbf{PD}^*(0) - \mathbf{I}] + \\ &+ \frac{p_1}{\mu_1} \frac{\partial \mathbf{R}(0)}{\partial z} \mathbf{PD}^*(0). \end{aligned}$$

Для момента первого порядка числа заявок первого типа в системе можно записать:

$$\overline{m}_1^{(1)}(\infty) = \mathbf{m}_1^{(1)}(\infty) \mathbf{E} = \frac{p_1}{\mu_1} \frac{\partial \mathbf{R}(0)}{\partial z} \mathbf{E} = \frac{p_1}{\mu_1} \lambda.$$

Аналогично для момента первого порядка заявок второго типа в системе получим равенство

$$\overline{m}_2^{(1)}(\infty) = \mathbf{m}_2^{(1)}(\infty) \mathbf{E} = \frac{p_2}{\mu_2} \frac{\partial \mathbf{R}(0)}{\partial z} \mathbf{E} = \frac{p_2}{\mu_2} \lambda.$$

### Моменты второго порядка

**Теорема 2.** Моменты второго порядка числа заявок первого и второго типа в системе  $MR/M/\infty$  имеют следующий вид:

$$\begin{aligned} \overline{m}_1^{(2)}(\infty) &= \frac{p_1}{\mu_1} \lambda + \frac{p_1^2 \lambda}{\mu_1} \mathbf{PD}^*(\mu_1) [\mathbf{I} - \mathbf{PD}^*(\mu_1)]^{-1}, \\ \overline{m}_2^{(2)}(\infty) &= \frac{p_2}{\mu_2} \lambda + \frac{p_2^2 \lambda}{\mu_2} \mathbf{PD}^*(\mu_2) [\mathbf{I} - \mathbf{PD}^*(\mu_2)]^{-1}, \end{aligned}$$

где  $\lambda = \frac{\mathbf{R}(0)}{\partial z}$ .

*Доказательство.* Из свойств характеристической функции имеем

$$\begin{aligned} \left. \frac{\partial^2 H(k, u, w, z)}{\partial u^2} \right|_{\substack{u=0 \\ w=0}} &= j^2 m_1^{(2)}(k, z), \\ \left. \frac{\partial^2 H(k, u, w, z)}{\partial w^2} \right|_{\substack{u=0 \\ w=0}} &= j^2 m_2^{(2)}(k, z). \end{aligned}$$

Продифференцировав (4) по  $u$  и положив  $u$  и  $w$  равными нулю, получаем

$$\frac{\partial \mathbf{m}_1^{(2)}(z)}{\partial z} + \frac{\partial \mathbf{m}_1^{(2)}(0)}{\partial z} [\mathbf{PD}(z) - \mathbf{I}] + 2p_1 \mathbf{PD}(z) \frac{\partial \mathbf{m}_1^{(1)}(0)}{\partial z} + p_1 \frac{\partial \mathbf{R}(0)}{\partial z} \mathbf{PD}(z) - 2\mu_1 \mathbf{m}_1^{(2)}(z) + \mu_1 \mathbf{m}_1^{(1)}(z) = 0. \quad (7)$$

Эту систему дифференциальных уравнений будем решать при помощи преобразования Лапласа–Стилтьеса [2]:

$$\varphi_1(\alpha) = \int_0^{\infty} e^{-\alpha z} d\mathbf{m}_1^{(1)}(z), \varphi_2(\alpha) = \int_0^{\infty} e^{-\alpha z} d\mathbf{m}_1^{(2)}(z), D^*(\alpha) = \int_0^{\infty} e^{-\alpha z} d\mathbf{D}(z).$$

Выполнив в (7) преобразование Лапласа–Стилтьеса, получим

$$\varphi_2(\alpha)(2\mu_1 - \alpha) = \frac{\partial \mathbf{m}_1^{(2)}(0)}{\partial z} [\mathbf{PD}^*(\alpha) - \mathbf{I}] + 2p_1 \frac{\partial \mathbf{m}_1^{(1)}(0)}{\partial z} \mathbf{PD}^*(\alpha) + p_1 \frac{\partial \mathbf{R}(0)}{\partial z} \mathbf{PD}^*(\alpha) + \mu_1 \varphi_1(\alpha). \quad (8)$$

Положив в (8)  $2\mu_1 = \alpha$ , имеем

$$\frac{\partial \mathbf{m}_1^{(2)}(0)}{\partial z} = \left[ 2p_1 \frac{\partial \mathbf{m}_1^{(1)}(0)}{\partial z} \mathbf{PD}^*(2\mu_1) + p_1 \frac{\partial \mathbf{R}(0)}{\partial z} \mathbf{PD}^*(2\mu_1) + \mu_1 \varphi_1(2\mu_1) \right] [\mathbf{I} - \mathbf{PD}^*(2\mu_1)]^{-1}.$$

Тогда для момента второго порядка числа заявок первого типа в системе можно записать:

$$\begin{aligned} \overline{m}_1^{(2)}(\infty) &= \mathbf{m}_1^{(2)}(\infty) \mathbf{E} = \varphi_2(0) \mathbf{E} = \\ &= \left[ \frac{p_1}{\mu_1} \frac{\partial \mathbf{R}(0)}{\partial z} + \frac{p_1^2}{\mu_1} \frac{\partial \mathbf{R}(0)}{\partial z} \mathbf{PD}^*(\mu_1) [\mathbf{I} - \mathbf{PD}^*(\mu_1)]^{-1} \right] \mathbf{E} = \\ &= \frac{p_1}{\mu_1} \lambda + \frac{p_1^2 \lambda}{\mu_1} \mathbf{PD}^*(\mu_1) [\mathbf{I} - \mathbf{PD}^*(\mu_1)]^{-1}. \end{aligned}$$

Аналогично для момента второго порядка заявок второго типа в системе получим равенство

$$\overline{m}_2^{(2)}(\infty) = \frac{p_2}{\mu_2} \lambda + \frac{p_2^2 \lambda}{\mu_2} \mathbf{PD}^*(\mu_2) [\mathbf{I} - \mathbf{PD}^*(\mu_2)]^{-1}.$$

### Заключение

В настоящей работе исследована система массового обслуживания MR/M/∞ с разнотипным обслуживанием, найдены первые и вторые начальные моменты, что позволяет определить среднюю загрузженность системы и в последующем дать оценку производительности системы.

## Литература

1. Erlang, A. K. The theory of probability and telephone conversations. 1911. Вып. 20. P. 33–39.
2. Назаров А. А., Терпугов А. Ф. Теория массового обслуживания. Томск: Изд-во НТЛ, 2004.
3. Моисеева С. П. Методы асимптотического анализа в теории массового обслуживания. Томск: Изд-во НТЛ, 2006. 112 с.
4. Artalejo J. R. Retrial queueing systems: A computational approach. Berlin: Springer, 2008. 318 p.

## АСИМПТОТИЧЕСКИЙ АНАЛИЗ ВТОРОГО ПОРЯДКА RQ-СИСТЕМЫ ММРР|M|1 В УСЛОВИИ БОЛЬШОЙ ЗАГРУЗКИ

*Е. А. Фёдорова*

*Национальный исследовательский*

*Томский государственный университет, Томск, Россия*

Системы массового обслуживания с источником повторных вызовов (ИПВ) или так называемые RQ-системы [1–2] являются телекоммуникационными системами, в которых присутствует эффект повторных вызовов. Исследованию таких систем посвящено большое количество работ, однако аналитические результаты получены лишь для систем с простейшим входящим потоком.

В предыдущих работах [3–4] для исследования процесса числа заявок в ИПВ в RQ-системах (в том числе с входящим ММРР-поток) был предложен метод асимптотического анализа в условии большой загрузки. В ходе исследования было показано, что асимптотическая характеристическая функция (первого порядка) имеет вид характеристической функции гамма-распределения. Однако область применения метода оказалась невелика: расстояние Колмогорова между асимптотическим и допредельным распределениями меньше 0.05 при значениях загрузки, равной 0.95. В связи с этим, в данной работе предлагается увеличить область применимости метода в виде получения асимптотики второго порядка.

### Описание математической модели

Рассмотрим однолинейную RQ-систему ММРР|M|1 (рис. 1).

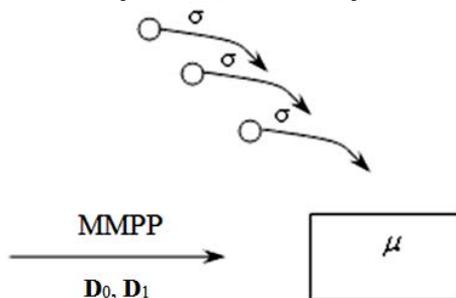


Рис. 1. RQ-система ММРР|M|1

На вход системы поступает ММРР-поток заявок (Markov Modulated Poisson Process), который является частным МАР-потока (Markovian Arrival Process) и описывается матрицами  $\mathbf{D}_0$  и  $\mathbf{D}_1$  [5].

Обозначим  $n(t)$  – цепь Маркова, управляющая ММРР-потоком. Матрица инфинитезимальных характеристик управляющего процесса  $n(t)$  равна  $\mathbf{Q} = \mathbf{D}_1 + \mathbf{D}_0$ . Матрица  $\mathbf{D}_1$  – диагональная матрица с элементами  $\rho\lambda_n$ , где  $\lambda_n$  – условные интенсивности ММРР-потока,  $\rho$  – параметр, характеризующий загрузку системы (определен ниже);  $n = 1, 2, \dots, N$ . Таким образом, можно ввести матрицу  $\mathbf{\Lambda} = \text{diag}\{\lambda_n\}$ , для которой выполняется  $\mathbf{D}_1 = \rho\mathbf{\Lambda}$ . Вектор-строка  $\mathbf{R}$  описывает стационарное распределение вероятностей состояний управляющего процесса  $n(t)$ . Очевидно, что тогда интенсивность входящего потока будет равна  $\lambda = \mathbf{R} \cdot \rho\mathbf{\Lambda} \cdot \mathbf{e}$ , где  $\mathbf{e}$  – единичный вектор-столбец.

Пусть параметры системы таковы, что выполняется  $\mathbf{R} \cdot \mathbf{\Lambda} \cdot \mathbf{e} = \mu$ , тогда загрузка системы определяется как  $\rho = \lambda / \mu = \lambda / \mathbf{R}\mathbf{\Lambda}\mathbf{e}$ .

Время обслуживания каждой заявки распределено по экспоненциальному закону с параметром  $\mu$ . Если поступившая заявка застаёт прибор свободным, то она занимает его для обслуживания. Если прибор занят, то заявка переходит в ИПВ, где осуществляет случайную задержку, продолжительность которой имеет экспоненциальное распределение с параметром  $\sigma$ .

Пусть  $i(t)$  – случайный процесс, характеризующий число заявок в ИПВ, а  $k(t)$  определяет состояние прибора следующим образом:  $k(t) = \{1, \text{если прибор занят}, 0, \text{если прибор свободен}\}$ .

Обозначим  $P\{k(t)=k, n(t)=n, i(t)=i\} = P(k, n, i, t)$  – вероятность того, что прибор в момент времени  $t$  находится в состоянии  $k$ , управляющая ММРР-потоком цепь Маркова – в состоянии  $n$  и в источнике повторных вызовов находится  $i$  заявок.

Для распределения вероятностей  $P(k, i, n, t)$  состояний рассматриваемой RQ-системы была составлена система дифференциальных уравнений Колмогорова:

$$\begin{cases} \frac{\partial P(0, n, i, t)}{\partial t} = \mu P(1, n, i, t) - (\rho\lambda_n + i\sigma - q_{nn})P(0, n, i, t) + \sum_{v \neq n} P(0, v, i, t)q_{vn}, \\ \frac{\partial P(1, n, i, t)}{\partial t} = -(\rho\lambda_n + \mu - q_{nn})P(1, n, i, t) + \rho\lambda_n P(0, n, i, t) + \\ + \rho\lambda_n P(1, n, i-1, t) + \sigma(i+1) \cdot P(0, n, i+1, t) + \sum_{v \neq n} P(1, v, i, t) \cdot q_{vn}. \end{cases} \quad (1)$$

Были введены вектор-строки

$$\mathbf{P}(k, i) = \{P(k, 1, i) \quad P(k, 2, i) \quad \dots \quad P(k, N, i)\}$$

и система (1) была записана в стационарном режиме в матричном виде:

$$\begin{cases} \mathbf{P}(0, i)(\mathbf{Q} - \rho\mathbf{\Lambda} - i\sigma \cdot \mathbf{I}) + \mu\mathbf{P}(1, i) = \mathbf{0}, \\ \mathbf{P}(1, i)(\mathbf{Q} - \rho\mathbf{\Lambda} - \mu\mathbf{I}) + \mathbf{P}(0, i)\rho\mathbf{\Lambda} + \mathbf{P}(1, i-1)\rho\mathbf{\Lambda} + \sigma(i+1) \cdot \mathbf{P}(0, i+1) = \mathbf{0}, \end{cases} \quad (2)$$

Далее был совершен переход к частичным характеристическим функциям и введены обозначения:  $u = \varepsilon w$ ,  $\mathbf{H}(0, u) = \varepsilon \mathbf{G}(w, \varepsilon)$ ,  $\mathbf{H}(1, u) = \mathbf{F}(w, \varepsilon)$ . Откуда из (2) была получена следующая система дифференциальных уравнений:

$$\begin{cases} \varepsilon \mathbf{G}(w, \varepsilon)(\mathbf{Q} - (1 - \varepsilon)\boldsymbol{\lambda}) + j\sigma \frac{\partial \mathbf{G}(w, \varepsilon)}{\partial w} + \mu \mathbf{F}(w, \varepsilon) = \mathbf{0}, \\ \mathbf{F}(w, \varepsilon)(\mathbf{Q} + (1 - \varepsilon)(e^{j\varepsilon w} - 1)\boldsymbol{\lambda} - \mu \mathbf{I}) + \varepsilon(1 - \varepsilon)\mathbf{G}(w, \varepsilon)\boldsymbol{\lambda} - j\sigma e^{-j\varepsilon w} \frac{\partial \mathbf{G}(w, \varepsilon)}{\partial w} = \mathbf{0}. \end{cases}$$

Для получения асимптотики второго порядка необходимо рассмотреть следующие разложения функций:

$$\mathbf{G}(w, \varepsilon) = \mathbf{G}(w) + \varepsilon \cdot \mathbf{g}(w) + \varepsilon^2 \cdot \mathbf{g}_2(w) + O(\varepsilon^3), \quad (3)$$

$$\mathbf{F}(w, \varepsilon) = \mathbf{F}(w) + \varepsilon \cdot \mathbf{f}(w) + \varepsilon^2 \cdot \mathbf{f}_2(w) + O(\varepsilon^3), \quad (4)$$

где  $O(\varepsilon^3)$  - бесконечно малая величина порядка  $\varepsilon^3$ .

Асимптотической характеристической функцией второго порядка будем называть функцию

$$h_2(u) = \mathbf{F}\left(\frac{u}{1-\rho}\right)\mathbf{E} + (1-\rho)\left\{\mathbf{G}\left(\frac{u}{1-\rho}\right)\mathbf{E} + \mathbf{f}\left(\frac{u}{1-\rho}\right)\mathbf{E}\right\},$$

где функции  $\mathbf{F}(w)$ ,  $\mathbf{G}(w)$  и  $\mathbf{f}(w)$  определяются разложениями (3) и (4).

Вид функции  $h_2(u)$  определяется следующей теоремой.

**Теорема.** *Асимптотическая характеристическая функция второго порядка имеет вид*

$$h_2(u) = \left(1 - \frac{j\mu}{(1-\rho)\beta}\right)^{-\alpha} \left\{1 + (1-\rho) \left[ \frac{j\mu}{1-\rho} \mathbf{V}\mathbf{E} - j \int_0^{\frac{u}{1-\rho}} \frac{a(y)}{(jy - \beta)} dy \right]\right\},$$

где

$$\alpha = 1 + \frac{\mu}{\sigma}\beta, \quad \beta = \frac{\mu}{\mathbf{V}\boldsymbol{\lambda}\mathbf{E} - \mu\mathbf{V}\mathbf{E} + \mu},$$

а вектор  $\mathbf{V}$  является решением неоднородной системы  $\mathbf{V}\mathbf{Q} = \mathbf{R}(\mu\mathbf{I} - \boldsymbol{\lambda})$ , функция  $a(w)$  определяется выражением

$$\begin{aligned} a(w) = & \frac{\alpha}{\beta} \left(1 - \frac{jw}{\beta}\right)^{-1} \left[ -jw \cdot \frac{(2\mathbf{V}\boldsymbol{\lambda}\mathbf{E} - \mu\mathbf{V}\mathbf{E})}{\mu} + (jw)^2 \frac{\delta}{\mu} \right] - \frac{(2\mathbf{V}\boldsymbol{\lambda}\mathbf{E} - \mu\mathbf{V}\mathbf{E})}{\mu} + \\ & + 2jw \frac{\delta}{\mu} - 2jw \frac{\mu}{\sigma} - 2 \left(1 + \frac{\mu}{\sigma}\right) \left(1 - \frac{jw}{\beta}\right) + jw \mathbf{V}\mathbf{E} \frac{\mu}{\sigma}, \end{aligned}$$

постоянная  $\delta$  равна

$$\delta = \mu\mathbf{V}\mathbf{E} + \mathbf{V}_1(\boldsymbol{\lambda}\mathbf{E} - \mu\mathbf{E}) - \frac{\mu}{2},$$

где вектор  $\mathbf{V}_1$  является решением неоднородной системы

$$V_1 Q = \frac{\mu}{\beta} R - \frac{1}{2} (R\lambda + \mu R) - (V\lambda - \mu V).$$

Используя асимптотическую характеристическую функцию, можно построить асимптотическое распределение вероятностей числа заявок с помощью обратного преобразования Фурье.

Для оценки области применимости метода проведем сравнение асимптотических распределений первого и второго порядков и распределения, полученного в результате применения численного алгоритма (таблица).

**Расстояние Колмогорова между асимптотическими и точным распределениями в RQ-системе MMPP|M|1**

Значения параметров системы	Асимптотика первого порядка	Асимптотика второго порядка
$\rho = 0,7$	0,295	0,119
$\rho = 0,8$	0,170	0,043
$\rho = 0,9$	0,097	0,029
$\rho = 0,95$	0,046	0,016

Таким образом, численное сравнение асимптотического (2-го порядка) и допредельного распределений показало, что область применения метода (в сравнении с первой асимптотикой) расширяется в 4 раза: при загрузке  $\rho > 0,8$  расстояние Колмогорова  $\Delta \leq 0,05$ .

Литература

1. Artalejo J. R., Gomez-Corral A. Retrial Queueing Systems: A Computational Approach. Berlin: Springer, 2008.
2. Falin G. I, Templeton J. G. C. Retrial queues. London: Chapman & Hall, 1997.
3. Назаров А. А., Моисеева Е. А. Исследование RQ-системы MMPP|M|1 методом асимптотического анализа в условии большой загрузки // Изв. ТПУ. Математика и механика. Физика. 322, 2. Томск: Изд-во ТПУ, 2013. С. 19-23.
4. Моисеева Е. А., Назаров А. А. Исследование RQ-системы MMPP|GI|1 методом асимптотического анализа // Вестник ТГУ. Управление, вычислительная техника и информатика. Т. 25, № 4. Томск: Изд-во Том. ун-та, 2013. С. 84-94.
5. Lucantoni D. M. New results on the single server queue with a batch Markovian arrival process // Stochastic Models. 1991. Vol. 7. P. 1-46.

# ОПТИМИЗАЦИОННЫЕ МОДЕЛИ И ИССЛЕДОВАНИЕ ОПЕРАЦИЙ

## DYNAMIC LOCALLY OPTIMAL CONTROL OF DISCRETE STATE DELAY SYSTEMS WITH RANDOM PARAMETERS

*O. O. Mukhina, V. I. Smagin*

*National Research Tomsk State University, Tomsk, Russia*

Control of the discrete systems with random parameters is considered in [1–4]. In this paper we propose for object with random parameters to realize synthesis of dynamic tracking system control on output based on optimization of the local criterion under indirect measurements taking into account state delay. To improve the quality control of objects practice of insertion in control law Luenberger observer [5] or dynamic feedback reduced dimensionality [6] is used. Control is defined as a function of the measured variables and the tracking signal. Asymptotic behavior of the system, construct estimates for asymptotic tracking accuracy are researched. An example is given to illustrate the usefulness of the proposed results.

**1. Problem statement.** Consider the following discrete-time system with time delay and random parameters:

$$x(k+1) = (A + \sum_{i=1}^r A_i \theta_i(k))x(k) + (\tilde{A} + \sum_{i=1}^r \tilde{A}_i \theta_i(k))x(k-h) + \\ + (B + \sum_{i=1}^r B_i \theta_i(k))u(k) + q(k); \\ x(\tau) = \varphi(\tau), \tau = -h, 1-h, 2-h, \dots, 0; k = 0, 1, 2, \dots, \quad (1)$$

$$y(k) = Sx(k) + v(k). \quad (2)$$

In (1), (2)  $x(k) \in R^n$  is state vector;  $h > 0$  is positive integer time delay;  $u(k) \in R^m$  is control input;  $y(k) \in R^l$  is observations vector;  $A, A_i, \tilde{A}, \tilde{A}_i, B, B_i, i = \overline{1, r}$  are constant matrices of appropriate dimensions;  $S$  is matrix of measurement channel; matrix  $B$  and  $S$  are of full rank; pairs of matrices  $(A, B)$  and  $(\tilde{A}, B)$  are controllable; pairs of matrices  $(S, A)$  and  $(S, \tilde{A})$  are observable;  $\varphi(\tau)$  is prescribed determinate function of initial conditions on the interval  $[-h, 1-h, \dots, -1]$ , as this takes place  $\varphi(0) = x(0) = x_0$  is random vector with characteristics:  $M\{x_0\} = \bar{x}_0$ ,  $M\{x_0 x_0^T\} = P_{x_0}$ ;  $q(k)$ ,  $v(k)$  are Gaussian random sequences of input disturbances and measurement errors with characteristics:  $M\{q(k)\} = 0$ ,  $M\{v(k)\} = 0$ ,  $M\{q(k)v^T(j)\} = 0$ ,  $M\{q(k)q^T(j)\} = Q(k)\delta_{kj}$ ,  $M\{v(k)v^T(j)\} = V(k)\delta_{kj}$  ( $\delta_{k,j}$  is Kronecker delta),  $Q(k) = Q^T(k) \geq 0$ ,

$V(k) = V^T(k) \geq 0$  are nonnegative definite matrices);  $\theta_i(k)$  is Gaussian random sequences ( $M\{\theta_i(k)\} = 0$ ;  $M\{\theta_i(k)\theta_i^T(j)\} = \delta_{k,j}$ ).

Local criterion described by:

$$I(k) = M\{(w(k+1) - z(k))^T C(w(k+1) - z(k)) + u^T(k)Du(k)\}, \quad (3)$$

where  $w(k) = Hx(k)$  is controlled output of the system ( $H$  is matrix of the output of the system);  $C = C^T, D = D^T \geq 0$  are weighting matrices;  $z(k) \in R^n$  is the reference input which described by the equation

$$z(k+1) = Fz(k) + q_z(k), \quad z(0) = z_0, \quad k = 0, 1, 2, \dots \quad (4)$$

In (4)  $q_z(k)$  is Gaussian random sequence with characteristics  $M\{q_z(k)\} = 0$ ,  $M\{q_z(k)q^T(j)\} = 0$ ,  $M\{q_z(k)v^T(j)\} = 0$ ,  $M\{q_z(k)q_z^T(j)\} = Q_z(k)\delta_{k,j}$ ,  $z_0$  are initial conditions ( $M\{z_0 z_0^T\} = P_{z_0}$ ,  $M\{z_0 x_0^T\} = P_{z_0 x_0}$ ,  $M\{x_0 z_0^T\} = P_{x_0 z_0}$ ),  $F$  is matrix of the dynamics model of the reference input.

It is required to construct a control of system (1), using observations (2) and minimizing the criterion (3).

**2. Locally criterion optimization.** Let the control law of the system (1) under observations (2) is defined as:

$$u(k) = K_0(k)\omega(k) + K_1(k)y(k) + K_2(k)y(k-h) + K_3(k)z(k) = K_0(k)\omega(k) + K_1(k)Sx(k) + K_1(k)v(k) + K_2(k)Sx(k-h) + K_2(k)v(k-h) + K_3(k)z(k), \quad (5)$$

where transfer coefficients are  $K_0(k), K_1(k), K_2(k), K_3(k)$  to be determined and variable  $\omega(k)$  is defined with by the reduced dimensionality equation [6].

**Theorem 1.** If for object (1), observations (2) and local criterion (3) matrices

$$\begin{aligned} \bar{C}(k) &= (B^T H^T C H B + D + \sum_{i=1}^r B_i^T H^T C H B_i) > 0, \\ \bar{P}(k) &= \begin{bmatrix} P_\omega(k) & SP_{x\omega}(k) \\ P_{\omega x}(k)S^T & SP_x(k)S^T + V(k) \\ P_{\omega x}(k, k-h)S^T & SP_x(k, k-h)S^T \\ P_{\omega z}(k) & SP_{xz}(k) \\ & SP_{x\omega}(k-h, k) & P_{z\omega}(k) \\ & SP_x(k-h, k)S^T & P_{zx}(k)S^T \\ & SP_x(k-h)S^T + V(k-h) & P_{zx}(k, k-h)S^T \\ & SP_{xz}(k-h, k) & P_z(k) \end{bmatrix} > 0 \quad (6) \end{aligned}$$

are positive definite for all  $k = 1, 2, \dots$ , then optimal in the sense of minimum criteria (3) transfer coefficients for control (5) are determined by the formulas:

$$K_0^*(k) = aK_1^*(k) + bK_2^*(k) + cK_3^*(k) + d; \quad (7)$$

$$K_1^*(k) = eK_0^*(k) + fK_1^*(k) + gK_2^*(k) + h; \quad (8)$$

$$K_2^*(k) = mK_0^*(k) + nK_1^*(k) + pK_3^*(k) + r; \quad (9)$$

$$K_3^*(k) = sK_0^*(k) + tK_1^*(k) + lK_2^*(k) + k, \quad (10)$$

where  $a, b, c, d, e, f, g, h, m, n, p, r, s, l, k$  are variables depending on  $P_z(k, r) = M\{z(k)z^T(r)\}$ ;  $P_\omega(k, r) = M\{\omega(k)\omega^T(r)\}$ ;  $P_{z\omega}(k, r) = P_{\omega z}^T(r, k) = M\{z(k)\omega^T(r)\}$ , which are defined by the system of difference matrix equations without delays with initial conditions:  $P_z(0) = P_{z_0}$ ;  $P_\omega(0) = P_{z\omega}(0) = P_{\omega z}(0) = 0$  and  $P_x(k, r) = M\{x(k)x^T(r)\}$ ;  $P_{zx}(k, r) = P_{xz}^T(r, k) = M\{z(k)x^T(r)\}$ , which are defined by the system of difference matrix equations with time delays with initial and boundary conditions:  $P_x(t-h, j-h) = \varphi(t-h)\varphi^T(j-h)$  for  $t, j = \overline{0, h-1}$ ;  $P_x(0, \tau) = \bar{x}_0\varphi^T(\tau)$ ;  $P_x(\tau, 0) = \varphi(\tau)\bar{x}_0^T$ ;  $P_{zx}(0, \tau) = \bar{z}_0\varphi^T(\tau)$ ;  $P_{xz}(\tau, 0) = \varphi(\tau)\bar{z}_0^T$ ;  $P_{x\omega}(\tau, 0) = P_{\omega x}(0, \tau) = 0$  для  $\tau = -h, 1-h, 2-h, \dots, -1$ ;  $P_x(0) = P_{x_0}$ ;  $P_{zx}(0) = P_{z_0x_0}$ ;  $P_{xz}(0) = P_{x_0z_0}$ ;  $P_{zx}(0) = P_{z_0x_0}$ ;  $P_{\omega x}(0) = P_{x\omega}(0) = 0$ .

The proof of the theorem is performed similar paper [7].

**3. Asymptotic behavior.** Asymptotic tracking accuracy for the object (1) is defined by calculating criterion estimation:

$$J = \lim_{k \rightarrow \infty} M\{\|x(k+1) - z\|^2\}, \quad (11)$$

where  $\|\cdot\|$  is Euclidean norm of vector,  $z$  is constant reference input.

**Theorem 2.** Let in description of system (1), observations (2), criterion (3) and model of reference input (4) matrices  $A, A_i, \tilde{A}, \tilde{A}_i, B, B_i, Q, S, V, C, D$ ,  $i = \overline{1, r}$  are constant;  $F = E$ ;  $q_z(k) = 0$ . Then, if the condition (6) theorem 1 is satisfied, there exist steady-state solutions of difference equations for  $P_x(t, j)$ ,  $P_{zx}(t, j)$ ,  $P_{xz}(t, j)$ ,  $P_{x\omega}(t, j)$ ,  $P_{\omega x}(t, j)$  and condition is satisfied:

$$\alpha_1^2 + \Phi_1^2 < 1, \quad (12)$$

then for criterion (11) estimate is valid:

$$J \leq \frac{[(G+R)^2 + (g+r_2)^2 + tr\tilde{Q}][(\alpha_1^2 + \Phi_1^2) + (\alpha_2^2 + \Phi_2^2)]}{1 - (\alpha_1^2 + \Phi_1^2)} + 2(\alpha_1\alpha_2 + \Phi_1\Phi_2) \frac{(G+R)^2 + (g+r_2)^2 + tr\tilde{Q}}{1 - (\alpha_1^2 + \Phi_1^2)} + (g+r_1)^2 + (G+R)^2 + 2(g+r_2) \frac{(\alpha_1 + \alpha_2)(g+r_1) + (\Phi_1 + \Phi_2)(G+R)}{1 - \alpha_1} + tr\tilde{Q}, \quad (13)$$

where

$$\begin{aligned}
\|\xi\|_s &= \alpha_1; \|\xi_2\|_s = \alpha_2; \|\varphi_1\|_s = \Phi; \|\varphi_2\|_s = \Phi_2; \xi = A + BK_1^*S; \xi_2 = \tilde{A} + BK_2^*S; \\
\xi_3 &= BK_3^* - E; \varphi_1 = \sum_{i=1}^r (A_i\theta_i(k) + B_i\theta_i(k)K_1^*S); \varphi_2 = \sum_{i=1}^r (\tilde{A}_i\theta_i(k) + B_i\theta_i(k)K_2^*S); \\
K_1^* &= \lim_{k \rightarrow \infty} K_1^*(k); K_2^* = \lim_{k \rightarrow \infty} K_2^*(k); K_3^* = \lim_{k \rightarrow \infty} K_3^*(k); r_1 = \|\xi_3 z\|; r_2 = \|BK_3^*z\|; \\
R &= \left\| \sum_{i=1}^r B_i\theta_i(k)K_3^*z \right\|; g = \|BK_0^*\omega\|; G = \left\| \sum_{i=1}^r B_i\theta_i(k)K_0^*\omega \right\|; \\
\tilde{Q}_1 &= BK_1^*VK_2^{*\top}B^\top + \sum_{i=1}^r B_iK_1^*VK_2^{*\top}B_i^\top; \\
\tilde{Q} &= Q + BK_1^*VK_1^{*\top}B^\top + \sum_{i=1}^r B_iK_1^*VK_1^{*\top}B_i^\top + BK_2^*VK_1^{*\top}B^\top + \sum_{i=1}^r B_iK_2^*VK_1^{*\top}B_i^\top.
\end{aligned}$$

**4. Illustrative example.** Let the system and local criterion are described by the following matrices and vectors:

$$\begin{aligned}
A &= \begin{bmatrix} 0,05 & 1 \\ -0,025 & 1 \end{bmatrix}; \tilde{A} = \begin{bmatrix} 0 & 0 \\ 0,03 & 0 \end{bmatrix}; B = \begin{bmatrix} 0,1 \\ 1 \end{bmatrix}; A_1 = \begin{bmatrix} 0,05 & 0 \\ 0 & 0 \end{bmatrix}; A_2 = \begin{bmatrix} 0 & 0 \\ 0,005 & 0 \end{bmatrix}; \\
A_3 &= \begin{bmatrix} 0 & 0,1 \\ 0 & 0 \end{bmatrix}; A_4 = \begin{bmatrix} 0 & 0 \\ 0 & 0,01 \end{bmatrix}; A_5 = A_6 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}; \tilde{A}_1 = \begin{bmatrix} 0,03 & 0 \\ 0 & 0 \end{bmatrix}; \\
\tilde{A}_2 &= \begin{bmatrix} 0 & 0 \\ 0,05 & 0 \end{bmatrix}; \tilde{A}_3 = \begin{bmatrix} 0 & 0,001 \\ 0 & 0 \end{bmatrix}; \tilde{A}_4 = \begin{bmatrix} 0 & 0 \\ 0 & 0,002 \end{bmatrix}; \tilde{A}_5 = \tilde{A}_6 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}; \\
B_1 = B_2 = B_3 = B_4 &= \begin{bmatrix} 0 \\ 0 \end{bmatrix}; B_5 = \begin{bmatrix} 0,005 \\ 0 \end{bmatrix}; B_6 = \begin{bmatrix} 0 \\ 0,025 \end{bmatrix}; F = 1; z = \begin{bmatrix} 10 \\ 10 \end{bmatrix}; \\
Q &= \begin{bmatrix} 0,02 & 0 \\ 0 & 0,02 \end{bmatrix}; S = [0 \ 1]; H = [1 \ 0]; C = 1; D = 0,2; h = 1.
\end{aligned}$$

Parameters  $\theta_i(k)$  are modeled as Gaussian random variable with mathematical expectation is 0 and mean-square deviation is 1.

To substantiate of the utility of introduction of the dynamic element  $\omega(k)$  in control law we modeled four control systems (algorithms): algorithm 1 is control built on the nominal values of the parameters; algorithm 2 is locally-optimal control calculated with regard to random parameters; algorithm 3 is control synthesized on the nominal values of the parameters and with the introduction of the dynamic element in the control law; algorithm 4 is locally-optimal control with the dynamic element, calculated taking into account the random parameters.

In the table the values of quality criterion of the convergence of the state vector  $x(k)$  to the reference input  $z(k)$  for the four control systems is cited.

**Average error**

Algorithm	1	2	3	4
$e$	2,325	2,103	1,806	1,638

The table shows that minimum average error deviation is reached at 4th algorithm with locally-optimal control with dynamic element, constructed on the transfer coefficients calculated with regard to the random parameters.

**Conclusion.** The problem of controlling the output of discrete systems with state delays and random parameters based on the synthesis of locally optimal linear tracking control system of discrete systems with indirect observations and introduction dynamic element in control law has been solved. The asymptotic behavior of the system has been analyzed. It is shown that introduction of the dynamic element in system greatly improves system performance and reduces the sensitivity to disturbing influences. Also it is shown that the optimal dynamic control system, built on random parameters with constant transfer coefficients has the property of robustness.

#### References

1. Bar-Shalom Y., Sivan R. On the optimal control of discrete-time linear systems with random parameters // IEEE Trans. Automatic Control. 1969. Vol. 14, no. 1. P. 3–8.
2. Lee J. H., Cooley B. L. Optimal feedback control strategies for state-space systems with stochastic parameters // IEEE Trans. Automatic Control. 1998. Vol. 43. P. 1469–1474.
3. Dombrovsky V. V., Dombrovsky D. V., Lyashenko E. A. Model predictive control of systems with random dependent parameters under constraints and its application to the investment portfolio optimization // Automation and Remote Control. 2006. Vol. 67, issue 12. P. 1927–1939.
4. Cannon M., Kouvaritakis B., Couchman P. Mean-variance receding horizon control for discrete time linear stochastic systems // Proc. of the 17th World Congress The International Federation of Automatic Control. Seoul, Korea, 2008. P. 15321–15326.
5. Luenberger D. G. An introduction to observers // IEEE Trans. Automatic Control. 1971. Vol. 16, issue 6. P. 596–603.
6. Dombrovskii V. V. Synthesis of Reduced-order Dynamic Regulators Under H-infinity-constraints // Automation and Remote Control. 1996. Vol. 57, issue 11. P. 1537–1543.
7. Mukhina O., Smagin V. Locally Optimal Control for Discrete Time Delay Systems with Interval Parameters / Dudin A., ed. // Communications in Computer and Information Science (CCIS 487). ITMM 2014. 2014. P. 301–311.

## ИССЛЕДОВАНИЕ ЭФФЕКТИВНОСТИ ГЕНЕТИЧЕСКОГО АЛГОРИТМА БЕЗУСЛОВНОЙ ОПТИМИЗАЦИИ

*А. В. Дягель*

*Сибирский федеральный университет, Красноярск, Россия*

Оптимизационные задачи заключаются в нахождении экстремума заданной целевой функции. Как правило, это сложная функция, зависящая от многих переменных (входных параметров). Требуется найти значения этих параметров, при которых целевая функция достигает своего экстремума.

Одним из наиболее часто встречающихся на практике типов задач являются задачи безусловной оптимизации. В задачах такого типа требуется

найти значение экстремума заданной целевой функции, на переменные которой не наложены ограничения (условия).

Существует множество различных традиционных оптимизационных методов, с помощью которых можно найти экстремальное значение целевой функции. Однако в случае многоэкстремальных целевых функций не все методы позволяют получить значение глобального оптимума. Кроме того, решение таких задач требует большого объема вычислений и получение решения занимает существенное время. При решении оптимизационных задач, для которых классические методы являются неприменимыми, предпочтение отдается стохастическим, в частности эволюционным алгоритмам, одним из которых и является генетический алгоритм (ГА) [1, 2]. ГА обладает важными свойствами, благодаря которым он показывает свое превосходство над традиционными методами оптимизации: работает одновременно с множеством решений, использует только значение целевой функции и не использует какую-либо дополнительную информацию, является стохастической оптимизационной процедурой с глобальным поведением. В то же время ГА обладает существенным недостатком: его эффективность существенно зависит от точного выбора конфигурации алгоритма, т.е. типов генетических операторов.

В данной работе был реализован стандартный ГА, включающий в себя основные генетические операторы, такие как селекция (турнирная, пропорциональная и ранговая), скрещивание (одноточечное, двухточечное и равномерное) и мутация (слабая, средняя и сильная). Учитывались два основных критерия эффективности ГА. Главным критерием является надежность, т.е. доля успешных запусков, когда алгоритм находил глобальный оптимум с заданной точностью. Если несколько конфигураций демонстрируют одинаковую надежность, то лучший алгоритм определялся по минимуму затрат на поиск, показателем чего является усредненный по успешным запускам номер поколения (второе число в ячейках таблиц), на котором впервые был найден искомый оптимум (алгоритм не останавливается в этом случае, а продолжает работу до конца, так как априорная информация об оптимуме ему не предоставляется).

Исследование эффективности ГА, т.е. выбора наилучших конфигураций, проводилось для задач безусловной оптимизации пятнадцати целевых функций, представляющих собой репрезентативное множество тестовых задач. Две функции одной вещественной переменной, а остальные – двух вещественных переменных:

- Функции 1,2 – многоэкстремальные функции одной переменной. Глобальные экстремумы расположены около границы допустимой области. Значения функций в точках глобального оптимума слабо отличаются от значений в точках локальных оптимумов.

- Функция 5 – функция Розенброка, одноэкстремальная, имеет протяженный овраг в области оптимума.

- Функции 9, 10 – функции Катковника, многоэкстремальные, имеют сложную систему локальных оптимумов, локальные оптимумы очень «глубокие».

- Функции 3, 4 – функции Растригина. Функция 3 – классический вариант, численные коэффициенты подобраны таким образом, что значение функций в точке глобального оптимума слабо отличается от значений в точках локальных оптимумов. Функция 4 – модифицированная функция 3. Нарушена регулярность расположения локальных оптимумов.

- Функции 6, 12, 13, 14 – многоэкстремальные функции с очень большим количеством локальных оптимумов.

- Функции 7, 8, 11, 15 – многоэкстремальные функции, имеющие в области определения множества постоянства. Функция 7 «Де Йонга 2», почти на всей области определения имеет равное значение, области притяжения локальных и глобального оптимумов очень малы. Функция 8 «Сомбреро», имеет глобальный оптимум в начале координат, значения функции в точках, равноудаленных от глобального оптимума, равны. Функция 15 Шекеля – «лисьи норы». Функция имеет в области определения 25 оптимумов с очень маленькими зонами притяжения, значения в локальных и глобальном оптимумах слабо различаются, в оставшейся части области определения функция имеет равное значение.

Как видно из описания, используемые тестовые функции реализуют многие свойства, затрудняющие оптимизацию: многоэкстремальность, несепарабельность, овражность, значения функции в глобальном оптимуме слабо отличаются от значений в локальных и др. [3].

Для каждой из этих функций было проверено 27 конфигураций алгоритма – три типа скрещивания, три уровня мутации и три типа селекции (турнирная селекция с размером турнира – 2). Число индивидов и размер популяции изменялись. Вероятность скрещивания (1) не изменялась. Все результаты усреднены по 1000 запускам для более точного получения результатов.

В таблице приведены лучшие конфигурации алгоритма для каждой из 15 тестовых функций, выделенный ресурс и используемая точность, а также усреднённые результаты тестирования. Используются следующие обозначения: в столбце «Результаты» первое число означает надежность, второе число – скорость сходимости (номер итерации, когда было найдено решение, усредненное по числу успешных запусков алгоритма).

### Оценки эффективности работы ГА

Номер функции	Конфигурация алгоритма	Вычислительный ресурс и используемая точность	Результаты
1	Пропорциональная селекция + двухточечное скрещивание + низкая мутация	Число индивидов: 60 Число поколений: 80 Точность: 0,01	0,777 21,965
2	Пропорциональная селекция + равномерное скрещивание + средняя мутация	Число индивидов: 60 Число поколений: 80 Точность: 0,01	0,898 12,819
3	Ранговая селекция + одноточечное скрещивание + низкая мутация	Число индивидов: 100 Число поколений: 250 Точность: 0,08	0,955 26,187
4	Пропорциональная селекция + двухточечное скрещивание + низкая мутация	Число индивидов: 100 Число поколений: 250 Точность: 0,08	0,799 57,261
5	Ранговая селекция + равномерное скрещивание + высокая мутация	Число индивидов: 100 Число поколений: 250 Точность: 0,01	0,583 109,212
6	Пропорциональная селекция + равномерное скрещивание + низкая мутация	Число индивидов: 100 Число поколений: 250 Точность: 0,01	0,471 189,947
7	Турнирная селекция + двухточечное скрещивание + низкая мутация	Число индивидов: 100 Число поколений: 250 Точность: 0,05	0.192 219.749
8	Пропорциональная селекция + равномерное скрещивание + низкая мутация	Число индивидов: 300 Число поколений: 500 Точность: 0,05	0.215 211.58
9	Турнирная селекция + равномерное скрещивание + высокая мутация	Число индивидов: 100 Число поколений: 250 Точность: 0,025	0.968 24.231
10	Турнирная селекция + равномерное скрещивание + высокая мутация	Число индивидов: 100 Число поколений: 250 Точность: 0,025	0.899 43.331
11	Ранговая селекция + равномерное скрещивание + средняя мутация	Число индивидов: 100 Число поколений: 250 Точность: 0,025	0.934 35.981
12	Ранговая селекция + равномерное скрещивание + высокая мутация	Число индивидов: 100 Число поколений: 250 Точность: 0,01	0.983 18.582
13	Ранговая селекция + двухточечное скрещивание + низкая мутация	Число индивидов: 100 Число поколений: 250 Точность: 0,01	0,957 27,038
14	Ранговая селекция + двухточечное скрещивание + низкая мутация	Число индивидов: 100 Число поколений: 250 Точность: 0,01	0,965 24,875
15	Турнирная селекция + двухточечное скрещивание + высокая мутация	Число индивидов: 100 Число поколений: 250 Точность: 0,25	0,972 30,371

Анализ результатов показал, что эффективность алгоритма для данного множества тестовых задач существенно зависит от точного выбора его настроек. Причем на разных функциях лучшие конфигурации существенно отличаются, при этом на всех задачах неудачный выбор конфигурации алгоритма может привести к его неэффективной работе, поэтому выбор настроек «наугад» неприемлем. В то же время существуют различные удачные конфигурации, дающие очень высокую надежность, хотя и при очень разных затратах на оптимизацию.

Таким образом, можно сделать вывод, что ГА являются универсальным методом оптимизации многоэкстремальных функций, что позволяет решать широкий спектр задач. Однако следует помнить, что применение ГА полезно лишь в тех случаях, когда для данной задачи нет подходящего специального алгоритма решения, так как установить надежную конфигурацию за приемлемое время не представляется возможным.

#### Литература

1. Семенкин Е. С., Семенкина О. Э., Коробейников С. П. Адаптивные поисковые методы оптимизации сложных систем. Красноярск: СИБУП, 1996. 358 с.
2. Семенкин Е. С., Лебедев В. А. Метод обобщенного адаптивного поиска для синтеза управления сложными объектами. М.: МАКС Пресс, 2002. 320 с.
3. Бежитский С. С., Семенкин Е. С., Семенкина О. Э. Гибридный эволюционный алгоритм для задач выбора эффективных вариантов систем управления // Автоматизация и современные технологии. 2005. № 11. С. 24.

## **ЗАДАЧА КОММИВОЯЖЕРА: СУБОПТИМАЛЬНЫЙ АЛГОРИТМ ПО МЕТОДУ ВЕТВЕЙ И ГРАНИЦ**

*Ю. Л. Костюк*

*Национальный исследовательский*

*Томский государственный университет, Томск, Россия*

### **Введение**

В известном алгоритме Литтла [1], реализующим метод ветвей и границ для задачи коммивояжера (ЗК), на каждом шаге обработки матрицы расстояний из каждой строки и столбца вычитается величина наименьшего в строке или столбце элемента, после чего в каждой строке и каждом столбце появляется не менее чем по одному нулю. Сумма величин всех вычтенных элементов – это оценка снизу оптимального решения, на основе которой производится отсечение бесперспективных вариантов.

В статье [2] описан модифицированный алгоритм, в котором выполняется следующее дополнительное преобразование матрицы: находятся такие группы строк, в которых имеется ровно по одному нулю в одном и том же столбце, и для каждой такой группы вычитается минимальный ненулевой элемент, а к столбцу, где находятся нули, добавляется такое же значение. Аналогичное действие выполняется для столбцов. В результате в

матрице появляются дополнительные нули, увеличивается оценка снизу оптимального маршрута, что улучшает процесс отсечения. Трудоемкость преобразования матрицы остается в пределах  $O(n^2)$ , как в алгоритме Литтла. Как показал вычислительный эксперимент [2], трудоемкость модифицированного алгоритма такая же, как у алгоритма Литтла, порядка  $O(n^2 \cdot c^n)$ , но в случае несимметричной матрицы константа  $c$  у него меньше.

В статье [3] описано дальнейшее улучшение модифицированного алгоритма, в нем на каждом шаге обработки матрицы вычисляется уточненная оценка снизу оптимального решения. Как показал вычислительный эксперимент [3], этот алгоритм имеет преимущество перед алгоритмом Литтла как на симметричных, так и на несимметричных матрицах. Тем не менее его трудоемкость остается экспоненциальной, поэтому единственным способом решения ЗК для большой размерности остается поиск приближенного решения. Однако в общем случае для несимметричной матрицы приближенное решение с гарантированной точностью можно находить лишь алгоритмами с экспоненциальной трудоемкостью [4–6].

В статье рассмотрен приближенный алгоритм с гарантированной точностью решения и с трудоемкостью, близкой к  $O(n^3)$ .

#### **Субоптимальный алгоритм**

Алгоритм реализует поиск вглубь на дереве решений, каждый узел дерева соответствует частично построенному маршруту. Пусть  $S_{\min}$  – стоимость наилучшего на текущий момент построенного маршрута;  $S_m$  – оценка снизу для частично построенного маршрута в текущий момент. В методе ветвей и границ проверка  $S_m \geq S_{\min}$  отсекает все бесперспективные варианты, которые можно построить на основе частичного маршрута. Если использовать проверку в виде  $S_m \cdot (1 + \varepsilon) \geq S_{\min}$ , то можно гарантировать, что наименьший найденный маршрут не будет превышать оптимального маршрута в  $(1 + \varepsilon)$  раз. В статье [2] рассмотрен вариант модифицированного алгоритма с такой проверкой, при этом время вычислений уменьшилось в десятки раз при средней погрешности, в 2–4 раза меньшей, чем  $\varepsilon$ . Чтобы среднюю погрешность сделать еще меньше при том же времени вычислений, проверка в субоптимальном алгоритме выполняется в виде

$$S_m \geq S_{\min} / (1 + \varepsilon(m)),$$

где функция  $\varepsilon(m)$  зависит от количества  $m$  отрезков частичного маршрута следующим образом:

$$\begin{aligned} \varepsilon(m) &= 0, \text{ если } m \leq n_0, \\ \varepsilon(m) &= \varepsilon \cdot (m - n_0) / (n - n_0), \text{ если } m > n_0, \end{aligned}$$

где  $n_0$  – параметр.

В алгоритме узлы дерева решений будут сцепляться в линейный список, представляющий собой обход частично построенного дерева решений по висячим узлам слева направо. Узлы в списке обрабатываются, начиная с корневого узла. После обработки текущего узла производится переход к левому или правому узлу в списке.

Далее приведен общий вид алгоритма.

1. Создание начального элемента списка – корневого узла дерева решений – размером  $n$ . Присваивание начальных значений для матрицы в корневом узле, вспомогательным массивам и глобальным переменным.

2. Создание всего списка узлов дерева решений, присоединяя их слева от корневого узла и задавая им размеры соответственно:  $n - 1, n - 2, \dots, 3$ .

3. Цикл, пока список узлов дерева решений не просмотрен до конца:

3.1. Если  $m \leq n_0$ , то для  $S_{mm}$  присваивается значение  $S_{min}$ .

3.1. Иначе для  $S_{mm}$  присваивается значение  $S_{min} / (1 + \varepsilon \cdot (m - n_0) / (n - n_0))$ .

3.2. Если  $S_m < S_{mm}$ , то для текущего узла в списке:

3.2.1. Вычитание из строк и столбцов матрицы минимальных значений с коррекцией  $S_m$ .

3.2.2. Если  $S_m < S_{mm}$ , то для текущего узла в списке:

3.2.2.1. Дополнительное преобразование матрицы с коррекцией  $S_m$ .

3.2.2.2. Если  $S_m < S_{mm}$ , то для текущего узла в списке:

3.2.2.2.1. Если  $m > 4$ , то вычисление улучшенной оценки снизу с коррекцией  $S_m$ .

3.2.2.2.2. Если  $S_m < S_{mm}$ , то для текущего узла в списке:

3.2.2.2.2.1. Выбор ребра для включения в маршрут.

3.2.2.2.2.2. Формирование матрицы размером  $m - 1$  в левом узле списка.

3.2.2.2.2.3. Коррекция матрицы в старом узле.

3.2.2.2.2.4. Если  $m = 4$ , то:

3.2.2.2.2.4.1. Выбор лучшего варианта замыкания маршрута в левом узле, и если его длина меньше  $S_{min}$ , то запоминание нового маршрута.

3.2.2.2.2.4.2. Текущий узел справа.

3.2.2.2.2.4. Иначе текущий узел слева.

3.2.2.2.2. Иначе текущий узел справа.

3.2.2.2. Иначе текущий узел справа.

3.2.2. Иначе текущий узел справа.

3.2. Иначе текущий узел справа.

4. Уничтожение списка узлов дерева решений.

Исходный текст модифицированного алгоритма, модифицированного алгоритма с улучшенной нижней границей и субоптимального алгоритма на языке Паскаль (Delphi) доступен в сети Интернет [7–9].

### Вычислительный эксперимент

В вычислительном эксперименте на модельных данных проверялась сравнительная трудоемкость следующих трех алгоритмов:

- 1) алгоритма Литтла;
- 2) улучшенного модифицированного алгоритма;
- 3) субоптимального алгоритма с заданной гарантированной точностью решения  $\varepsilon$ .

Алгоритмы были написаны на языке Паскаль в системе Delphi, вычисления производились на компьютере с процессором Atlon, (частота 2,9 ГГц). В табл. 1 и 2 представлено среднее время вычисления маршрута (в секундах), полученное по одним и тем же сгенерированным экземплярам матриц для всех трех алгоритмов. Элементы матриц – равномерно распределенные независимые случайные целые числа из диапазона от 0 до 1000, матрицы несимметричные.

Таблица 1

**Результаты эксперимента для  $n$  от 40 до 100**

$n$	40	50	60	70	80	90	100
Алгоритм 1	0,077	0,69	5,2	–	–	–	–
Алгоритм 2	0,010	0,036	0,12	0,40	1,24	3,65	8,1
Алгоритм 3 $\varepsilon = 0,05$	0,0055 $\bar{\varepsilon}=0,05\%$	0,0113 $\bar{\varepsilon}=0,08\%$	0,021 $\bar{\varepsilon}=0,11\%$	0,043 $\bar{\varepsilon}=0,14\%$	0,078 $\bar{\varepsilon}=0,16\%$	0,12 $\bar{\varepsilon}=0,21\%$	0,18 $\bar{\varepsilon}=0,21\%$
Алгоритм 3 $\varepsilon = 0,1$	0,0039 $\bar{\varepsilon}=0,18\%$	0,0058 $\bar{\varepsilon}=0,31\%$	0,0085 $\bar{\varepsilon}=0,46\%$	0,013 $\bar{\varepsilon}=0,55\%$	0,016 $\bar{\varepsilon}=0,61\%$	0,021 $\bar{\varepsilon}=0,75\%$	0,025 $\bar{\varepsilon}=0,75\%$
Алгоритм 3 $\varepsilon = 0,15$	0,0032 $\bar{\varepsilon}=0,3\%$	0,0046 $\bar{\varepsilon}=0,5\%$	0,0062 $\bar{\varepsilon}=0,7\%$	0,0091 $\bar{\varepsilon}=0,8\%$	0,011 $\bar{\varepsilon}=0,9\%$	0,0147 $\bar{\varepsilon}=1,1\%$	0,0182 $\bar{\varepsilon}=1,1\%$

Для каждого размера матриц  $n$  генерировалось от 4000 (для минимальных  $n$ ) до 400 (для максимальных  $n$ ) вариантов матриц. В таблицах отсутствуют только те результаты вычислений, для получения которых потребовалось бы слишком много времени. В результате подбора было решено задавать параметр  $n_0$  равным 25, это означает, что если частичный маршрут состоит не более чем из 25 отрезков, то для всех вариантов, строящихся на основе данного частичного, ищется в точности оптимальный маршрут.

Таблица 2

**Результаты эксперимента для  $n$  от 150 до 1000**

$n$	150	200	300	400	500	700	1000
Алгоритм 3 $\varepsilon = 0,1$	0,074	0,15	0,47	10	–	–	–
Алгоритм 3 $\varepsilon = 0,15$	0,053 $\Delta\bar{\varepsilon}=0,32\%$	0,11 $\Delta\bar{\varepsilon}=0,27\%$	0,36 $\Delta\bar{\varepsilon}=0,26\%$	0,84 $\Delta\bar{\varepsilon}=0,23\%$	1,6	4,6	18

В табл. 1 для алгоритма 3 (субоптимального) представлены также средние значения погрешности  $\bar{\varepsilon}$  в процентах. В табл. 2 для алгоритма 3 при заданной величине  $\varepsilon = 0,15$  приведена разность  $\Delta\bar{\varepsilon}$  между средними значениями  $\bar{\varepsilon}$  при  $\varepsilon = 0,15$  и  $\varepsilon = 0,1$ .

Из табл. 1 видно, что для случайных матриц улучшенный модифицированный алгоритм имеет существенное преимущество над алгоритмом Литтла, при  $n = 60$  его время работы в 40 раз меньше. При этом субоптимальный алгоритм выполняется быстрее модифицированного в 20 раз при средней погрешности 0,7%. Для размерности  $n = 100$  превышение быстродействия субоптимального над модифицированным алгоритмом достигает 300 раз при средней погрешности 0,75%.

Из табл. 1 и 2 видно, что при фиксированной заданной погрешности  $\varepsilon$  и увеличении размерности  $n$  средняя погрешность вначале растет, а затем стабилизируется. Так, при  $\varepsilon = 0,15$  она не растет выше 1,1%. Из табл. 2 также видно, что субоптимальный алгоритм при  $\varepsilon = 0,15$  и размерности  $n = 1000$  способен вычислять маршрут за среднее время 18 с.

Используя данные из приведенных таблиц, можно подобрать наиболее приемлемое значение параметра  $\varepsilon$  для заданной размерности  $n$ .

В статье [3] приведены оценки константы  $c$  в формуле трудоемкости  $O(n^2 \cdot c^n)$  алгоритма Литтла и улучшенного модифицированного алгоритма: они равны соответственно 1,1857 и 1,0885. В предположении, что для субоптимального алгоритма формула трудоемкости имеет вид  $O(n^p)$ , по методу наименьших квадратов получены следующие оценки константы  $p$ :

$$\begin{aligned} p &= 3,8 \text{ при } \varepsilon = 0,05 \text{ для диапазона } n \text{ от } 40 \text{ до } 100, \\ p &= 2,37 \text{ при } \varepsilon = 0,1 \text{ для диапазона } n \text{ от } 40 \text{ до } 300, \\ p &= 2,7 \text{ при } \varepsilon = 0,15 \text{ для диапазона } n \text{ от } 40 \text{ до } 1000. \end{aligned}$$

Если расширить диапазон  $n$  в сторону увеличения, то величина константы  $p$  будет больше приведенных значений.

### Заключение

Как показал вычислительный эксперимент, предложенный субоптимальный алгоритм способен находить приближенное решение ЗК со средней погрешностью, близкой к 1%, в диапазоне до  $n = 1000$  для случайной несимметричной матрицы расстояний. Его порядок трудоемкости –  $O(n^p)$ , где значение константы  $p$  меньше чем 3. Это позволяет находить решения ЗК, почти не уступающие оптимальным, за приемлемое время для большей размерности  $n$ , чем с помощью других известных алгоритмов.

Применение предложенного алгоритма для матрицы расстояний с другими (неравномерными) распределениями требует дальнейшего исследования.

### Литература

1. Little J. D. C., Murty K. G., Sweeney D. W., Karel C. An algorithm for the Traveling Salesman Problem // Operations Research. 1963. Vol. 11. P. 972–989.
2. Костюк Ю. Л. Эффективная реализация алгоритма решения задачи коммивояжера методом ветвей и границ // Прикладная дискретная математика. 2013. № 2 (20). С. 78–90.
3. Костюк Ю. Л. Задача коммивояжера: улучшенная нижняя граница в методе ветвей и границ // Прикладная дискретная математика. 2013. № 4 (22). С. 73–81.

4. Меламед И.И., Сергеев С. И., Сигал И. Х. Задача коммивояжера. Вопросы теории // Автоматика и телемеханика. 1989. № 9. С. 3–33.
5. Меламед И. И., Сергеев С. И., Сигал И. Х. Задача коммивояжера. Точные алгоритмы // Автоматика и телемеханика. 1989. № 10. С. 3–29.
6. Меламед И. И., Сергеев С. И., Сигал И. Х. Задача коммивояжера. Приближенные алгоритмы // Автоматика и телемеханика. 1989. № 11. С. 3–26.
7. Костюк Ю. Л. Модифицированный алгоритм решения задачи коммивояжера методом ветвей и границ. Версия 2. Программа и модуль с описанием класса: язык Паскаль в системе Delphi [Электронный ресурс]. <http://inf.tsu.ru/Decanat/Staff.nsf/people/KostjukJuL>, 2013.
8. Костюк Ю. Л. Модифицированный алгоритм решения задачи коммивояжера методом ветвей и границ с улучшенной нижней границей. Программа и модуль с описанием класса: язык Паскаль в системе Delphi [Электронный ресурс]. <http://inf.tsu.ru/Decanat/Staff.nsf/people/KostjukJuL>, 2013.
9. Костюк Ю. Л. Модифицированный алгоритм решения задачи коммивояжера методом ветвей и границ с улучшенной нижней границей при заданной гарантированной погрешности решения. 2014, июнь. Программа и модуль с описанием класса: язык Паскаль в системе Delphi [Электронный ресурс]. <http://inf.tsu.ru/Decanat/Staff.nsf/people/KostjukJuL>, 2014.

## **ТАБЛИЧНЫЙ СПОСОБ ГЕНЕРАЦИИ ОБРАТНОЙ ПОЛЬСКОЙ СТРОКИ ДЛЯ LL(1)-ГРАММАТИКИ**

*Ю. Л. Костюк*

*Национальный исследовательский*

*Томский государственный университет, Томск, Россия*

Для грамматического анализа контекстно-свободных языков известны различные алгоритмы [1–3], в частности, базирующиеся на LL(1)-грамматиках. При этом для построения транслятора, генерирующего обратную польскую строку (ОПС), применяют такие методы, как атрибутные грамматики [3, 4], синтаксически управляемый перевод [2] и др. В соответствии с этими методами требуется разработать набор специальных программ, генерирующих отдельные элементы ОПС. Разработка этих программ является сложной и объемной задачей при создании транслятора.

В статье предлагается способ построения генератора ОПС, большинство действий в котором кодируется символами, помещенными в таблицу. Это позволяет систематизировать описание языка и упростить разработку алгоритма генерации ОПС.

### **Табличный генератор ОПС**

Для более эффективной реализации алгоритма распознавания LL(1)-грамматики порождающие правила предварительно преобразуются в нестрогую нормальную форму Грейбах [1]. В такой форме все правые части порождающих правил либо пусты, либо начинаются с терминальных символов. Затем строится таблица LL(1)-распознавателя. Для всех порождающих правил грамматики, имеющих вид

$$A \rightarrow a\gamma,$$

где  $A$  – нетерминал,  $a$  – терминальный символ,  $\gamma$  – цепочка из терминальных и нетерминальных символов, правая часть правила  $a\gamma$  заносится на пересечение строки, помеченной нетерминалом  $A$ , и столбца, помеченного терминалом  $a$  (лексемой). Если же порождающее правило имеет вид

$$A \rightarrow \lambda,$$

где  $\lambda$  – пустая цепочка, то во все клетки строки, помеченной нетерминалом  $A$  и не занятые другими правилами, записывается  $\lambda$ .

До начала работы в стек анализатора записывается два символа: ограничитель  $\perp$  и начальный нетерминальный символ. На каждом шаге анализатор проверяет корректность очередного входного символа, после чего выполняет одно из двух действий:

1) если на вершине стека нетерминал, то, в зависимости от того, каков очередной входной символ, этот нетерминал заменяется в стеке символами правой части соответствующего правила, причем символы записываются в обратном порядке. Если для очередного входного символа в таблице записана пустая строка, то нетерминал просто удаляется из стека, а если в таблице пустая клетка, то анализатор фиксирует ошибку во входной цепочке;

2) если на вершине стека терминал, то он сравнивается с очередным входным символом. При совпадении терминал удаляется из стека и делается переход к следующему символу входной цепочки. При несовпадении анализатор фиксирует ошибку.

Работа анализатора прекращается, когда входная цепочка символов, дополненная в конце ограничителем  $\perp$ , будет полностью просмотренной. Входная цепочка символов считается правильной, если стек окажется пустым.

Генератор ОПС работает с таблицей семантических правил. Ее размеры в точности совпадают с таблицей анализатора. Более того, для каждой непустой клетки таблицы анализатора, где находится правая часть какого-либо правила порождения, в таблицу семантических правил помещается правило из символов, количество которых в точности равно длине правой части соответствующего правила порождения.

Генератор ОПС работает синхронно с LL(1)-анализатором, в стек генератора он записывает семантические правила. Символы в правилах кодируют действия, генерирующие элементы ОПС в моменты извлечения символов из стека генератора.

В качестве примера рассмотрим грамматику простых арифметических выражений [1], в которой символы  $S$ ,  $T$ ,  $F$  – нетерминалы ( $S$  – начальный нетерминал), символы плюс, звездочка, круглые скобки и  $a$  – терминалы:

$$S \rightarrow S + T \mid T, \quad T \rightarrow T * F \mid F, \quad F \rightarrow (S) \mid a.$$

Символ  $a$  обозначает простой операнд в выражении – константу или переменную, которые распознаются лексическим анализатором. Верти-

кальная черта разделяет различные правые части при одинаковом нетерминале в левой части правила.

Символы в семантических правилах этой грамматики:  $a$  – запись в ОПС операнда из входной цепочки (переменной или константы), плюс – запись в ОПС сложения, звездочка – запись в ОПС умножения,  $\square$  – пустое действие.

В табл. 1 представлены преобразованные к нестрогой нормальной форме Грейбах порождающие и семантические правила грамматики. После преобразования в грамматике появились еще два нетерминала:  $U$  и  $V$ .

Таблица 1

**Порождающие и семантические правила грамматики  
простых арифметических выражений**

Порождающие правила	Семантические правила
$S \rightarrow (S)VU \mid aVU$	$\square\square\square\square \mid a\square\square$
$U \rightarrow + TU \mid \lambda$	$\square\square+ \mid \lambda$
$T \rightarrow (S)V \mid aV$	$\square\square\square \mid a\square$
$V \rightarrow * FV \mid \lambda$	$\square\square* \mid \lambda$
$F \rightarrow (S) \mid a$	$\square\square \mid a$

До начала работы генератора ОПС в стек генератора записываются два пустых символа. На каждом шаге работы анализатора генератор ОПС делает следующее:

– если верхний символ в стеке анализатора нетерминал, то в стеке генератора верхний символ удаляется и вместо него записывается соответствующее семантическое правило. Если при этом удаляемый символ из стека генератора не пустой, а, например, операция, то выполняется запись в ОПС операции;

– если верхний символ в стеке анализатора терминал, то его тип должен соответствовать типу терминала на входе (в противном случае – ошибка), и тогда из стека генератора удаляется верхний символ, и, если он не пустой, а, например операнд, то выполняется запись в ОПС операнда в виде ссылки на таблицу констант или переменных. При этом также выполняется поиск в таблице констант или переменных и при неудачном поиске в таблицу заносится новая константа или переменная.

Заметим, что на каждом шаге работы в обоих стеках содержится строго одинаковое количество символов.

Таблица 2 совмещает обе таблицы: анализатора и генератора ОПС.

Таблица 2

## Совмещенная таблица анализатора и генератора ОПС

	+	*	(	)	<i>a</i>	⊥
<i>S</i>			( <i>S</i> ) <i>VU</i> □□□□		<i>aVU</i> <i>a</i> □□	
<i>U</i>	+ <i>TU</i> □□+	λ	λ	λ	λ	λ
<i>T</i>			( <i>S</i> ) <i>V</i> □□□□		<i>aV</i> <i>a</i> □	
<i>V</i>	λ	* <i>FV</i> □□*	λ	λ	λ	λ
<i>F</i>			( <i>S</i> ) □□□		<i>a</i> <i>a</i>	

Пусть на вход LL(1)-анализатора и генератора ОПС, заданного табл. 2, поступает цепочка

$$a^*(c + d) \perp$$

в которой буквами *a*, *c*, *d* обозначены различные переменные, которые все соответствуют одинаковому терминалу *a* в грамматике. Ниже приведена последовательность состояний стека анализатора и стека генератора ОПС. В стеке анализатора нетерминал, выделенный жирным шрифтом, на следующем шаге заменяется на правую часть порождающего правила:

<i>S</i> ⊥ □□	<i>aVU</i> ⊥ <i>a</i> □□□	* <i>FVU</i> ⊥ □□*□□	( <i>S</i> ) <i>VU</i> ⊥ □□□*□□	<i>aVU</i> ) <i>VU</i> ⊥ <i>c</i> □□□*□□	+ <i>TU</i> ) <i>VU</i> ⊥ □□+□*□□	<i>aVU</i> ) <i>VU</i> ⊥ <i>d</i> □+□*□□	) <i>VU</i> ⊥ □*□□	⊥ □
------------------	------------------------------	-------------------------	------------------------------------	---	--------------------------------------	---	-----------------------	--------

На выходе будет сформирована таблица переменных (*a*, *c*, *d*) и сгенерирована следующая ОПС:

$$a c d + *$$

## Генерация ОПС условных операторов, циклов и присваиваний

Генерацию ОПС для условных операторов, циклов и присваиваний рассмотрим на примере простой грамматики, которая кроме порождающих и семантических правил из табл. 1 дополнительно содержит правила, представленные в табл. 3.

Таблица 3

## Дополнительные порождающие и семантические правила

Порождающие правила	Семантические правила
$P \rightarrow \mathbf{begin} AB \mathbf{end}$	□□□□
$A \rightarrow a = SZ \mid \mathbf{if} SD \mathbf{then} ABE \mid \mathbf{while} SD \mathbf{do} AB \mathbf{end}$	<i>a</i> □□=   □□□1□□□   4□□1□□5
$B \rightarrow ; AB \mid \lambda$	□□□   λ
$E \rightarrow \mathbf{else} AB \mathbf{end} \mid \mathbf{end}$	2□□3   3
$D \rightarrow >SZ \mid <SZ \mid \neq SZ$	□□>   □□<   □□≠
$Z \rightarrow \lambda$	λ

Начальный нетерминал в этой грамматике –  $P$ . В порождающих правилах жирным шрифтом выделены зарезервированные служебные слова, они, наряду со знаками «равно», «больше», «меньше», «не равно» и «точка с запятой», являются терминалами. Служебные слова, также как переменные и константы, распознаются лексическим анализатором.

Для простоты будем считать, что переменные описаны по умолчанию. ОПС для условных операторов и циклов содержит операнды-метки, и операторы – операции сравнения ( $>$ ,  $<$ ,  $\neq$ ), а также условные и безусловные переходы. Число в семантическом правиле задает вызов соответствующей семантической программы.

В семантических программах используется еще один стек – стек меток, а также переменная  $k$  – счетчик (номер) очередного генерируемого элемента ОПС. Этот счетчик увеличивается на 1 всякий раз, как только генерируется очередной элемент ОПС. Операнд-метка в ОПС представляется своим типом и номером того элемента ОПС, на который будет выполняться переход при вычислении (интерпретации) ОПС.

Программа 1.

1. В стек меток записывается  $k$ .
2. В ОПС записывается пустой элемент – место для будущей метки.
3. В ОПС записывается операция **jf** – переход по условию false.

Программа 2.

1. Через верхний элемент стека меток как ссылку на ранее заготовленное место для метки в ОПС записывается  $k + 2$ .
2. В стек меток записывается  $k$ .
3. В ОПС записывается пустой элемент – место для будущей метки.
4. В ОПС записывается операция **j** – безусловный переход.

Программа 3.

1. Через верхний элемент стека меток как ссылку на ранее заготовленное место для метки в ОПС записывается  $k$ .

Программа 4.

1. В стек меток записывается  $k$ .

Программа 5.

1. Через верхний элемент стека меток как ссылку на ранее заготовленное место для метки в ОПС записывается  $k + 2$ .
2. В ОПС записывается метка, ее значение читается из стека меток.
3. В ОПС записывается операция **j** – безусловный переход.

Конец семантических программ.

Благодаря использованию стека меток будет генерироваться корректная ОПС для вложенных условных операторов и циклов. Так, например, по входной цепочке символов

**begin  $x=b-c$ ; while  $x\neq c$  do if  $x<c$  then  $x=x+b$  else  $x=x+d$  end end end**

будет сгенерирована следующая ОПС:



метрики пространства. В этой работе мы дадим описание многошаговых релаксационных субградиентных методов оптимизации и приведем примеры успешного их применения для решения гладких и негладких задач оптимизации.

В релаксационных субградиентных методах  $\varepsilon$ -субградиентного типа (PCM) [1] последовательные приближения решения задачи  $f: R^n \rightarrow R^1$ .  $f(x) \rightarrow \min$  строятся по формулам

$$x_{k+1} = x_k - \gamma_k s_k, \quad \gamma_k = \arg \min_{\gamma \in R^1} f(x_k - \gamma s_k), \quad (1)$$

где выбор направления спуска  $s_k$  производится из множества  $S(G) = \{s \in R^n \mid \min_{g \in G} (s, g) > 0\}$  решений неравенств [1, 2]:

$$(s^*, g) > 0, \quad \forall g \in G, \quad (2)$$

где  $G = \partial_\varepsilon f(x_k)$  –  $\varepsilon$ -субградиентное множество в точке  $x_k$ .

Дадим описание многошаговых субградиентных методов, используемых в дальнейшем для решения прикладных задач оптимизации.

### **Многошаговый субградиентный метод минимизации (PCMM)**

Основываясь на работах [2, 3], сформулируем для задачи (1) алгоритм минимизации, содержащий в своей схеме многошаговый метод решения неравенств.

Алгоритм PCMM

*Задать начальное приближение  $x_0 \in R^n$  и  $s_0 \in R^n$ , целые  $k = i = 0$ .*

*Положить  $k = k+1$ ,  $q_k = i$ ,  $s_i = 0$ ,  $p_{i-1} = 0$ ,  $\sigma_i = 0$ .*

*Задать  $\varepsilon_k, m_k$ .*

*Вычислить субградиент  $g_i \in \partial f(x_i)$ , удовлетворяющий условию  $(s_i, g_i) \leq 0$ , если  $g_i = 0$ , то закончить работу алгоритма.*

*Получить новое приближение  $s_{i+1} = s_i + \frac{1 - (s_i, g_i)}{(p_i, g_i)} p_i$ ,*

$$\text{где } p_i = \begin{cases} g_i, & \text{если } (g_i, p_{i-1}) \geq 0, \\ g_i - \frac{(g_i, p_{i-1})}{\|p_{i-1}\|^2} p_{i-1}, & \text{если } (g_i, p_{i-1}) < 0. \end{cases}$$

*Вычислить новое приближение критерия  $\sigma_{i+1} = \sigma_i + (g_i, g_i)^{-1}$ .*

*Вычислить новое приближение точки минимума:*

$$x_{i+1} = x_i - \gamma_i s_{i+1}, \quad \gamma_i = \arg \min_{\gamma \in R} f(x_i - \gamma s_{i+1}).$$

*Положить  $i = i+1$ .*

*Если  $1/\sqrt{\sigma_i} < \varepsilon_k$  или  $i - q_k > m_k$ , то перейти на пункт 2, иначе перейти на пункт 4.*

## Алгоритм минимизации с коррекцией векторов спуска на основе пары обучающих соотношений (РСПМ)

Сформулируем алгоритм минимизации, основанный на методе решения неравенств с коррекцией векторов спуска, на основе пары обучающих соотношений [3, 4].

Алгоритм РСПМ

1. Задать начальное приближение  $x_0 \in R^n$ , целые  $k = i = 0$ .

Положить  $k = k + 1$ ,  $q_k = i$ ,  $s_i = 0$ ,  $g_{i-1} = 0$ ,  $\Sigma_i = 0$ .

Задать  $\varepsilon_k, m_k$ .

Вычислить субградиент  $g_i \in \partial f(x_i)$ , удовлетворяющий условию  $(s_i, g_i) \leq 0$ , если  $g_i = 0$ , то закончить работу алгоритма.

Получить новое приближение  $s_{i+1} = s_i + \frac{1 - (s_i, g_i)}{(p_i, g_i)} p_i$ , где

$$p_i = \begin{cases} g_i, & \text{если } (g_i, g_{i-1}) \geq 0, \\ g_i - \frac{(g_i, g_{i-1})}{\|g_{i-1}\|^2} g_{i-1}, & \text{если } (g_i, g_{i-1}) < 0. \end{cases}$$

Вычислить новое приближение критерия  $\Sigma_{i+1} = \Sigma_i + (g_i, g_i)^{-1}$ .

Вычислить новое приближение точки минимума:

$$x_{i+1} = x_i - \gamma_i s_{i+1}, \quad \gamma_i = \arg \min_{\gamma \in \mathbf{R}} f(x_i - \gamma_i s_{i+1}).$$

Положить  $i = i + 1$ .

Если  $1/\sqrt{\Sigma_i} < \varepsilon_k$ , то перейти на пункт 2.

Если  $i - q_k > m_k$ , то перейти на пункт 2, иначе перейти на пункт 4.

В следующей ниже большеразмерной задаче негладкой оптимизации возникает необходимость применения методов негладкой оптимизации.

**Задача максимизации прибыли совокупности предприятий при ограничениях на выбросы вредных веществ** [5] возникает в системах экологического мониторинга чистоты атмосферы для населенных мест и состоит в распределении объемов выбросов в атмосферу вредных веществ для предприятий с целью максимизации их прибыли при ограничениях на загрязнение окружающей среды в областях проживания [5]. По причинам зашумленности данных задачи, плохой ее обусловленности и вытекающей из этих факторов неоднозначности решения для нее целесообразно использовать специализированные методы, обеспечивающие равноценное распределение объемов выбросов предприятий.

Известные подходы к этому распределению приводят к математической постановке в виде задачи линейного программирования. В силу неполноты и неточности данных, плохой обусловленности, большой размерности задачи линейного программирования имеет место проблема неединственности решения и его неустойчивости к возмущениям данных задачи.

Для решения подобных задач используются подходы поиска регуляризованного решения.

Формализация данной задачи приводит к следующей математической постановке [5]:

$$\min_{a \leq x \leq d} [ (c, x) + \frac{1}{2} \sum_{i=1}^n \varepsilon_i x_i^2 ], \quad Ax \leq b, \quad \varepsilon_i > 0, \quad i = 1, 2, \dots, n,$$

где  $c_i = -1$ ;  $x$  – вектор неизвестных параметров,  $n$  – количество предприятий;  $\varepsilon_i$  – индивидуальный параметр регуляризации;  $A_{ji}$  – удельная (на единицу выброса) концентрация от  $i$ -го источника в  $j$ -й точке покрытия,  $b_j$  – предельно допустимая концентрация в  $j$ -й точке покрытия.

Использовались следующие алгоритмы:

1. Релаксационный субградиентный метод (РСМ) на основе алгоритма Качмажа (МО) [2].

РСМ с ортогонализацией векторов спуска на основе пары обучающих соотношений (РСПМ) [4].

РСМ с растяжением пространства в направлении субградиента (МН0) [2].

РСМ с растяжением-сжатием пространства (МН1) [2].

В табл. 1 приведены затраты количества итераций ( $it$ ), вычислений значений целевой функции и градиента ( $nfg$ ), а также параметр точности  $q = |f - f^*| / f^*$  уклонения полученного экстремума  $f = (c, \hat{x})$  от истинного  $f = (c, x^*)$ .

Таблица 1

**Прикладная задача распределения нагрузок**

Прикладная задача (ограничения-неравенства), $n = 1098$ , $\varepsilon = 0.1$			
Метод	$it$	$nfg$	$q$
МН0	328	649	5.9E-7
РСПМ	4126	8270	1.7E-4
МО	1685	3387	3.4E-4
МН1	493	989	2.6E-6

Все использованные методы позволяют найти значение оптимума задачи, практически совпадающее с теоретическим. Представленная задача имеет высокую степень сложности.

**Задача решения системы нелинейных уравнений гидродинамики**

Основываясь на работе [6], в ограниченной области  $\Omega$  рассматривается следующая стационарная задача Навье–Стокса:

$$\bar{u} \cdot \nabla \bar{u} + \nabla p = \nu \cdot \Delta \bar{u}, \quad (3)$$

$$div \bar{u} = 0, \quad (4)$$

$$l(\bar{u}) = 0, \quad (5)$$

где  $\Gamma$  – граница  $\Omega$ ;  $\bar{u}$  – двухмерный или трехмерный вектор скоростей;  $p$  – давление;  $\nu > 0$  – коэффициент кинематической вязкости,  $l(\bar{u})$  – некоторые краевые условия.

Для решения задачи (3)–(4) использован метод сеток, который заключается в построении сетки в области  $\Omega$  и замене слагаемых в (3) разностными соотношениями. В итоге получаем с учетом краевых условий (5) систему нелинейных алгебраических уравнений (СНАУ), размерность которой пропорциональна числу точек в  $\Omega \cup \Gamma$ , в которых необходимо вычислять скорость и давление.

Один из подходов решения систем нелинейных уравнений, аппроксимирующих уравнения гидродинамики (3)–(5), состоит в сведении задачи решения системы уравнений

$$r_i(u) = 0, \quad i = 1, 2, \dots, m, \quad u \in R^m \quad (6)$$

к задаче минимизации нормы невязки

$$\|r\|^2 = f(u) = \sum_{i=1}^m r_i^2(u), \quad u \in R^m. \quad (7)$$

Тестирование проводилось также на одномерной задаче Бюргерса. Использовались следующие алгоритмы: MBFGS (квазиньютоновский метод, основанный на формуле BFGS) [7], метод сопряженных градиентов (МСГ) [7, 8], АРП (алгоритм с растяжением пространства) [2], МШ (метод Шора) [9], РСМК (РСМ Качмажа) [2] и РСММ (многошаговый субградиентный метод). Функция и градиент во всех методах вычислялись одновременно. Останов алгоритма осуществлялся при достижении требуемой точности  $\varepsilon$  по правилу  $|f_k - f^*| < \varepsilon$ . В табл. 2 указаны размерность  $n$ , точность  $\varepsilon$ , затраты количества итераций ( $It$ ), числа вычислений функции и градиента ( $Nfg$ ), необходимые для достижения заданной точности  $\varepsilon$  для решаемой задачи.

Таблица 2

### Результаты для задачи Бюргерса

		<i>MBFGS</i>	АРП	МШ	РСМК	РСММ	МСГ
$n = 500,$ $\varepsilon = 10^{-15}$	<i>It</i>	705	700	4486	28759	40570	59659
	<i>Nfg</i>	1651	1240	8650	57452	68000	119421
$n = 1000,$ $\varepsilon = 10^{-15}$	<i>It</i>	1242	1362	7830	245907	358411	578891
	<i>Nfg</i>	2778	2109	14792	491831	592019	1158354

Как следует из таблицы, субградиентные методы успешно справляются с поставленной задачей, не уступая в скорости сходимости методам, предназначенным для минимизации гладких функций.

В результате проделанных вычислительных экспериментов выяснилось также, что используемый для решения различных разностных задач высокой размерности, аппроксимирующих системы уравнений Навье–Стокса (3)–(4), многошаговый метод РСММ успешно справился с их ре-

шением за разумное время и с достаточно высокой точностью, что выявлено посредством сравнения с известными результатами.

### Заключение

Идея многошаговых алгоритмов заключается в привлечении дополнительной обучающей информации и формализуется в виде ортогонализации последовательных субградиентов. С одной стороны, это ускоряет сходимость релаксационных субградиентных методов минимизации, строящихся на многошаговых алгоритмах решения неравенств, с другой, в силу отсутствия в их схеме матриц метрики, позволяет решать задачи высокой размерности в сравнении с методами с растяжением пространства.

### Литература

1. Демьянов, В. Ф. Недифференцируемая оптимизация / В. Ф. Демьянов, Л. В. Васильев. М.: Наука, 1972. 368 с.
2. Крутиков В. Н. Обучающиеся методы безусловной оптимизации и их применение / В. Н. Крутиков. Томск: Изд-во Том. гос. пед. ун-та, 2008. 264 с.
3. Крутиков В. Н. Многошаговый субградиентный метод для решения негладких задач минимизации высокой размерности / В. Н. Крутиков, Я. Н. Вершинин // Вестник Томского государственного университета. Математика и механика. 2014. Вып. 3 (29). С. 5–19.
4. Крутиков В. Н. Субградиентный метод минимизации с коррекцией векторов спуска на основе пар обучающих соотношений / В. Н. Крутиков, Я. Н. Вершинин // Вестник КемГУ. 2014. Вып. 1 (57), т. 1. С. 46–55.
5. Вершинин Я. Н. О решении субградиентными методами регуляризованной задачи линейного программирования в системе экологического мониторинга / Я. Н. Вершинин, А. А. Быков, В. Н. Крутиков, В. В. Мешечкин // Вестник КемГУ. 2014. Вып. 1 (57), т. 1. С. 35–41.
6. Захаров Ю. Н. О решении нелинейных задач гидродинамики методами оптимизации / Ю. Н. Захаров, В. Н. Крутиков, Я. Н. Вершинин, А. И. Ерунов // Информационные технологии и математическое моделирование: матер. XII Всерос. науч.-практ. конф. с междунар. участием. Томск: Изд-во Том. ун-та, 2013. Ч. 1. С. 22–25.
7. Гилл Ф. Практическая оптимизация / Ф. Гилл, У. Мюррей, М. Райт. М.: Мир, 1985. 509 с.
8. Поляк Б. Т. Введение в оптимизацию / Б. Т. Поляк. М.: Наука, 1983. 384 с.
9. Шор Н. З. Методы минимизации недифференцируемых функций и их приложения / Н. З. Шор. Киев: Наук. думка, 1979. 199 с.

# МНОГОШАГОВЫЙ МЕТОД МИНИМИЗАЦИИ НЕГЛАДКИХ ФУНКЦИЙ ПРИ ИЗВЕСТНОМ ЗНАЧЕНИИ МИНИМУМА

*В. Н. Крутиков, Н. С. Самойленко*

*Кемеровский государственный университет, Кемерово, Россия*

## Введение

Во многих случаях функции, которые нужно минимизировать, являются недифференцируемыми. Подобные функции возникают в задачах декомпозиции, при использовании штрафных функций, в задачах наилучшего приближения и многих других задачах минимизации. Субградиентные методы являются расширением градиентных методов на недифференцируемые функции [1–3].

Градиентные методы линейной алгебры [4, 5], с точки зрения дальнейшего анализа их аналогий в методах оптимизации, можно подразделить на два вида. В первом из них минимизируется непосредственно квадратичная функция, т.е. соблюдается релаксационность метода относительно минимизируемой функции. К методам этого типа относятся метод скорейшего спуска и метод сопряженных градиентов. Второй вид методов обеспечивает релаксационность убывания расстояния до экстремума. К таким методам линейной алгебры можно отнести метод минимальных ошибок и метод минимальных итераций [4, 5].

Градиентные аналоги релаксационных по функции методов линейной алгебры получили широкое развитие в оптимизации [1–3, 6–8] и, в частности, в недифференцируемой оптимизации [1–3, 6–15]. Субградиентные методы, основанные на минимизации расстояния до экстремума [1–2], в оптимизации появились независимо от методов линейной алгебры, и аналогии с линейной алгеброй в этих методах не устанавливались. Установление аналогий и заимствование полезно сказываются на развитии теории минимизации. В свое время это позволило получить метод сопряженных градиентов и развить квазиньютоновские методы [2].

В этой работе мы преобразуем метод минимальных итераций в многошаговый метод минимизации, в котором операции с матрицами будут заменены на операции с характеристиками минимизируемой функции. Полученный в работе метод принадлежит к классу методов негладкой оптимизации, основное свойство которых заключается в сокращении расстояния до экстремума минимизируемой функции  $f(x)$ ,  $x \in R^n$ .

Итерация простого субградиентного метода имеет вид

$$x_{k+1} = x_k - \gamma_k g(x_k).$$

Здесь  $x_k$  – текущее приближение на  $k$ -й итерации;  $\gamma_k$  – шаговый множитель;  $g(x_k)$  – произвольный субградиент субградиентного множества  $\partial f(x_k)$  в точке  $x_k$ , т.е.  $g(x_k) \in \partial f(x_k)$ . В работах Б. Т. Поляка [1, 2] предложен и исследован вариант субградиентного метода минимизации для вы-

пуклой, не обязательно дифференцируемой функции, в котором длина шага вычисляется (а не находится подбором) на основании известного минимального значения функции  $f^*$ :

$$x_{k+1} = x_k - \frac{\gamma [f(x_k) - f^*]}{(g(x_k), g(x_k))} g(x_k). \quad (1)$$

Здесь  $\gamma$  – шаговый множитель, определяемый свойствами минимизируемой функции  $f(x)$ ;  $g(x_k)$  – субградиент функции в точке  $x_k$ ;  $f^*$  – минимальное значение функции  $f(x)$ , которое считается известным. Представленный алгоритм (1) является аналогом метода минимальных ошибок в оптимизации [16, 17]. Аналогично (1) мы сформулируем многошаговый субградиентный метод – аналог метода минимальных итераций, применимый для минимизации произвольных выпуклых функций.

### Аналог метода минимальных итераций в оптимизации

Рассмотрим метод сопряженных градиентов для решения системы уравнений

$$Bx = b, \quad (2)$$

где  $B$  – невырожденная матрица. В силу применимости метода для решения систем уравнений с симметричной и положительно определенной матрицей образуем вспомогательную систему уравнений посредством второй трансформации Гаусса, где после замены  $B^T y = x$  получаем

$$BB^T y = b. \quad (3)$$

Применим метод сопряженных градиентов для системы уравнений (3) относительно  $y$ :

$$y_{k+1} = y_k - \gamma_k s_k, \quad (4)$$

$$s_k = r_k + b_k s_{k-1}, \quad (5)$$

$$\gamma_k = \arg \min_{\gamma > 0} f(x_k - \gamma s_k) = \frac{(s_k, r_k)}{(s_k, A s_k)} = \frac{(r_k, r_k)}{(A s_k, r_k)}, \quad (6)$$

$$b_k = -\frac{(r_k, A s_{k-1})}{(A s_{k-1}, s_{k-1})} = \frac{(r_k, r_k)}{(r_{k-1}, r_{k-1})} > 0. \quad (7)$$

Здесь матрица  $A = BB^T$ ,  $r = Bx - b$ . Введем квадратичную функцию

$$f(x) = \frac{(Bx - b, Bx - b)}{2} + f^* = \frac{(r, r)}{2} + f^*, \quad (8)$$

точка минимума  $x^*$  которой совпадает с решением системы (2). Ее градиент  $g(x) = \nabla f(x) = B^T (Bx - b) = B^T r$ .

Обозначим  $B^T s_k = p_k$ . Так как  $B^T y = x$ , умножая на  $B^T$  (4), (5), получим расчетные формулы в старой системе координат:

$$x_{k+1} = x_k - \gamma_k p_k, \quad (9)$$

$$p_k = g_k + b_k p_{k-1}. \quad (10)$$

Используя равенства  $A = BB^T$ ,

$$\begin{aligned}(As_k, r_k) &= (BB^T s_k, r_k) = (B^T s_k, B^T r_k) = (p_k, g_k), \\ (As_k, s_k) &= (BB^T s_k, s_k) = (B^T s_k, B^T s_k) = (p_k, p_k), \\ (r_k, r_k) &= 2(f(x_k) - f^*)\end{aligned}$$

преобразуем выражения (6) и (7):

$$\gamma_k = \frac{(r_k, r_k)}{(As_k, r_k)} = \frac{(r_k, r_k)}{(p_k, g_k)} = \frac{(r_k, r_k)}{(p_k, p_k)} = \frac{2(f(x_k) - f^*)}{(p_k, g_k)} = \frac{2(f(x_k) - f^*)}{(p_k, p_k)}, \quad (11)$$

$$b_k = -\frac{(r_k, As_{k-1})}{(As_{k-1}, s_{k-1})} = -\frac{(g_k, p_{k-1})}{(p_{k-1}, p_{k-1})} > 0. \quad (12)$$

Расчетные формулы метода минимальных итераций можно записать в виде

$$x_{k+1} = x_k - \gamma_k p_k, \quad \gamma_k = \frac{(r_k, r_k)}{(p_k, g_k)} = \frac{(r_k, r_k)}{(p_k, p_k)}, \quad (13)$$

$$p_k = g_k + b_k p_{k-1}, \quad b_k = -\frac{(g_k, p_{k-1})}{(p_{k-1}, p_{k-1})} > 0. \quad (14)$$

Отметим, что в (14) в явном виде производится ортогонализация векторов  $p_k$ .

Пусть требуется найти минимум функции  $f(x)$ . Мы построим метод, который в отличие от (13), (14) использует только характеристики функции и в случае квадратичных функций реализует метод минимальных итераций, который, как известно, сходится, как и метод сопряженных градиентов, за конечное число шагов. Отметим, что мы предполагаем, что информация о матрице, как это было при решении систем линейных уравнений, нам не доступна. В формулах метода минимальных итераций (13)–(14) содержится следующий ниже метод, который использует только значения функции  $f(x)$ , градиента  $\nabla f(x)$  или субградиента, который мы будем иногда также обозначать  $g$ ,  $g(x)$  или  $g_f$ ,  $g_f(x)$ , и ее минимальное значение  $f^*$ . Расчетные формулы эквивалентны формулам (13), (14):

$$x_{k+1} = x_k - \gamma_k p_k, \quad (15)$$

$$p_k = g_k + \beta_k p_{k-1}, \quad (16)$$

$$\gamma_k = \gamma \frac{(f(x_k) - f^*)}{(p_k, g_k)} = \gamma \frac{(f(x_k) - f^*)}{(p_k, p_k)}, \quad (17)$$

$$\beta_k = -\frac{(g_k, p_{k-1})}{(p_{k-1}, p_{k-1})} > 0. \quad (18)$$

Выражение (17) запишем в более общем виде:

$$\gamma_k = \gamma \frac{(f(x_k) - f^*)}{(p_k, g_k)} = \gamma \frac{(f(x_k) - f^*)}{(p_k, p_k)}. \quad (19)$$

Для квадратичных функций величина  $\gamma = 2$ .

Приведенный метод не будет работоспособным, поскольку неравенство в (18) не всегда выполняется для функций, отличных от квадратичных. Приведем его модификацию, которая будет сходиться на квадратичных функциях при  $\gamma = 2$  за конечное число шагов и в то же время будет пригоден для минимизации выпуклых функций

$$x_{k+1} = x_k - \gamma_k p_k, \quad (20)$$

$$p_k = g_k + \beta_k p_{k-1}, \quad (21)$$

$$\gamma_k = \gamma \frac{(f(x_k) - f^*)}{(p_k, p_k)}, \quad (22)$$

$$\beta_k = \begin{cases} -\frac{(g_k, p_{k-1})}{(p_{k-1}, p_{k-1})}, & \text{при } (g_k, p_{k-1}) < 0, \\ 0, & \text{в противном случае.} \end{cases} \quad (23)$$

Для квадратичных функций условие в (23) не разрушает процесса (15)–(18) поскольку для них всегда выполняется неравенство в (18). Параметр  $\gamma$  определяется свойствами функций и задается пользователем. Для квадратичных функций  $\gamma = 2$ .

Таким образом, в данной работе метод минимальных итераций представлен в терминах минимизируемой функции, что позволяет его использовать для минимизации выпуклых функций при известном значении минимума. Для квадратичных функций величина шагового множителя в (19)  $\gamma = 2$ . При  $\gamma = 1$  метод (20)–(23) применяется для решения выпуклых задач минимизации общего вида. Выбор величины шагового множителя  $\gamma$  в алгоритме (19) определяется свойствами функции.

#### Литература

1. Поляк Б. Т. Минимизация негладких функционалов // Журн. вычислит. математики и матем. физики. 1969. Т.9, № 3. С. 507–521.
2. Поляк Б. Т. Введение в оптимизацию / Б. Т. Поляк. М.: Наука, 1983. 384 с.
3. Шор Н. З. Методы минимизации недифференцируемых функций и их приложения / Н. З. Шор. Киев: Наук. думка, 1979. – 199 с.
4. Фадеев Д.К. Вычислительные методы линейной алгебры / Д. К. Фадеев, В. Н. Фадеева. М.: Физматгиз, 1960. 656 с.
5. Воеводин В. В., Кузнецов Ю. А. Матрицы и вычисления. М.: Наука, 1984. 329 с.
6. Демьянов В. Ф. Недифференцируемая оптимизация / В. Ф. Демьянов, Л. В. Васильев. М.: Наука, 1981. 384 с.
7. Немировский А. С. Сложность задач и эффективность методов оптимизации / А. С. Немировский Д. Б. Юдин. М.: Наука, 1979. 340 с.
8. Крутиков В. Н. Обучающиеся методы безусловной оптимизации и их применение / В. Н. Крутиков. Томск: Изд-во Том. гос. пед. ун-та, 2008. 264 с.
9. Крутиков В. Н. Релаксационные методы безусловной оптимизации, основанные на принципах обучения / В. Н. Крутиков. Кемерово: КемГУ, 2004. 171 с.
10. Крутиков В. Н. Релаксационный метод минимизации с растяжением пространства в направлении субградиента / В. Н. Крутиков, Т. В. Петрова // Экономика и мат. методы. 2003. Т. 39, вып. 1. С. 33–49.

11. Крутиков В. Н. Семейство релаксационных субградиентных методов с двухранговой коррекцией матриц метрики / В. Н. Крутиков, Т. А. Горская // Экономика и мат. методы. 2009. Т. 45, № 4. С. 37–80.
12. Крутиков, В. Н. Алгоритмы обучения на основе ортогонализации последовательных векторов / В. Н. Крутиков, Я. Н. Вершинин // Вестник КемГУ. 2012. Вып. 2 (50). С. 37–42.
13. Крутиков В. Н. Многошаговый субградиентный метод для решения негладких задач минимизации высокой размерности / В.Н. Крутиков, Я.Н. Вершинин // Вестник Томского государственного университета. Сер. Математика и механика. 2014. № 3. С. 5–19.
14. Крутиков В. Н. Субградиентный метод минимизации с коррекцией векторов спуска на основе пар обучающих соотношений / В. Н. Крутиков, Я. Н. Вершинин // Вестник КемГУ. 2014. Вып. 1 (1). С. 46–54.
15. Крутиков В. Н. Методы оптимизации: учеб. пособие с грифом КемГУ / В. Н. Крутиков. Кемерово, 2011. 93 с.
16. Самойленко Н. С., Крутиков В. Н., Мешечкин В. В. Об аналогии между методом минимальных ошибок и субградиентным методом Поляка // Научное творчество молодежи. Математика. Информатика // Материалы XVIII Всероссийской научно-практической конференции 24–25 апреля 2014 г. Ч. 1. Томск: Изд-во Том. ун-та, 2014. С. 67–68.
17. Самойленко Н. С., Крутиков В. Н., Мешечкин В. В. Об оценке сходимости субградиентного метода // Образование, наука, инновации: вклад молодых исследователей: матер. IX (XLI) Междунар. науч. конф. студентов, аспирантов и молодых ученых. Кемерово, 2014. Вып. 15. С. 1518–1519.

## СИНТЕЗ И РЕАЛИЗАЦИЯ РЕЛЕЙНОГО УПРАВЛЕНИЯ СИСТЕМАМИ ТЕПЛОПОТРЕБЛЕНИЯ

*А. А. Шилин, В. Г. Букреев*

*Национальный исследовательский Томский политехнический университет,  
Томск, Россия*

### Введение

В работе рассматривается актуальная задача синтеза и реализации трехпозиционного релейного управления сложными динамическими объектами, на примере систем теплоснабжения зданий и сооружений. Линеаризованная математическая модель такого объекта в точке равновесия относительно положения штока регулирующего клапана в операторной форме может быть записана следующим образом [1]:

$$W(s) = \frac{k_o(t, x)}{T_v s} \cdot \frac{\exp(-\tau(t, x)s)}{1 + T_o(t, x)s}, \quad (1)$$

где  $T_v$  – длительность перемещения штока клапана;  $s$  – оператор преобразования Лапласа. Длительность  $\tau(t, x)$  транспортного запаздывания, коэффициент передачи  $k_o(t, x)$  и постоянная  $T_o(t, x)$  времени объекта имеют сложные зависимости от переменных  $x(t)$  состояния и обычно являются нестационарными.

В качестве альтернативных направлений рассмотрены методы робастного управления в скользящем режиме и оптимального по быстродействию управления. Основываясь на результатах исследований систем робастного управления в скользящем режиме [2] можно использовать условие притяжения к траектории скольжения в граничных режимах и предположить существование оптимальных по быстродействию режимов. Так, объект (1) без учета запаздывания является частным случаем математической модели

$$\begin{cases} \dot{x}_1(t) = x_2(t), \\ \dot{x}_2(t) = -h(x,t) \cdot x_2 + g(x,t) \cdot u(t), \end{cases} \quad (2)$$

где  $g(x,t)$  и  $h(x,t)$  – положительно определенные функциональные параметры объекта. Для этой модели существует управление

$$u = -\text{sign}(a_1 \cdot x_1 + x_2), \quad (3)$$

где  $a_1$  – параметр, определяющий траекторию скольжения. Известно условие

$$\left| \frac{a_1 \cdot (x_2 + h(x))}{g(h)} \right| \leq 1, \quad (4)$$

при котором обеспечивается притяжение к траектории скольжения [3].

На рис. 1 представлен случай, когда параметр траектории скольжения подобран из условия притяжения при повторном достижении траектории движения объекта в фазовой плоскости.

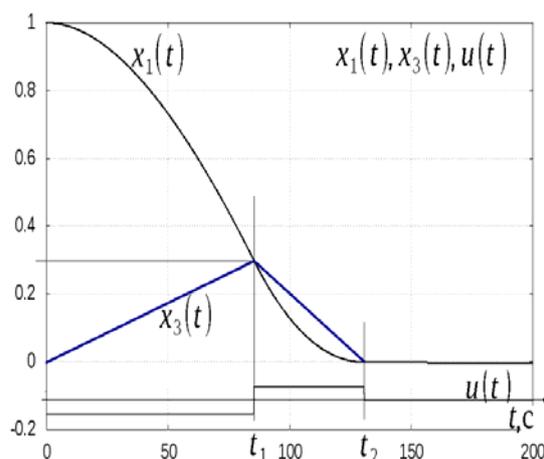
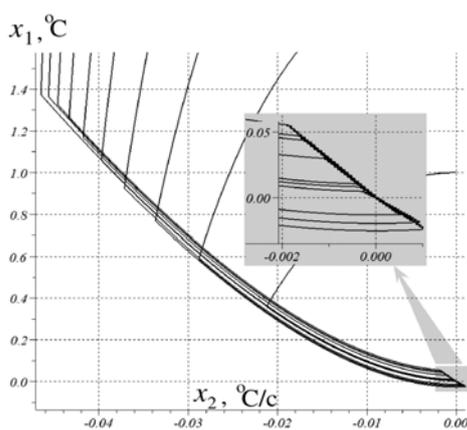


Рис. 1. Оптимальная по быстродействию траектория для управления в скользящем режиме

Рис. 2. Оптимальный по быстродействию переходный процесс

Такая реализация регулятора с переменной структурой соответствует оптимальному управлению объектом второго порядка (1) при условии, что

второе пересечение траектории скольжения обеспечивается в точке равновесия системы.

Известны теоретические подходы [4], позволяющие выполнить операцию дифференцирования на релейном элементе с интегрирующей обратной связью без жестких требований к качеству измерения. Покажем, что программная реализация релейного элемента с обратной связью создает хорошие предпосылки для реализации оптимального управления и имеет значительный ресурс развития.

Введем дополнительную переменную состояния  $x_3$ , физически связанную с переменными  $x_1$  и  $x_2$  через управление  $u$ , в математическую модель (2) объекта второго порядка. Далее для вычисления управления вместо (3) воспользуемся уравнениями

$$\begin{cases} \dot{x}_3 = \begin{cases} -a_3 \cdot x_3, & \text{если } (u = 0), \\ -k_3 \cdot u, & \text{если } (u \neq 0), \end{cases} \\ u = \eta((x_3 - x_1), \delta), \end{cases} \quad (5)$$

где сигнатурная функция заменена функцией  $\eta(\dots)$  релейного элемента с зоной гистерезиса  $\delta$ .

В работе [5] показано, что для замкнутой системы (1) и (4) возможно существование траектории скольжения вида

$$s = a_s(t, x(t_0)) \cdot x_1 + x_2 = 0, \quad (6)$$

где  $a_s(t, x(t_0)) > 0 \forall t \in (t \geq t_0)$ , если выполнены необходимые условия

$$\begin{aligned} x_1(t_0) &= x_3(t_0), \\ a_3 &\geq a_s(t). \end{aligned} \quad (7)$$

Авторами показано, что для объекта (1) при стационарных параметрах  $g$ ,  $h$  и известном начальном состоянии  $\{x_1(t_0) = C_1 \in (C_{\min}, C_{\max}); x_2 = 0\}$  возможно организовать оптимальное по быстродействию управление [4]. Доказательство основано на теореме о количестве переключений управляющего воздействия для устойчивого объекта второго порядка, в соответствии с которой обеспечивается оптимальное управление двумя разнополярными воздействиями. Решение замкнутой системы управления объекта (1) и регулятора (2) отражается в виде переходных процессов на рис. 2. Согласно организованной программным образом отрицательной обратной связи релейного элемента траектория переменной  $x_3$  всегда стремится к траектории переменной  $x_1$ .

Скорость изменения переменной  $x_3$  определяется функцией  $k_3(x, u)$ , обеспечивающей пересечение переменных  $x_1$  и  $x_3$  в точке переключения оптимального управления по быстродействию. В частности, для заданного начального состояния ( $x_1 = 1, x_2 = 0$ ) объекта управления функция  $k_3(x, u)$  равна абсолютному значению тангенсов углов между фазовой траекторией  $x_3$  и осью времени для двух случаев, определяемых знаком произведения  $x_1$

и  $u = \text{sign}(u \cdot x_1)$ . Очевидно, задача решается для любого начального положения ( $x_1=x_1(t_0)$ ,  $x_2=0$ ). По результатам анализа процессов для организации оптимального управления по быстродействию предлагается метод подстройки параметра траектории скольжения. Запишем управление (2) с учетом зависимости перенастраиваемого параметра  $k_3$  от начального состояния и знака:

$$\begin{cases} \dot{x}_3 = \begin{cases} -a_3 \cdot x_3, & \text{если } (u = 0), \\ -\frac{u}{T_{p1}(C_1)}, & \text{если } (\text{sign}(u \cdot x_1) > 0), \\ -\frac{u}{T_{p2}(C_1)}, & \text{если } (\text{sign}(u \cdot x_1) < 0), \end{cases} \\ u = \eta((x_3 - x_1), \delta), \end{cases} \quad (8)$$

где зависимости  $T_{p1}(C_1)$  и  $T_{p2}(C_1)$  определяются для различных начальных условий из заданного множества  $C_1 \in (C_{\min}, C_{\max})$ . Алгоритм определения зависимостей приведен в работе [6].

Параметрическая связь объекта и регулятора предполагает вычисление параметров  $T_{p1}(C_1, g, h)$  и  $T_{p2}(C_1, g, h)$  в многомерном пространстве. В этом случае целесообразно применить замену переменных  $T_o = h^{-1}$ ;  $T_v = h \cdot g^{-1}$  и рассматривать зависимости в виде

$$C_{\max} = \frac{k_{cmx} T_o}{T_v}, k_{cmx} \approx 4.49, T_{p2} = \left( k_{p20} + k_{p21} \frac{C_{\max}}{C_1} \right) T_v, T_{p1} = \left( k_{p10} + k_{p11} \frac{C_1}{C_{\max}} \right) T_{p2}, \quad (9)$$

где  $k_{p10}$ ,  $k_{p11}$ ,  $k_{p20}$ ,  $k_{p21}$  – коэффициенты аппроксимации.

Данная реализация не может быть классифицирована в рамках существующих общепринятых ПИД- или  $LQR$ -методов управления, так как в предложенном алгоритме наряду с дифференциальными составляющими имеются программные решения и сложный релейный элемент. Также рассмотренный алгоритм не может быть определен как регулятор с переменной структурой, так как вычисления выполняются не в фазовой плоскости. При этом траектории переменных  $x_1$  и  $x_3$  на временной плоскости (рис. 2) напоминают греческий символ « $\lambda$ », соответственно предлагается такой регулятор именовать как  **$\lambda$ -регулятор**.

**Моделирование.** Для сравнения выбираем  $LQR$ -регулятор [7], где на этапе моделирования приняты следующие допущения: все переменные состояния линеаризованной системы [8] доступны для измерения. Значения управления для  $LQR$ -регулятора в отличие от релейного управления могут включать все множество значений в диапазоне  $[-1, +1]$ . В решении уравнения Риккати весовая матрица  $Q$  является единичной

матрицей, а весовая матрица  $R$  представляет собой скалярный элемент, значение которого подбирается из условия ограничения на амплитуду управления. Результаты моделирования линеаризованной системы с  $LQR$ - и  $\lambda$ -регулятором представлены на рис. 3.

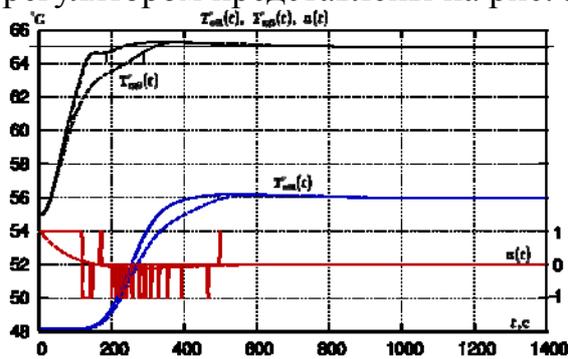


Рис. 3. Результаты моделирования переходных процессов системы теплотребления с  $LQR$ - и  $\lambda$ -регуляторами

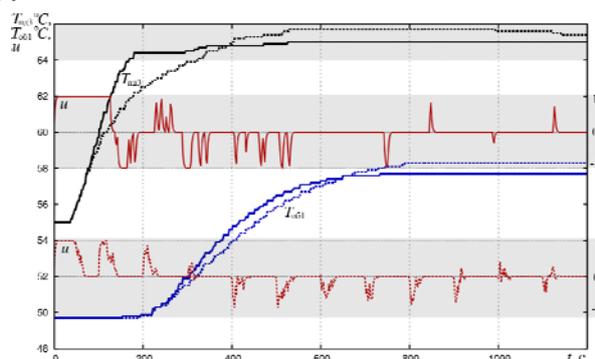


Рис. 4. Оптимальное управление с релейным и ПИ-регуляторами

**Исследование на действующих системах.** Работа алгоритма  $\lambda$ -регулятора сравнивается с работой  $LQR$ -регулятора, который с учетом астатизма объекта реализует ПИ-закон регулирования. Для согласования с трехпозиционным управлением используется широтно-импульсная модуляция. В результате исследований преимущества  $\lambda$ -регулятора можно наблюдать на рис. 4, где отчетливо выражены два импульса воздействия в начале переходного процесса.

### Заключение

Для динамических объектов с астатизмом и трехпозиционным управлением использование  $\lambda$ -регулятора дает преимущества в быстродействии и робастных свойствах замкнутой системы по сравнению с традиционными методами реализации ПИД-регулирования. Синтез  $\lambda$ -регулятора основан на принципах организации скользящих режимов, обладающих свойствами оптимального управления по быстродействию.

Полученный алгоритм реализации  $\lambda$ -регулятора успешно используется в течение нескольких лет на более чем 100 объектах теплоснабжения. В результате алгоритм  $\lambda$ -регулятора может быть реализован в ПЛК в качестве альтернативы ПИД-регулятору в системах с трехпозиционным управлением.

### Литература

1. Шилин А. А. Упрощение модели сложной теплообменной системы для решения задачи релейного управления / А. А. Шилин, В. Г. Букреев // Теплоэнергетика. 2014. № 9. С. 56–63.
2. Уткин В. И. Скользящие режимы в задачах оптимизации и управления / В. И. Уткин. М.: Наука, 1981. 368 с.
3. Khalil Hassan K. Nonlinear Systems, Third Edition. Prentice Hall, 2002. 750 p.

4. Ключев А. С. Оптимизация автоматических систем управления по быстрдействию / А. С. Ключев, А. А. Колесников. М.: Энергоиздат, 1982. 240 с.
5. Шилин А. А. Исследование оптимального и скользящего режимов управления с релейным элементом / А. А. Шилин, В. Г. Букреев // Вестник Том. гос. ун-та. Управление, вычислительная техника и информатика. 2014. № 3(28). С. 12–19.
6. Шилин А. А. Оптимальное управление динамическим объектом на основе аппроксимации параметров релейного регулятора / А. А. Шилин, В. Г. Букреев // Приборы и системы. Управление, контроль, диагностика. 2014. № 10. С. 73–79.
7. Anderson B. D. O., Moore J. B., Optimal Control: Linear Quadratic Methods. Prentice Hall, 1990. 368 p.
8. Шилин А. А. Линеаризация нелинейной модели теплообменной системы с аппроксимацией транспортного запаздывания / А. А. Шилин, В. Г. Букреев // Теплоэнергетика. 2014. № 10. С. 49–54.

## **РЕШЕНИЕ ЗАДАЧИ ОРТОГОНАЛЬНОГО РАСКРОЯ НА ОСНОВЕ ГРУППОВОГО ДЕКОДЕРА И ГЕНЕТИЧЕСКОГО АЛГОРИТМА**

***В. С. Шилов***

*Пермский национальный исследовательский  
политехнический университет, Пермь, Россия*

Задачи раскроя-упаковки возникают во множестве отраслей промышленности: от текстильной до машиностроения. В то же время для данного класса задач не найдено полиномиального алгоритма решения, поскольку они относятся к классу NP-трудных [1–3]. Таким образом, разработка новых алгоритмов решения является важной и актуальной задачей.

Среди всех задач раскроя-упаковки наиболее часто встречаются задачи ортогонального раскроя. В статье описываются новые алгоритмы решения данной задачи.

Будем рассматривать задачу раскроя полубесконечной полосы *1.5-DBPP* [3]. Задана полоса, имеющая ширину  $W$  и бесконечную длину, а также  $n$  деталей размерами  $w_i \times l_i$ . Детали должны быть размещены на полосе таким образом, чтобы занятая длина полосы была минимальна, а детали не пересекались друг с другом и краями полосы.

Хорошо зарекомендовавшим себя подходом в задачах ортогонального раскроя является использование генетических алгоритмов. Сам по себе генетический алгоритм позволяет получить лишь закодированное решение, поэтому совместно с ним необходимо использовать алгоритм-декодер, который генерирует по закодированному решению карту раскроя.

Большим недостатком существующих алгоритмов-декодеров является то, что они размещают детали по одной, это в итоге приводит к образованию лишних свободных пространств. Поэтому предлагается алгоритм, выбирающий группу деталей, которая позволяет максимально заполнить текущий рассматриваемый блок. Данный алгоритм может быть записан в виде следующей последовательности действий:

1. Текущий вертикальный блок  $C=1$ .
2. Если нет упакованных деталей, то КОНЕЦ.
3. Пока нет нерассмотренных пустых участков в блоке  $C$ , выполнять  $C=C+1$ .
4. Выбрать следующий нерассмотренный участок в блок  $C$ .
5. Найти группу объектов, максимально заполняющую текущий участок.
6. Если группа не найдена, то перейти на шаг 2.
7. Разместить группу объектов в текущем блоке.
8. Перейти на шаг 2.

Для выбора группы объектов на шаге 5 необходимо решить задачу о рюкзаке. Поскольку на практике размеры объектов являются целыми числами, а ширина листа составляет порядка нескольких тысяч миллиметров, точное решение данной задачи может быть найдено с помощью динамического программирования. Единственное отличие данной задачи от классической задачи о рюкзаке заключается в том, что деталь может быть повернута на 90 градусов. Таким образом, конструирование решения производится по следующему рекуррентному соотношению:

$$f(h, k) = \begin{cases} \text{False,} & \text{при } h < 0 \\ \text{False,} & \text{при } l \neq 0 \text{ и} \\ & k = 0, \\ \text{True,} & \text{при } h = 0 \\ f(h - w_k, k - 1) \vee & \text{при } k > 0 \text{ и} \\ f(h - l_k, k - 1) \vee & h \geq 0, \\ f(h, k - 1) & \end{cases}$$

где  $h$  – текущая суммарная высота;

$k$  – количество уже рассмотренных на текущем шаге объектов;

$h_k$  – высота  $k$ -го объекта;

$l_k$  – длина  $k$ -го объекта.

Данная задача может иметь несколько решений. Какое именно решение будет выбрано, зависит от того, в каком порядке детали поступают на вход алгоритма. Для определения такого порядка принято использовать метаэвристики. В данной статье рассматривается генетический алгоритм.

Генетический алгоритм осуществляет поиск такой перестановки деталей, при которой достигается наилучшее значение коэффициента раскроя (ГОСТ 27782–88). При этом, если в задании на раскрой есть повторяющиеся детали, некоторые перестановки окажутся равнозначными. Так, если есть 10 групп деталей, каждая из которых состоит из 10 одинаковых деталей, то всего существует  $100!$  перестановок, в то время как различных ва-

риантов всего  $\frac{100!}{(10!)^{10}}$ . Поэтому предлагается в генетическом алгоритме

осуществлять поиск перестановок с повторениями. Для этого необходимо разработать новый алгоритм скрещивания и уточнить операции мутации.

Одним из наиболее популярных алгоритмов скрещивания, работающих с перестановками, является частично-отображаемое скрещивание [4]. Рассмотрим его подробнее, а после на его основе сформулируем алгоритм скрещивания для перестановок с повторениями.

Имеется две хромосомы, которые представляют собой перестановку целых чисел. Оператор должен осуществить их скрещивание таким образом, чтобы результатом были хромосомы, представляющие перестановку. Пусть есть две хромосомы:

$$\begin{array}{cccc|cc} p_1 & 1 & 3 & 4 & 2 & 5 \\ p_2 & 2 & 1 & 3 & 5 & 4 \end{array}$$

Линией показана точка, по которой происходит скрещивание. Оператор скрещивания выполняет копирование левой части первой хромосомы в результат; при этом запоминается отображение генов первой хромосомы на соответствующие гены второй, т.е. следующее соответствие 1->2, 3->1, 4->2. Затем копируется правая часть второй хромосомы в результат; если при этом происходит попытка записать в результат ген, который уже в нем есть, вместо данного гена берется ген из полученного соответствия. Аналогично выполняется операция для левой части второй хромосомы и правой части первой хромосомы. Таким образом, будет получен следующий результат операции скрещивания:

$$\begin{array}{cccc|cc} p_1' & 1 & 3 & 4 & 5 & 2 \\ p_2' & 2 & 1 & 3 & 4 & 5 \end{array}$$

Такой подход может быть использован и для перестановок с повторениями. В хромосоме будут храниться индексы деталей, причем каждый индекс должен встречаться ровно столько раз, сколько таких деталей в задании на раскрой. В такой ситуации необходимо строить не отображение гена на ген, а отображение гена на массив генов. А при генерации потомка обращаться к данному отображению только в том случае, если уже использовано ровно столько таких генов, сколько должно быть. Выбрав ген из массива, необходимо сдвигать указатель в массиве на следующий элемент. Рассмотрим данный алгоритм на следующем примере:

$$\begin{array}{cccc|cccc} p_1 & 1 & 2 & 3 & 1 & 2 & 3 & 1 \\ p_2 & 1 & 1 & 2 & 3 & 2 & 3 & 1 \end{array}$$

Для первой хромосомы будет использовано соответствие 1->{3,1}, 2->{3}; для второй – 3->{1,2}, 1->{1}. Результатом операции скрещивания будут следующие хромосомы:

$$\begin{array}{cccc|cccc} p_1' & 1 & 2 & 1 & 2 & 1 & 3 & 3 \\ p_2' & 3 & 3 & 1 & 1 & 2 & 2 & 1 \end{array}$$

Кроме того, требует уточнения оператор мутации. Наиболее распространенные операции мутации – одноточечная, двуточечная и операция инверсии. В случае одноточечной и двуточечной мутации выполняется простой обмен двух значений местами, а это значит, что такую операцию имеет смысл выполнять, только если меняемые местами значения различны. Операция инверсии выполняется, если рассматриваемый участок хромосомы не является палиндромом.

Оценка предложенных алгоритмов производилась на группе задач из работы Е. Хоппер [5]. В таблице приведено сравнение результатов предложенных алгоритмов с результатами, полученными в [5]. В первом столбце таблицы приведены результаты для группового декодера с классическим генетическим алгоритмом, во втором – для группового декодера с генетическим алгоритмом с перестановками с повторениями, в третьем – результаты, полученные Е. Хоппер. Для оценки использовалось отклонение полученного решения от оптимального (в процентах).

**Результаты вычислительных экспериментов**

Задача	Алгоритм 1	Алгоритм 2	Алгоритм 3
<b>C1</b>	4	<b>0</b>	4
<b>C2</b>	11	<b>4</b>	6
<b>C3</b>	6	<b>4</b>	4
<b>C4</b>	3	<b>2</b>	3
<b>C5</b>	2	<b>1</b>	3
<b>C6</b>	1,4	<b>0,8</b>	3
<b>C7</b>	1,1	<b>0,4</b>	4

Необходимо отметить, что чем больше номер задачи, тем больше деталей в задании на раскрой. Таким образом, можно сделать вывод, что групповой декодер с классическим генетическим алгоритмом в сравнении с алгоритмами Е. Хоппер проигрывает на малых задачах, но на больших показывает лучшие результаты. Применение же в генетическом алгоритме перестановки с повторениями позволяет значительно улучшить получаемые решения.

В статье предложено решение задачи ортогонального раскроя материала на основе группового декодера и перестановок с повторениями. Решение задачи о рюкзаке на каждом шаге группового декодера позволяет оптимально заполнять рассматриваемый блок. Применение перестановок с повторениями в генетическом алгоритме позволяет производить поиск по меньшему количеству вариантов, что увеличивает вероятность нахождения оптимального решения. Проведены вычислительные эксперименты, показавшие эффективность предложенных алгоритмов.

#### Литература

1. Мухачева А. С., Ширгазин Р. Р. Задачи упаковки прямоугольников: рандомизированная эвристика на базе двойственной схемы локального поиска оптимума // Информационные технологии. 2003. № 5. С. 18–22.

2. Мурзакаев Р. Т., Шилов В. С., Буркова А. В. Основные методы решения задачи фигурной нерегулярной укладки плоских деталей [Электронный ресурс] // Инженерный вестник Дона. 2013. No. 4. (<http://www.ivdon.ru/magazine/archive/n4y2013/2043>). Проверено: 26.05.2014.

3. Application of the Group Decoder for Solving the Orthogonal Materials Cutting Problem. [Электронный ресурс] / R. A. Fayzrakhmanov [и др.] // World Applied Sciences Journal. 2013. Vol. 10, no. 28. ([http://www.idosi.org/wasj/wasj28\(10\)13/4.pdf](http://www.idosi.org/wasj/wasj28(10)13/4.pdf)). Проверено 12.03.2015.

4. Емельянов В. В., Курейчик В. В., Курейчик В. М. Теория и практика эволюционного моделирования. М.: Физматлит, 2003.

5. Hopper E., Turton V. C. H. An empirical investigation of metaheuristic and heuristic algorithms for a 2D-packing problem // European Journal of Operational Research. 2000. No. 128(1). P. 34–57.

## СОДЕРЖАНИЕ

### ВЕРОЯТНОСТНЫЕ МЕТОДЫ И МОДЕЛИ

<i>Burkatovskaya Yu., Kabanova T., Vorobeychikov S.</i> CUSUM algorithms for parameter estimation in queueing systems with jump intensity of the arrival process.....	3
<i>Callegari C., Giordano S., Pagano M.</i> Statistical approaches to intrusion detection: an overview of selected methods.....	9
<i>Czachorski T., Domanska J., Pagano M.</i> On stochastic models of internet traffic.....	13
<i>Gulev S. K., Korolev V. U., Gorshenin A. K., Belyaev K. P.</i> The analysis of heat fluxes variability by moving separation of finite normal mixtures.....	18
<i>Бахолдина М. А., Горцев А. М.</i> Совместная плотность вероятностей длительностей интервалов модулированного обобщенного полусинхронного потока событий в условиях его неполной наблюдаемости.....	23
<i>Горцев А. М., Сиротина М. Н.</i> Совместная плотность вероятностей длительности интервалов модулированного синхронного потока событий в условиях неполной наблюдаемости.....	28
<i>Данилюк Е. Ю.</i> Вероятностный анализ для исследования европейского экзотического опциона на фондовый индекс.....	35
<i>Дмитриев Ю. Г., Тарасенко П. Ф.</i> О $k$ -адаптивных комбинированных оценках вероятности.....	40
<i>Киселева А. И.</i> Применение методов статистического анализа в исследовании влияния некоторых факторов на себестоимость сельскохозяйственной продукции в сельскохозяйственных предприятиях Кемеровской области.....	45
<i>Кошечкина Т. О., Дмитриев Ю. Г.</i> Комбинированные оценки доли.....	50
<i>Лившиц К. И., Ульянова Е. С.</i> Релейно-гистерезисное управление потоком моментов продажи скоропортящейся продукции.....	55
<i>Нежелская Л. А.</i> Плотность вероятностей длительности интервала модулированного МАР-потока событий в условиях непродлевающего мёртвого времени.....	60
<i>Симахин В. А., Черепанов О. С.</i> Адаптивная фильтрация изображений на основе взвешенного метода максимального правдоподобия.....	67

### МОДЕЛИ И МЕТОДЫ МАССОВОГО ОБСЛУЖИВАНИЯ

<i>Nazarov A., Chernikova Ya.</i> A $M^{(2)} M^{(2)} 1$ retrial queueing system with $r$ -persistent exclusion of alternative customers.....	72
<i>Samouylov K. E., Sopin E. S., Vikhrova O. G.</i> Analyzing blocking probability in LTE wireless Network via queueing system with finite amount of resources.....	76
<i>Vashkevich N. P., Zinkin S. A., Pashchenko D. V.</i> Discrete-event simulation of closed queueing systems based on logical-algebraic approach (on the example of resource using in infocommunication system).....	81

<b>Багирова С. А.</b> Об одной модели системы обслуживания-запасания с ограниченной очередью нетерпеливых заявок.....	86
<b>Баймеева Г. В., Назаров А. А.</b> Система $M / GI / \infty$ в полумарковской случайной среде.....	91
<b>Гайдамака Ю. В., Зарипова Э. Р.</b> Сравнительный анализ дисциплин обслуживания для систем поллинга с двумя очередями.....	96
<b>Даммер Д. Д.</b> Исследование математической модели страховой компании в виде системы массового обслуживания с неограниченным количеством приборов с учетом неявной рекламы.....	100
<b>Задиранова Л. А.</b> Исследование потока повторных обращений в СМО с повторным обслуживанием заявок методом асимптотического анализа.....	105
<b>Задиранова Л. А., Моисеева С. П.</b> Исследование потока суммарных обращений в системе $MMPP M \infty$ с повторным обслуживанием требований.....	110
<b>Задорожный В. Н.</b> Особенности имитационного моделирования фрактальных систем с очередями.....	114
<b>Зорин А. В., Зорин В. А.</b> О потоке повторных требований в тандеме циклических систем обслуживания в стационарном режиме.....	119
<b>Крук Ю. С., Дудовская Ю. Е.</b> Инвариантность стационарного распределения вероятностей состояний замкнутой сети с неактивными заявками и многорежимным обслуживанием.....	122
<b>Лисовская Е. Ю., Моисеева С. П.</b> Исследование времени ожидания обслуживания клиентов call-центра.....	127
<b>Любина Т. В., Азева В. О.</b> Исследование динамической RQ-системы с входящим MAP-потокком заявок.....	132
<b>Малинковский Ю. В.</b> Стационарное функционирование сетей массового обслуживания с групповым поступлением отрицательных заявок.....	137
<b>Моисеев А. Н., Назаров А. А.</b> Анализ многофазной системы обслуживания $SM/(GI/\infty)^K$ в условии высокой интенсивности входящего потока.....	142
<b>Науменко В. В., Маталыцкий М. А.</b> Исследование G-сетей с многолинейными системами и доходами.....	147
<b>Панкратова Е. В.</b> Исследование системы массового обслуживания $GI/GI/\infty$ с двумя типами заявок.....	152
<b>Пауль С. В.</b> Выходящий поток заявок в системе с прогулками.....	157
<b>Русилко Т. В.</b> Асимптотический анализ замкнутой структуры массового обслуживания с заявками двух классов.....	162
<b>Убонова Е. Г.</b> Исследование системы массового обслуживания $MR/M/\infty$ с разнотипным обслуживанием.....	168
<b>Фёдорова Е. А.</b> Асимптотический анализ второго порядка RQ-системы $MMPP M 1$ в условии большой загрузки.....	173

## ОПТИМИЗАЦИОННЫЕ МОДЕЛИ И ИССЛЕДОВАНИЕ ОПЕРАЦИЙ

<i>Mikhina O. O., Smagin V. I.</i> Dynamic locally optimal control of discrete state delay systems with random parameters.....	177
<i>Дягель А. В.</i> Исследование эффективности генетического алгоритма безусловной оптимизации.....	181
<i>Костюк Ю. Л.</i> Задача коммивояжера: субоптимальный алгоритм по методу ветвей и границ.....	185
<i>Костюк Ю. Л.</i> Табличный способ генерации обратной польской строки для LL(1)-грамматики.....	190
<i>Крутиков В. Н., Вершинин Я. Н.</i> Приложения многошаговых субградиентных алгоритмов.....	195
<i>Крутиков В. Н., Самойленко Н. С.</i> Многошаговый метод минимизации негладких функций при известном значении минимума.....	201
<i>Шилин А. А., Букреев В. Г.</i> Синтез и реализация релейного управления системами теплопотребления.....	205
<i>Шилов В. С.</i> Решение задачи ортогонального раскроя на основе группового декодера и генетического алгоритма.....	210

Научное издание

**Информационные технологии  
и математическое моделирование  
(ИТММ–2015)**

Материалы XIV Международной конференции  
имени А. Ф. Терпугова  
18–22 ноября 2015 г.  
Часть 1

Составитель Т. В. Любина  
Редактор В. Г. Лихачева

---

Подписано в печать 26.10.2015.

Формат 60×84<sup>1</sup>/<sub>16</sub>. Бумага офсетная № 1. Печать офсетная.

Печ. л. 13,6; усл. печ. л. 12,6; уч.-изд. л. 12,5. Тираж 300 экз. Заказ 1263

---

ООО «Издательство ТГУ», 634029, г. Томск, ул. Никитина, 4.

ООО «Новые Печатные Технологии», 634040, г. Томск, ул. Высоцкого, 28, стр. 1.