



Национальный исследовательский
ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

Национальный исследовательский
КАЗАНСКИЙ ТЕХНОЛОГИЧЕСКИЙ УНИВЕРСИТЕТ
РОССИЙСКИЙ УНИВЕРСИТЕТ ДРУЖБЫ НАРОДОВ

ИНСТИТУТ ПРОБЛЕМ УПРАВЛЕНИЯ
им. В.А. ТРАПЕЗНИКОВА РАН

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ И МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ (ИТММ-2017)

**МАТЕРИАЛЫ
XVI Международной конференции
имени А. Ф. Терпугова
29 сентября – 3 октября 2017 г.**

Часть 1



ТОМСК
«Издательство НТЛ»
2017

УДК 519
ББК 22.17
И74

И74 Информационные технологии и математическое моделирование
(ИТММ-2017): Материалы XVI Международной конференции
имени А.Ф. Терпугова (29 сентября – 3 октября 2017 г.). – Томск:
Изд-во НТЛ, 2017. – Ч. 1. – 318 с.

ISBN 978-5-89503-603-7

В часть 1 вошли материалы докладов по теории массового обслуживания, теории телетрафика и прикладному вероятностному анализу, представленные на XVI Международной конференции имени А.Ф. Терпугова «Информационные технологии и математическое моделирование».

УДК 519
ББК 22.17

Редколлегия:

С.П. Моисеева, доктор физико-математических наук, доцент,
А.А. Назаров, доктор технических наук, профессор,
С.П. Сущенко, доктор технических наук, профессор.

*Конференция проводится при поддержке
Российского фонда фундаментальных исследований
(проект № 17-07-20543-г)*

ISBN 978-5-89503-603-7

© Авторы. Текст, 2017
© Оформление. Дизайн.
ООО «Издательство НТЛ», 2017

МОДЕЛИ И МЕТОДЫ МАССОВОГО ОБСЛУЖИВАНИЯ

ASYMPTOTIC ANALYSIS OF RETRIAL QUEUE M/GI/1 WITH IMPATIENT CALLS UNDER HEAVY LOAD CONDITION*

Ekaterina Fedorova

Tomsk State University, Tomsk, Russia

Retrial queueing systems are mathematical model of real systems, where unserved calls perform repeated attempts to get a service after a random time. The main results and comprehensive description of retrial queues are contained in the books [1, 2].

In this paper, we use the asymptotic analysis method, that gives analytical result for different types of queueing systems and networks, in particular with non Poisson arrivals [3, 4]. Here we expand results obtained previously for retrial queues without loss (patience time equals infinity), where the gamma form of the probability distribution of number of calls in the orbit under heavy load limit condition was proved [5, 6].

Mathematical model

Let us consider a retrial queueing system of M/GI/1 type with impatient calls. The structure of the model is presented in Figure 1.

The input process is Poisson with a rate λ . There is one server with the service time distribution function $B(x)$. If a call arrives when the server is free, the call occupies it for the service. Otherwise, the call goes to the orbit, where it stays during a random time distributed exponentially with a rate σ . After the delay, the call makes an attempt to reach the server again. If it is free, the call occupies it, otherwise the call instantly returns back to the orbit. From the orbit calls can leave the system after a random time distributed

* The publication was financially supported by RFBR according to the research project No.16-31-00292mol-a.

exponentially with dynamical rate α/i , where i is a number of calls in the orbit at this moment.

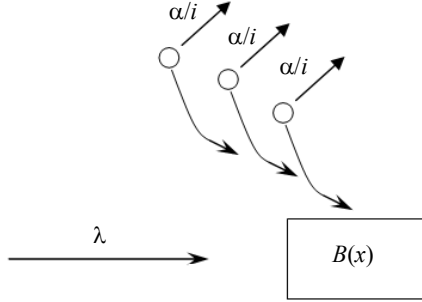


Fig. 1. Retrial queue M/GI/1 with impatient calls

Denote the random process that describes the number of calls in the orbit as $i(t)$, the random process of remaining service time as $z(t)$, and the random process that defines server states as $k(t) = \{0, \text{if the server is free}; 1, \text{if the server is busy}\}$.

The problem is to find the probability distribution of the number of calls in the orbit. The process $i(t)$ is not Markovian, therefore we consider the multidimensional process $\{k(t), i(t), z(t)\}$.

Denote the stationary probability distribution of the system states as $P(k, i, z) = P\{k(t) = k, i(t) = i, z(t) < z\}$.

Considered process is Markovian, so the following system of Kolmogorov equations for $P(k, i, z)$ can be written:

$$\begin{cases} \frac{\partial P_1(0,0)}{\partial z} - \lambda P_0(0) + \alpha P_0(1) = 0, \\ \frac{\partial P_1(0,z)}{\partial z} - \frac{\partial P_1(0,0)}{\partial z} - \lambda P_1(0,z) + \sigma P_0(1)B(z) + \lambda P_0(0)B(z) + \alpha P_1(1,z) = 0, \\ \frac{\partial P_1(i,0)}{\partial z} - (\lambda + i\sigma + \alpha)P_0(i) + \alpha P_0(i+1) = 0, \text{ for } i \geq 1, \\ \frac{\partial P_1(i,z)}{\partial z} - \frac{\partial P_1(i,0)}{\partial z} - (\lambda + \alpha)P_1(i,z) + \lambda P_1(i-1,z) + \\ \lambda P_0(i)B(z)(i+1)\sigma \cdot P_0(i+1)B(z) + \alpha P_1(i+1,z) = 0. \end{cases} \quad (1)$$

Asymptotic analysis under a heavy load condition

We introduce partial characteristic functions:

$$H_0(u) = \sum_i e^{ju i} P_0(i), \quad H_1(u, z) = \sum_i e^{ju i} P_1(i, z), \quad (2)$$

where j is an imaginary unit.

Substituting functions (2) into Eq. (1), the following equation system is obtained:

$$\begin{cases} b \frac{\partial H_1(u, 0)}{\partial z} - (\rho + \alpha b) H_0(u) + j \sigma b \frac{\partial H_0(u)}{\partial u} + \alpha b e^{-ju} H_0(u) = \\ = \alpha b P_0(0)(e^{-ju} - 1), \\ b \frac{\partial H_1(u, z)}{\partial z} - b \frac{\partial H_1(u, 0)}{\partial z} - e^{-ju} j \sigma b \frac{\partial H_0(u)}{\partial u} B(z) + \\ + \rho H_0(u) B(z) + \rho(e^{ju} - 1) H_1(u, z) + \alpha b(e^{-ju} - 1) H_1(u, z) = \\ = \alpha b P_1(0, z)(e^{-ju} - 1), \end{cases} \quad (3)$$

where $\rho = \lambda b$ is the system load parameter, b is a mean of the service time.

System (3) is solved by the asymptotic analysis method under limit condition of a heavy load $\rho \rightarrow S$, where S is the system throughput (the supremum of the load value when the stationary regime exists for the retrial queue).

Theorem. Let $i(t)$ be a number of calls in the orbit in the retrial queue M/GI/1 with impatient calls in the stationary regime, then

$$\lim_{\rho \rightarrow S} M e^{jw(S-\rho)t(t)} = \left(1 - \frac{jw}{\beta(S-\rho)} \right)^{-s},$$

where $\beta = \frac{2b^2}{b_2 + 2b^2(S-1)}$, $s = 1 + \frac{\beta}{\sigma b}$, b is the mean of the service time, b_2

is the second moment of the service time and $S = 1 + \alpha b$ is the system throughput.

To prove the theorem, we introduce the following notations for $\varepsilon \rightarrow 0$:

$$\rho = S - \varepsilon, \quad u = \varepsilon w, \quad H_0(u) = \varepsilon F_0(w, \varepsilon), \quad H_1(u, z) = F_1(w, \varepsilon, z),$$

$$P_0 = \varepsilon \pi_0, \quad P_1 = \varepsilon \pi_1(z). \quad (4)$$

Substituting notations (4) into system (3), we obtain

$$\left\{ \begin{aligned} & b \frac{\partial F_1(w, 0, \varepsilon)}{\partial z} + j\sigma b \frac{\partial F_0(w, \varepsilon)}{\partial w} - (S - \varepsilon + \alpha b)\varepsilon F_0(w, \varepsilon) + \\ & + \alpha b e^{-j\varepsilon w} \varepsilon F_0(w, \varepsilon) = \alpha b \varepsilon \pi_0 (e^{-j\varepsilon w} - 1), \\ & b \frac{\partial F_1(w, z, \varepsilon)}{\partial z} - b \frac{\partial F_1(w, 0, \varepsilon)}{\partial z} - e^{-j\varepsilon w} j\sigma b \frac{\partial F_0(w, \varepsilon)}{\partial w} B(z) + \\ & + (S - \varepsilon)\varepsilon F_0(w, \varepsilon) B(z) + (S - \varepsilon)(e^{j\varepsilon w} - 1)F_1(w, z, \varepsilon) + \\ & + \alpha b (e^{-j\varepsilon w} - 1)F_1(w, z, \varepsilon) = \alpha b \varepsilon \pi_1(z) (e^{-j\varepsilon w} - 1). \end{aligned} \right. \quad (5)$$

Denoting limits $F_k(w, z) = \lim_{\varepsilon \rightarrow 0} F_k(w, z, \varepsilon)$, $F_1(w) = \lim_{z \rightarrow \infty} F_1(w, z)$ and using expansions

$$F_k(w, z, \varepsilon) = F_k(w, z) + \varepsilon \cdot f_k(w, z) + O(\varepsilon^2), \quad (6)$$

the following system of asymptotic equations can be derived from system (5):

$$\left\{ \begin{aligned} & b \frac{\partial F_1(w, 0)}{\partial z} = -j\sigma b F_0'(w), \\ & b \frac{\partial F_1(w, z)}{\partial z} - b \frac{\partial F_1(w, 0)}{\partial z} (1 - B(z)) = 0, \\ & b \frac{\partial f_1(w, 0)}{\partial z} - S F_0(w) + j\sigma b f_0'(w) = 0, \\ & b \frac{\partial f_1(w, z)}{\partial z} - b \frac{\partial f_1(w, 0)}{\partial z} + jw j\sigma b F_0'(w) B(z) - \\ & - j\sigma b f_0'(w) B(z) + S F_0(w) B(z) + (S - \alpha b) jw F_1(w, z) = 0. \end{aligned} \right. \quad (7)$$

Then summing up Eq. (5), we obtain the following equation for $z \rightarrow \infty$:

$$-j\sigma b \frac{\partial F_0(w, \varepsilon)}{\partial w} + \alpha b \varepsilon F_0(w, \varepsilon) - (S - \varepsilon) e^{j\varepsilon w} F_1(w, \varepsilon) + \alpha b F_1(w, \varepsilon) = \alpha b \varepsilon (\pi_1 + \pi_0).$$

Substitute expansions (6) and write equalities for members with equal powers of ε :

$$\left\{ \begin{aligned} & -j\sigma b F_0'(w) - S F_1(w) + \alpha b F_1(w) = 0, \\ & -j\sigma b f_0'(w) + \alpha b F_0(w) + (1 - S jw) F_1(w) - \\ & - S f_1(w) + \alpha b f_1(w) = \alpha b (\pi_1 + \pi_0). \end{aligned} \right. \quad (8)$$

The asymptotic characteristic function of probability distribution of number of calls in the orbit $h(u)$ under the heavy load condition can be presented as

$$h(u) = M e^{jw(S-\rho)it} = F_1(w) + O(\varepsilon) \approx F_1(w). \quad (9)$$

So, it is necessary to obtain the function $F_1(w)$ from Eqs. (7), (8).

The derivation is performed in four stages.

Stage 1. Let the function $F_1(w, z)$ have the form

$$F_1(w, z) = A(z) \cdot \Phi(w). \quad (10)$$

Thus $F_1(w) = \Phi(w)$.

From the second equation of system (7), it is easy to show that

$$A(z) = A'(0) \int_0^z (1 - B(x)) dx,$$

where $A'(0) = 1/b$.

Stage 2. From the first equation of System (7), we have the following expression:

$$j\sigma F_0'(w) = -\frac{1}{b} \Phi(w). \quad (11)$$

Substituting formulas (10), (11) into the first equation of system (8), the value of the system throughput is obtained

$$S = 1 + \alpha b. \quad (12)$$

So, the stationary regime for this retrial queue exists, when $\rho < 1 + \alpha b$ or $\lambda < \alpha + 1/b$.

Stage 3. From the third equation of system (7), we have

$$j\sigma b f_0'(w) = -b \frac{\partial f_1(w, 0)}{\partial z} + S F_0(w).$$

Perform some transformations in the fourth equation of system (7):

$$f_1(w, z) = \frac{\partial f_1(w, 0)}{\partial z} \int_0^z (1 - B(x)) dx - \frac{jw}{b} \Phi(w) \int_0^z (A(x) - B(x)) dx.$$

Let us find the solution $f_1(w, z)$ in the form

$$f_1(w, z) = \frac{jw}{b} \Phi(w) v(z).$$

Then we obtain

$$f_1(w) = jw\Phi(w)v'(0) - \frac{jw}{b}\Phi(w)\int_0^\infty (A(x) - B(x))dx, \quad (13)$$

where $\int_0^\infty (A(x) - B(x))dx = b - \frac{1}{2b}b_2$.

Thus, on the one hand,

$$f_1(w) = jw\Phi(w)v'(0) - \frac{jw}{b}\Phi(w)\left(b - \frac{1}{2b}b_2\right).$$

On the other hand,

$$f_1(w) = f_1(w, \infty) = \frac{jw}{b}\Phi(w)v(\infty).$$

Comparing these expressions, it can be obtained that

$$bv'(0) - v(\infty) = b - \frac{1}{2b}b_2. \quad (14)$$

Stage 4. Substitute formulas (11) – (14) into the last equation of system (8):

$$(\alpha b - S)F_0(w) + (1 - Sjw)\Phi(w) + \frac{jw}{b}\Phi(w)\left(b - \frac{1}{2b}b_2\right) = \alpha b(\pi_1 + \pi_0).$$

Then we differentiate this equation. After some transformations, it is easy to obtain the following equation

$$\Phi'(w)(\beta - jw) - j\Phi(w)\frac{s}{\beta} = 0,$$

where

$$\beta = \frac{2b^2}{b_2 + 2b^2(S-1)}, \quad s = 1 + \frac{\beta}{\sigma b}.$$

The solution of this equation has the form

$$\Phi(w) = C\left(1 - \frac{jw}{\beta}\right)^{-s}.$$

Turning back to expressions (10) and (9), we finally obtain the following function:

$$h(u) = C \left(1 - \frac{ju}{\beta(S - \rho)} \right)^{-s},$$

where $C = 1$ due to the normalization requirement, q.e.d.

Conclusions

In this regard, the mathematical model of the retrial queue M/GI/1 with impatient calls (impatience time is distributed exponentially with dynamical rate α/i), is considered in the paper. For its studying, we propose the asymptotic method under a heavy load condition. It is proved the theorem about the gamma form of the asymptotic probability distribution of the number of calls in the orbit. In addition, the expression for the system throughput is derived.

REFERENCES

1. *Artalejo J.R., Gomez-Corral A.* Retrial Queueing Systems: A Computational Approach. – Springer, 2008.
2. *Falin G.I., Templeton J.G.C.* Retrial queues. – London: Chapman & Hall, 1997.
3. *Nazarov A.A., Lyubina T.V.* The non-Markov dynamic RQ system with the incoming MMP flow of requests // Automation and Remote Control. – 2013. – V. 74. – No. 7. – P. 1132–1143.
4. *Pankratova E., Moiseeva S.* Queueing System MAP/M/ ∞ with n Types of Customers // Communications in Computer and Information Science. – 2014. – V. 487. – P. 356–366.
5. *Moiseeva E., Nazarov A.* Asymptotic analysis of RQ-systems M/M/1 on heavy load condition // Proceedings of the IV International Conference Problems of Cybernetics and Informatics. – Baku, Azerbaijan, 2012. – P. 164–166.
6. *Моисеева Е.А., Назаров А.А.* Исследование RQ-системы MMPP|GI|1 методом асимптотического анализа // Вестник ТГУ. Управление, вычислительная техника и информатика. – 2013. – Т. 25. – № 4. – С. 84–94.

TOTAL CUSTOMERS' CAPACITY IN THE MULTI SERVER QUEUES*

Ekaterina Lisovskaya

Tomsk State University, Tomsk, Russia

In this paper we consider queueing system with N servers and unlimited queue. Below we present a review of the work on the studies of resource queueing systems or queueing systems with random capacity of customers.

In the paper [1] an efficient analytical model that evaluates the behavior of the downlink LTE channel with CLA is presented. Since video traffic is resource intensive, it is a challenging issue to stream video over low bandwidth networks, whereas video communication over LTE becomes an open research topic nowadays due to LTE's high throughput capabilities.

The paper [2] considered a model of a multi-server queueing system with losses caused by lack of resources necessary to service claims. A claim accepted for servicing occupies a random amount of resources of several types with given distribution functions. Random vectors that define the requirements of claims for resources are independent of the processes of customer arrivals and servicing, mutually independent, and identically distributed. Under the assumptions of a Poisson arrival process and exponential service times, we analytically find the joint distribution of the number of customers in the system and the vector of amounts of resources occupied by them. We show sample computations that illustrate an application of the model to analyzing the characteristics of a videoconferencing service in an LTE wireless network.

In the paper [3] authors consider queueing systems, in which customers occupy some resources that are released after customer departure. Arriving customers are lost if there are not enough free resources required for their servicing. In such systems for each customer it is necessary to record vector of occupied resources until its departure.

In the paper [4], multi-server M/M/ n -type queueing system with a bounded total volume and finite queue size is considered. An AQM

* The publication was financially supported by RFBR according to the research project No.16-31-00292mol-a.

algorithm with the “accepting” function is being used to control the arrival process of incoming packets. The stationary queue-size distribution and the loss probability are derived. Numerical examples illustrating theoretical results are attached as well.

The paper [5] investigated single server queueing systems with batch Poisson arrivals and without demands losses under assumption that each demand has some random capacity (generally, each demand is characterized by 1-dimensional indication vector). Service time of the demand arbitrary depends on its capacity (indications). The Laplace – Stieltjes transform of total capacities (random vector of sum of indications) of demands that were served during a busy period of the system is determined.

Mathematical model

We consider queueing system (QS) $M/GI/N/\infty$. The system arrival process is distributed by Poisson law with rate λ . The system has N servers. Service times on each server are i.i.d. with distribution function $A(x)$. The arriving customer occupies any free server or goes to the queue in case of all servers are busy. Let each customer has some random capacity $v > 0$ with distribution function $G(v)$. Customers' capacities and service times are mutually independent and do not dependent on the epochs of customers' arrivals.

Denote by $i(t)$ and $V(t)$ the number of customers in the system at time t and their total capacity, respectively.

Approximation of probability distribution of the customers' number in the system

Let $P(i) = P\{i(t) = i\}$ is the stationary probability distribution of the number of customers in the system. Denoted π_i as an approximation of the probability distribution which is defined as a composite distribution [6]

$$\pi_i = \begin{cases} C_1 P_1(i), & 0 \leq i \leq N, \\ C_2 P_2(i - N + 1), & i \geq N. \end{cases} \quad (1)$$

The probabilities $P_1(i)$, where $0 \leq i \leq N$, are the probabilities of the number of occupied servers in N -server QS with customer's losses ($M/GI/N/0$), when all servers are busy. Then it can be determined by the Erlang formulas:

$$P_1(i) = \frac{(\lambda a)^i}{i!} \left(\sum_{k=0}^N \frac{(\lambda a)^k}{k!} \right)^{-1},$$

where a is mean of service time.

The probabilities $P_2(i)$ are defined when all servers are busy. In this case, the block of occupied servers is considered as a single and its service has distribution function $B(x)$. Therefore, the probabilities $P_2(i)$, where $i = 0, 1, \dots$ are defined as the probabilities of the number of customers in the single-server system M/GI/1/ ∞ with waiting. It can be determined by the Pollaczek – Khinchin formula [6] and we can write

$$P_2(i) = (1 - \lambda b) \sum_{k=0}^i \alpha_k b_{i-k},$$

where coefficients of expansion

$$\alpha_0 = \frac{1}{\beta_0}, \quad \alpha_n = \frac{1}{\beta_0} \left[\alpha_{n-1} - \sum_{k=0}^{n-1} \alpha_k \beta_{n-k} \right],$$

$$b_0 = \beta_0, \quad b_n = \beta_n - \beta_{n-1},$$

$$\beta_n = \int_0^{\infty} e^{-\lambda z} \frac{(\lambda z)^n}{n!} dB(z),$$

and the distribution function $B(x)$ has the form

$$B(x) = 1 - (1 - A(x)) \left(1 - \frac{1}{a} \int_0^x (1 - A(z)) dz \right)^{N-1},$$

b is mean of a random value with the distribution function $B(x)$.

Constants C_1 and C_2 from (1) founded from the normalization condition and the conditions of «stitching» [6]. So the expression (1) has the form

$$\pi_i = \begin{cases} \frac{P_2(1)}{P_2(1) + P_1(N)(1 - (P_2(0) + P_2(N)))} P_1(i), & 0 \leq i \leq N, \\ \frac{P_1(N)}{P_2(1) + P_1(N)(1 - (P_2(0) + P_2(N)))} P_2(i - N + 1), & i \geq N. \end{cases} \quad (2)$$

Characteristic function for the total capacity

We write the characteristic function of the total capacity probability distribution in the form

$$\begin{aligned} h(u) &= M \left\{ e^{juV(t)} \right\} = M \left\{ M \left\{ e^{ju \sum_{k=1}^i v_k} \middle| i(t) = i \right\} \right\} = \\ &= \sum_{i=0}^{\infty} M \left\{ e^{ju \sum_{k=1}^i v_k} \right\} P \{ i(t) = i \} = \sum_{i=0}^{\infty} \left(M \left\{ e^{juv} \right\} \right)^i P \{ i(t) = i \}. \end{aligned}$$

Then, using approximation (2):

$$h(u) = \sum_{i=0}^{\infty} \left(M \left\{ e^{juv} \right\} \right)^i \pi_i.$$

Using the inverse Fourier transform, we obtain an approximation of the density function of the customers' total capacity in the system M/GI/N/∞

$$f_V(x) = \int_{-\infty}^{\infty} e^{-jux} h(u) du. \quad (3)$$

Simulation and numerical examples

Simulation was performed in the same way as [7]. Its result is the distribution function of the total capacity of customers in this system. Also we obtained the distribution function by (3). Let us compare this results. We will use the Kolmogorov distance as a measure between simulation and approximation results

$$\Delta = \sup_x |F_V(x) - D(x)|.$$

Here $F_V(x)$ is the approximation based on (3), and $D(x)$ is the cumulative distribution function build on the basis of simulation results. As typically done in the literature [8], we suppose that an approximation is applicable if its Kolmogorov distance is less than 0,03.

We can note, that the parameters for the arriving and service of customers are selected in such a way that the condition for the stationary regime existence is met ($N > \lambda a$).

Example 1. Let the rate of arrivals $\lambda = 25$, the number of servers is $N = 10$, the distribution function of service time is

$$A(x) = \begin{cases} 1 - e^{-\mu x}, & x \geq 0, \\ 0, & x < 0, \end{cases}$$

the distribution function of customers' capacity is

$$G(y) = \begin{cases} 0, & y < 0, \\ \frac{y-a}{b-a}, & a \leq y \leq b, \\ 1, & y > b. \end{cases}$$

In this example $\mu = 5$; $a = 0$; $b = 1$. In this case, we obtain that $\Delta = 0,012$.

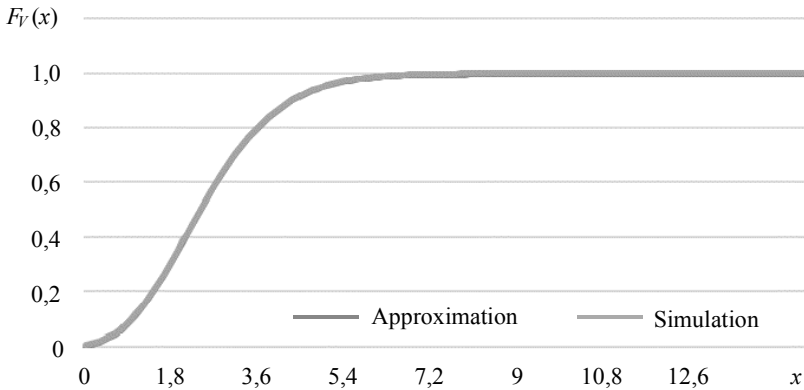


Fig. 1. Distributions of the total capacity

Example 2. Let the rate of arrivals $\lambda = 25$, the number of servers is $N = 10$, the distribution function of service time is

$$A(x) = \begin{cases} \frac{\gamma(\alpha, \beta x)}{\Gamma(\alpha)}, & x \geq 0, \\ 0, & x < 0, \end{cases}$$

the distribution function of customers' capacity is

$$G(y) = \begin{cases} 0, & y < 0, \\ \frac{y-a}{b-a}, & a \leq y \leq b, \\ 1, & y > b. \end{cases}$$

In this example $\alpha = 0,5$; $\beta = 2,5$; $a = 0$; $b = 1$. In this case, we obtain that $\Delta = 0,007$.

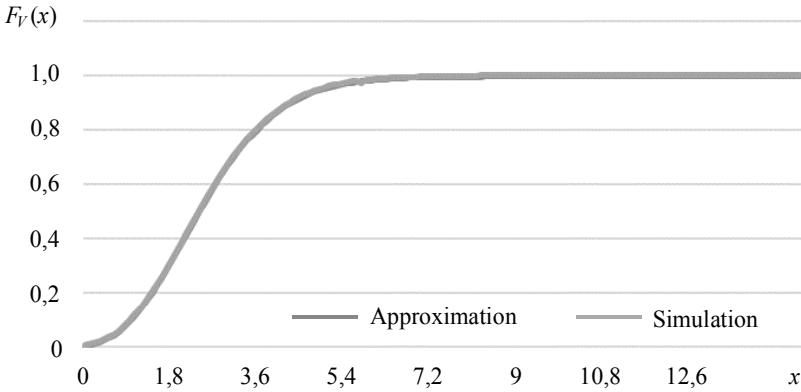


Fig. 2. Distributions of the total capacity

Example 3. Let the rate of arrivals $\lambda = 45$, the number of servers is $N = 6$, the distribution function of service time is

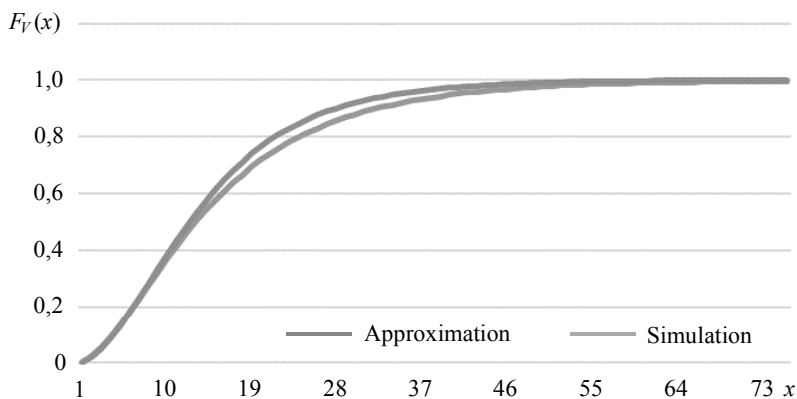
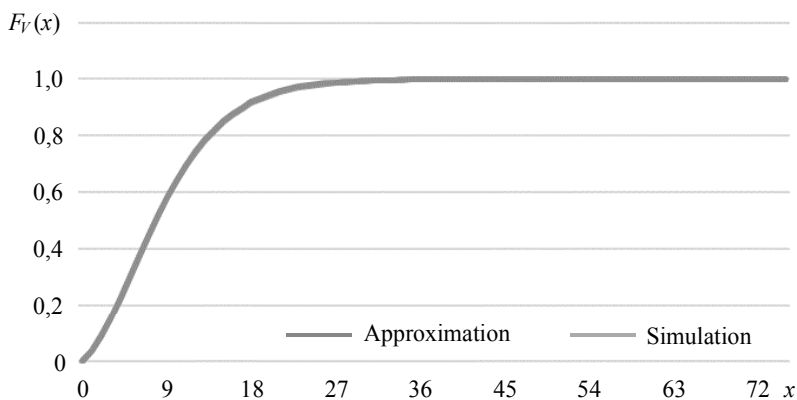
$$A(x) = \begin{cases} \frac{\gamma(\alpha, \beta x)}{\Gamma(\alpha)}, & x \geq 0, \\ 0, & x < 0, \end{cases}$$

the distribution function of customers' capacity is

$$G(y) = P\{v = y\} = p(1-p)^y.$$

In this example $\alpha = 3,5$; $\beta = 29,7$; $p = 0,4$. In this case, we obtain that $\Delta = 0,048$.

We will simulate this queue at the same parameters, but increase the servers' number. For $N = 7$, we obtain that $\Delta = 0,016$, for $N = 8$, we obtain that $\Delta = 0,005$.

Fig. 3. Distributions of the total capacity ($N = 6$)Fig. 4. Distributions of the total capacity ($N = 8$)

We can conclude that the accuracy of the total capacity approximation has a wide application and improves with the increase the number of servers in the system.

Conclusions

We studied the total capacity of customers in the system $M/GI/N/\infty$. The characteristic function of this distribution was obtained. Simulation has been carried out, it is shown that the approximation has a wide application.

REFERENCES

1. Efimushkina T., Gabbouj M., Samuylov K. Analytical Model in Discrete Time for Cross-Layer Video Communication over LTE // Automatic Control and Computer Sciences. – 2014. – V. 48. – No. 6. – P. 345–357.
2. Naumov V.A., Samuilov K.E., Samuilov A.K. On the total amount of resources occupied by serviced customers // Automation and Remote Control. – 2016. – V. 77. – No. 8. – P. 1419–1427.
3. Naumov V.A., Samuilov K.E. On Modeling Queueing Systems with Multiple Resources // Vestn. Ross. Univ. Druzhby Narodov, Ser. Mat. Informatika. Fiz. – 2014. – No. 3. – P. 60–64.
4. Tikhonenko O., Kempa W.M. On the queue-size distribution in the multi-server system with bounded capacity and packet dropping // Kybernetika. – 2013. – V. 49. – No. 6. – P. 855–867.
5. Tikhonenko O., Kawecka M. Busy Period Characteristics for Single Server Queue with Random Capacity Demands // Communications in Computer and Information Science. – 2012. – V. 291. – P. 393–400.
6. Lisovskaya E., Moiseeva S. Study of the Queuing Systems $M/GI/N/\infty$ // Communications in Computer and Information Science. – 2015. – V. 564. – P. 175–184.
7. Lisovskaya E., Pagano M. Imitacionnoe modelirovanie sistemy massovogo obsluzhivaniya trebovanij sluchajnogo ob"ema // Problemy optimizacii slozhnyh sistem Trudy 12-j Mezhdunarodnoj Aziatskoj shkoly-seminara. Pod redakciej S.I. Kabanikhina, A.V. Kel'manova, A.S. Rodionova. – 2016. – P. 352–357.
8. Moiseev A., Sinyakov M. Razrabotka ob'ektno-orientirovannoj modeli sistemy imitacionnogo modelirovaniya processov massovogo obsluzhivaniya // Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naya tekhnika i informatika. – 2010. – V. 1. – P. 89–93.
9. Kempa W.M. On main characteristics of the $M/M/1/N$ queue with single and batch arrivals and the queue size controlled by AQM algorithms // Kybernetika. – 2011. – V. 47. – No. 6. – P. 930–943.

OPTIMIZATION OF PIPELINING AND DATA PROCESSING

P.A. Mikheev, A.A. Pichugina, S.P. Suschenko, R.V. Tkachev

Tomsk State University, Tomsk, Russia

Introduction

The most important indicator of efficiency of functioning of the network packet switching is the transmission time of user data between the communicating subscribers [1, 2]. Functions for the delivery of message flow to the user and compensation of overhead in the transmission of packets that may occur in the communication network are performed by the transport layer protocol [2]. The basis of a reliable transport protocol is the principle of decision feedback. The delay in subscriber traffic in a virtual connection depends to a large extent on the characteristics of the individual links of the connecting path, the length of the data transmission path, the size of the user messages, the intensity of the network streams and the protocol parameters, among which the most important is the packet size, which actually determines the power of the pipeline effect [3–6]. It should also be noted that the connecting path of the virtual channel in the packet switching network is used in by many interacting subscribers. This leads to the fact that the load on various parts of the data path along which the virtual connection goes can be significantly different. Then the effective bandwidth of individual links for the traffic of this virtual connection will be reduced by the corresponding parts of "external" flows, as a result of which the time of packet transfer over inter-node connections even of a uniform virtual channel can be substantially different [6–8].

Simulation of the transport connection and analysis of its operational characteristics under various loading conditions is performed in [3–8]. A wide range of studies [9–13] are aimed at optimizing protocol parameters by various criteria and adapting protocol parameters to the changing network load, the level of losses, the activity of interacting subscribers. A key indicator of the quality of service for network subscribers is the message delivery time, which is determined by the pipelining effect. This indicator is also very important for pipelined implementation of the instruction processing [14, 15]. The development of the results of [9–15] is to optimize the size of the protocol data units when sending the subscription message via

the transport connection and the structure of the data transmission path. In the proposed method of this work the partitioning the subscriber messages into packets of optimum size and the conditions the feasibility of converting transmission path and pipelined processor to a uniform appearance.

The virtual connection model

Consider an nonuniform virtual connection consisting of D links of data transmission. Define the time of the message transmission from the N packets, according to the deterministic virtual connection in the data transfer phase. We believe that the flow control procedure carried by the virtual connection provides end-to-end confirmation of the delivery of individual messages, and each virtual connection node can simultaneously perform data reception and transmission, however, packet transmission can be started only after its reception is completed. All message packets have the same length, except the last one, which can contain the remainder of dividing the message into fragments and can be smaller. We believe that there is no competing traffic and there are no packet queues at the switching nodes to the output communication channels. Then the delay of the subscriber's message in the data transmission path will be [10]

$$T(D, N) = (N - 1)\tau_m + \sum_{d=1}^D \tau_d, \quad \tau_m = \max_{d=1, D} \tau_d, \quad (1)$$

where $\tau_d, d = \overline{1, D}$, – packet delay in d sector of the hops.

The optimal partitioning of the message into the packets

For further analysis, we express explicitly the time of packet transmission in inter-node connection through the parameters of data transmission link. Suppose that the transmission rate and time of node packet processing is independent of the size of the package. In fact, the assumption of rate is true only for absolutely reliable inter-node communication channels included in the virtual connection.

Then the packet delay at the d link of the transmission path taking into account the previously introduced notation we can write as: $\tau_d = L/C_d + t_d$. Here $\tau_d, d = \overline{1, D}$, the packet processing time in the receiver node of the d data link. Substituting this relation in (1) and taking into account that $L = B/N + H$, where B is the size of the transmitted message, we get

$$T(D, N) = (N-1) \left[\frac{B/N + H}{C_m} + t_m \right] + \sum_{d=1}^D \left[\frac{B/N + H}{C_d} + t_d \right],$$

$$\frac{B/N + H}{C_m} + t_m = \max_{d=1, D} \left[\frac{B/N + H}{C_d} + t_d \right]. \quad (2)$$

Obviously, when transmitting a message in the form of a sequence of packets, it is possible to reduce the time of its delivery significantly over a virtual connection in comparison to its transmission by one packet. This gain is due to the pipelining effect [10], as a result of which the different parts of the message are simultaneously in the transmission state at different parts of the path. On the one hand, the number of packets in the sequence should be increased in order to enhance the pipelining effect and thereby reduce the message delivery time. On the other hand, sequence growth leads to an increase in the volume of the transferred service information and the processing time of packets by nodes. Hence it follows that the dependence (2) is unimodal from the argument N . Using (2), we determine the benefit in time from the transmission of a message over a virtual connection of length D by a sequence of $N > 1$ packets in comparison with its delivery as a whole:

$$\Delta(D, N) = T(D, 1) - T(D, N) = (N-1) \left(\frac{B}{N} \sum_{\substack{d=1 \\ d \neq m}}^D \frac{1}{C_d} - \frac{H}{C_m} - t_m \right). \quad (3)$$

For uniform virtual connection, $C_d = C, t_d = t, d = \overline{1, D}$, the benefit (3) is converted to the form

$$\Delta(D, N) = \frac{N-1}{C} \left(\frac{B}{N} (D-1) - H - Ct \right). \quad (4)$$

The relations (3), (4) define unimodal of the argument $N \geq 1$ functions with asymptotes

$$\Delta(D, N) = -(N-1) \left(\frac{H}{C_m} + t_m \right) + B \sum_{\substack{d=1 \\ d \neq m}}^D \frac{1}{C_d}$$

and

$$\Delta(D, N) = -\frac{N-1}{C} (H + Ct) + \frac{B(D-1)}{C}$$

accordingly. It can be seen from (3), (4) that it is expedient to split/divide

the message into packets only for long $D > 1$ virtual connections and if condition $B \sum_{\substack{d=1 \\ d \neq m}}^D \frac{1}{C_d} > \frac{H}{C_m} + t_m$ is fulfilled, and benefit (3) is positive

for partitions that satisfy the inequality $1 < N < BC_m \sum_{\substack{d=1 \\ d \neq m}}^D \frac{1}{C_d} \bigg/ (H + C_m t_m)$.

For a uniform virtual connection, this inequality, which determines the set of expedient partitions, takes the form of $1 < N < \frac{B(D-1)}{H + Ct}$. When

$B \sum_{\substack{d=1 \\ d \neq m}}^D \frac{1}{C_d} > \frac{H}{C_m} + t_m$ splitting $N > BC_m \sum_{\substack{d=1 \\ d \neq m}}^D \frac{1}{C_d} \bigg/ (H + C_m t_m)$ lead to the fact

that the losses from transmission and processing of the sequence of packets prevail over the benefit from the pipeline effect. When $B \sum_{\substack{d=1 \\ d \neq m}}^D \frac{1}{C_d} < \frac{H}{C_m} + t_m$

splitting $N > 1$ increase the negative effect of exceeding the overhead on user information.

On virtual connections of a single length, there is no pipelining effect, and $N > 1$ partitions lead to an increase in the multiplex packet delay due to an increase in the amount of overhead transfer of the service information and the node packet processing. From the size of the subscriber message B the benefit (3), (4) has a linear dependence. When transferring over the uniform virtual connection, the benefit (4) also grows linearly with the path length, and the values $\Delta(D, N)$ are positive for $D > 1 + \frac{N(H + Ct)}{B}$.

From (3) we find that the partition $N_o = \sqrt{\frac{B}{H/C_m + t_m} \sum_{\substack{d=1 \\ d \neq m}}^D \frac{1}{C_d}}$, $N_o \geq 1$

maximizes the benefit (3). Substituting the relation for N in (3) we obtain:

$$\Delta(D, N_o) = 2 \left\{ \frac{1}{2} \left[\frac{H}{C_m} + t_m + B \sum_{\substack{d=1 \\ d \neq m}}^D \frac{1}{C_d} \right] - \sqrt{\left(\frac{H}{C_m} + t_m \right) \left[B \sum_{\substack{d=1 \\ d \neq m}}^D \frac{1}{C_d} \right]} \right\}.$$

Hence it is easy to see that the optimal benefit is equal to twice the difference between the arithmetic mean and geometric mean values $B \sum_{\substack{d=1 \\ d \neq m}}^D \frac{1}{C_d}$ and

$H/C_m + t_m$, which correspond respectively to the transmission time of the information part of the message as a single unit over a virtual connection without a narrow link and the overhead of a narrow link in the form of the time for transmitting the service part of the packet and the processing time of the packet. For a uniform virtual connection, the optimal benefit is:

$\Delta(D, N_o) = \left\{ \sqrt{H + Ct} - \sqrt{B(D-1)} \right\}^2$. Since $N \geq 1$, we can conclude that the

area of definition $\Delta(D, N_o)$ for the uniform virtual connection is the length of the paths that satisfy the inequality $D \geq (1 + H + Ct)/B$. Knowing the optimal ratio of splitting N , it is easy to determine the packet size L_o , that minimizes the delivery time of a message over the virtual connection:

$L_o = H + B/N_o = H + \sqrt{B(H/C_m + t_m) / \sum_{\substack{d=1 \\ d \neq m}}^D \frac{1}{C_d}}$. On the uniform connection the

expression for the optimal packet size is simplified: $L_o = H + \sqrt{\frac{B(H + Ct)}{D-1}}$.

The conditions of feasibility of unification of the non-uniform phases in the pipeline

One of the most important conditions for achieving minimum latency on the line as in (1) is to eliminate the most time-consuming stages of processing (transmission) data through their pipeline. Most often, this approach of bringing the individual phases of the pipeline to a uniform duration of stages is used in the processor of the data processing or telecommunication systems to eliminate low-speed, geographically distributed network sections. In this case, the narrow phases of the pipeline are broken (if possible) into sub-phases of the minimal complexity of the input of the original or desired conveyor, which leads to an increase in its length.

Let us analyze the conditions under which such a partition of complex phases into sub-phases reduces the processing time of the data stream. Consider fully ununiformed pipeline, which should lead to a uniform with

the duration of the phases equal to $\tau \leq \tau_d, d = \overline{1, D}$. We assume that each phase of an ununiformed pipeline has a duration $\tau_d = l_d \tau, d = \overline{1, D}$, where $l_d \geq 1$ – is an integer. Then every d phase of the source pipeline should be pipelined in the form $l_d \geq 1$ of stages of the same duration τ . The normalized delay in the original ununiformed pipeline length is

$$t_n(D, N) = \frac{T(D, N)}{\tau} = (N - 1)l_m + \sum_{d=1}^D l_d, l_m = \max_{d=\overline{1, D}} l_d,$$

and in the uniformed – with the resulting number of stages, equal $\sum_{d=1}^D l_d$, will take the form

$$t_o\left(\sum_{d=1}^D l_d, N\right) = \frac{T\left(\sum_{d=1}^D l_d, N\right)}{\tau} = N - 1 + \sum_{d=1}^D l_d.$$

We determine normalized to the complexity of the uniformed phase τ the benefit from the unification of pipelining N applications in the form of a difference of the processing times of the original ununiformed pipeline and extended uniformed pipeline:

$$\nabla(D, N) = t_n(D, N) - t_o\left(\sum_{d=1}^D l_d, N\right) = (N - 1)(l_m - 1).$$

Hence it follows that the positive values of the benefit are invariant to the complexity l_d of all stages of the pipeline, except the most labor-intensive and possible when $N > 1$ and $l_m > 1$.

Conclusions

The paper proposes the method of optimal partitioning of subscriber messages into protocol data units by the transport layer according to the criterion of delays in the multi-hop transmission path. Analytical dependence for the optimal packet size from the structure of network traffic and settings of the virtual connections. The conditions of feasibility of super pipelining, ununiformed pipeline and bring it to a uniform by splitting time-consuming stages into uniformed phase. The direction of the further development of

research on the unification of the pipeline should be distinguished by the task of analyzing the delay in conditions of rebooting of the pipeline with repeated transmissions of distorted data in networks or incorrect branch prediction processed by the processor of instruction stream.

REFERENCES

1. Fall K., Stevens R. TCP/IP Illustrated, Vol. 1: The Protocols (2nd Edition). – Addison-Wesley Professional Computing Series, 2012. – 1017 p.
2. Boguslavsky L.B. Upravlenie potokami dannykh v setyakh EVM (Controlling Data Flows in Computer Networks). – Moscow: Energoatomizdat, 1984.
3. Boguslavsky L.B., Gelenbe E. Analytical models transmission link control procedures for data computer networks with packet // Automation and Remote Control. – 1980. – No 7. – P. 181–192.
4. Gelenbe E., Labetoulle J., Pujolle G. Performance Evaluation of the HDLC Protocol // Computer Networks, 1978. – V. 2. – No. 4/5. – P. 409–415.
5. Kokshenev V.V., Sushchenko S.P. Analysis of the asynchronous performance management procedures link transmission data // Computational Technologies. – 2008. – V. 15. – No. 5. – P. 61–65.
6. Kokshenev V.V., Mikheev P.A., Sushchenko S.P. Comparative analysis of the performance of selective and group repeat transmission models in a transport protocol // Automation and Remote Control. – 2017. – V. 78. No. 2. – P. 247–261.
7. Kokshenev V.V., Mikheev P.A., Sushchenko S.P. Analysis of selective denial mode transport protocol loaded in the transmit path, data // Vestnik TSU. Series Control, Computer Facilities and Computer Science. – 2013. – No. 3(24). – P. 78–94.
8. Sushchenko S.P. On the effect of buffer memory locks on the performance of a multi-tier data transmission path // Avtomatiki i Telemekhanika. – 1999. – No. 7. – P. 66–79.
9. Sushchenko S.P. Method of rational choice of packet length of packet switching network // Automation and Computer Engineering. – 1984. – No. 3. – P. 24–28.
10. Sushchenko S.P. Parametric optimization of a packet switching network // Automation and Computer Engineering. – 1985. – No. 2. – P. 43–49.
11. Sushchenko S.P. Analysis of end-to-end message delay in a multi-tier virtual channel // Automation and Computer Technology. – 1989. – No. 3. – P. 52–64.
12. Sushchenko S.P. Influence of the duration of a through time-out on the delay of data in a virtual channel // Automation and Computer Technology. – 1991. – No. 6. – P. 36–40.
13. Sushchenko S.P. Analysis of the influence of the duration of a through time-out on the operational characteristics of a virtual channel // Automation and Computer Technology. – 1995. – No. 4. – P. 43–66.
14. Patterson D.A., Hennessy J.L. Computer Organization and Design. – Amsterdam: Elsevier, 2014. – 793 p.
15. Gorshenin A.K., Zamkovets S.V., Zakharov V.N. Parallelism in microprocessors // Systems and Means of Informatics. – 2014. – V. 24. – No. 1. – P. 46–60.

CONTROLLABLE QUEUEING SYSTEMS FROM THE VERY BEGINNING UP TO NOWADAYS*

V.V. Rykov

Gubkin Russian Oiland Gas University, Moscow, Russia

The theory of Controllable Queueing Systems (CQS) is a special direction of investigations of a general theory of controllable stochastic processes from one side, and of a Queueing Theory (QT) from another side. The theory of controllable stochastic processes is a special topic, which we will not touch here, and will fix on CQS. Some papers devoted to the problems of Queueing Systems (QS) control have been arisen almost simultaneously with the first works about QS, but the special approach to CQS has been done by Rykov [1] in 1975. Several monographs devoted to this problem arisen thereafter [2–5].

The present paper represents a review of the Controllable Queueing Systems (CQS) theory development from the very beginning up to nowadays. A main stages of this theory development are considered. Some new problems are mentioned.

It is destined for those who want to know the theory of CQS with the origin and its development from its generation up to now and to understand its development tendency and its new directions and new problems

The lecture contains several Sections:

- Introduction.
- CQS. Definition and main properties.
- Examples.
- Discrete time controllable semi-regenerative processes.
- Optimal strategy construction.
- Qualitative properties of optimal policies.
- Arrival control.
- Service mechanism control.
- System structure control.

* The publication has been financially supported by the Ministry of Education and Science of the Russian Federation (the Agreement number 02.A03.21.0008) and by the Russian Foundation for Basic Research according to the research projects No. 17-07-00142 and No. 17-01-00633.

- Service discipline control.
- Conclusion and further problems.

Because the theory of QCS based on the theory Discrete time controllable semi-regenerative processes (DTC SRP) then after the Controllable Queueing System definition the short review of this theory will be proposed. As a result of this theory application it follows the main theorem about Markov property of optimal strategies for the CQS. Moreover, it is known that the Markov strategy is determined with the optimal policy, which can be found as a solution of optimality (or Bellmen) equation. This result allows to turn the theoretical investigations about general properties of optimal policies to algorithmic and computing sphere and concentrate on the construction of the algorithms for real calculation of optimal policies. Some of such algorithms are considered. Moreover, the optimality principle allows to investigate some qualitative properties of optimal policies before their determination. The last approach allows extremely simplify real calculation of optimal policies for many examples of CQS's. Some examples of CQS's investigation based on the proposed approach will be considered. The lecture closed with some new problems settings and new direction of investigations in the theory of CQS.

REFERENCES

1. Rykov V.V. Controllable queueing systems // Itogi nauki i tehniki. Teoria verojatn. Matem. Statist. Teoretich. Kibern. – 1975. – V. 12. – P. 45–152.
2. Heyman D.P., Sobel M. Stochastic Models in Operation Research. V. 1. – New York: McGraw-Hill, 1982.
3. Heyman D.P., Sobel M. Stochastic Models in Operation Research. V. 2. – New York: McGraw-Hill, 1984.
4. Kitaev V.Yu. and Rykov V.V. Controlled queueing systems. – New York: CRC Press, 1995.
5. Sennott L.I. Stochastic Dynamic Programming and the Control of Queueing Systems. – New York: John Wiley & Sons, 1999. – 358 p.
6. Tilms H. Stochastic Models. An Algorithmic Approach. – New York: Wiley, 1994.

ОБ ОДНОЙ МОДЕЛИ СИСТЕМЫ УПРАВЛЕНИЯ ЗАПАСАМИ С МГНОВЕННЫМ ОБСЛУЖИВАНИЕМ ЗАЯВОК

С.А. Багирова

*Бакинский государственный университет,
г. Баку, Азербайджан*

В последние годы интенсивно исследуются модели систем управления запасами (СУЗ) при различных допущениях относительно схемы обслуживания заявок, политики пополнения запасов (ППЗ) и времени доставки заказов [1–3]. Подробный обзор известных результатов можно найти в работах [4–5]. Анализ литературы показал, что почти не исследованы модели СУЗ, в которых объемы заказов являются переменными величинами, зависящими от текущего уровня запасов системы [6]. Отметим, что, с одной стороны, изучение моделей СУЗ с переменным объемом заказов представляет определенный теоретический интерес, так как они в ряде случаев обобщают известные ППЗ; а с другой стороны, с практической точки зрения использование таких ППЗ расширяет область определения надлежащей политики, которая, в свою очередь, позволяет увеличить эффективность работы системы.

Исходя из этого, в данной работе предложен класс ППЗ, основанной на переменном объеме заказов. Предложены точный и приближенный методы расчета характеристик модели с мгновенным обслуживанием и повторными заявками.

Максимальный размер склада изучаемой СУЗ равен S , $0 < S < \infty$. В эту систему поступает пуассоновский поток заявок с интенсивностью λ . Рассматривается модель с мгновенным обслуживанием заявок. Неудовлетворенные заявки, т.е. первичные заявки (p -заявки), поступившие в моменты отсутствия запасов системы, окончательно не теряются, а образуют источник повторных заявок (r -заявки). При этом если в момент поступления p -заявки в системе отсутствуют запасы, то она либо с вероятностью $H_p > 0$ уходит на орбиту для повторения своего запроса, либо с дополнительной вероятностью $1 - H_p$ окончательно уходит из системы.

Времена пребывания в орбите r -заявок являются независимыми друг от друга случайными величинами (с.в.) с общей показательной функцией распределения (ф.р.) с конечным параметром η , $0 < \eta < \infty$. После завершения времени пребывания на орбите r -заявки независимо друг от друга повторяют свои запросы. Если в момент поступления r -заявки на складе имеется хотя бы одна единица запаса, то она мгновенно обслуживается и уходит из орбиты; иначе, r -заявка либо с вероятностью $H_r > 0$ покидает орбиту, либо с дополнительной вероятностью $1 - H_r$ остается там для дальнейшего повторения своего запроса.

Каждая заявка (первичная или повторная) требует запас единичного размера. Здесь рассматривается модель с ограниченным размером орбиты для повторных заявок, т.е. максимальное r -число заявок на орбите может быть равно R , $R < \infty$. Если p -заявка поступила на орбиту в момент, когда в ней уже имеется R заявок, то она теряется.

Политика пополнения запасов определяется следующим образом. Отпуск запасов по заявкам продолжается, пока склад системы не является пустым. Определяется некоторая пороговая величина s , $s < S$, и если уровень запасов на складе выше этой величины, то система не делает заказов для пополнения запасами; а когда текущий уровень запасов системы m становится меньшим или равным величине s , делается заказ такого объема, чтобы полностью заполнить склад системы.

Сделанный заказ выполняется с некоторой задержкой, при этом для общности здесь предполагается, что указанная задержка имеет показательную ф.р. с параметром $v_n(m)$, который в общем случае зависит от текущего уровня запасов m на складе системы, $m = 0, 1, \dots, s$, и числа r -заявок на орбите, $n = 0, 1, \dots, R$.

Задача исследования состоит в определении совместного распределения уровня запасов системы и числа заявок на орбите. Решение этой задачи позволяет вычислить характеристики изучаемых систем обслуживания-запасания (СОЗ): средний уровень запасов на складе (S_{av}), вероятности потери p -заявок (P_p) и r -заявок (P_r), а также среднее число r -заявок на орбите (L_o).

Работа системы описывается двумерной конечной цепью Маркова (ЦМ) с состояниями вида (m, n) , где m означает уровень запасов сис-

темы, а n указывает число r -заявок на орбите. Фазовое пространство состояний (ФПС) определяется так:

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

Интенсивность перехода из состояния $(m_1, n_1) \in E$ в другое состояние $(m_2, n_2) \in E$ обозначим как $q((m_1, n_1), (m_2, n_2))$. Переходы между состояниями ФПС (1) связаны со следующими событиями: поступлением p -заявок, уходом p -заявок на орбиту, поступлением r -заявки из орбиты и поступлением запасов со склада верхнего уровня.

Элементы Q-матрицы данной модели определяются так:

случай $m_1 > s$:

$$q((m_1, n_1), (m_2, n_2)) = \begin{cases} \lambda, & \text{если } m_2 = m_1 - 1, n_2 = n_1, \\ n_1 \eta, & \text{если } m_2 = m_1 - 1, n_2 = n_1 - 1, \\ 0, & \text{в остальных случаях;} \end{cases} \quad (2)$$

случай $0 < m_1 \leq s$:

$$q((m_1, n_1), (m_2, n_2)) = \begin{cases} \lambda, & \text{если } m_2 = m_1 - 1, n_2 = n_1, \\ n_1 \eta, & \text{если } m_2 = m_1 - 1, n_2 = n_1 - 1, \\ v_{n_1}(m_1), & \text{если } m_2 = S, n_2 = n_1, \\ 0, & \text{в остальных случаях;} \end{cases} \quad (3)$$

случай $m_1 = 0$:

$$q((0, n_1), (m_2, n_2)) = \begin{cases} \lambda H_p, & \text{если } m_2 = 0, n_2 = n_1 + 1, \\ n_1 \eta H_r, & \text{если } m_2 = 0, n_2 = n_1 - 1, \\ v_{n_1}(0), & \text{если } m_2 = S, n_2 = n_1, \\ 0, & \text{в остальных случаях.} \end{cases} \quad (4)$$

Стационарную вероятность состояния $(m, n) \in E$ обозначим через $p(m, n)$. Эти величины находятся в результате решения системы уравнения равновесия (СУР), которая известным способом составляется на основе соотношений (2) – (4).

Нахождение стационарных вероятностей состояний позволяет вычислить характеристики исследуемой системы. Действительно, исполь-

зуя результаты работ [4, 5], можно показать, что искомые величины определяются так:

$$S_{av} = \sum_{m=1}^S m \sum_{n=0}^R p(m, n); \quad (5)$$

$$L_o = \sum_{n=1}^R n \sum_{m=0}^S p(m, n); \quad (6)$$

$$P_p = p(0, N) + (1 - H_p) \sum_{n=0}^{R-1} p(0, n); \quad (7)$$

$$P_r = H_r \sum_{n=1}^R p(0, n). \quad (8)$$

Отметим, что СУР данной системы имеет матрицу сложной структуры и потому не удалось найти ее аналитическое решение, т.е. возникает необходимость использования стандартных численных методов теории марковских процессов. Указанные методы являются работоспособными для моделей умеренной размерности и становятся бесполезными для моделей большой размерности.

Ниже предлагается приближенный метод расчета стационарного распределения данной модели. Он использует подход, разработанный в [4, 5].

Этот метод имеет высокую точность для моделей СУЗ, в которых интенсивность p -заявок намного превосходит интенсивности r -заявок. При выполнении данного допущения рассмотрим следующее расщепление исходного ФПС (1):

$$E = \bigcup_{n=0}^R E_n, E_n \cap E_{n'} = \emptyset, n \neq n', \quad (9)$$

где $E_n = \{(m, n) \in E : m = 0, 1, \dots, S\}, n = 0, 1, \dots, R$.

Далее в исходном ФПС (1) определяется следующая функция укрупнения:

$$U(m, n) = \langle n \rangle, \quad (10)$$

где $\langle n \rangle$ означает укрупненное состояние, которое объединяет в себе класс состояний $E_n, n = 0, 1, \dots, R$. Обозначим $\Omega = \{\langle n \rangle : n = 0, 1, \dots, R\}$.

Через $\rho_n(m)$ обозначается вероятность состояния (m, n) внутри расщепленной модели с пространством состояний $E_n, n = 0, 1, \dots, R$. Тогда из соотношений (2) – (4) получаем, что эти вероятности вычисляются так:

$$\rho_n(m) = \begin{cases} \frac{1}{\lambda^m} \prod_{i=1}^m (v_n(i-1) + \lambda) \rho_n(0), & \text{если } 1 \leq m \leq s+1, \\ \frac{1}{\lambda^{s+1}} \prod_{i=1}^{s+1} (v_n(i-1) + \lambda) \rho_n(0), & \text{если } s+1 < m \leq S, \end{cases} \quad (11)$$

где $\rho_n(0)$ находится из условия нормировки, т.е. $\sum_{m=0}^S \rho_n(m) = 1$.

Укрупненная модель описывается процессом размножения и гибели. Интенсивность перехода из укрупненного состояния $\langle n_1 \rangle$ в другое укрупненное состояние $\langle n_2 \rangle$ обозначим через $q(\langle n_1 \rangle, \langle n_2 \rangle)$. Тогда с учетом (2) – (4) и (11) после определенных математических преобразований получаем, что указанные величины вычисляются так:

$$\begin{aligned} q(\langle i \rangle, \langle i+1 \rangle) &= \lambda H_p \rho_n(0), \quad i = 0, 1, \dots, R-1; \\ q(\langle i \rangle, \langle i-1 \rangle) &= i \eta (1 - (1 - H_r) \rho_n(0)), \quad i = 1, 2, \dots, R; \\ q(\langle i \rangle, \langle j \rangle) &= 0, \quad \text{если } |i - j| > 1. \end{aligned} \quad (12)$$

Тогда из системы соотношений (12) находим, что вероятности укрупненных состояний $\pi(\langle n \rangle), \langle n \rangle \in \Omega$ вычисляются следующим образом:

$$\pi(\langle n \rangle) = \prod_{i=0}^{n-1} \frac{q(i, i+1)}{q(i+1, i)} \pi(\langle 0 \rangle), \quad n = 1, \dots, R, \quad (13)$$

где вероятность $\pi(\langle 0 \rangle)$ находится из условия нормировки, т.е.

$$\sum_{n=0}^R \pi(\langle n \rangle) = 1.$$

Вероятности состояний исходной модели приближенно определяются так:

$$p(m, n) \approx \rho_n(m) \pi(\langle n \rangle). \quad (14)$$

Тогда из соотношений (11) – (14) находим следующие формулы для приближенного расчета характеристик (5) – (8) исследуемой системы:

$$S_{av} \approx \sum_{m=1}^S m \sum_{n=0}^R \rho_n(m) \pi(< n >); \quad (15)$$

$$L_o \approx \sum_{n=1}^R n \pi(< n >); \quad (16)$$

$$P_p \approx (1 - H_p) \sum_{n=0}^{R-1} \rho_n(0) \pi(< n >) + \rho_R(0) \pi(< R >); \quad (17)$$

$$P_r \approx H_r \sum_{n=1}^R \rho_n(0) \pi(< n >). \quad (18)$$

Таким образом, изучена модель системы управления запасами с мгновенным обслуживанием и повторными заявками. В этой системе используется политика пополнения запасов, согласно которой объемы заказов являются переменными величинами, при этом их значения зависят от текущего уровня запасов и числа повторных заявок на орбите. Изучена модель с ограниченным размером орбиты. Получены явные формулы для расчета характеристик изучаемой системы.

ЛИТЕРАТУРА

1. Schwarz M., Daduna H. Queuing systems with inventory management with random lead times and with backordering // Mathematical Methods of Operations Research, 2006. – V. 64. – No. 3. – P. 383–414.
2. Schwarz M., Sauer C., Daduna H., Kulik R., Szekli R. M/M/1 queueing systems with inventory // Queuing systems. Theory and applications, 2006. – V. 54. – No. 1. – P. 55–78.
3. Рыжиков Ю.И. Теория очередей и управление запасами. СПб.: Питер, 2001. – 384 с.
4. Melikov A.Z., Ponomarenko L.A., Bagirova S.A. Analysis of queueing-inventory systems with impatience consume customers // J. Automation and Information Sciences, 2016. – V. 48. – No. 1. – P. 53–68.
5. Melikov A.Z., Ponomarenko L.A., Bagirova S.A. Models of queueing-inventory systems with randomized lead policy // J. Automation and Information Sciences, 2016. – V. 48. – No. 10. – P. 23–35.
6. Demchenko I. Yu. Stationary distribution in inventory control models // Cybernetics and Systems Analysis. – 2016. – V. 52. – No. 2. – P. 319–322.

ПРИМЕНЕНИЕ МЕТОДА РАСЩЕПЛЕНИЯ ДЛЯ ИМИТАЦИОННОГО МОДЕЛИРОВАНИЯ УПРАВЛЯЕМОЙ ДЕГРАДИРУЮЩЕЙ СИСТЕМЫ*

А.В. Бородина^{1,2}, Д.В. Ефросинин^{3,4}, Е.В. Морозов^{1,2}

¹*Институт прикладных математических исследований
КарНЦ РАН, г. Петрозаводск, Россия*

²*Петрозаводский государственный университет, г. Петрозаводск, Россия*

³*Российский университет дружбы народов, г. Москва, Россия*

⁴*Институт проблем управления РАН, г. Москва, Россия*

В последние годы модели старения и деградации для технических и биологических объектов достаточно популярны [1–3]. Модели старения и деградации предполагают изучение систем с постепенными отказами, для которых было разработано большое количество многоуровневых моделей надежности (см., например, [4, 5]). Предлагается метод ускоренного имитационного моделирования для оценивания основных характеристик процесса деградации в системе с постепенными и внезапными отказами.

Процесс деградации

Рассмотрим процесс деградации $X = \{X(t)\}_{t \geq 0}$ с фиксированным пространством состояний $E = \{0, 1, \dots, L, M, \dots, K, F\}$, который представляют стадии деградации системы (рис. 1.) Процесс стартует из состояния $X(0) = 0$ и затем переходит от одной стадии к другой (*постепенный* или *наблюдаемый* отказ), время пребывания на стадии i определяется значением с.в. T_i . Переход из состояния i в j происходит за время

$$S_{ij} = \sum_{k=i}^{j-1} T_k,$$

где $0 \leq i \leq K-1$, $j > i$, $\{T_k\}$ предполагаются н.о.р. с.в.

* Исследование выполнено при финансовой поддержке РФФИ в рамках научных проектов № 15-07-02341, 15-07-02354, 15-07-02360, 16-37-60072, 16-37-60072 mol_a_dk. Работа поддержана Министерством образования РФ (номер 02.a03.21.0008 от 24 июня 2016 г.)

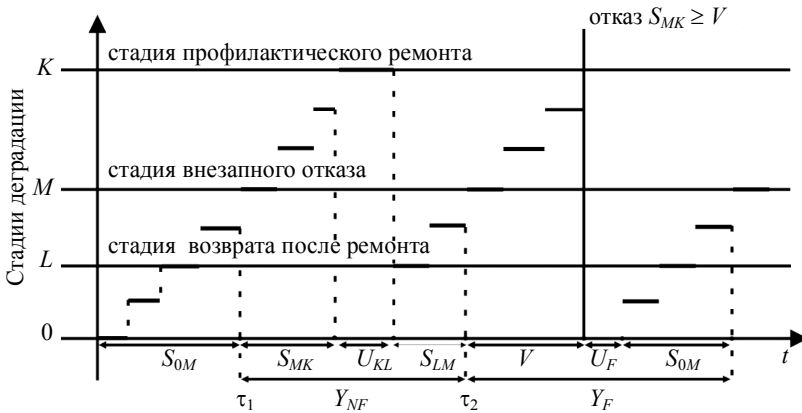


Рис. 1. Динамика процесса деградации с двумя типами циклов регенерации

После достижения стадии M либо произойдет переход в состояние K за время

$$S_{MK} = \sum_{i=M}^{K-1} T_i$$

и далее будет выполнен профилактический ремонт, либо случится отказ системы за случайное время $V \leq S_{MK}$. Процесс X является регенерирующим с моментами регенерации $\{\tau_i\}$ (момент перехода на стадию M) и имеет два типа циклов регенерации: с отказом и без. Необходимо вычислить вероятность внезапного отказа

$$p_F = P[S_{MK} \geq V], \quad (1)$$

а также такие характеристики, как время до случайного отказа, среднюю длину цикла регенерации, средние длины циклов с отказом и без отказа, определить значение асимптотической функции надежности в заданной точке и т. д.

Поскольку оценить аналитически параметры деградирующего процесса и вероятность внезапного отказа, как правило, не представляется возможным, то необходимо предложить метод имитационного моделирования, который также будет эффективен для случаев, когда отказ является редким событием (в этом случае метод Монте-Карло неэффективен).

Метод расщепления для процесса деградации

Для построения циклов регенерации будем применять модификацию стандартного динамического метода расщепления [6–8], которая объединяет технику расщепления и регенеративное имитационное моделирование.

Расщепление траекторий процесса деградации начинается в моменты регенерации $\{\tau_i\}$. Стадии деградации мы используем в качестве порогов для метода расщепления. Таким образом, для нашей задачи пороги являются жестко заданными, в отличие от стандартного алгоритма расщепления, где значения порогов (уровней) выбираются специальным образом. Каждый момент регенерации с учетом расщепления дает группу из $D = R_M \cdot \dots \cdot R_{K-1}$ зависимых циклов регенерации, где R_i – это количество расщеплений на каждой стадии деградации. Общее количество отказов в i -й группе

$$A_i = \sum_{j=(i-1) \cdot D+1}^{i \cdot D} I^{(j)}, i = 1, \dots, R_{M-1},$$

где $I^{(j)} = 1$ ($I^{(j)} = 0$) для цикла с отказом (без отказа) и группы являются н.о.р. Оценка \hat{p}_F для вероятности (1) является несмещенной и строго состоятельной, т.е. при $R_i \rightarrow \infty$ имеет место сходимость с в. 1:

$$\hat{p}_F = \frac{\sum_{j=1}^{R_{M-1}} A_j}{R_{M-1} \cdot D} \rightarrow \frac{E \sum_{j=1}^D I^{(j)}}{D} = p_F.$$

Кроме того, регенеративные свойства последовательности $\{I^{(j)}, j \geq 1\}$ дают основания для построения доверительного интервала для (1) с использованием регенеративного метода [9–11].

Заключение

Результаты имитационного моделирования показали, что при равном числе циклов регенерации время, затраченное на моделирование методом расщепления, значительно меньше, чем у метода Монте-Карло. При этом, метод расщепления дает хорошую близость оценок характе-

ристик деградирующего процесса к аналитическим значениям, в однородном случае для экспоненциально распределенных с.в. $\{T_i\}$.

ЛИТЕРАТУРА

1. Esary J.D., Marshall A.W., Proshan F. Shock models and wear processes // The Annals of Probability. – 1973. – V. 1. – No. 1. – P. 627–649.
2. Kopnov V.A., Timashev S.A. Optimal deatch process control in two-level policies // Proc. 4th Vilnius Conf. on Probability Theory and Statistics. – 1985. – V. 4. – P. 308–309.
3. Murphy D.N.P., Iskandar B.P. A new shock damage model: Part II – Optimal maintenance policies // Reliability of Engineering System Safety. – 1991. – V. 31. – P. 211–231.
4. Lisnuansky A., Levitin G. Multi-state system reliability. assessment, optimization and application. – New Jersey, London, Singapore, Hong-Kong: World Scientific, 2003.
5. Rykov V., Dimitrov B. On multi-state reliability systems // Proc. of Seminar Applied Stochastic Models and Information Processes. – 2002. – P. 128–135.
6. Garvels M. PhD Thesis. The splitting method in rare event simulation. – The University of Twente, The Netherlands, May, 2000.
7. Glasserman P., Heidelberger P., Shahabuddin P., Zajic T. A look at multilevel splitting. In H. Niederreiter, editor, Monte Carlo and Quasi Monte Carlo Methods // Lecture Notes in Statistics. – Springer Verlag, 1996. – V. 127. – P. 99–108.
8. Rubinstein R.Y., Kroese D.P. Simulation and the Monte Carlo method. – Hoboken, New Jersey: John Wiley & Sons, Inc., 2017.
9. Glynn P.W. Some topics in regenerative steady state simulation // Acta Applic. Math. – 1994. – V. 34. – P. 225–236.
10. Glynn P.W., Iglehart D.L. Conditions for the applicability of the regenerative method // Management Science, 1993. – V. 39. – P. 1108–1111.
11. Glynn P.W., Iglehart D.L. A joint central limit theorem for the sample mean and regenerative variance estimator // Annals of Operations Research. – 1987. – V. 8. – P. 41–55.

АЛГОРИТМ РАСЧЕТА ОДНОЙ СХЕМЫ РАЗДЕЛЕНИЯ КАНАЛОВ В СЕТЯХ С ШИРОКОПОЛОСНЫМИ ВЫЗОВАМИ ДАННЫХ

Г.М. Велиджанова

*Институт систем управления НАН Азербайджана,
г. Баку, Азербайджан*

Проблемы распределения общего пула каналов между разнотипными вызовами подробно изучены для моделей, в которых предполагается, что разнотипные вызовы требуют одинаковых полос пропускания (каналов). Вместе с тем в сетях сотовой связи (ССС) вызовы данных для их передачи требуют одновременно несколько каналов (т.е. вызовы данных являются широкополосными), в то время как для передачи вызовов речи достаточно лишь одного свободного канала (т.е. вызовы речи являются узкополосными). Анализ доступной литературы показал, что задачи разделения общего пула каналов между разнотипными вызовами в таких моделях недостаточно изучены (см. [1–4] и список их литературы).

Автором предлагается одна схема изолированного разделения каналов. Она идейно близка к схеме [5], но отличается от нее по следующим моментам. Во-первых, в отличие от указанной схемы, здесь различаются новые и хэндовер-вызовы данных; во-вторых, здесь предполагается, что общая зона каналов может быть использована лишь хэндовер-вызовами данных и речи. Главное отличие этой работы от [5] состоит в том, что здесь для исследования этой схемы предложен аналитический подход. Подобный подход ранее был использован в работе [6], где изучены модели СССР, в которых разнотипные вызовы требуют одинаковых полос пропускания.

Базовая станция изолированной соты сети имеет $N > 1$ радиоканалов. В этих каналах обрабатываются четыре пуассоновских потока вызовов: хэндовер-речевые вызовы (hv -вызовы), новые речевые вызовы (ov -вызовы), хэндовер-вызовы данных (hd -вызовы) и новые вызовы данных (od -вызовы).

Вызовы речи (новые и хэндовер) требуют для обработки лишь один канал, а для обработки вызовов данных (новых и хэндовер) требуются

одновременно b , $b > 1$, каналов. Весь пул каналов разделен на три группы: N_v каналов из общего числа N каналов выделяются лишь для речевых вызовов (v -вызовов), N_d каналов из общего числа N каналов выделяются лишь для вызовов данных (d -вызовы), а остальные $N_{vd} = N - N_v - N_d$ каналов используются совместно вызовами речи и данных. Иными словами, весь пул из N каналов делится на три зоны: индивидуальную зону из N_v каналов (v -зона лишь для v -вызовов), индивидуальную зону из N_d каналов (d -зона лишь для d -вызовов) и общую из N_{vd} каналов (vd -зона для v - и d -вызовов). Изолированность распределения каналов означает, что ни один канал не может быть переведен из одной зоны в другую. При этом для эффективного использования d -зоны каналов считается, что b является делителем N_d , т.е. $N_d = K_d b$, где K_d является максимально возможным числом d -вызовов в указанной зоне каналов.

Интенсивность x -вызовов равна λ_x , $x \in \{hv, ov, hd, od\}$. Функции распределения времени занятости каналов разнотипными вызовами являются экспоненциальными, среднее время занятости канала для одного речевого вызова (нового или хэндовер) равно $1/\mu_v$, а соответствующий показатель для вызовов данных (новых или хэндовер) равен $1/\mu_d$. Идентичность средних времен занятия каналов новых и хэндовер-вызовов обоих типов объясняется отсутствием памяти экспоненциального распределения. Функция распределения времени занятости канала v -вызовом (d -вызовом) каждого типа в vd -зоне также является экспоненциальной с тем же средним.

Разнотипные вызовы используют каналы соответствующих индивидуальных зон согласно полнодоступной схеме (Complete Sharing, CS), т.е. если в момент поступления вызова речи (данных) любого типа имеется хотя бы один свободный канал в соответствующей индивидуальной зоне, то он занимает любой свободный канал этой зоны; иначе этот свободный канал ищется в общей зоне. Это означает, что поступивший вызов речи (новый или хэндовер) принимается в vd -зону, если в этой зоне имеется хотя бы один свободный канал; в противном случае, поступивший в vd -зону вызов речи теряется. Точно так же поступивший вызов данных (новый или хэндовер) принимается в vd -зону, если в этой зоне имеются хотя бы b свободных каналов; в противном случае, т.е.

если число свободных каналов в общей зоне меньше, чем b , то поступивший в vd -зону вызов данных теряется.

Основными показателями QoS разнотипных вызовов являются вероятности их потери, $P_x, x \in \{hv, ov, hd, od\}$, и коэффициент использования каналов общей зоны, C_u . Проблема состоит в нахождении этих параметров.

Согласно данной схеме, v -вызовы и d -вызовы, которые получили отказ в своей (индивидуальной) зоне каналов, поступают в общую зону. Потому интенсивность поступления v -вызовов ($\tilde{\lambda}_v$) и d -вызовов ($\tilde{\lambda}_d$) в общую зону определяются так:

$$\tilde{\lambda}_v = \lambda_v P_v^{(i)}, \quad \tilde{\lambda}_d = \lambda_d P_d^{(i)}, \quad (1)$$

где $\lambda_v = \lambda_{ov} + \lambda_{hv}$, $\lambda_d = \lambda_{od} + \lambda_{hd}$ и $P_v^{(i)}$ и $P_d^{(i)}$ обозначают соответственно вероятности потери ov -вызовов и hv -вызовов в v -зоне каналов.

Очевидно, что величины $P_v^{(i)}$ и $P_d^{(i)}$ определяются как вероятности потери в классических моделях Эрланга $M/M/N_v/0$ и $M/M/K_d/0$ с нагрузками $v_v = \lambda_v/\mu_v$ и $v_d = \lambda_d/\mu_d$ соответственно, т.е. с помощью В-формулы Эрланга:

$$P_v^{(i)} = \frac{v_v^{N_v}}{N_v!}; \quad P_d^{(i)} = \frac{v_d^{K_d}}{K_d!}. \quad (2)$$

Потерянные в индивидуальных зонах разнотипные вызовы поступают в общую зону, которая имеет $N_{vd} > 1$ каналов. Поэтому далее нужно исследовать систему обслуживания с двумя пуассоновскими потоками с интенсивностями $\tilde{\lambda}_v$ и $\tilde{\lambda}_d$.

Поскольку средние времена обработки новых и хэндовер вызовов каждого типа являются одинаковыми, то состояние данной системы в произвольный момент времени определяется двумерным вектором $\mathbf{n} = (n_d, n_v)$, где n_d и n_v указывают соответственно суммарное число d -вызовов и v -вызовов в общей зоне. Разнотипные вызовы равноправно используют каналы общей зоны, поэтому фазовое пространство состоя-

ний (ФПС) соты определяется так:

$$S = \{ \mathbf{n} : n_d = 0, 1, \dots, [N_{vd} / b], n_v = 0, 1, \dots, N_{vd}; n_d b + n_v \leq N_{vd} \}, \quad (3)$$

где $[x]$ означает целую часть x .

Элементы производящей матрицы данной двумерной цепи Маркова (ЦМ) обозначим через $q_1(\mathbf{n}, \mathbf{n}')$. Эти величины определяются из следующих соотношений:

$$q(\mathbf{n}, \mathbf{n}') = \begin{cases} \tilde{\lambda}_d, & \text{если } \mathbf{n}' = \mathbf{n} + \mathbf{e}_1, \\ \tilde{\lambda}_v, & \text{если } \mathbf{n}' = \mathbf{n} + \mathbf{e}_2, \\ n_d \mu_d, & \text{если } \mathbf{n}' = \mathbf{n} - \mathbf{e}_1, \\ n_v \mu_v, & \text{если } \mathbf{n}' = \mathbf{n} - \mathbf{e}_2, \\ 0, & \text{в остальных случаях,} \end{cases} \quad (4)$$

где $\mathbf{e}_1 = (1, 0)$, $\mathbf{e}_2 = (0, 1)$.

Вероятность состояния (n_d, n_v) обозначим через $p(n_d, n_v)$.

Утверждение. Стационарное распределение вероятностей состояний системы имеет следующий вид:

$$p(n_d, n_v) = p(0, 0) \frac{\tilde{v}_d^{n_d} \tilde{v}_v^{n_v}}{n_d! n_v!}, \quad (5)$$

где $\tilde{v}_d = \tilde{\lambda}_d / \mu_d$, $\tilde{v}_v = \tilde{\lambda}_v / \mu_v$, $p_1(0, 0)$ находится из условия нормировки, т.е. $\sum_{(n_d, n_v) \in S} p(n_d, n_v) = 1$.

Доказательство. Из теоремы Колмогорова [7] об обратимости двумерных цепей Маркова получаем, что в данной системе существует нулевая циркуляция между ее состояниями, т.е. удовлетворяется условие локального баланса. Тогда удастся выразить вероятность любого состояния (n_d, n_v) через вероятность состояния $(0, 0)$, при этом можно выбирать любой путь между этими двумя состояниями в соответствующем графе переходов между состояниями цепи. Таким образом, выбирая путь $(0, 0), (1, 0), \dots, (n_d, 0), (n_d, 1), \dots, (n_d, n_v)$ между состояниями $(0, 0)$ и (n_d, n_v) , получим, что мультипликативное представление (5) справедливо для нахождения вероятностей состояний изучаемой модели.

Поскольку речевые вызовы теряются тогда, когда все каналы общей зоны заняты, то для вычисления вероятности потери таких вызовов получим следующую формулу:

$$P_v = \sum_{(n_d, n_v) \in S} p(n_d, n_v) I(n_v + bn_d = N_{vd}), \quad (6)$$

где $I(A)$ – индикаторная функция события A .

Вызовы данных теряются тогда, когда число свободных каналов общей зоны меньше, чем b , следовательно вероятность потери таких вызовов вычисляется из следующей формулы:

$$P_d = \sum_{(n_d, n_v) \in S} p(n_d, n_v) I(N_{vd} - n_v - bn_d < b). \quad (7)$$

Коэффициент использования каналов общей зоны определяется так:

$$C_u = \left(\sum_{(n_d, n_v) \in S} p(n_d, n_v) (n_v + bn_d) \right) / N_{vd}. \quad (8)$$

Выводы

Предложена одна схема разделения каналов в интегральных сотовых сетях связи, в которых осуществляется обработка речевых вызовов и вызовов данных. Предполагается, что вызовы данных являются широкополосными. Весь пул каналов изолированно разделен между разнотипными вызовами, при этом имеется общая зона каналов для обслуживания разнотипных вызовов. Индивидуальные зоны и общая зона каналов используются равноправно вызовами любого типа. Разработан алгоритм расчета показателей QoS предложенной схемы разделения каналов при заданных значениях числа каналов и нагрузок разнотипных вызовов. С помощью этого алгоритма можно проводить расчет характеристик системы.

ЛИТЕРАТУРА

1. *Ogbonmwan S.E., Wei L.* Multi-threshold bandwidth reservation scheme of an integrated voice/data wireless network // *Computer Communications.* – 2006. – V. 29. – No. 9. – P. 1504–1515.
2. *Меликов А.З., Фаттахова М.И., Велиджанова Г.М.* Метод расчета параметров интегральной сотовой сети связи с изолированным разделением каналов // *Электронное моделирование.* – 2014. – Т. 36. – № 5. – С. 37–48.

3. Melikov A.Z., Ponomarenko L.A., Velidzanova G.M. Analysis of integrated cellular network model with virtual partitioning of channels // *Cybernetics and System Analysis*. – 2014. – V. 50. – No. 6. – P. 884–890.
4. Melikov A.Z., Ponomarenko L.A. Analysis of models of the integral cellular communication network with different schemes of channel partitioning // *Automation and Remote Control*, 2016. – V. 77. – No. 5. – P. 802–816.
5. Li B., Li L., Li B., Sivalingam K.M., Cao X.R. Call admission control for voice/data integrated cellular networks: Performance analysis and comparative study // *IEEE Journal of Selected Areas on Communications*. – 2004. – V. 22. – No. 4. – P. 706–718.
6. Меликов А.З., Пономаренко Л.А., Велиджанова Г.М. Многопараметрические гибридные схемы изолированного разделения каналов в интегральных сотовых сетях связи // *Управляющие системы и машины*. – 2016. – № 1. – С. 62–72.
7. Kolmogorov A. Zum theorie der Markoffschen ketten // *Mathematische Annalen*. – 1936. – V. B112. – P. 155–160.

ИССЛЕДОВАНИЕ ДВУХФАЗНОЙ БЕСКОНЕЧНОЛИНЕЙНОЙ СМО С ММРР-ВХОДЯЩИМ ПОТОКОМ ТРЕБОВАНИЙ СЛУЧАЙНОГО ОБЪЕМА*

А.А. Галилейская, Е.Ю. Лисовская

Томский государственный университет, г. Томск, Россия

Рассматривается двухфазная ресурсная система массового обслуживания (СМО), которая характеризуется наличием буфера, имеющего неограниченный размер. Каждое поступающее в систему требование занимает некоторый случайный объем памяти в буфере, а по окончании обслуживания освобождает его. Объем занимаемой памяти не меняется в течение пребывания требования в системе [1].

Данные системы представляют интерес для практических приложений, в частности для систем передачи данных при их проектировании или оптимизации [2], а также для моделирования процессов в социально-экономических системах (страховых компаний, банков, негосударственных пенсионных фондов).

В работах [3, 4] были получены асимптотические характеристические функции первого и второго порядков в условии высокой интенсивности непуассоновских входящих потоков, показано, что распределение вероятностей суммарного объема требований в системе является асимптотически гауссовским.

Математическая модель

Рассмотрим двухфазную СМО с неограниченным числом приборов, с поступающим на вход ММРР-поток заявок. Данный поток, управляемый цепью Маркова $k(t) \in 1, 2, \dots, K$, задается матрицей инфинитезимальных характеристик $Q = \|q_{ij}\|$ размера $K \times K$ и набором неотрицательных величин $\lambda_1, \dots, \lambda_K$. Предполагаем, что каждое требование характеризуется некоторым случайным объемом $v > 0$. $G(y)$ – функция

* Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 16-31-00292 мол_а.

распределения случайной величины v , $B_1(x)$ и $B_2(x)$ – функции распределения продолжительности обслуживания заявок на первой фазе и второй фазах соответственно. После обслуживания на первой фазе заявка с тем же объемом переходит на вторую фазу, после обслуживания на второй фазе заявка покидает систему и «уносит» с собой свой объем.

Пусть $i_k(t)$ – число заявок на k -й фазе в момент времени t , $V_k(t)$ – суммарный объем заявок на k -й фазе в момент времени t , где $k = 1, 2$ – номер фазы.

Поставим задачу нахождения характеристик четырехмерного случайного процесса $\{i_1(t), V_1(t), i_2(t), V_2(t)\}$. Отметим, что исследуемый процесс не является марковским, поэтому для его исследования применим метод динамического просеивания.

Метод динамического просеивания

Более подробно просеивание заявок рассматривается в работах [3–5]. Отметим лишь, динамические вероятности просеивания заявок входящего потока.

Вероятность того, что заявка входящего потока, поступившая в систему в момент времени $t > t_0$ к моменту времени T не закончит обслуживание на первой фазе, то есть просеется на первую ось, равна $S_1(t) = 1 - B_1(T - t)$. Вероятность того, что заявка входящего потока, поступившая в систему в момент времени $t > t_0$, к моменту времени T закончит обслуживание на первой фазе и не закончит на второй, то есть просеется на вторую ось, равна $S_2(t) = B_1(T - t) - B_2^*(T - t)$, где $B_2^*(\tau) = (B_1 * B_2)(\tau)$ – свертка функций распределения $B_1(x)$, $B_2(x)$ длительности обслуживания на фазах системы. Причем заявка может не просеяться ни на одну из фаз с вероятностью $S_0(t) = 1 - S_1(t) - S_2(t)$, то есть к моменту времени T заявка закончит обслуживание на обеих фазах и покинет систему.

Обозначим через $n_k(t)$ – число событий, наступивших на k -й оси просеянного потока до момента t ; $W_k(t)$ – суммарный объем просеянных заявок на k -ю ось.

Нетрудно показать, что для процесса $\{i_1(t), V_1(t), i_2(t), V_2(t)\}$ справедливо

$$\begin{aligned} &P\{i_1(T) = m_1, V_1(T) < z_1, i_2(T) = m_2, V_2(T) < z_2\} = \\ &= P\{n_1(T) = m_1, W_1(T) < z_1, n_2(T) = m_2, W_2(T) < z_2\}, \end{aligned} \quad (1)$$

для любых $m_1, m_2 = 0, 1, 2, \dots$ и $z_1, z_2 \geq 0$. Будем использовать равенство (1) для исследования процесса $\{i_1(t), V_1(t), i_2(t), V_2(t)\}$ с помощью исследования процесса $\{n_1(t), W_1(t), n_2(t), W_2(t)\}$.

Система дифференциальных уравнений Колмогорова

Введем обозначение распределения вероятностей марковского процесса $\{k(t), n_1(t), W_1(t), n_2(t), W_2(t)\}$:

$$P(k, n_1, w_1, n_2, w_2, t) = P\{k(t) = k, n_1(t) = n_1, W_1(t) < w_1, n_2(t) = n_2, W_2(t) < w_2\},$$

где $k(t)$ – состояние управляющей цепи Маркова в момент времени t . Для этого распределения составим Δt -методом прямую систему дифференциальных уравнений Колмогорова:

$$\begin{aligned} \frac{\partial P(k, n_1, w_1, n_2, w_2, t)}{\partial t} = & -\lambda_k (S_1(t) + S_2(t)) P(k, n_1, w_1, n_2, w_2, t) + \\ & + \lambda_k S_1(t) \int_0^{w_1} P(k, n_1 - 1, w_1 - y, n_2, w_2, t) dG(y) + \\ & + \lambda_k S_2(t) \int_0^{w_2} P(k, n_1, w_1, n_2 - 1, w_2 - y, t) dG(y) + \\ & + \sum_v q_{vk} P(v, n_1, w_1, n_2, w_2, t), \end{aligned} \quad (2)$$

при $k = 1, \dots, K, n_1, n_2 = 0, 1, 2, \dots, w_1, w_2 > 0$.

Начальное условие определим в виде

$$P(k, n_1, w_1, n_2, w_2, t_0) = \begin{cases} r(k), & n_1 = w_1 = n_2 = w_2 = 0, \\ 0, & \text{иначе,} \end{cases}$$

где $r(k)$ – стационарная вероятность того, что цепь Маркова будет находиться в состоянии k .

Характеристические функции

Введем характеристическую функцию вида

$$h(k, u_1, v_1, u_2, v_2, t) = \sum_{n_1=0}^{\infty} e^{ju_1 n_1} \int_0^{\infty} e^{jv_1 w_1} \sum_{n_2=0}^{\infty} e^{ju_2 n_2} \int_0^{\infty} e^{jv_2 w_2} P(k, n_1, dw_1, n_2, dw_2, t),$$

где $j = \sqrt{-1}$ – мнимая единица. Тогда из (2) можем записать следующие уравнения:

$$\frac{\partial h(k, u_1, v_1, u_2, v_2, t)}{\partial t} = \lambda_k h(k, u_1, v_1, u_2, v_2, t) \times \\ \times \left[S_1(t) \left(e^{ju_1} G^*(v_1) - 1 \right) + S_2(t) \left(e^{ju_2} G^*(v_2) - 1 \right) \right] + \sum_v q_{vk} h(v, u_1, v_1, u_2, v_2, t),$$

при $k = 1, \dots, K$, где $G^*(v) = \int_0^\infty e^{jvy} dG(y)$.

Перепишем полученную систему в виде матричного уравнения:

$$\frac{\partial \mathbf{h}(u_1, v_1, u_2, v_2, t)}{\partial t} = \mathbf{h}(u_1, v_1, u_2, v_2, t) \cdot \\ \cdot \left[\Lambda \left(S_1(t) \left(e^{ju_1} G^*(v_1) - 1 \right) + S_2(t) \left(e^{ju_2} G^*(v_2) - 1 \right) \right) + \mathbf{Q} \right], \quad (3)$$

с начальным условием $\mathbf{h}(u_1, v_1, u_2, v_2, t_0) = \mathbf{r}$, где

$$\mathbf{h}(u_1, v_1, u_2, v_2, t) = [h(1, u_1, v_1, u_2, v_2, t), \dots, h(K, u_1, v_1, u_2, v_2, t)],$$

$\mathbf{r} = [r(1), \dots, r(K)]$ – вектор стационарного распределения вероятностей управляющей цепи Маркова, удовлетворяющий следующей системе:

$$\begin{cases} \mathbf{rQ} = \mathbf{0}, \\ \mathbf{re} = 1, \end{cases}$$

где \mathbf{e} – единичный вектор-столбец.

Метод асимптотического анализа

Точное решение уравнения (3) в общем случае невозможно, но его можно решить при асимптотическом условии. В данной работе рассматривается асимптотическое условие высокой интенсивности входящего потока и предельно частых изменений состояний управляющей цепи Маркова. Подставим в уравнение (3) $\Lambda = N\Lambda_1$ и $\mathbf{Q} = N\mathbf{Q}_1$, где $N \rightarrow \infty$ – некоторый параметр. Тогда можно записать

$$\frac{1}{N} \frac{\partial \mathbf{h}(u_1, v_1, u_2, v_2, t)}{\partial t} = \mathbf{h}(u_1, v_1, u_2, v_2, t) \times \\ \times \left[\Lambda_1 \left(S_1(t) \left(e^{ju_1} G^*(v_1) - 1 \right) + S_2(t) \left(e^{ju_2} G^*(v_2) - 1 \right) \right) + \mathbf{Q}_1 \right] \quad (4)$$

с начальным условием $\mathbf{h}(u_1, v_1, u_2, v_2, t_0) = \mathbf{r}$. Решение уравнения (4) будем находить методом асимптотического анализа, алгоритм решения изложен подробно в работах [3, 4, 6].

Лемма. Асимптотическая характеристическая функция первого порядка четырехмерного случайного процесса $\{n_1(t), W_1(t), n_2(t), W_2(t)\}$ имеет вид

$$\mathbf{h}(u_1, v_1, u_2, v_2, t) = \mathbf{r} \exp \left\{ jN\kappa_1 (u_1 + v_1 a_1) \int_{t_0}^t S_1(\tau) d\tau + jN\kappa_1 (u_2 + v_2 a_1) \int_{t_0}^t S_2(\tau) d\tau \right\},$$

где $\kappa_1 = \mathbf{r} \mathbf{\Lambda}_1 \mathbf{e}$ – средняя интенсивность входящего потока, a_1 – средний объем заявки.

Теорема. Асимптотическая характеристическая функция второго порядка четырехмерного случайного процесса $\{n_1(t), W_1(t), n_2(t), W_2(t)\}$ имеет вид

$$\begin{aligned} \mathbf{h}(u_1, v_1, u_2, v_2, t) = \mathbf{r} \exp & \left\{ jN\kappa_1 (u_1 + v_1 a_1) \int_{t_0}^t S_1(\tau) d\tau + \right. \\ & + jN\kappa_1 (ju_2 + jv_2 a_1) \int_{t_0}^t S_2(\tau) d\tau + \frac{(ju_1)^2}{2} \left(N\kappa_1 \int_{t_0}^t S_1(\tau) d\tau + N\kappa_2 \int_{t_0}^t S_1^2(\tau) d\tau \right) + \\ & + \frac{(jv_1)^2}{2} \left(N\kappa_1 a_2 \int_{t_0}^t S_1(\tau) d\tau + N\kappa_2 a_1^2 \int_{t_0}^t S_1^2(\tau) d\tau \right) + \\ & + ju_1 jv_1 \left(N\kappa_1 a_1 \int_{t_0}^t S_1(\tau) d\tau + N\kappa_2 a_1 \int_{t_0}^t S_1^2(\tau) d\tau \right) + \\ & + \frac{(ju_2)^2}{2} \left(N\kappa_1 \int_{t_0}^t S_2(\tau) d\tau + N\kappa_2 \int_{t_0}^t S_2^2(\tau) d\tau \right) + \\ & \left. + \frac{(jv_2)^2}{2} \left(N\kappa_1 a_2 \int_{t_0}^t S_2(\tau) d\tau + N\kappa_2 a_1^2 \int_{t_0}^t S_2^2(\tau) d\tau \right) \right\} \end{aligned}$$

$$\begin{aligned}
& +ju_2 jv_2 \left(N\kappa_1 a_1 \int_{t_0}^t S_2(\tau) d\tau + N\kappa_2 a_1 \int_{t_0}^t S_2^2(\tau) d\tau \right) + ju_1 ju_2 N\kappa_1 \int_{t_0}^t S_1(\tau) S_2(\tau) d\tau + \\
& + jv_1 jv_2 N\kappa_2 a_1^2 \int_{t_0}^t S_1(\tau) S_2(\tau) d\tau + (ju_1 jv_2 + ju_2 jv_1) N\kappa_2 a_1 \int_{t_0}^t S_1(\tau) S_2(\tau) d\tau \Big\},
\end{aligned}$$

где $\kappa_2 = 2\mathbf{g}(\mathbf{\Lambda}_1 - \kappa_1 \mathbf{I})\mathbf{e}$, $a_2 = \int_0^\infty y^2 dG(y)$, и \mathbf{g} – вектор-строка, удовлетворяющая системе матричных уравнений

$$\begin{cases} \mathbf{g}\mathbf{Q}_1 = \mathbf{r}(\kappa_1 \mathbf{I} - \mathbf{\Lambda}_1), \\ \mathbf{g}\mathbf{e} = const. \end{cases}$$

Следствие. Для характеристической функции процесса $\{i_1(t), V_1(t), i_2(t), V_2(t)\}$ в стационарном режиме получим

$$\begin{aligned}
h(u_1, v_1, u_2, v_2) &= \exp \{ jN\kappa_1 (u_1 + v_1 a_1) b_1 + \\
& + jN\kappa_1 (ju_2 + jv_2 a_1) b_2 + \frac{(ju_1)^2}{2} (N\kappa_1 b_1 + N\kappa_2 \beta_1) + \\
& + \frac{(jv_1)^2}{2} (N\kappa_1 a_2 b_1 + N\kappa_2 a_1^2 \beta_1) + ju_1 jv_1 (N\kappa_1 a_1 b_1 + N\kappa_2 a_1 \beta_1) + \\
& + \frac{(ju_2)^2}{2} (N\kappa_1 b_2 + N\kappa_2 \beta_2) + \frac{(jv_2)^2}{2} (N\kappa_1 a_2 b_2 + N\kappa_2 a_1^2 \beta_2) + \\
& + ju_2 jv_2 (N\kappa_1 a_1 b_2 + N\kappa_2 a_1 \beta_2) + ju_1 ju_2 N\kappa_1 b + \\
& + jv_1 jv_2 N\kappa_2 a_1^2 b + (ju_1 jv_2 + ju_2 jv_1) N\kappa_2 a_1 b \},
\end{aligned}$$

где

$$\begin{aligned}
b_1 &= \int_0^\infty (1 - B_1(\tau)) d\tau, \beta_1 = \int_0^\infty (1 - B_1(\tau))^2 d\tau, \\
b_2 &= \int_0^\infty \left(B_1(\tau) - \int_0^\tau B_2(\tau - y) dB_1(y) \right) d\tau, \beta_2 = \int_0^\infty \left(B_1(\tau) - \int_0^\tau B_2(\tau - y) dB_1(y) \right)^2 d\tau, \\
b &= \int_0^\infty (1 - B_1(\tau)) \left(B_1(\tau) - \int_0^\tau B_2(\tau - y) dB_1(y) \right) d\tau.
\end{aligned}$$

Очевидно, что случайный процесс $\{i_1(t), V_1(t), i_2(t), V_2(t)\}$ является асимптотически гауссовским с вектором математических ожиданий $\mathbf{a} = N\kappa_1 \cdot [b_1 \quad a_1 b_1 \quad b_2 \quad a_1 b_2]$ и ковариационной матрицей

$$\mathbf{K} = N \cdot \begin{bmatrix} \kappa_1 b_1 + \kappa_2 \beta_1 & \kappa_1 a_1 b_1 + \kappa_2 a_1 \beta_1 & \kappa_1 b & \kappa_2 a_1 b \\ \kappa_1 a_1 b_1 + \kappa_2 a_1 \beta_1 & \kappa_1 a_2 b_1 + \kappa_2 a_1^2 \beta_1 & \kappa_2 a_1 b & \kappa_2 a_1^2 b \\ \kappa_1 b & \kappa_2 a_1 b & \kappa_1 b_2 + \kappa_2 \beta_2 & \kappa_1 a_1 b_2 + \kappa_2 a_1 \beta_2 \\ \kappa_2 a_1 b & \kappa_2 a_1^2 b & \kappa_1 a_1 b_2 + \kappa_2 a_1 \beta_2 & \kappa_1 a_2 b_2 + \kappa_2 a_1^2 \beta_2 \end{bmatrix}.$$

Заключение

Проведено исследование двухфазной ресурсной СМО с бесконечным числом приборов. Показано, что совместное распределение вероятности числа заявок на фазах и их суммарного объема на каждой фазе сходится к четырехмерному гауссовскому распределению в асимптотическом условии высокой интенсивности входящего потока.

ЛИТЕРАТУРА

1. Тихоненко О.М. Системы обслуживания требований случайной длины с ограничениями // Автоматика и телемеханика. – 1991. – № 10. – С. 126–134.
2. Наумов В.А., Самуйлов К.Е., Самуйлов А.К. О суммарном объеме ресурсов, занимаемых обслуживаемыми заявками // Автоматика и телемеханика. – 2016. – № 8. – С. 125–132.
3. Lisovskaya E., Moiseeva S., Pagano M. The total capacity of customers in the MMPP/GI/ ∞ queueing system // Distributed computer and communication networks: control, computation, communications. – Москва: Изд-во РУДН, 2016. – С. 313–325.
4. Лисовская Е.Ю., Моисеева С.П. Асимптотический анализ немарковской бесконечнолинейной системы обслуживания требований случайного объема с входящим рекуррентным потоком // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. – 2017. – № 39. – С. 30–38.
5. Моисеев А.Н. Исследование математических моделей систем и сетей массового обслуживания с высокоинтенсивными непуассоновскими входящими потоками: дис. ... д-р физ.-мат. наук. – Томск, 2016. – 333 с.
6. Назаров А.А., Моисеева С.П. Метод асимптотического анализа в теории массового обслуживания. – Томск: Изд-во НТЛ, 2006. – 112 с.

ИССЛЕДОВАНИЕ МАТЕМАТИЧЕСКОЙ МОДЕЛИ СТРАХОВОЙ КОМПАНИИ В ВИДЕ СИСТЕМЫ МАССОВОГО ОБСЛУЖИВАНИЯ С ВЫСОКОИНТЕНСИВНЫМ ПОТОКОМ ВХОДЯЩИХ РИСКОВ И В СЛУЧАЙНОЙ СРЕДЕ

Д.Д. Даммер, А.А. Назаров

Томский государственный университет, г. Томск, Россия

Исследования в области потоков событий в различных предметных областях показывают, что классические модели потоков (например, пуассоновские) не совсем точно моделируют реальные потоки. В связи с этим достаточно актуальной является задача исследования моделей экономических систем с учетом этого аспекта. Например, интенсивность поступления рисков в страховую компанию не является постоянной величиной, а зависит от воздействия каких-то внешних случайных факторов, таких, как время года, политика государства (закон об обязательном страховании автогражданской ответственности), вероятность стихийных бедствий и т.д. В основном во всех работах, посвященных исследованию математических моделей страховых компаний, находятся их характеристики при пуассоновском потоке входящих рисков. Так, в [1, 2] получено двумерное распределение числа застрахованных рисков и требований на выплату единовременных страховых сумм в модели с ограниченным и неограниченным полем. Автором [3] исследуется модель с учетом неявной рекламы. В данной работе рассматривается математическая модель страховой компании в случайной среде, то есть в случае, когда интенсивность входящего потока, интенсивность наступления страховых случаев и время действия договора страхования не являются постоянными величинами, а зависят от воздействия внешних факторов и меняются со временем.

Рассмотрим модель страховой компании с неограниченным страховым полем [4] в виде системы массового обслуживания с неограниченным числом обслуживающих приборов (рис. 1).

Пусть в компанию поступают риски (заявки), образуя высокоинтенсивный модулированный пуассоновский поток событий [5], управляемый случайным процессом $k(t)$, который является цепью Маркова

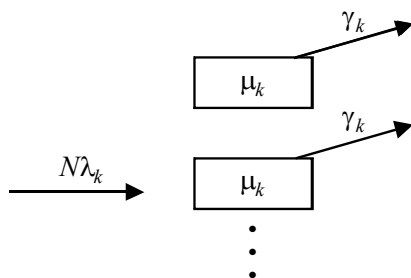


Рис. 1. Модель страховой компании в виде системы массового обслуживания с неограниченным количеством приборов в случайной среде

с непрерывным временем, определяемым матрицей NQ инфинитезимальных характеристик $Nq_{k_1 k_2}$, $k_1 = 1 \dots K$, $k_2 = 1 \dots K$. При этом выполняются условия

$$P\{i(t + \Delta t) = i + 1 \mid i(t) = i, k(t) = k\} = N\lambda_k \Delta t + o(\Delta t),$$

$$P\{i(t + \Delta t) > i + 1 \mid i(t) = i, k(t) = k\} = o(\Delta t).$$

Определим диагональную матрицу $N\Lambda$ с элементами $N\lambda_k$ на главной диагонали. Здесь $N\lambda_k \geq 0$ – интенсивность поступления рисков в компанию, когда цепь Маркова находится в состоянии $k = 1 \dots K$, N – параметр высокой интенсивности потока, λ_k – фиксированная величина. Таким образом, состояние цепи Маркова $k(t)$ определяет состояние случайной среды.

При поступлении в компанию риска (заявки) заключается договор страхования. Срок действия договора соответствует длительности обслуживания заявки на приборе. Каждый риск, находящийся в компании, на протяжении длительности действия договора страхования независимо от других рисков генерирует с интенсивностью γ_k требование на выплату страховых сумм. Эти интенсивности также зависят от состояния среды и образуют диагональную матрицу Γ . Требования на выплату страховых сумм также образуют 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 μ_k , который также зависит от состояния среды. Эти величины образуют диагональную матрицу \mathbf{M} .

Обозначим: $n(t)$ – число требований на выплату за интервал времени $[0, t]$, $i(t)$ – число страховых рисков, находящихся в компании в момент времени t , $P_k(i, n, t) = P\{i(t) = i, n(t) = n, k(t) = k\}$ – 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 , среда находится в состоянии k . Задача состоит в том, чтобы найти распределение вероятностей двумерного процесса $\{i(t), n(t)\}$.

Используя Δt -метод, составим систему дифференциальных уравнений Колмогорова [6] для распределения вероятностей $P_k(i, n, t)$. Сначала запишем допредельные равенства:

$$\begin{aligned} P_k(i, n, t + \Delta t) = & P_k(i, n, t)(1 - N\lambda_k \Delta t)(1 - i\mu_k \Delta t)(1 - i\gamma_k \Delta t)(1 + Nq_{kk} \Delta t) + \\ & + P_k(i - 1, n, t)N\lambda_k \Delta t + P_k(i + 1, n, t)(i + 1)\mu_k \Delta t + \\ & + P_k(i, n - 1, t)i\gamma_k \Delta t + \sum_{v \neq k} P_v(i, n, t)Nq_{vk} \Delta t + o(\Delta t). \end{aligned}$$

Перейдем к пределу при $\Delta t \rightarrow 0$, получим систему дифференциальных уравнений

$$\begin{aligned} \frac{\partial P_k(i, n, t)}{\partial t} = & -P_k(i, n, t)(N\lambda_k + i\mu_k + i\gamma_k) + P_k(i - 1, n, t)N\lambda_k + \\ & + P_k(i + 1, n, t)(i + 1)\mu_k + P_k(i, n - 1, t)i\gamma_k + \sum_{v=1}^K P_v(i, n, t)Nq_{vk}. \end{aligned} \quad (1)$$

Для решения системы (1) определим частичные характеристические функции $H_k(u_1, u_2, t) = \sum_{i, n=0}^{\infty} e^{ju_1 i} e^{ju_2 n} P_k(i, n, t)$, $k = 1 \dots K$, где j – мнимая единица. Тогда из системы (1) с учетом свойств характеристических функций получим дифференциальное уравнение в частных производных первого порядка относительно функции $H_k(u_1, u_2, t)$:

$$\begin{aligned} \frac{\partial H_k(u_1, u_2, t)}{\partial t} = j \frac{\partial H_k(u_1, u_2, t)}{\partial u_1} & \left[\mu_k - \mu_k e^{-ju_1} + \gamma_k - \gamma_k e^{ju_2} \right] + \\ & + N \lambda_k (e^{ju_1} - 1) H_k(u_1, u_2, t) + \sum_{v=1}^K H_v(u_1, u_2, t) N q_{vk}, \quad k = 1 \dots K. \end{aligned} \quad (2)$$

Рассмотрим векторную характеристическую функцию

$$\mathbf{H}(u_1, u_2, t) = \{H_1(u_1, u_2, t), H_2(u_1, u_2, t), \dots, H_K(u_1, u_2, t)\},$$

тогда уравнение (2) относительно этой функции можно записать так:

$$\begin{aligned} \frac{\partial \mathbf{H}(u_1, u_2, t)}{\partial t} = j \frac{\partial \mathbf{H}(u_1, u_2, t)}{\partial u_1} & \left[(1 - e^{-ju_1}) \mathbf{M} + (1 - e^{ju_2}) \mathbf{\Gamma} \right] + \\ & + \mathbf{H}(u_1, u_2, t) \left[N \mathbf{\Lambda} (e^{ju_1} - 1) + N \mathbf{Q} \right]. \end{aligned} \quad (3)$$

Полученное уравнение будем решать методом асимптотического анализа [7] в предельном условии высокой интенсивности потока входящих рисков и предельно частых изменений состояний среды. Сначала рассмотрим асимптотику первого порядка. Для этого в уравнении (3) выполним замены

$$\varepsilon = \frac{1}{N}, \quad u_1 = \varepsilon w_1, \quad u_2 = \varepsilon w_2, \quad \mathbf{H}(u_1, u_2, t) = \mathbf{F}(w_1, w_2, t, \varepsilon),$$

получим

$$\begin{aligned} \varepsilon \frac{\partial \mathbf{F}(w_1, w_2, t, \varepsilon)}{\partial t} = j \frac{\partial \mathbf{F}(w_1, w_2, t, \varepsilon)}{\partial w_1} & \left[(1 - e^{-jw_1 \varepsilon}) \mathbf{M} + (1 - e^{jw_2 \varepsilon}) \mathbf{\Gamma} \right] + \\ & + \mathbf{F}(w_1, w_2, t, \varepsilon) \left[\mathbf{\Lambda} (e^{jw_1 \varepsilon} - 1) + \mathbf{Q} \right]. \end{aligned} \quad (4)$$

Обозначим $\lim_{\varepsilon \rightarrow 0} \mathbf{F}(w_1, w_2, t, \varepsilon) = \mathbf{F}(w_1, w_2, t)$, и эту функцию будем называть асимптотическим решением уравнения (4). В этом же уравнении сделаем предельный переход $\varepsilon \rightarrow 0$, получим $\mathbf{F}(w_1, w_2, t) \mathbf{Q} = \mathbf{0}$, т.е. функция $\mathbf{F}(w_1, w_2, t)$ является решением однородной системы линейных алгебраических уравнений (5). Решение этой системы имеет вид

$$\mathbf{F}(w_1, w_2, t) = \mathbf{R} \Phi(w_1, w_2, t), \quad (6)$$

где $\Phi(w_1, w_2, t)$ – некоторая скалярная функция, а \mathbf{R} – вектор стационарного распределения вероятностей значений цепи Маркова $k(t)$, опре-

деляемой системой уравнений $\mathbf{RQ} = \mathbf{0}$ и условием нормировки $\mathbf{RE} = 1$, где \mathbf{E} – единичный вектор-столбец размерности K . Чтобы найти функцию $\Phi(w_1, w_2, t)$, просуммируем уравнения системы (4), с учетом условия $\mathbf{F}(w_1, w_2, t)\mathbf{Q} = \mathbf{0}$, получим

$$\varepsilon \frac{\partial \mathbf{F}(w_1, w_2, t, \varepsilon)}{\partial t} \mathbf{E} = j \frac{\partial \mathbf{F}(w_1, w_2, t, \varepsilon)}{\partial w_1} \left[(1 - e^{-jw_1\varepsilon}) \mathbf{ME} + (1 - e^{jw_2\varepsilon}) \mathbf{GE} \right] + \\ + \mathbf{F}(w_1, w_2, t, \varepsilon) (e^{jw_1\varepsilon} - 1) \mathbf{LE}. \quad (7)$$

Разделим выражение (7) на ε , совершив предельный переход $\varepsilon \rightarrow 0$, с учетом (6) и условия $\mathbf{RE} = 1$, получим уравнение

$$\frac{\partial \Phi(w_1, w_2, t)}{\partial t} = -w_1 \frac{\partial \Phi(w_1, w_2, t)}{\partial w_1} \mathbf{RME} + w_2 \frac{\partial \Phi(w_1, w_2, t)}{\partial w_1} \mathbf{RGE} + \\ + jw_1 \Phi(w_1, w_2, t) \mathbf{RAE}, \quad (8)$$

с начальным условием

$$\Phi(w_1, w_2, 0) = \Phi(w_1) = \exp \left\{ j \frac{\mathbf{RAE}}{\mathbf{RME}} w_1 \right\}.$$

Решение дифференциального уравнения (5) определяется решением системы обыкновенных дифференциальных уравнений для характеристических кривых [8]:

$$\frac{dt}{1} = \frac{dw_1}{w_1 \kappa_1 - w_2 \kappa_2} = \frac{d\Phi(w_1, w_2, t)}{j\Phi(w_1, w_2, t) w_1 \kappa}, \quad (9)$$

где $\kappa = \mathbf{RAE}$, $\kappa_1 = \mathbf{RME}$, $\kappa_2 = \mathbf{RGE}$. Решение уравнение (8) с учетом начального условия будет иметь вид

$$\Phi(w_1, w_2, t) = \exp \left\{ j \frac{\kappa}{\kappa_1} w_1 + j \frac{\kappa \kappa_2}{\kappa_1} w_2 t \right\}. \quad (10)$$

Рассмотрим асимптотику второго порядка. Обозначим через $\mathbf{H}_2(u_1, u_2, t)$ векторную функцию, удовлетворяющую выражению

$$\mathbf{H}(u_1, u_2, t) = \mathbf{H}_2(u_1, u_2, t) \exp \left\{ j \frac{N\kappa}{\kappa_1} u_1 + j \frac{N\kappa \kappa_2}{\kappa_1} u_2 t \right\}. \quad (11)$$

Подставим это выражение в (3), получим уравнение относительно функции $\mathbf{H}_2(u_1, u_2, t)$:

$$\frac{\partial \mathbf{H}_2(u_1, u_2, t)}{\partial t} = j \frac{\partial \mathbf{H}_2(u_1, u_2, t)}{\partial u_1} \left[(1 - e^{-ju_1}) \mathbf{M} + (1 - e^{ju_2}) \mathbf{\Gamma} \right] + \mathbf{H}_2(u_1, u_2, t) \times \\ \times \left[N \mathbf{\Lambda} (e^{ju_1} - 1) + N \mathbf{Q} - j \frac{N \kappa \kappa_2}{\kappa_1} \mathbf{I} u_2 - \frac{N \kappa}{\kappa_1} \left[(1 - e^{-ju_1}) \mathbf{M} + (1 - e^{ju_2}) \mathbf{\Gamma} \right] \right], \quad (12)$$

где \mathbf{I} – диагональная единичная матрица.

В (12) выполним замены

$$\varepsilon^2 = \frac{1}{N}, \quad u_1 = \varepsilon w_1, \quad u_2 = \varepsilon w_2, \quad \mathbf{H}_2(u_1, u_2, t) = \mathbf{F}_2(w_1, w_2, t, \varepsilon).$$

Тогда уравнение (12) с учетом этих обозначений будет иметь вид

$$\varepsilon^2 \frac{\partial \mathbf{F}_2(w_1, w_2, t, \varepsilon)}{\partial t} = \mathbf{F}_2(w_1, w_2, t, \varepsilon) \times \\ \times \left[\mathbf{\Lambda} (e^{jw_1\varepsilon} - 1) + \mathbf{Q} - j \frac{\kappa \kappa_2}{\kappa_1} \mathbf{I} \varepsilon w_2 - \frac{\kappa}{\kappa_1} \left[(1 - e^{-jw_1\varepsilon}) \mathbf{M} + (1 - e^{jw_2\varepsilon}) \mathbf{\Gamma} \right] \right] + \\ + j \varepsilon \frac{\partial \mathbf{F}_2(w_1, w_2, t, \varepsilon)}{\partial w_1} \left[(1 - e^{-jw_1\varepsilon}) \mathbf{M} + (1 - e^{jw_2\varepsilon}) \mathbf{\Gamma} \right]. \quad (13)$$

Обозначим $\lim_{\varepsilon \rightarrow 0} \mathbf{F}_2(w_1, w_2, t, \varepsilon) = \mathbf{F}_2(w_1, w_2, t)$, и эту функцию мы будем называть асимптотическим решением уравнения (13). Далее в этом же уравнении выполним предельный переход $\varepsilon \rightarrow 0$, получим уравнение $\mathbf{F}_2(w_1, w_2, t) \mathbf{Q} = \mathbf{0}$, решение которого запишем в виде $\mathbf{F}_2(w_1, w_2, t) = \mathbf{R} \mathbf{\Phi}_2(w_1, w_2, t)$, где скалярная функция $\Phi_2(w_1, w_2, t)$ будет определена ниже. Решение уравнения (13) будем искать в виде разложения

$$\mathbf{F}_2(w_1, w_2, t, \varepsilon) = \mathbf{\Phi}_2(w_1, w_2, t) \left(\mathbf{R} + j w_1 \varepsilon \mathbf{f}_1 + j w_2 \varepsilon \mathbf{f}_2 + O(\varepsilon^2) \right), \quad (14)$$

подставив которое в (13) и с учетом $\mathbf{R} \mathbf{Q} = \mathbf{0}$, получим систему уравнений для определения векторов $\mathbf{f}_1, \mathbf{f}_2$:

$$\mathbf{f}_1 \mathbf{Q} = \frac{\kappa}{\kappa_1} \mathbf{R} \mathbf{M} - \mathbf{R} \mathbf{\Lambda}, \quad \mathbf{f}_2 \mathbf{Q} = \frac{\kappa}{\kappa_1} \mathbf{R} \kappa_2 - \frac{\kappa}{\kappa_1} \mathbf{R} \mathbf{\Gamma}. \quad (15)$$

Для нахождения функции $\Phi_2(w_1, w_2, t)$ просуммируем все уравнения системы (13), подставим разложение (14), получим уравнение

$$\begin{aligned} & \frac{\partial \Phi_2(w_1, w_2, t)}{\partial t} + \frac{\partial \Phi_2(w_1, w_2, t)}{\partial w_1} (w_1 \kappa_1 - w_2 \kappa_2) = \\ & = \Phi_2(w_1, w_2, t) \left(w_1^2 (\mathbf{f}_1 A_1 - \kappa) + w_2^2 \left(\mathbf{f}_2 A_2 - \frac{\kappa \kappa_2}{2 \kappa_1} \right) + w_1 w_2 (\mathbf{f}_2 A_1 + \mathbf{f}_1 A_2) \right), \end{aligned} \quad (16)$$

с начальным условием

$$\Phi_2(w_1, w_2, 0) = \exp \left\{ \frac{A_1 \mathbf{f}_1 - \kappa}{\kappa_1} \cdot \frac{w_1^2}{2} \right\},$$

где $A_1 = \left(\frac{\kappa}{\kappa_1} \mathbf{M} - \Lambda \right) \mathbf{E}$, $A_2 = \left(\frac{\kappa}{\kappa_1} \kappa_2 \mathbf{I} - \frac{\kappa}{\kappa_1} \Gamma \right) \mathbf{E}$, \mathbf{I} – единичная матрица.

Решение дифференциального уравнения (16) будем искать в виде

$$\Phi_2(w_1, w_2, t) = \exp \left\{ -\frac{1}{2} (K_{11} w_1^2 + 2K_{12}(t) w_1 w_2 + K_{22}(t) w_2^2) \right\}, \quad (17)$$

где $K_{11}, K_{22}(t)$ – дисперсии процессов $i(t), n(t)$ соответственно, $K_{12}(t)$ – ковариация $i(t), n(t)$. Подставим предполагаемое решение (17) в (16) и с учетом $K_{12}(0) = 0, K_{22}(0) = 0$ получим $K_{11}, K_{22}(t)$ и $K_{12}(t)$ в виде

$$K_{11} = \frac{\kappa - A_1 \mathbf{f}_1}{\kappa_1},$$

$$K_{12}(t) = \left(\frac{\kappa - A_1 \mathbf{f}_1}{\kappa_1^2} \kappa_2 - \frac{A_1 \mathbf{f}_2 + A_2 \mathbf{f}_1}{\kappa_1} \right) (1 - e^{-\kappa_1 t}),$$

$$\begin{aligned} K_{22}(t) = & 2t \left(\frac{\kappa - A_1 \mathbf{f}_1}{\kappa_1^2} \kappa_2^2 - \frac{A_1 \mathbf{f}_2 + A_2 \mathbf{f}_1}{\kappa_1} \kappa_2 - \left(A_2 \mathbf{f}_2 - \frac{\kappa \kappa_2}{2 \kappa_1} \right) \right) + \\ & + 2(1 - e^{-\kappa_1 t}) \left(\frac{\kappa - A_1 \mathbf{f}_1}{\kappa_1^3} \kappa_2^2 - \frac{A_1 \mathbf{f}_2 + A_2 \mathbf{f}_1}{\kappa_1^2} \kappa_2 \right). \end{aligned}$$

Теперь запишем выражение для функции $\mathbf{H}(u_1, u_2, t)$:

$$\mathbf{H}(u_1, u_2, t) = \mathbf{R} \exp \left\{ -\frac{1}{2} \left(K_{11} (Nu_1)^2 + 2K_{12}(t) N^2 u_1 u_2 + K_{22}(t) (Nu_2)^2 \right) + \right. \\ \left. + j \frac{N\kappa}{\kappa_1} u_1 + j \frac{N\kappa_2}{\kappa_1} u_2 t \right\}.$$

Таким образом, характеристическая функция $h(u_1, u_2, t) = \mathbf{H}(u_1, u_2, t) \mathbf{E}$ двумерного процесса числа застрахованных в компании рисков и числа требований на выплату страховых сумм в указанных выше предельных условиях имеет вид двумерной характеристической функции гауссовского распределения. Полученные результаты могут быть полезны для анализа деятельности страховых компаний.

ЛИТЕРАТУРА

1. Даммер Д.Д. Математическая модель страховой компании в виде системы массового обслуживания с неограниченным количеством приборов с учетом единовременных страховых выплат // Материалы XV Международной конференции имени А.Ф. Терпугова «Информационные технологии и математическое моделирование». – Томск: Изд-во Том. ун-та, 2016. – С. 18–23.
2. Даммер Д.Д. Исследование общей суммы единовременных страховых выплат в модели с ограниченным страховым полем // Материалы V молодежной научно-практической конференции «Математическое и компьютерное моделирование в экономике, страховании и управлении рисками». – Саратов, 2016. – С. 51–56.
3. Dammer D. Research of mathematical model of insurance company in the form of queueing system with unlimited number of devices considering implicit advertising // Information Technologies and Mathematical Modeling. – Springer, 2015. – V. 564. – P. 163–175.
4. Гафуров Ш.Р., Гугнин В.И., Аманов С.Н. Язык бизнеса. – Ташкент: Шарк, 1995. – 738 с.
5. Моисеев А.Н., Назаров А.А. Бесконечнолинейные системы и сети массового обслуживания. – Томск: Изд-во НТЛ, 2015. – 240 с.
6. Назаров А.А., Терпугов А.Ф. Теория массового обслуживания. – Томск: Изд-во НТЛ, 2005. – 228 с.
7. Назаров А.А., Моисеева С.П. Метод асимптотического анализа в теории массового обслуживания. – Томск: Изд-во НТЛ, 2006. – 112 с.
8. Эльсгольц Л.Э. Дифференциальные уравнения и вариационное исчисление. – Москва: Наука, 1969. – 424 с.

МНОГОПРОДУКТОВАЯ МОДЕЛЬ БЫСТРО ПОРТЯЩИХСЯ ЗАПАСОВ С ЗАВИСЯЩИМ ОТ ЦЕНЫ СПРОСОМ

А.В. Китаева, К.И. Лившиц, Е.С. Ульянова

Томский государственный университет, г. Томск, Россия

Одной из классических задач теории управления запасами является проблема управления запасами, имеющими ограниченный срок годности, которая интенсивно изучается в последнее время [1, 2]. Первоначально исследовалась задача реализации единственного товара, имеющего фиксированную цену, затем аналогичные задачи начали рассматриваться для многопродуктовых моделей [2]. Целью данной работы является определение оптимального размера многопродуктовой партии товара, имеющего ограниченный срок реализации, и оптимальных розничных цен, которые обеспечивают продавцу наибольшую среднюю прибыль.

1. Постановка задачи и асимптотическая плотность вероятностей спроса

Продавец, располагая средствами размера S , приобретает партию товара $q = [q_1, q_2 \dots q_m]^T$ по оптовым ценам $d = [d_1, d_2 \dots d_m]^T$ и перепродает её по розничным ценам $c = [c_1, c_2 \dots c_m]^T$. Считается, что время реализации ограничено. По истечению времени T товар не может быть реализован, а продавец несет дополнительные затраты $b = [b_1, b_2 \dots b_m]^T$, связанные с утилизацией непроданной части товара. Необходимо определить оптимальный размер партии товара q и розничные цены c , которые обеспечивают продавцу максимальную среднюю прибыль.

Предполагается, что поток покупателей – пуассоновский поток с известной интенсивностью $\lambda(t, c)$, где t – текущий момент времени. Интенсивность потока покупателей зависит, в частности, от розничных цен c и может меняться, например, при изменении розничных цен. Далее предполагается, что покупатели приобретают товар независимо друг от друга. Объем покупки – случайная величина $z = [z_1, z_2 \dots z_m]^T$ с вектором

средних значений $a = [a_1, a_2 \dots a_m]^T$, ковариационной матрицей $R = [R_{pq}]$ и плотностью распределения $p(z_1, z_2 \dots z_m)$.

Пусть T – длительность торговой сессии и n – число покупателей за время T . При сделанных предположениях число покупателей распределено по закону Пуассона

$$P(n) = \frac{(\Lambda(T))^n}{n!} \exp(-\Lambda(T)), \quad (1)$$

где $\Lambda(T, c) = \int_0^T \lambda(t, c) dt$. Пусть $z_j = [z_{1j}, z_{2j} \dots z_{mj}]^T$ – объем j -й покупки. Тогда если было совершено (или могло быть совершено) n покупок, то объем спроса x^n в течение торговой сессии

$$x^n = z_1 + z_2 + \dots + z_n. \quad (2)$$

Если $P(x/n)$ – плотность вероятностей величины x^n и $x = [x_1, x_2 \dots x_m]^T$ – общий объем спроса на товар в течение торговой сессии, то плотность распределения x

$$P(x) = \sum_{n=0}^{\infty} P(x/n) \frac{(\Lambda(T, c))^n}{n!} \exp(-\Lambda(T, c)). \quad (3)$$

Теорема 1. Если при $T \rightarrow \infty$ $\Lambda(T, c) \rightarrow \infty$, то случайные величины

$$y_i = \frac{x_i - \Lambda(T) a_i}{\sqrt{\Lambda(T)}}, \quad i = \overline{1, m},$$

имеют совместное нормальное распределение с нулевым средним и ковариационной матрицей R .

Пример. Пусть $m = 1$, то есть продается один вид товара. Объемы покупок z – случайные величины, имеющие гамма-распределение с параметрами α и $n-1$, где n – целое. Тогда из (3), учитывая свойства гамма-распределения, получим, что плотность распределения спроса x в течение торговой сессии имеет вид

$$P(x) = \delta(x) e^{-\Lambda(T, c)} + \sum_{k=1}^{\infty} \frac{x^{kn-1}}{\alpha^{kn} (kn-1)!} \cdot \frac{\Lambda(T, c)^k}{k!} e^{-\frac{x}{\alpha} - \Lambda(T, c)}. \quad (4)$$

В табл. 1 приведены значения расстояния Колмогорова

$$\Delta = \max_x \left| P(x) - N \left(\frac{x - \alpha n \Lambda(T, c)}{\sqrt{\alpha^2 n(n+1) \Lambda(T, c)}} \right) \right|$$

между истинным распределением $P(x)$ и его нормальной аппроксимацией при $n = 2$ и $\alpha = 0,5$.

Т а б л и ц а 1

**Расстояние Колмогорова между истинным распределением
и нормальной аппроксимацией**

$\Lambda(T)$	5	10	50	100	200	1000
$\Delta, n = 2$	0,031	0,014	$2,604 \cdot 10^{-3}$	$1,277 \cdot 10^{-3}$	$6,303 \cdot 10^{-4}$	0
$\Delta, n = 3$	$0,019 \cdot 10^{-3}$	$8,756 \cdot 10^{-3}$	$1,615 \cdot 10^{-3}$	$7,941 \cdot 10^{-4}$	0	0

Зная плотность распределения объема спроса, можно определить вероятность $P(q, T)$ того, что партия товара объема q будет продана за время T :

$$P(q, T) = \int_{q_1}^{\infty} \dots \int_{q_m}^{\infty} p(x_1, x_2 \dots x_m) dx_1 dx_2 \dots dx_m.$$

Или в асимптотическом случае:

$$P(q, T) = \frac{(\det R)^{\frac{1}{2}}}{(2\pi\Lambda(T, c))^{\frac{m}{2}}} \int_{q_1 - \Lambda(T, c)a_1}^{\infty} \dots \int_{q_m - \Lambda(T, c)a_m}^{\infty} \exp\left(-\frac{x^T W x}{2\Lambda(T, c)}\right) dx_1 dx_2 \dots dx_m, \quad (5)$$

где матрица $W = R^{-1}$.

В общем случае интеграл (5) через элементарные функции не выражается. Однако наибольший интерес представляет случай, когда объемы партий товара q_i растут с ростом T . При этом возможны два варианта.

В первом случае при $T \rightarrow \infty$ $S_i = (\Lambda(T, c)a_i - q_i) / \sqrt{\Lambda(T, c)} \rightarrow \infty$, т.е. размер партии товара, q_i , растет медленнее, чем $\Lambda(T)$. Тогда при $T \gg 1$ вероятность

$$P(q, T) \approx \frac{e^{-\frac{1}{2} \sum_{i,j=1}^m W_{ij} \frac{(q_i - \Lambda(T, c) a_i)(q_j - \Lambda(T, c) a_j)}{\Lambda(T, c)}}}{(2 \cdot \pi)^{\frac{m}{2}} \det R^{\frac{1}{2}} \prod_{i=1}^m \sum_{j=1}^m W_{ij} \frac{q_j - \Lambda(T, c) a_j}{\sqrt{\Lambda(T, c)}}}.$$

Во втором случае при $T \rightarrow \infty$ $S_i = \frac{(\Lambda(T, c) a_i - q_i)}{\sqrt{\Lambda(T, c)}} \rightarrow -\infty$, т.е. размер партии товара q_i растет быстрее, чем $\Lambda(T, c)$. Тогда можно показать, что при $T \gg 1$ $P(q, T) \ll 1$.

2. Плотность вероятностей длительности продаж при постоянной интенсивности потока покупателей

Будем теперь предполагать, что интенсивность потока покупателей $\lambda(t, c) = \lambda$ не зависит от времени. Пусть S – стоимость товара, который покупатель может приобрести за одну покупку. Обозначим через $t(S)$ – случайную длительность продажи партии товара стоимости S .

Введем случайную величину

$$z = \frac{t(s) - q^2}{\delta q},$$

где

$$q = \frac{S}{\lambda m_1}, \quad \delta = \sqrt{\frac{\sigma^2 + m_1^2}{\lambda m_1^2}},$$

$$m_1 = M\{S\} = \sum_{i=1}^m c_i a_i, \quad \sigma^2 = Var\{S\} = \sum_{i=1}^m R_{ij} c_i c_j.$$

Теорема 2. При $S \rightarrow \infty$ случайная величина z имеет асимптотически стандартное нормальное распределение.

Поэтому можно считать, что при $S \gg 1$ длительность продажи партии товара $t(S)$ имеет нормальное распределение с параметрами

$$T_1(S) = \frac{S}{\lambda m_1}, \quad D(S) = \frac{S(m_1^2 + \sigma^2)}{\lambda^2 m_1^3}.$$

Зная плотность вероятностей длительности продаж, можно определить стоимость партии товара, которая может быть реализована за заданное время T с заданной вероятностью α . Имеем

$$\alpha = P\{t(S) \leq T\} = \frac{1}{\sqrt{2\pi D}} \int_{-\infty}^T \exp\left(-\frac{(t-T_1)^2}{2D}\right) dt,$$

где T_1 и D определяются выше. Откуда

$$S = \frac{\lambda m_1}{4} \left[\sqrt{\frac{m_2}{\lambda m_1^2} \psi(\alpha)^2 + 4T} + \sqrt{\frac{m_2}{\lambda m_1^2} \psi(\alpha)} \right]^2,$$

где $\psi(\alpha)$ – функция, обратная интегралу вероятностей.

3. Определение средней прибыли

Обозначим через W среднюю прибыль продавца за время торговой сессии. Тогда

$$W = -S + \sum_{j=1}^m \left[c_j q_j \int_{q_j}^{\infty} p_j(x) dx + \int_0^{q_j} (c_j x - b_j(q_j - x)) p_j(x) dx \right], \quad (6)$$

где $p_j(x)$ – плотность вероятностей спроса на j -й товар и S – затраченный на приобретение товара капитал:

$$S = \sum_{j=1}^m d_j q_j.$$

При фиксированных ценах c_j и d_j и затратах за утилизацию b_j оптимальные объемы партий товара q_j , максимизирующие среднюю прибыль, определяются выражениями

$$q_j = \psi_j \left(1 - \frac{b_j + d_j}{b_j + c_j} \right), \quad (7)$$

где $\psi_j(z)$ – функция, обратная $\int_0^z p_j(x) dx$. Соответствующие (7) значение средней прибыли (6)

$$W = \sum_{i=1}^m (c_i + b_i) \int_0^{q_i} x p_i(x) dx. \quad (8)$$

Соотношения (7), (8) получены в предположении, что на закупку партии товара могли быть затрачены неограниченные средства S . Предположим теперь, что на закупку товара могут быть затрачены некоторые ограниченные средства S_0 . Тогда оптимальное значение q_j должны максимизировать W (6) при дополнительном ограничении

$$S = \sum_{j=1}^m d_j q_j = S_0. \quad (9)$$

Решение получившейся задачи на условный экстремум (6), (9) приводит к соотношениям

$$q_j = \Psi_j \left(1 - \frac{b_j + \mu d_j}{b_j + c_j} \right), \quad (10)$$

где μ – неопределенный множитель Лагранжа, определяемый из условия

$$f(\mu) = \sum_{j=1}^m d_j \Psi_j \left(1 - \frac{b_j + \mu d_j}{b_j + c_j} \right) = S_0. \quad (11)$$

Как было показано ранее, при $\Lambda(T) \gg 1$ плотность вероятностей спроса $p_j(x)$ может быть аппроксимирована нормальным распределением с параметрами $\Lambda(T)a_i$ и $\Lambda(T)R_{ii}$. При этом соотношение (7) принимает вид

$$q_j = \Lambda(T)a_j + \sqrt{\Lambda(T)R_{jj}} \Psi \left(1 - \frac{b_j + d_j}{b_j + c_j} \right), \quad (12)$$

а соотношение (10) переходит в соотношение

$$q_j = \lambda(T)a_j + \sqrt{\lambda(T)R_{jj}} \Psi \left(1 - \frac{b_j + \mu d_j}{b_j + c_j} \right), \quad (13)$$

где $\Psi(x)$ – функция, обратная интегралу вероятностей.

Относительная ошибка аппроксимации $\Delta = |q - q_{as}|/q$ при $m = 1$ для случая гамма-распределения объема покупок с параметрами $\alpha = 0,5$, $n - 1 = 1$ в зависимости от величины $\Lambda(T, c)$, где q вычисляется по формуле (7), а q_{as} по формуле (12) при $b = 0,1$, $d = 2$, $c = 4$, приведена в табл. 2.

Т а б л и ц а 2

Относительная погрешность аппроксимации

$\Lambda(T, c)$	5	10	20	50	100	200
Δ	0,074	0,035	0,017	$6,7 \cdot 10^{-3}$	$3,36 \cdot 10^{-3}$	$1,67 \cdot 10^{-3}$

Как следует из приведенных данных, ошибка аппроксимации практически линейно уменьшается с ростом $\Lambda(T, c)$.

4. Определение оптимальной розничной цены

Будем в дальнейшем считать, что $T \gg 1$ и оптимальные объемы партий товара определяются соотношениями (12). Предположим так же, что по всем товарам устанавливается одинаковая торговая наценка, то есть $c_i = \gamma d_i$, где $\gamma \geq 1$. Далее будем предполагать, что $\Lambda(T, c) = \lambda_0 T f(\gamma)$, где $f(\gamma)$ – монотонно убывающая функция: $f(0) = 1$ и функция $\gamma f(\gamma) \rightarrow 0$ при $\gamma \rightarrow \infty$, так как по бесконечно большой цене никто покупать не будет. Условие $\Lambda(T, c) \gg 1$ переходит теперь в условие $\lambda_0 T \gg 1$, а задача определения оптимальных розничных цен сводится к задаче определения оптимальной наценки γ . Средняя прибыль (8) переписывается в виде

$$W(\gamma) = \lambda_0 T \left(\sum_{i=1}^m a_i d_i f(\gamma)(\gamma - 1) - \frac{\sqrt{f(\gamma)}}{\sqrt{2\pi\lambda_0 T}} \sum_{i=1}^m \sqrt{R_{ii}} (\gamma d_i + b_i) e^{-\frac{1}{2} \psi \left(1 - \frac{b_i + d_i}{b_i + \gamma d_i} \right)^2} \right). \quad (14)$$

Так как второе слагаемое в (14) имеет порядок $\frac{1}{\sqrt{\lambda_0 T}}$, то при $\lambda_0 T \gg 1$ решение задачи можно искать в виде

$$\gamma = \gamma_0 + \gamma_1 \frac{1}{\sqrt{\lambda_0 T}},$$

где γ_0 определяется из условия $f(\gamma)(\gamma-1) = \max$, а γ_1 – при найденном γ_0 из условия $\frac{dW(\gamma_1)}{d\gamma_1} = 0$.

Заключение

Рассмотрена задача розничной продажи ассортимента продукции, имеющей ограниченный срок годности. Найдены асимптотические распределения спроса и длительности продаж при большой интенсивности потока покупателей. Получены соотношения, позволяющие определить оптимальный размер партии товара, предназначенного для продажи, и определить оптимальную торговую наценку.

ЛИТЕРАТУРА

1. *Petruzzi N.C., Dada M.* Pricing and the newsvendor problem: A review with extensions // *Operations Research*. – 1999. – V. 47. – No. 2. – P. 183–194.
2. *Chen X., Simchi-Levi D.* Joint Pricing and Inventory Management // *The Oxford Handbook of Pricing Management* / ed. by Ö. Özer and R. Phillips. – London: Oxford University Press, 2012.
3. *Turken N., Tan Y., Vakharia A.J., Wang L., et al.* The Multi-Product Newsvendor Problem: Review, Extension, and Directions for Future Research // *Handbook of Newsvendor Problems: Models, Extensions and Applications*. – 2012. – V. 176. – P. 3–39.

НЕНАДЕЖНАЯ СИСТЕМА МАССОВОГО ОБСЛУЖИВАНИЯ С МАРКОВСКИМ ВХОДЯЩИМ ПОТОКОМ И ПОВТОРНЫМИ ВЫЗОВАМИ*

В.И. Клименок

*Белорусский государственный университет,
г. Минск, Республика Беларусь*

Характерной особенностью некоторых телекоммуникационных сетей и систем является то, что поступившие в полностью занятую систему заявки не буферизируются и не покидают систему навсегда, а повторяют попытки занять обслуживающий прибор через случайные интервалы времени. Адекватными математическими моделями таких систем являются системы массового обслуживания с повторными вызовами. Исследование таких систем является интересным с практической и теоретической точек зрения и в ряде случаев может быть выполнено с привлечением матрично-аналитических методов. С положением дел в этой области можно ознакомиться, например, в [1, 2]. На сегодняшний день достаточно хорошо изучены системы с повторными вызовами со стационарными пуассоновскими входными потоками. В то же время известно, что потоки в современных телекоммуникационных сетях имеют коррелированный взрывной характер, что требует для их описания более сложных моделей. Одной из наиболее известных моделей коррелированного входного потока является групповой марковский поток (*Batch Markovian Arrival Process, BMAP*). В данной работе рассматривается система массового обслуживания *BMAP/M/N* с повторными вызовами. Рассматриваемая система является ненадёжной. После поломки прибора сразу же начинается его ремонт. Заявка, обслуживание которой было прервано поломкой, с некоторой вероятностью помещается на орбиту и с дополнительной вероятностью теряется.

Для рассматриваемой системы получено достаточное условие существования стационарного режима, приведён алгоритм вычисления стационарных вероятностей и выведены формулы для основных характеристик производительности.

* Исследование выполнено при финансовой поддержке Белорусского республиканского фонда фундаментальных исследований в рамках научного проекта № Ф16Р-017.

Математическая модель

Рассматривается N -линейная система массового обслуживания с $ВМАР$ -потоком. $ВМАР$ задаётся управляющим процессом $v_t, t \geq 0$, который является неприводимой цепью Маркова с непрерывным временем и конечным пространством состояний $\{0, \dots, W\}$ и матричной про-

изводящей функцией $D(z) = \sum_{k=0}^{\infty} D_k z^k, |z| \leq 1$. Группы первичных зая-

вок поступают в систему только в моменты скачков цепи Маркова $v_t, t \geq 0$. Матрицы $D_k, k \geq 1$ и недиагональные элементы матрицы D_0

задают интенсивности переходов процесса $v_t, t \geq 0$, сопровождающихся поступлением группы из k заявок. Матрица $D(1)$ является инфините-

зимальным генератором процесса $v_t, t \geq 0$. Интенсивность λ поступления

заявок в $ВМАР$ вычисляется по формуле $\lambda = \theta D'(1) \mathbf{e}$, где вектор θ

является единственным решением системы $\theta D(1) = \mathbf{0}, \theta \mathbf{e} = 1$. Интен-

сивность λ_b поступления групп заявок определяется как $\lambda_b = \theta(-D_0) \mathbf{e}$.

Здесь и далее \mathbf{e} – вектор-столбец, состоящий из единиц, размерность которого определяется из контекста. Коэффициент вариации интервалов между поступлениями групп заявок задаётся соотношением

$c_{\text{var}}^2 = 2\lambda_b \theta(-D_0)^{-1} \mathbf{e} - 1$, коэффициент корреляции интервалов между

последовательными поступлениями групп заявок вычисляется как

$c_{\text{corr}} = (\lambda_b \theta(-D_0)^{-1} (D(1) - D_0)(-D_0)^{-1} \mathbf{e} - 1) / c_{\text{var}}^2$. Подробную информацию о

$ВМАР$, его свойствах и частных случаях можно найти, например, в [3].

Если в момент поступления группы первичных заявок в системе имеется достаточное количество свободных приборов, то каждая из заявок группы занимает один обслуживающий прибор и начинает обслуживаться. Время обслуживания любой заявки имеет экспоненциальное распределение с параметром μ . Если свободных приборов недостаточно, то остаток группы помещается на орбиту. Если свободные приборы отсутствуют, то на орбиту помещается вся группа заявок. Заявки с орбиты делают попытки занять обслуживающий прибор через случайные интервалы времени. Если на орбите в некоторый момент времени t на-

ходится i заявок, то вероятность того, что в течение интервала времени

$(t, t + \Delta t)$ произойдёт повторная попытка занять обслуживающий прибор, равна $\alpha_i \Delta t + o(\Delta t)$, $i \geq 1$. Объём орбиты предполагается неограниченным. Конкретный вид зависимости α_i от i не фиксируется, требуется лишь выполнение условия $\lim_{i \rightarrow \infty} \alpha_i = \infty$. Заметим, что этому условию удовлетворяют, в частности, классическая стратегия повторов ($\alpha_i = i\alpha$) и линейная стратегия ($\alpha_i = i\alpha + \gamma$).

Помимо входного потока в систему поступает стационарный пуассоновский поток поломок с интенсивностью h . Событие из этого потока с равной вероятностью вызывает поломку любого занятого прибора (заметим, что для данной системы это эквивалентно тому, что событие из потока поломок вызывает поломку произвольно выбранного занятого прибора). Поломка, в момент поступления которой ни один прибор не занят обслуживанием, игнорируется. Немедленно после поломки прибора начинается его ремонт, в ходе которого прибор не может обслуживать заявки. Продолжительность ремонта имеет экспоненциальное распределение с интенсивностью k и не зависит от процессов обслуживания и ремонтов на других приборах. Заявка, обслуживание которой было прервано поломкой, с вероятностью p , $0 \leq p \leq 1$, помещается на орбиту и с дополнительной вероятностью покидает систему.

Процесс изменения состояний системы

Пусть i_t – число заявок на орбите, $i_t \geq 0$, n_t – суммарное число занятых приборов и приборов, находящихся на ремонте, $n_t = \overline{0, N}$, r_t – число занятых приборов, $r_t = \overline{0, n_t}$, v_t – состояние управляющего процесса $ВМАР$, $v_t = \overline{0, W}$, в момент времени t , $t \geq 0$.

Процесс изменения состояний системы описывается многомерной неприводимой регулярной цепью Маркова с непрерывным временем

$$\xi_t = \{i_t, n_t, r_t, v_t\}, t \geq 0.$$

Лемма 1. Инфинитезимальный генератор Q цепи Маркова ξ_t , $t \geq 0$, имеет блочную структуру $Q = (Q_{ij})_{i,j \geq 0}$, где ненулевые блоки Q_{ij} порядка $\overline{W}(N+1)(N+2)/2$ задаются следующим образом:

$$Q_{i,i-1} = \alpha_i \begin{pmatrix} O_{\overline{W}} & B_0 & O & \cdots & O \\ O & O & B_1 & \cdots & O \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ O & O & O & \cdots & B_{N-1} \\ O & O & O & \cdots & O \end{pmatrix},$$

$$Q_{i,i+k} = \begin{pmatrix} O & \cdots & O & D_{0,N+k} \\ O & \cdots & O & D_{1,N+k} \\ \vdots & \ddots & \vdots & \vdots \\ O & \cdots & O & D_{N,N+k} \end{pmatrix} + \delta_{1,k} (1-p)H,$$

$$Q_{i,i} = \begin{pmatrix} C_0 & D_{0,1} & D_{0,2} & \cdots & D_{0,N-1} & D_{0,N} \\ A_1 & C_1 & D_{1,2} & \cdots & D_{1,N-1} & D_{1,N} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ O & O & O & \cdots & C_{N-1} & D_{N-1,N} \\ O & O & O & \cdots & A_N & C_N \end{pmatrix} + pH - \alpha_i \overline{I},$$

где

$$A_n = A_n \otimes I_{\overline{W}}, \quad A_n = \begin{pmatrix} n\kappa & 0 & \cdots & 0 \\ \mu & (n-1)\kappa & \cdots & 0 \\ 0 & 2\mu & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & n\mu \end{pmatrix}, \quad n = \overline{1, N},$$

$$B_n = B_n \otimes I_{\overline{W}}, \quad B_n = (\mathbf{0}_{n+1} \quad I_{n+1}), \quad n = \overline{0, N-1},$$

$$C_n = C_n \oplus D_0, \quad C_n = -diag \{ r\mu + (n-r)\kappa + (1-\delta_{r,0})h, r = \overline{0, n} \}, \quad n = \overline{0, N},$$

$$D_{n,n+l} = (O_{(n+1)\overline{W} \times \min\{N-n, l\}\overline{W}} \quad I_{n+1} \otimes D_l), \quad n = \overline{0, N}, l \geq 0,$$

$$H = \begin{pmatrix} O_{\overline{W}} & O & \cdots & O \\ O & H_1 & \cdots & O \\ \vdots & \vdots & \ddots & \vdots \\ O & O & \cdots & H_N \end{pmatrix}, \quad H_n = H_n \otimes I_{\overline{W}}, \quad H_n = \begin{pmatrix} 0 & 0 & \cdots & 0 & 0 \\ h & 0 & \cdots & 0 & 0 \\ 0 & h & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & h & 0 \end{pmatrix}, \quad n = \overline{1, N},$$

Здесь матрицы H_n имеют порядок $n+1$; $\overline{W} = W + 1$; \overline{I} – диагональная матрица, последние $(N+1)\overline{W}$ диагональных элементов которой нули, а остальные – единицы; \otimes – символ Кронекерова умножения матриц; $\delta_{n,r}$ – символ Кронекера.

Следствие 1. Цепь Маркова ξ_t , $t \geq 0$ принадлежит классу асимптотически квазитеплицевых цепей Маркова с непрерывным временем (АКТЦМ), см. [4].

Доказательство. В соответствии с данным в [4] определением, цепь Маркова ξ_t , $t \geq 0$, с инфинитезимальным генератором $Q = (Q_{ij})_{i,j \geq 0}$ принадлежит классу АКТЦМ с непрерывным временем, если существуют пределы $Y_k = \lim_{i \rightarrow \infty} R_i^{-1} Q_{i,i+k-1}$, $k = 0, 2, 3, \dots$; $Y_1 = \lim_{i \rightarrow \infty} R_i^{-1} Q_{i,i} + I$, причём

матрица $\sum_{k=0}^{\infty} Y_k$ является стохастической. Здесь R_i – диагональные матрицы, ненулевыми элементами которых являются модули диагональных элементов матриц $Q_{i,i}$, $i \geq 0$.

Легко видеть, что диагональные элементы матрицы R_i , соответствующие первым N блочным строкам матрицы $Q_{i,i}$, содержат элемент α_i , остальная же часть диагонали от i не зависит. С учётом этого пределы, задающие матрицы Y_k , $k \geq 0$, вычисляются следующим образом:

$$Y_0 = \begin{pmatrix} O & B_0 & O & \cdots & O \\ O & O & B_1 & \cdots & O \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ O & O & O & \cdots & B_{N-1} \\ O & O & O & \cdots & O \end{pmatrix}, \quad Y_1 = \begin{pmatrix} O & \cdots & O & & O \\ \vdots & \ddots & \vdots & & \vdots \\ O & \cdots & O & & O \\ O & \cdots & R^{-1}A_N & R^{-1}(C_N + pH_N) + I \end{pmatrix},$$

$$Y_k = \begin{pmatrix} O & \cdots & O & & O \\ \vdots & \ddots & \vdots & & \vdots \\ O & \cdots & O & & O \\ O & \cdots & O & R^{-1}(D_{N,N+k-1} + \delta_{1,k-1}(1-p)H_N) \end{pmatrix}, \quad k > 1,$$

где R – диагональная матрица, диагональными элементами которой являются модули диагональных элементов матрицы C_N .

Легко убедиться, что сумма этих матриц является стохастической матрицей. Таким образом, процесс $\xi_t, t \geq 0$, принадлежит классу АКТЦМ с непрерывным временем, что и требовалось доказать.

Важную роль в исследовании АКТЦМ выполняет производящая функция $Y(z)$ матриц $Y_k, k \geq 0$. В нашем случае эта функция имеет следующий вид:

$$Y(z) = \begin{pmatrix} O & B_0 & \cdots & O & O \\ O & O & \cdots & O & O \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ O & O & \cdots & O & B_{N-1} \\ O & O & \cdots & R^{-1}A_N z & R^{-1}F(z) + zI \end{pmatrix},$$

где $F(z) = z(C_N - D_{N,N} + pH_N + D_N(z) + (1-p)H_N z)$.

Стационарное распределение

Теорема 1. Для существования стационарного распределения цепи Маркова $\xi_t, t \geq 0$, достаточно, чтобы выполнялось следующее неравенство:

$$\lambda + (1-p)(1-\delta_0)h < \sum_{r=0}^N \delta_r [(N-r)\kappa + r\mu], \quad (1)$$

где
$$\delta_r = \frac{C_N^r r! (\kappa/h)^r}{\sum_{l=0}^N C_N^l l! (\kappa/h)^l}, \quad r = \overline{0, N}.$$

Далее предполагается, что неравенство (1) выполнено.

Введём стационарные вероятности состояний системы $p(i, n, r, v)$, $i \geq 0, n = \overline{0, N}, r = \overline{0, n}, v = \overline{0, W}$, и векторы-строки $\mathbf{p}_i, i \geq 0$, упорядоченных в лексикографическом порядке стационарных вероятностей, соответствующих значению i счётной компоненты цепи Маркова $\xi_t, t \geq 0$.

Обозначим $\mathbf{p} = (\mathbf{p}_0, \mathbf{p}_1, \dots)$ и $\mathbf{P}(z) = \sum_{i=1}^{\infty} \mathbf{p}_i z^i, |z| \leq 1$.

Для вычисления векторов $\mathbf{p}_i, i \geq 0$, используется приведённый в [3] алгоритм, разработанный для вычисления стационарного распределения асимптотически квазитеплицевых цепей Маркова.

Характеристики производительности

Вычислив стационарные вероятности состояний системы, можно найти ряд важных стационарных характеристик производительности системы. Ниже приводятся наиболее важные из них:

- Среднее число заявок на орбите $L_{\text{orb}} = \sum_{i=1}^{\infty} i \mathbf{p}_i \mathbf{e}$.
- Среднее число занятых приборов $N_{\text{busy}} = \mathbf{P}(1) \text{diag} \{J_n, n = \overline{0, N}\} \mathbf{e}$,

где

$$J_n = \text{diag} \{r I_{\overline{W}}, r = \overline{0, n}\}, n = \overline{0, N}.$$

- Среднее число заявок в системе $L = L_{\text{orb}} + N_{\text{busy}}$.
- Среднее число приборов, находящихся на ремонте

$$N_{\text{repair}} = \mathbf{P}(1) \text{diag} \{n I - J_n, n = \overline{0, N}\} \mathbf{e}.$$

- Среднее число доступных приборов

$$N_{\text{idle}} = \mathbf{P}(1) \text{diag} \{(N - n) I_{(n+1)\overline{W}}, n = \overline{0, N}\} \mathbf{e}.$$

- Совместная вероятность того, что в системе r занятых приборов $n - r$ приборов, находящихся на ремонте, и i заявок на орбите

$$p_i(n, r) = \mathbf{p}_i I^{(n, r)} \mathbf{e}, r = \overline{0, n}, n = \overline{0, N}, i \geq 0,$$

где $I^{(n, r)} = \begin{pmatrix} O_{a\overline{W} \times \overline{W}} \\ I_{\overline{W}} \\ O_{d\overline{W} \times \overline{W}} \end{pmatrix}, a = n(n+1)/2 + r, d = N(N+1)/2 - a - 1.$

Здесь матрица $I^{(n, r)}$ выделяет часть вектора \mathbf{p}_i , соответствующую числу r занятых приборов и числу $n - r$, находящихся на ремонте.

- Совместная вероятность того, что в системе j свободных приборов и i заявок на орбите

$$p_i^{(\text{idle})}(j) = \mathbf{p}_i \sum_{r=0}^{N-j} I^{(N-j, r)} \mathbf{e}, j = \overline{0, N}, i \geq 0.$$

– Вероятность того, что в момент поступления произвольной заявки в системе r занятых приборов $n - r$ приборов, находящихся на ремонте, и i заявок на орбите

$$p_i^{(a)}(n, r) = \lambda^{-1} \mathbf{p}_i I^{(n, r)} \left(\sum_{k=1}^{\infty} k D_k \right) \mathbf{e}, r = \overline{0, n}, n = \overline{0, N}, i \geq 0.$$

– Вероятность того, что произвольная поступившая заявка немедленно начнёт обслуживаться

$$P_{\text{imm}} = \lambda^{-1} \mathbf{P}(1) \sum_{j=1}^N \sum_{r=0}^{N-j} I^{(N-j, r)} \left(\sum_{k=0}^j (k-j) D_k \right) \mathbf{e}.$$

– Вероятность того, что произвольная заявка будет потеряна

$$P_{\text{loss}} = \lambda^{-1} \mathbf{P}(1) \mathbf{e} p h.$$

– Среднее число прерываний обслуживания на каждую поступившую в систему заявку

$$N_{\text{break}} = p^{-1} P_{\text{loss}}.$$

Заключение

В данной работе исследована многолинейная ненадёжная система обслуживания с групповым марковским входным потоком и повторными вызовами. Найдено в явном виде достаточное условие существования стационарного распределения, приведён алгоритм вычисления стационарных вероятностей состояний системы. Получены формулы для ряда важных стационарных характеристик производительности системы.

ЛИТЕРАТУРА

1. Gomez-Corral A. Bibliographical guide to the analysis of retrial queues through matrix analytic techniques // Annals of Operations Research. – 2006. – V. 141. – P. 163–191.
2. Artalejo J.R., Gomez-Corral A. Retrial Queueing Systems: A Computational Approach. – Springer, 2008.
3. Lucantoni D. New results on the single server queue with a batch Markovian arrival process // Commun. Statist.-Stochastic Models. – 1991. – V. 7. – P. 1–46.
4. Klimenok V.I., Dudin A.N. Multi-dimensional asymptotically quasi-toeplitz Markov chains and their application in queueing theory // Queueing System. – 2006. – V. 54. – No. 4. – P. 245–259.

ИССЛЕДОВАНИЕ В НЕСТАЦИОНАРНОМ РЕЖИМЕ СЕТИ С ПОЛОЖИТЕЛЬНЫМИ И ОТРИЦАТЕЛЬНЫМИ ЗАЯВКАМИ РАЗЛИЧНЫХ ТИПОВ

Д.Я. Копать, М.А. Маталыцкий

*Гродненский государственный университет им. Я. Купалы,
г. Гродно, Республика Беларусь*

Описание сети. Постановка задачи

Рассмотрим открытую G-сеть массового обслуживания [1] с n однолинейными системами массового обслуживания (СМО), в которую поступают положительные и отрицательные заявки r типов. В i -ю СМО из внешней среды поступает простейший поток обычных заявок (положительных) с интенсивностью λ^+ и дополнительный поток отрицательных заявок, который также является простейшим с интенсивностью λ^- , $i = \overline{1, n}$. Все поступающие потоки независимы. Каждая положительная заявка входного потока независимо от других заявок направляется в i -ю СМО как заявка типа c с вероятностью p_{0ic}^+ , $\sum_{i=1}^n \sum_{c=1}^r p_{0ic}^+ = 1$. Длительности обслуживания положительных заявок в i -й СМО c -го типа распределены по экспоненциальному закону с параметром μ_{ic} .

В сети циркулируют не только положительные заявки, но и отрицательные [2]. Каждая отрицательная заявка входного потока независимо от других отрицательных заявок направляется в i -ю СМО как отрицательная заявка типа c с вероятностью p_{0ic}^- , $\sum_{i=1}^n \sum_{c=1}^r p_{0ic}^- = 1$ и уничтожает одну положительную заявку типа c . После окончания обслуживания положительной заявки типа c в i -й СМО она направляется в j -ю СМО с вероятностью p_{icjs}^+ опять как положительная заявка типа s , а с вероятностью p_{icjs}^- — как отрицательная заявка типа s и с вероятностью

$$p_{ic0} = 1 - \sum_{j=1}^n \sum_{s=1}^r (p_{icjs}^+ + p_{icjs}^-) \text{ уходит из сети, } i, j = \overline{1, n}.$$

Под состоянием сети в момент времени t будем понимать вектор $(\vec{k}, t) = (k_{11}, k_{12}, \dots, k_{1r}, k_{21}, k_{22}, \dots, k_{2r}, \dots, k_{n1}, k_{n2}, \dots, k_{nr}, t)$, где k_{ic} , l_{ic} – соответственно количество положительных и отрицательных заявок типа c в i -й СМО, который образует однородную случайную цепь Маркова с непрерывным временем и счётным числом состояний. Требуется найти вероятности состояний сети в переходном режиме, зависящие от времени.

Следует отметить, что стационарные вероятности состояний такой сети в форме произведения найдены в работе [3]. В [4] была рассмотрена G-сеть с разнотипными положительными и отрицательными заявками и различными дисциплинами их обслуживания в системах; отрицательные заявки одних типов могут уничтожать положительные заявки других типов. Установлен также мультипликативный вид стационарных вероятностей состояний. Методики нахождения нестационарных вероятностей состояний G-сетей с однотипными положительными и отрицательными заявками, основанные на применении метода последовательных приближений и многомерных производящих функций, описаны в монографии [5].

В качестве примера практического использования сетей МО со многими типами положительных и отрицательных заявок можно привести упрощенную (примитивную) модель воздействия некоторых файловых вирусов (например, overwiting) на ресурсы пользовательских компьютеров или ресурсы серверов. (на персональный компьютер (ПК) пользователя). Под сетью МО в этом случае можно понимать совокупность таких узлов (ПК пользователей (рабочее место пользователя), серверы и т.п.), которые получают доступ, к примеру, к файловому серверу. В качестве внешней среды будем понимать устройства, которые соединяются с компьютером, а также сеть Интернет, либо запросы пользователей на доступ к определенным ресурсам и обработке необходимых файлов на сервере. Условно разобьём файловые вирусы на типы по воздействию на некоторые типы файлов. В качестве положительных заявок первого типа могут выступать управленческие файлы семейства ОС Windows с расширением .exe и .bat, а второго типа, к примеру, файлы офисных приложений с расширением .doc \ docx, .xls \ xlsx, в качестве файлов третьего типа – графические файлы формата .JPEG, .GIF и т.п., в качестве файлов четвёртого типа – справки с расширением .HLP.

При попадании файлового вируса он записывает свой код вместо кода заражаемого файла, уничтожая его содержимое. Естественно, что при этом файл перестает работать и не восстанавливается (другими словами, он безвозвратно уничтожает файлы своего типа). К вирусам, уничтожающим положительные заявки первого типа, относятся, к примеру, Vienna.648, SVC.3103, Win32.Neshta, к вирусам, уничтожающим заявки второго типа, – Macro.Word97. Ethan, Word.Macro.DMV, X97. Divi [6], к вирусам уничтожающим положительные заявки третьего типа относятся, к примеру, W32/Perrun, ACAD.Star и т.п., к вирусам, уничтожающим файлы четвертого типа, – WinHLP.Demo, WinHLP.Lucky. Под переходом положительной заявки из одной системы в другую можно понимать перемещение обрабатываемого файла между узлами некоторой информационной сети (например, корпоративной сети предприятия) либо от определённого пользователя ПК в файловый сервер и наоборот.

Под перемещением положительной заявки (после завершения ее обслуживания в одной из СМО) в другую систему в качестве отрицательной заявки можно понимать описанную выше маршрутизацию, но уже как заражённого файла.

Система разностно-дифференциальных уравнений Колмогорова для вероятностей состояний сети

Пусть I_{ic} – вектор размерности $n \times r$, состоящий из нулей, за исключением компоненты с номером $r(i-1)+c$, которая равна единице, I_{00} – $n \times r$ -вектор, состоящий из нулей, а $P(\vec{k}, t)$ – вероятность состояния $\vec{k}(t)$, $u(x) = \begin{cases} 1, x > 0, \\ 0, x \leq 0 \end{cases}$ – функция Хевисайда.

Возможны следующие переходы нашего случайного марковского процесса в состояние $(\vec{k}, t + \Delta t)$ за время Δt :

1) из состояния $(\vec{k} - I_{js}, t)$, в этом случае в j -ю СМО за время Δt поступит положительная заявка типа s с вероятностью

$$\lambda^+ p_{0,js}^+ u(k_{js}) \Delta t + o(\Delta t), \quad j = \overline{1, n}, \quad s = \overline{1, r};$$

2) из состояния $(\bar{k} + I_{ic}, t)$, в этом случае в i -ю СМО за время Δt поступит отрицательная заявка типа c или после завершения обслуживания положительная заявка типа c уходит из сети или переходит в j -ю СМО как заявка типа s , но не обнаруживает там положительных заявок данного типа с вероятностью

$$(\lambda^- p_{0ic}^- + \mu_{ic} p_{ic0} + \mu_{ic} p_{icjs}^- (1 - u(k_{ic}))) \Delta t + o(\Delta t), i = \overline{1, n}, c = \overline{1, r};$$

3) из состояния $(\bar{k} + I_{ic} - I_{js}, t)$, в этом случае из i -й СМО за время Δt положительная заявка типа c после обслуживания переходит в j -ю СМО как положительная заявка типа s с вероятностью

$$\mu_{ic} p_{icjs}^+ u(k_{js}) \Delta t + o(\Delta t), i, j = \overline{1, n}, s, c = \overline{1, r};$$

4) из состояния $(\bar{k} + I_{ic} + I_{js}, t)$, в этом случае из i -й СМО за время Δt положительная заявка типа c после обслуживания переходит в j -ю СМО как отрицательная заявка типа s , уничтожив положительную заявку данного типа с вероятностью $\mu_{ic} p_{icjs}^- \Delta t + o(\Delta t), i, j = \overline{1, n}, s, c = \overline{1, r};$

5) из состояния (\bar{k}, t) , в этом случае за промежуток времени Δt состояние сети не изменилось с вероятностью

$$\left(1 - \left[\lambda^+ + \lambda^- + \sum_{i=1}^n \sum_{c=1}^r \mu_{ic} u(k_{ic}) \right] \right) \Delta t + o(\Delta t), i = \overline{1, n}, c = \overline{1, r};$$

6) из остальных с вероятностью $o(\Delta t)$.

Используя формулу полной вероятности, получаем, что нестационарные вероятности состояний рассматриваемой сети удовлетворяют следующей системе разностно-дифференциальных уравнений (РДУ):

$$\begin{aligned} \frac{dP(\bar{k}, t)}{dt} = & - \left[\lambda^+ + \lambda^- + \sum_{i=1}^n \sum_{c=1}^r \mu_{ic} u(k_{ic}) \right] P(\bar{k}, t) + \sum_{j=1}^n \sum_{s=1}^r \lambda^+ p_{0js}^+ u(k_{js}) P(\bar{k} - I_{js}, t) + \\ & + \sum_{i=1}^n \sum_{c=1}^r (\lambda^- p_{0ic}^- + \mu_{ic} p_{ic0} + \mu_{ic} p_{icjs}^- (1 - u(k_{ic}))) P(\bar{k} + I_{ic}, t) + \\ & + \sum_{i,j=1}^n \sum_{s,c=1}^r \mu_{ic} p_{icjs}^- P(\bar{k} + I_{ic} + I_{js}, t) + \sum_{i,j=1}^n \sum_{c,s=1}^r (\mu_{ic} p_{icjs}^+ u(k_{js})) P(\bar{k} + I_{ic} - I_{js}, t). \quad (1) \end{aligned}$$

Нахождение вероятностей состояний G-сети методом последовательных приближений

Систему РДУ (1) можно представить в виде

$$\begin{aligned} \frac{dP(\vec{k}, t)}{dt} = & -\Lambda(\vec{k})P(\vec{k}, t) + \sum_{i,j=1}^n \sum_{c,s=1}^r \Phi_{icjs}^{+-}(\vec{k})P(\vec{k} + I_{ic} - I_{js}, t) + \\ & + \sum_{i,j=1}^n \sum_{c,s=1}^r \Phi_{icjs}^{++}(\vec{k})P(\vec{k} + I_{ic} + I_{js}, t), \end{aligned} \quad (2)$$

где

$$\begin{aligned} \Lambda(\vec{k}) = & \lambda^+ + \lambda^- + \sum_{i=1}^n \sum_{c=1}^r \mu_{ic} u(k_{ic}), \quad \Phi_{icjs}^{++}(\vec{k}) = \mu_{ic} p_{icjs}^-, \\ \Phi_{icjs}^{+-}(\vec{k}) = & \delta_{0j} \delta_{s0} (\lambda^- p_{0ic}^- + \mu_{ic} p_{ic0}) + (\mu_{ic} p_{icjs}^- (1 - u(k_{ic}))) + \delta_{0i} \delta_{c1} \lambda^+ p_{0js}^+ u(k_{js}) + \\ & + \delta_{0i} \delta_{c0} \lambda^+ p_{0js}^+ u(k_{js}) + \mu_{ic} p_{icjs}(k_{js}), \quad \delta_{ij} = \begin{cases} 1, i = j, \\ 0, i \neq j. \end{cases} \end{aligned}$$

Из (2) следует, что

$$\begin{aligned} P(\vec{k}, t) = & e^{-\Lambda(\vec{k})t} \left(P(\vec{k}, 0) + \int_0^t e^{\Lambda(\vec{k})x} \left(\sum_{i,j=1}^n \sum_{c,s=1}^r \Phi_{icjs}^{+-}(\vec{k})P(\vec{k} + I_{ic} - I_{js}, x) + \right. \right. \\ & \left. \left. + \sum_{i,j=1}^n \sum_{c,s=1}^r \Phi_{icjs}^{++}(\vec{k})P(\vec{k} + I_{ic} + I_{js}, x) \right) dx \right). \end{aligned} \quad (3)$$

Пусть $P_q(\vec{k}, t)$ – приближение $P(\vec{k}, t)$ на q -й итерации, $P_{q+1}(k, t)$ – решение системы (2), полученное методом последовательных приближений. Тогда из (3) вытекает, что

$$\begin{aligned} P_{q+1}(\vec{k}, t) = & e^{-\Lambda(\vec{k})t} \left(P_q(\vec{k}, 0) + \int_0^t e^{\Lambda(\vec{k})x} \left(\sum_{i,j=1}^n \sum_{c,s=1}^r \Phi_{icjs}^{+-}(\vec{k})P_q(\vec{k} + I_{ic} - I_{js}, x) + \right. \right. \\ & \left. \left. + \sum_{i,j=1}^n \sum_{c,s=1}^r \Phi_{icjs}^{++}(\vec{k})P_q(\vec{k} + I_{ic} + I_{js}, x) \right) dx \right). \end{aligned} \quad (4)$$

В качестве начального приближения возьмём стационарное распределение $P_0(\vec{k}, t) = P(\vec{k}) = \lim_{t \rightarrow \infty} P(\vec{k}, t)$, которое удовлетворяет соотношению

$$\begin{aligned} \Lambda(\vec{k})P(\vec{k}) &= \sum_{i,j=1}^n \sum_{c,s=1}^r \Phi_{icjs}^{+-}(\vec{k})P(\vec{k} + I_{ic} - I_{js}) + \\ &+ \sum_{i,j=1}^n \sum_{c,s=1}^r \Phi_{icjs}^{++}(\vec{k})P(\vec{k} + I_{ic} + I_{js}). \end{aligned} \quad (5)$$

Можно доказать, что для последовательных приближений справедливы следующие утверждения.

Теорема 1. Последовательные приближения $P_q(\vec{k}, t), q = 0, 1, 2, \dots$, сходятся при $t \rightarrow \infty$ к стационарному решению системы уравнений (2).

Теорема 2. Последовательность $\{P_q(\vec{k}, t)\}, q = 0, 1, 2, \dots$, построенная по схеме (4), при любом ограниченном по t нулевом приближении $P_0(\vec{k}, t), 0 \leq P_0(\vec{k}, t) \leq 1$, сходится при $m \rightarrow \infty$ к единственному решению системы (2).

Теорема 3. Любое приближение $P_q(\vec{k}, t), q \geq 1$ представимо в виде сходящегося степенного ряда

$$P_q(\vec{k}, t) = \sum_{l=0}^{\infty} d_{ql}^{+-}(\vec{k}) t^l, \quad (6)$$

коэффициенты которого удовлетворяют рекуррентным соотношениям:

$$\begin{aligned} d_{q+l}^{+-}(\vec{k}) &= \frac{-\Lambda(\vec{k})^l}{l!} \left\{ P(\vec{k}, 0) + \sum_{u=0}^{l-1} \frac{(-1)^{u+1} u!}{\Lambda(\vec{k})^{u+1}} D_{qu}^{+-}(\vec{k}) \right\}, \quad l \geq 0, \\ d_{q0}^{+-}(\vec{k}) &= P(\vec{k}, 0), \quad d_{0l}^{+-}(\vec{k}) = P(\vec{k}, 0) \delta_{l0}, \end{aligned} \quad (7)$$

$$D_{ql}^{+-}(\vec{k}) = \sum_{i,j=1}^n \left[\sum_{s,c=1}^r \Phi_{icjs}^{+-}(\vec{k}) d_{ql}^{+-}(\vec{k} + I_{ic} - I_{js}) + \Phi_{icjs}^{++}(\vec{k}) d_{ql}^{+-}(\vec{k} + I_{ic} + I_{js}) \right].$$

Аналогично как в [6] можно показать, что радиус сходимости степенного ряда (6) равен $+\infty$.

ЛИТЕРАТУРА

1. *Gelenbe E.* Product form queueing networks with negative and positive customers // *J. Applied Probability*. – 1991. – V. 28. – No. 1. – P. 656–663.
2. *Gelenbe E.* G-networks with triggered customer movement // *J. Applied Probability*. – 1993. – V. 30. – No. 1. – P. 742–748.
3. *Якубович О.В., Евдокимович В.Е.* Сеть массового обслуживания со случайным временем пребывания различных типов положительных, отрицательных заявок и сигналов // *Проблемы физики, математики и техники*. – 2010. – № 4. – С. 63–67.
4. *Fourneau J.M., Gelenbe E., Suros R.* G-networks with multiple classes of negative and positive customers // *Theoretical Computer Science*. – 1996. – V. 155. – P.141–156.
5. *Матальцкий М.А., Науменко В.В.* Стохастические сети с нестандартными перемещениями заявок: монография. – Гродно: ГрГУ, 2016. – 348 с.
6. *Копать Д.Я., Матальцкий М.А.* Исследование в переходном режиме G-сети с сигналами и групповым удалением заявок методом последовательных приближений // *Вестник ГрГУ*. Сер. 2. – 2017. – № 3.

АНАЛИЗ ПОТОКОВ ПЕРВИЧНЫХ ТРЕБОВАНИЙ В ТАНДЕМЕ ПРИ ЦИКЛИЧЕСКОМ УПРАВЛЕНИИ С ПРОДЛЕНИЕМ

В.М. Кочеганов, А.В. Зорин

*Нижегородский государственный университет им. Н.И. Лобачевского,
г. Н. Новгород, Россия*

Существует ряд исследований, посвященных управлению конфликтными потоками на автомобильных перекрестках. В литературе, как правило, изучаются следующие виды алгоритмов управления: циклический алгоритм с фиксированной длительностью, циклический алгоритм с петлей, циклический алгоритм со сменой режимов и т.д. [1–8]. Наибольший интерес, однако, представляет ситуация с несколькими последовательными перекрестками, поскольку является более близкой к реальности. А именно, выходной поток автомобилей одного перекрестка создает входной поток другого перекрестка. Таким образом, вероятностная структура некоторых входных потоков априори неизвестна и существенно зависит от алгоритма обслуживания перекрестков.

На данный момент существуют несколько работ, посвященных тандемам перекрестков. В работе [8] приведена система моделирования смежных перекрестков с использованием компьютера. В работе [9] исследована модель тандема перекрестков, управляемая циклическим алгоритмом, а также были найдены условия существования стационарного режима системы. В работах [10–12] учитывается немгновенность перемещения между перекрестками и продление по приоритетному направлению на втором перекрестке. Данная статья продолжает исследования работ [10–12] и излагает результаты, которые касаются потоков первичных требований, поступающих в систему из внешней среды.

Постановка задачи на содержательном уровне

Рассмотрим систему массового обслуживания следующего вида (рис. 1). Пусть в систему с одним обслуживающим устройством поступают потоки Π_1, Π_2, Π_3 и Π_4 . Требования по потоку Π_j становятся в соответствующую очередь O_j с неограниченной вместимостью, $j \in \{1, 2, 3, 4\}$. Для $j \in \{1, 2, 3\}$ дисциплина очереди O_j , поддерживае-

мая устройством δ_j , имеет тип FIFO (First In First Out). Таким образом, для обслуживания из соответствующей очереди выбирается то требование, которое пришло раньше. Дисциплина очереди O_4 будет описана ниже. Будем предполагать, что входные потоки Π_1 и Π_3 формируются внешней средой, которая имеет только одно состояние, то есть вероятностная структура потоков не меняется с течением времени. Требования потоков Π_1 и Π_3 формируют независимые между собой неординарные пуассоновские потоки. Интенсивности соответствующих простейших потоков для Π_1 и Π_3 будем обозначать λ_1 и λ_3 , а распределение числа заявок в группе по потоку Π_j будем описывать производящей

функцией $f_j(z) = \sum_{v=1}^{\infty} p_v^{(j)} z^v$, которая предполагается аналитической при

$|z| < (1 + \varepsilon)$ для некоторого $\varepsilon > 0$. Величина $p_v^{(j)}$ определяет вероятность того, что по потоку Π_j число требований в группе равно v , $j \in \{1, 3\}$. Обслуженные требования потока Π_1 поступают на повторное обслуживание, формируя при этом поток Π_4 . Потоки Π_2 и Π_3 являются конфликтными, что означает запрет на одновременное обслуживание требований этих потоков и, следовательно, исследование системы не может быть сведено к задаче с меньшим числом потоков.

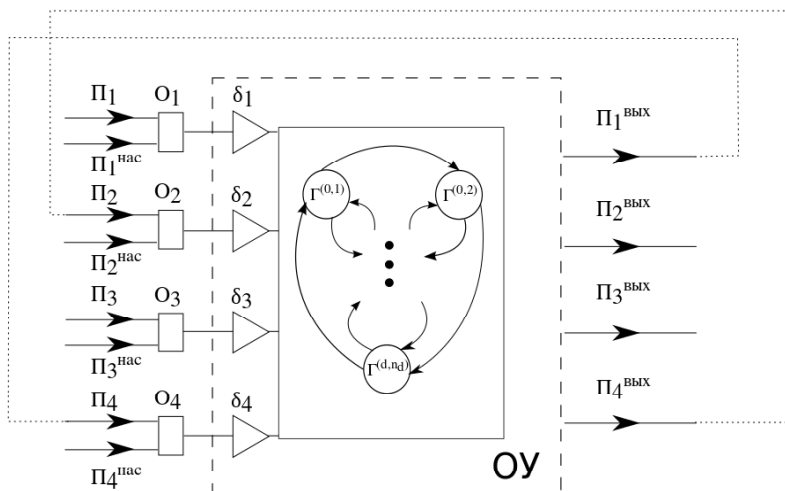


Рис. 1. Структурная схема системы обслуживания

В каждый момент времени обслуживающее устройство находится в одном из конечного множества состояний $\Gamma = \{\Gamma^{(k,r)}: k=0, 1, \dots, d; r=1, 2, \dots, n_k\}$ с заданными натуральными числами d, n_0, n_1, \dots, n_d . В каждом состоянии $\Gamma^{(k,r)}$ обслуживающее устройство находится в течение времени $T^{(k,r)}$. Будем предполагать, что для каждого фиксированного $k^* \geq 1$ цикл состояний $\{\Gamma^{(k^*,r)}: r=1, 2, \dots, n_{k^*}\} = C_{k^*}^N \cup C_{k^*}^O \cup C_{k^*}^I$ может быть представлен как объединение непересекающихся множеств нейтральных, выходных и входных состояний соответственно. Для $k^* = 0$ состояния из множества $\Gamma = \{\Gamma^{(0,r)}: r=1, 2, \dots, n_0\}$ будем называть состояниями продления. Более подробно алгоритм смены состояний обслуживающего устройства описан в статье [12].

В качестве наглядной физической интерпретации можно привести тандем из двух перекрестков. В роли потоков требований, формируемых внешней средой, выступают потоки прибывающих на перекрестки машин: конфликтные потоки Π_1, Π_5 на первом перекрестке, а также поток Π_3 на втором. Каждая машина из потока Π_1 , проезжая первый перекресток, становится в очередь O_4 потока Π_4 и затем с некой вероятностью (а именно, с вероятностью $p_{k,r}$ для состояния $\Gamma^{(k,r)}$ обслуживающего устройства) доезжает до следующего перекрестка или же не успевает это сделать и остается в очереди O_4 до следующего такта обслуживания. В случае, если машина из очереди O_4 успевает доехать до второго перекрестка, она становится в очередь O_2 и ждет своей очереди для его прохождения. Такая пара перекрестков является примером более общей модели системы массового обслуживания, описанной выше.

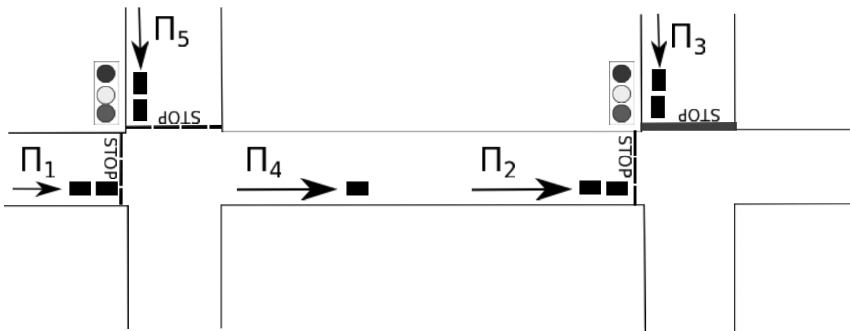


Рис. 2. Тандем перекрестков

Математическая модель

Описанная в предыдущем разделе на содержательном уровне система массового обслуживания должна рассматриваться как кибернетическая управляющая система обслуживания (см. работу [10]). Схема управляющей системы приведена на (рис. 1). На схеме присутствуют следующие блоки: 1) внешняя среда с одним состоянием; 2) входные полюса первого типа – входные потоки $\Pi_1, \Pi_2, \Pi_3, \Pi_4$; 3) входные полюса второго типа – потоки насыщения $\Pi_1^{\text{нас}}, \Pi_2^{\text{нас}}, \Pi_3^{\text{нас}}, \Pi_4^{\text{нас}}$; 4) внешняя память – очереди O_1, O_2, O_3, O_4 ; 5) устройство по переработке информации внешней памяти – устройства по поддержанию дисциплины очереди $\delta_1, \delta_2, \delta_3, \delta_4$; 6) внутренняя память обслуживающего устройства – обслуживающее устройство (ОУ); 7) устройство по переработке информации во внутренней памяти – граф смены состояний; 8) выходные полюса $\Pi_1^{\text{вых}}, \Pi_2^{\text{вых}}, \Pi_3^{\text{вых}}, \Pi_4^{\text{вых}}$. Координатой блока является номер этого блока на схеме.

Введем следующие величины и элементы, а также укажем множества их возможных значений. В качестве дискретной временной шкалы выберем последовательность $\tau_0 = 0, \tau_1, \tau_2, \dots$ моментов смены состояний обслуживающего устройства. Обозначим Γ_i из множества Γ состояние обслуживающего устройства в течение времени $(\tau_{i-1}; \tau_i]$, количество $\kappa_{j,i} \in Z_+$ требований в очереди O_j в момент времени τ_i , количество $\eta_{j,i} \in Z_+$ требований, поступивших в очередь O_j по потоку Π_j в течение времени $(\tau_i; \tau_{i+1}]$, количество $\xi_{j,i} \in Z_+$ требований по потоку насыщения $\Pi_j^{\text{нас}}$ в течение времени $(\tau_i; \tau_{i+1}]$, количество $\bar{\xi}_{j,i} \in Z_+$ реально обслуженных требований по потоку Π_j в течение времени $(\tau_i; \tau_{i+1}]$, $j \in \{1, 2, 3, 4\}$.

Закон изменения состояния обслуживающего устройства будем предполагать заданным соотношением $\Gamma_{i+1} = h(\Gamma_i, \kappa_{3,i})$, где

$$h(\Gamma^{(k,r)}, x) = \begin{cases} \Gamma^{(k, r \oplus_k 1)}, & \text{если } \Gamma^{(k,r)} \in C_k \setminus C_k^O; \\ \Gamma^{(k, r \oplus_k 1)}, & \text{если } \Gamma^{(k,r)} \in C_k^O \text{ и } x > L; \\ \Gamma^{(k, h_1(\Gamma^{(k,r)}))}, & \text{если } \Gamma^{(k,r)} \in C_k^O \text{ и } x \leq L; \\ \Gamma^{(0, h_2(r))}, & \text{если } k = 0 \text{ и } x \leq L; \\ h_3(r), & \text{если } k = 0 \text{ и } x > L. \end{cases}$$

Определим значения функций $\varphi_1(\cdot, \cdot)$, $\varphi_3(\cdot, \cdot)$ с помощью разложений

$$\sum_{x=0}^{\infty} z^x \varphi_j(x, t) = \exp(\lambda_j t (f_j(z) - 1)), \quad \text{где функции } f_j(z) \text{ определены выше,}$$

$j \in \{1, 3\}$. По своему смыслу число $\varphi_j(x, t)$ есть вероятность поступления $x = 0, 1, \dots$ требований за время $t \geq 0$ по потоку Π_j . Если $x < 0$, то значение функции $\varphi_j(x, t)$ положим равным нулю.

Более детально математическая модель рассмотрена в работе [12].

Потоки первичных требований

В этой работе основное внимание будет уделено стохастической последовательности $\{(\Gamma_i, \kappa_{1,i}, \kappa_{3,i}); i = 0, 1, \dots\}$, которая включает в себя состояния $(\kappa_{1,i}, \kappa_{3,i})$ очередей первичных требований и состояние Γ_i обслуживающего устройства в момент τ_i . Приведем ниже несколько результатов, касающихся этой последовательности.

Теорема 1. Пусть $\Gamma_0 = \Gamma^{(k,r)} \in \Gamma$ и $(\kappa_{1,0}, \kappa_{3,0}) = (x_1, x_3) \in Z_+^2$ фиксированы. Тогда стохастическая последовательность $\{(\Gamma_i, \kappa_{1,i}, \kappa_{3,i}); i = 0, 1, \dots\}$ является однородной счетной цепью Маркова. ■

Обозначим для $\gamma \in \Gamma$ и $(x_1, x_3) \in Z_+^2$

$$Q_{1,i}(\gamma, x_1, x_3) = P(\Gamma_i = \gamma, \kappa_{1,i} = x_1, \kappa_{3,i} = x_3).$$

Введем множество

$$H_{-1}(\Gamma^{(k,r)}, x_3) = \begin{cases} \{\Gamma^{(k_1, r_1)}, \Gamma^{(0, r_0-1)}\}, & (k=0) \wedge (x_3 \leq L); \\ \{\Gamma^{(k_1, r-k_1)}, \Gamma^{(0, r_2)}\}, & (\Gamma^{(k,r)} \in C_k^I) \wedge (x_3 > L); \\ \{\Gamma^{(k_1, r-k_1)}\}, & (\Gamma^{(k,r)} \in C_k^O) \vee (\Gamma^{(k,r)} \in C_k^N); \\ \emptyset, & (k=0 \wedge x_3 > L) \vee (\Gamma^{(k,r)} \in C_k^I \wedge x_3 \leq L); \end{cases}$$

где k_1, r_1 таковы, что $h_1(\Gamma^{(k_1, r_1)}) = r$ и r_2 таково, что $h_3(r_2) = \Gamma^{(k,r)}$. Отображения $h_1(\bullet)$ и $h_3(\bullet)$ являются частью алгоритма изменения состояния обслуживающего устройства и по своему смыслу задают правила перехода из выходного состояния и состояния продления соответственно (см. работу [12]).

Теорема 2. Пусть $\gamma = \Gamma^{(k,r)} \in \Gamma$ и $(x_1, x_3) \in Z_+^2$. Тогда для переходных вероятностей $\{Q_{1,i}(\bullet, \bullet, \bullet)\}_{i \geq 0}$ марковской цепи $\{(\Gamma_i, \kappa_{1,i}, \kappa_{3,i}); i = 0, 1, \dots\}$ имеют место следующие рекуррентные соотношения:

для $x_1 = 0, x_3 = 0$:

$$\begin{aligned} Q_{1,i+1}(\gamma, x_1, x_3) &= \sum_{\bar{x}_1=0}^{\ell(k,r,1)} \sum_{\bar{x}_3=0}^{\ell(k,r,3)} \sum_{\bar{\gamma} \in H_{-1}(\gamma, x_3)}^{\ell(k,r,3)} Q_{1,i}(\bar{\gamma}, \bar{x}_1, \bar{x}_3) \times \\ &\times \sum_{a=0}^{\ell(k,r,3)-\bar{x}_3} \varphi_3(a, T^{(k,r)}) \times \sum_{a=0}^{\ell(k,r,1)-\bar{x}_1} \varphi_1(a, T^{(k,r)}); \end{aligned}$$

для $x_1 = 0, x_3 > 0$:

$$\begin{aligned} Q_{1,i+1}(\gamma, x_1, x_3) &= \sum_{\bar{x}_1=0}^{\ell(k,r,1)} \sum_{\bar{x}_3=0}^{x_3 + \ell(k,r,3)} \sum_{\bar{\gamma} \in H_{-1}(\gamma, x_3)}^{\ell(k,r,3)} Q_{1,i}(\bar{\gamma}, \bar{x}_1, \bar{x}_3) \times \\ &\times \varphi_3(x_3 + \ell(k,r,3) - \bar{x}_3, T^{(k,r)}) \times \sum_{a=0}^{\ell(k,r,1)-\bar{x}_1} \varphi_1(a, T^{(k,r)}); \end{aligned}$$

для $x_1 > 0, x_3 = 0$:

$$\begin{aligned} Q_{1,i+1}(\gamma, x_1, x_3) &= \sum_{\bar{x}_1=0}^{x_1 + \ell(k,r,1)} \sum_{\bar{x}_3=0}^{\ell(k,r,3)} \sum_{\bar{\gamma} \in H_{-1}(\gamma, x_3)}^{\ell(k,r,3)} Q_{1,i}(\bar{\gamma}, \bar{x}_1, \bar{x}_3) \times \\ &\times \sum_{a=0}^{\ell(k,r,3)-\bar{x}_3} \varphi_3(a, T^{(k,r)}) \times \varphi_1(x_1 + \ell(k,r,1) - \bar{x}_1, T^{(k,r)}); \end{aligned}$$

для $x_1 > 0, x_3 > 0$:

$$\begin{aligned} Q_{1,i+1}(\gamma, x_1, x_3) &= \sum_{\bar{x}_1=0}^{x_1 + \ell(k,r,1)} \sum_{\bar{x}_3=0}^{x_3 + \ell(k,r,3)} \sum_{\bar{\gamma} \in H_{-1}(\gamma, x_3)}^{\ell(k,r,3)} Q_{1,i}(\bar{\gamma}, \bar{x}_1, \bar{x}_3) \times \\ &\times \varphi_3(x_3 + \ell(k,r,3) - \bar{x}_3, T^{(k,r)}) \times \varphi_1(x_1 + \ell(k,r,1) - \bar{x}_1, T^{(k,r)}). \blacksquare \end{aligned}$$

Заключение

Полученные результаты позволяют провести дальнейшее изучение свойств последовательности $\{(\Gamma_i, \kappa_{1,i}, \kappa_{3,i}); i = 0, 1, \dots\}$. С помощью рекуррентных соотношений для переходных вероятностей могут быть найдены рекуррентные соотношения для частных производных

функций при фиксированном состоянии обслуживающего устройства. После чего, с помощью итеративно-мажорантного метода, могут быть получены необходимые и достаточные условия существования стационарного режима марковской цепи $\{(\Gamma_i, \kappa_{1,i}, \kappa_{3,i}); i = 0, 1, \dots\}$.

ЛИТЕРАТУРА

1. Haight F.A. Mathematical Theories of Traffic Flow. – New York: Academic, 1963.
2. Inose H., Hamada T. Road Traffic Control. – Tokyo: Univ. of Tokyo Press, 1975.
3. Drew D.R. Traffic Stream Theory and Control. – New York: McGraw-Hill, 1968.
4. Fedotkin M.A. On a class of stable algorithms for control of conflicting flows or arriving airplanes // Problems of Control and Information Theory. – 1977. – V. 6. – No. 1. – P. 13–22.
5. Fedotkin M.A. Construction of a model and investigation of nonlinear algorithms for control of intense conflict flows in a system with variable structure of servicing demands // Lithuanian Mathematical Journal. – 1977. – V. 7. – No. 1. – P. 129–137.
6. Litvak N.V., Fedotkin M.A. A probabilistic model for the adaptive control of conflict flows // Automation and Remote Control. – 2000. – V. 61. – No. 5. – P. 777–784.
7. Proidakova E.V., Fedotkin M.A. Control of output flows in the system with cyclic servicing and readjustments // Automation and Remote Control. – 2008. – V. 69. – No. 6. – P. 993–1002.
8. Yamada K., Lam T.N. Simulation analysis of two adjacent traffic signals // Proceedings of the 17th Winter Simulation Conference. ACM, New York, 1985. – P. 454–464.
9. Zorine A.V. Stability of a tandem of queueing systems with Bernoulli noninstantaneous transfer of customers // Theory of Probability and Mathematical Statistics. – 2012. – V. 84. – P. 173–188.
10. Кочеганов В.М., Зорин А.В. Вероятностная модель тандема систем массового обслуживания с циклическим управлением с продлением // Теория вероятностей, случайные процессы, математическая статистика и приложения: материалы Междунар. науч. конф., посвящ. 80-летию проф., д-ра физ.-мат. наук Г.А. Медведева, Минск, 23–26 февр. 2015. – С. 94–99.
11. Kochegarov V.M., Zorine A.V. Low-priority queue fluctuations in tandem of queueing systems under cyclic control with prolongations // Distributed Computer and Communication Networks. Ser. Communications in Computer and Information Science. – 2016. – V. 601. – P. 268–279.
12. Kochegarov V.M., Zorine A.V. Low-priority queue and server's steady-state existence in a tandem under prolongable cyclic service // Distributed Computer and Communication Networks. DCCN 2016. Communications in Computer and Information Science, 2016. – V. 678. – P. 210–221.
13. Кочеганов В.М., Зорин А.В. Достаточное условие существования стационарного режима низкоприоритетной очереди в тандеме систем массового обслуживания // Вестник Волжской государственной академии водного транспорта. Вып. 50. – Н. Новгород: Изд-во ФГБОУ ВО «ВГУВТ», 2017. – С. 47–55.

УПРАВЛЕНИЕ ТЕМПОМ ПРОИВОДСТВА И ЦЕНОЙ ПРОДАЖИ СКОРОПОРТЯЩЕЙСЯ ПРОДУКЦИИ

К.И. Лившиц, Е.С. Ульянова

Томский государственный университет, г. Томск, Россия

В теории управления запасами с ограниченным сроком годности рассматриваются различные модели [1, 2]. При этом, как правило, предполагается, что задан некоторый начальный объем запаса и решаются задачи об оптимальном выборе начального объема и построении оптимального алгоритма его расходования. В данной работе в продолжение работ [3–5] предлагается и анализируется модель одновременно-го производства и сбыта скоропортящейся продукции, рассматривается релейное управление темпом производства и интенсивностью продаж.

1. Математическая модель задачи

Пусть $S(t)$ – количество продукции в момент времени t . Будем считать, что продукция производится с некоторой скоростью $c(S)$, зависящей от текущего запаса $S(t)$, так что за время Δt поступает $c(S)\Delta t$ единиц продукции. При хранении продукция непрерывно портится. Считается, что за малое время Δt потери равны $kS(t)\Delta t$. Будем считать далее, что продажа осуществляется партиями случайного объёма x , где величины покупок x – 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 λ зависит от цены продажи b . Считается, что интенсивность потока продаж λ монотонно убывает с ростом цены b .

При фиксированной цене продажи b и, следовательно, интенсивности потока продаж λ и скорости производства c среднее количество продукции, согласно [3], определяется соотношением

$$\bar{S}(t) = S(0)e^{-kt} + \frac{c - \lambda a}{k}(1 - e^{-kt}).$$

При $t \gg 1$ $\bar{S}(t)$ ограничено сверху величиной $(c - \lambda a)/k$, при $c - \lambda a > 0$ и $t \gg 1$ образуется постоянный запас нереализованной продукции, что нежелательно. При $c - \lambda a \leq 0$ появляется неудовлетворённый спрос. В этом случае считается, что $S(t) < 0$ и заказы удовлетворяются в порядке их поступления. Поэтому необходимо организовать управление темпом производства продукции c или (и) темпом продажи в зависимости от текущего запаса продукции.

Предполагается, что управление темпом производства и продажами осуществляется следующим образом. Устанавливается пороговое значение допустимого запаса продукции S_0 . В области $S(t) \leq S_0$ скорость производства $c(S) = c_0$ и назначается цена продажи b_0 , в области $S(t) > S_0$ скорость производства $c(S) = c_1$ и назначается цена продажи $b_1 < b_0$. Так как интенсивность λ зависит от цены продукции b , то можно записать

$$\lambda(S), c(S) = \begin{cases} \lambda_0, c_0, & S < S_0, \\ \lambda_1, c_1, & S \geq S_0, \end{cases}$$

где $c_1 < c_0$ и $\lambda_1 > \lambda_0$.

Основной целью работы является определение параметров, позволяющих максимизировать прибыль предприятия.

Обозначим

$$P_0(S, t) dS = \Pr \{ S \leq S(t) < S + dS, c(t) = c_0, \lambda(t) = \lambda_0 \}, \quad S < S_0,$$

$$P_1(S, t) dS = \Pr \{ S \leq S(t) \leq S + dS, c(t) = c_1, \lambda(t) = \lambda_1 \}, \quad S \geq S_0.$$

Плотность распределения количества продукции $P(S, t)$ определяется соотношением

$$P(S, t) = \begin{cases} P_0(S, t), & S < S_0, \\ P_1(S, t), & S \geq S_0. \end{cases}$$

Аналогично [3, 4] можно показать, что плотность распределения количества продукции определяется системой уравнений (1) и (2):

$$\begin{aligned} \frac{\partial P_1(S, t)}{\partial t} = & -\lambda_1 P_1(S, t) - \frac{\partial}{\partial S} [(c_1 - kS) P_1(S, t)] + \\ & + \lambda_1 \int_0^{\infty} P_1(S + x, t) \varphi(x) dx, \quad S \geq S_0; \end{aligned} \quad (1)$$

$$\begin{aligned} \frac{\partial P_0(S, t)}{\partial t} = & -\lambda_0 P_0(S, t) - \frac{\partial}{\partial S} ((c_0 - kSI(S))P_0(S, t)) + \\ & + \lambda_0 \int_0^{S_0-S} P_0(S+x, t)\varphi(x)dx + \lambda_1 \int_{S_0-S}^{\infty} P_1(S+x, t)\varphi(x)dx, \quad S < S_0. \end{aligned} \quad (2)$$

Решение системы уравнений (1), (2), очевидно, должно удовлетворять условию нормировки $\int_{-\infty}^{S_0} P_0(S, t)dS + \int_{S_0}^{\infty} P_1(S, t)dS = 1$.

2. Плотность распределения количества продукции при экспоненциальном распределении величин покупок

В общем виде решение системы уравнений (1), (2) найти не удастся даже в стационарном режиме. Рассмотрим случай, когда величины покупок имеют экспоненциальное распределение

$$\varphi(x) = \frac{1}{a} e^{-\frac{x}{a}}.$$

Тогда в стационарном режиме уравнения (1), (2) принимают вид:

$$-\lambda_1 P_1(S) - \frac{d}{dS} [(c_1 - kS)P_1(S)] + \frac{\lambda_1}{a} \int_S^{\infty} P_1(x) e^{-\frac{x}{a}} dx = 0, \quad S \geq S_0; \quad (3)$$

$$\begin{aligned} & -\lambda_0 P_0(S) - \frac{d}{dS} [(c_0 - kSI(S))P_0(S)] + \\ & + \frac{\lambda_0}{a} e^{\frac{S}{a}} \int_0^{S_0-S} P_0(x) e^{-\frac{x}{a}} dx + \frac{\lambda_1}{a} e^{\frac{S}{a}} \int_{S_0}^{\infty} P_1(x) e^{-\frac{x}{a}} dx = 0, \quad S < S_0. \end{aligned} \quad (4)$$

Решение системы уравнений (3), (4) имеет вид

$$P(S) = \begin{cases} A e^{\frac{S}{a}} (c_1 - kS)^{\frac{\lambda_1}{k}-1}, & S > S_0, \\ B e^{\frac{S}{a}} (c_0 - kS)^{\frac{\lambda_0}{k}-1}, & 0 \leq S \leq S_0, \\ B c_0^{\frac{\lambda_0}{k}-1} e^{\frac{c_0 - \lambda_0 a}{c_0 a} S}, & S < 0. \end{cases} \quad (5)$$

Связь между константами A и B определяется выражением

$$B = A \frac{(c_1 - kS_0)^{\frac{\lambda_1}{k}-1}}{(c_0 - kS_0)^{\frac{\lambda_0}{k}-1}}. \quad (6)$$

Константа A находится из условия нормировки.

3. Определение средней прибыли при экспоненциальном распределении величин покупок

Рассмотрим далее случай, когда интенсивность темпа продаж фиксированная: $\lambda_0 = \lambda_1 = \lambda$. Тогда выражение для средней прибыли определяется следующим образом (себестоимость единицы продукции считается равной 1):

$$R = \lambda ab \int_0^{\infty} P(S) dS - \int_{-\infty}^{\infty} c(S) P(S) dS. \quad (7)$$

На рис. 1 представлена зависимость прибыли от порога S_0 при следующих значениях параметров $k = 0,05$, $a = 1,5$, $b = 2$:

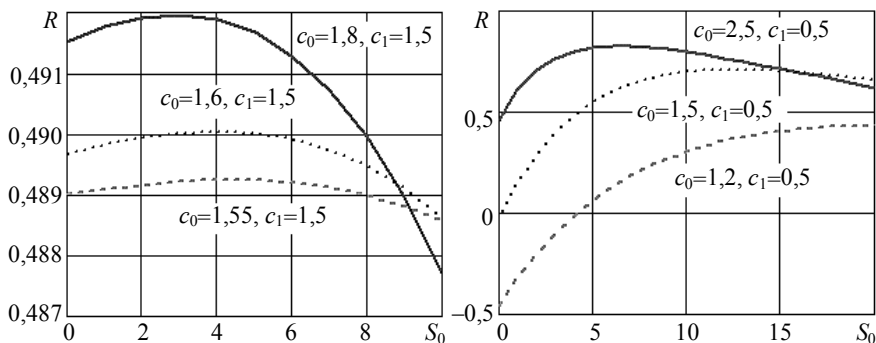


Рис. 1. Прибыль в зависимости от порогового значения количества запасов

В связи с большим количеством параметров, аналитическое решение задачи максимизации прибыли найти не удастся. В таблице представлена зависимость прибыли R от параметров S_0 , c_0 , c_1 , когда цена продажи $b = 2$, коэффициент порчи $k = 0,02$, $a = 1,5$.

Зависимость прибыли R от параметров c_0, c_1

Значения параметров		$c_1 = 0,1$	$c_1 = 0,3$	$c_1 = 0,5$	$c_1 = 0,7$	$c_1 = 0,9$	$c_1 = 1,1$	$c_1 = 1,3$	$c_1 = 1,5$
$c_0 = 6,0$	R	0,889	0,888	0,885	0,876	0,852	0,793	0,672	0,496
	S_0	4,1	3,7	3,3	2,7	1,9	1,1	0,1	0,1
$c_0 = 5,5$	R	0,885	0,884	0,881	0,873	0,849	0,791	0,672	0,496
	S_0	4,3	3,9	3,5	2,9	2,1	1,3	0,3	0,1
$c_0 = 4,5$	R	0,875	0,874	0,871	0,863	0,842	0,786	0,670	0,496
	S_0	4,7	4,5	3,9	3,5	2,7	1,7	0,7	0,1
$c_0 = 3,5$	R	0,857	0,859	0,856	0,849	0,830	0,778	0,667	0,495
	S_0	4,9	5,3	4,7	4,3	3,5	2,5	1,3	0,3
$c_0 = 2,5$	R	0,806	0,829	0,827	0,822	0,806	0,763	0,661	0,494
	S_0	4,9	6,7	6,3	5,7	4,9	3,7	2,5	1,3
$c_0 = 2,0$	R	0,728	0,797	0,796	0,792	0,780	0,745	0,654	0,493
	S_0	4,9	8,5	8,1	7,5	6,5	5,3	3,7	2,3
$c_0 = 1,5$	R	0,485	0,712	0,712	0,710	0,705	0,687	0,627	
	S_0	4,9	12,9	12,5	11,9	11,1	9,5	7,3	
$c_0 = 1,4$	R	0,375	0,671	0,670	0,670	0,667	0,655	0,610	
	S_0	4,9	14,9	14,7	14,1	13,3	11,7	9,3	
$c_0 = 1,3$	R	0,218	0,589	0,599	0,599	0,598	0,593		
	S_0	4,9	14,9	18,3	17,9	17,1	15,7		
$c_0 = 1,2$	R	-0,018	0,399	0,441	0,441	0,441	0,441		
	S_0	4,9	14,9	24,9	26,1	25,7	24,7		
$c_0 = 1,15$	R	-0,181	0,226	0,275	0,276	0,276	0,276		
	S_0	4,9	14,9	24,9	34,7	34,3	33,7		

Таким образом, из таблицы видно, что максимальное значение прибыли достигается при увеличении темпа производства c_0 . При этом оптимальное значение порога, при котором происходит переключение темпа производства, имеет вполне определенное значение. Другими словами, если нет запаса, то его нужно образовывать, в соответствии с производственными возможностями. При этом темп производства c_1 нужно поддерживать на минимальном уровне.

4. Асимптотические выражения для плотности распределения количества продукции и прибыли при $c_1 = 0$

Построим приближенное решение для плотности распределения количества продукции, когда $\varphi(x)$ – произвольная функция.

Введем параметр $\theta \ll 1$ и сделаем следующие предположения:

$$c_0 = \lambda a(1 + \theta), k = \theta^2 k_0. \quad (8)$$

Первое соотношение означает, что количество продукции, производимой в единицу времени, почти совпадает со средним сбытом в единицу времени. Второе – продукция портится достаточно медленно. Далее, будем считать, что при $\theta \rightarrow 0$ $S_0 = S_0(\theta) \rightarrow \infty$, но существует конечный предел $\lim \theta S_0(\theta) = z_0$ при $\theta \rightarrow 0$.

Можно показать аналогично [4], что при сделанных предположениях

$$P(S) = \begin{cases} D \theta e^{-\frac{\lambda a^2}{a_2 k_0} \int_0^{z_0} e^{\frac{(\lambda a - k_0 x)^2}{\lambda a_2 k_0}} dx} e^{\frac{2a}{a_2} \theta S} + o(\theta), & S \leq 0; \\ D \theta \int_{\theta S}^{z_0} e^{\frac{(\lambda a - k_0 x)^2}{\lambda a_2 k_0}} dx e^{-\frac{(\lambda a - k_0 \theta S)^2}{\lambda a_2 k_0}} + o(\theta), & 0 < S \leq S_0, \end{cases} \quad (9)$$

где постоянная D определяется из условия нормировки

$$D = \left(\frac{a_2}{2a} e^{-\frac{\lambda a^2}{a_2 k_0} \int_0^{z_0} e^{\frac{(\lambda a - k_0 x)^2}{\lambda a_2 k_0}} dx} + \int_0^{z_0} e^{\frac{(\lambda a - k_0 x)^2}{\lambda a_2 k_0}} \int_0^x e^{-\frac{(\lambda a - k_0 y)^2}{\lambda a_2 k_0}} dy dx \right)^{-1}. \quad (10)$$

Средняя прибыль в единицу времени в данном случае определится выражением

$$R = R(\theta S_0, \theta) = R(z_0, \theta) = \lambda a b \int_0^{S_0} P(S) dS - \lambda a(1 + \theta). \quad (11)$$

Можно показать, что функция $R(\theta S_0, \theta) = R(z_0, \theta)$ монотонно возрастает с ростом $z_0 = \theta S_0$ при любом $\theta \neq 0$. Поэтому оптимальные параметры нужно выбирать следующим образом. Задать предельный размер запаса S_0 , который, очевидно, всегда ограничен. Затем выбрать величину θ из условия максимума величины $R(\theta S_0, \theta)$. Зависимость зна-

чения прибыли от параметра θ представлена на рис. 2. Значения параметров: $b = 2$, $a = 1$, $a_2 = 2$, $\lambda / k_0 = 10$.

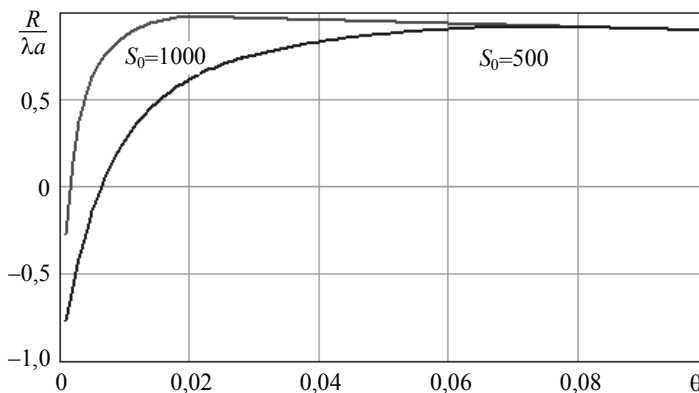


Рис. 2. Зависимость прибыли от параметра θ

Заключение

Получены уравнения, определяющие плотность распределения количества скоропортящейся продукции при релейном управлении скоростью производства и интенсивностью продаж. Найдены условия, определяющие максимальную среднюю прибыль при экспоненциальном распределении величин покупок и в общем случае, когда темп производства «почти совпадает» со средним сбытом продукции.

ЛИТЕРАТУРА

1. Goyal S.K., Giri B.C. Recent trends in modeling of deteriorating inventory // Eur. J. Oper. Res. – 2001. – V. 134. – No. 1. – P. 1–16.
2. Bakker M., Riezebos J., Teunter R.H. Review of inventory systems with deterioration since 2001 // Eur. J. Oper. Res. – 2012. – V. 221. – P. 275–284.
3. Livshits K., Ulyanova E. Switch-hysteresis control of the selling times flow in a model with perishable goods // CCIS. – 2015. – V. 564. – P. 263–274.
4. Livshits K., Ulyanova E. Switch-hysteresis control of the production process in a model with perishable goods // CCIS. – 2016. – V. 638. – P. 192–206.
5. Лившиц К.И., Ульянова Е.С. Диффузионная аппроксимация процесса производства и сбыта скоропортящейся продукции // Изв. вузов. Физика. – 2015. – Т. 58. – № 11/2. – С. 281–285.

К АНАЛИЗУ ВЕРОЯТНОСТИ УСПЕШНОГО УСТАНОВЛЕНИЯ СОЕДИНЕНИЯ ПО РАДИОКАНАЛУ СЛУЧАЙНОГО ДОСТУПА

Е.Г. Медведева

Российский университет дружбы народов, г. Москва, Россия

Изучается передача малых данных по сетям LTE (Long-Term Evolution). Актуальной является задача уменьшения нагрузки, создаваемой пользователями, а также обслуживающей сигнальной нагрузки. Для этого разрабатывают и внедряют алгоритмы упрощенного взаимодействия между пользовательскими устройствами и сетью с целью минимизации затраченных ресурсов сети. Данная работа является продолжением исследований процедуры установления соединения между устройством и базовой станцией LTE с помощью использования радиоканала случайного доступа RACH (Random Access Channel). Для этого строится математическая модель в виде цепи Маркова, получены формулы для расчета вероятности успешного установления соединения с учетом повторной передачи сигнальных сообщений. Численные результаты показывают точность предложенной аналитической модели.

Процедура установления соединения

Ввиду роста числа устройств, подключенных к Интернет и непрерывно передающих данные (как правило, это различные датчики и счетчики, которые помогают обеспечивать безопасность жизнедеятельности человека), анализ своевременной доставки данных и гарантии низких задержек интересен разработчикам [1].

Процедура установления соединения между пользователем и базовой станцией состоит из последовательной передачи четырех сообщений по радиоканалу случайного доступа [1–3] и начинается с передачи преамбулы от оборудования пользователя к базовой станции – Msg1. После этого оборудование пользователя ожидает от базовой станции ответ в виде Msg2 (RAR, Random-Access Response). Если время ожидания Msg2 истекло и сообщение не получено, значит, произошла коллизия по причине отправки несколькими устройствами одной и той же преамбулы в один момент времени. В случае коллизии оборудование поль-

зователя повторяет попытку передачи преамбулы после заранее заданного интервала ожидания. При превышении порога попыток передачи преамбулы процедура установления соединения по радиоканалу RACH признается неудачной. При успешной передаче преамбулы в виде Msg1 и получения ответа Msg2 следует сообщение HARQ (Hybrid Automatic Repeat request), или Msg3. Ответное сообщение Msg4 является последним в процедуре установления соединения. Считается, что ответной сообщением Msg2 всегда гарантированно передается. В предыдущих исследованиях предполагалось, что Msg4 также гарантированно передается, однако в данной работе на повторную передачу Msg4 устанавливается порог, при превышении которого необходимо заново передавать преамбулу Msg1 и последующие сообщения. Данная работа продолжает исследования [4–8]. В отличие от предыдущих работ сообщение Msg4 может быть повторно передано N_4 раз. Эти характеристики являются важными вероятностно-временными характеристиками и рекомендованы к исследованию в стандарте [1, разд. 6.3.1].

Введем вероятностные события $A_1 = \{\text{Msg1 передан успешно}\}$,

$$\overline{A_1} = \{\text{коллизия Msg1}\},$$

$$A_3 = \{\text{Msg3 передан успешно}\},$$

$$\overline{A_3} = \{\text{коллизия Msg3}\},$$

$$A_4 = \{\text{Msg4 передан успешно}\},$$

$$\overline{A_4} = \{\text{коллизия Msg4}\}.$$

Обозначим вероятность события $P(A_1) = 1 - p_1$, тогда $P(\overline{A_1}) = p_1$. Аналогично $P(A_3) = 1 - p_3$, $P(\overline{A_3}) = p_3$, $P(A_4) = 1 - p_4$, $P(\overline{A_4}) = p_4$.

Опишем процедуру установления соединения в виде цепи Маркова с дискретным временем $\{\xi_i, i = 0, \dots, N_1 \cdot N_3 \cdot N_4 + 1\}$, где N_i – число возможных повторных передач Msg(i), $i = \{1, 3, 4\}$. Построим пространство состояний:

$$\begin{aligned} \mathbf{X} = & \left\{ \begin{pmatrix} x_1, x_2, x_3, x_4, x_5 \\ y_1, y_2, y_3 \\ z_1, z_2 \end{pmatrix} \right. \left. \begin{array}{l} x_1 = \overline{0, N_1}, \quad 0 \leq x_2 \leq x_1, \quad 0 \leq y_2 \leq 1, \\ y_1 = \overline{0, N_3}, \quad 0 \leq x_i \leq x_2, \quad i \in \{3, 4\}, \quad x_4 \leq y_3 \leq x_4 \cdot N_3 \\ z_1 = \overline{0, N_4}, \quad x_3 + x_4 + x_5 = x_2, \quad 0 \leq z_2 \leq 1, \quad x_5 \in \{0, 1\} \end{array} \right\} = \\ & = \{(\mathbf{x}, \mathbf{y}, \mathbf{z})\}, \end{aligned}$$

где x_1 – общее число переданных Msg1; x_2 – число успешно переданных Msg1; x_3 – число успешно переданных Msg1, после которых соединение не установилось по причине N_3 неуспешных передач Msg3; x_4 – число успешно переданных Msg1, после которых соединение не установилось по причине N_4 неуспешных передач Msg4; x_5 – индикатор, равный «1» при успешной последней передаче Msg1, «0» – при коллизии Msg1; y_1 – общее число переданных Msg3 при последней успешной передаче Msg1; y_2 – число успешно переданных Msg3 при последней успешной передаче Msg1; y_3 – число переданных Msg3 (успешных и неуспешных) после которых N_4 раза произошли коллизии при передаче Msg4; z_1 – общее число переданных Msg4 при последней успешной передаче Msg1 и Msg3; z_2 – число успешных Msg4 при последнем успешном Msg1 и Msg3.

Цепь Маркова $\{\xi_i\}$ попадет в состояние $(\mathbf{x}, \mathbf{y}, \mathbf{z})$ из состояния $(\mathbf{0}, \mathbf{0}, \mathbf{0})$ с вероятностью

$$P_{(\mathbf{x}, \mathbf{y}, \mathbf{z})} = C_{x_1-1}^{x_3+x_4} C_{x_3+x_4}^{x_3} p_1^{x_1-x_2} (1-p_1)^{x_2} p_3^{N_3 x_3 - x_4 + y_1 - y_2 + y_3} (1-p_3)^{x_4+y_2} \times \\ \times p_4^{N_4 x_4 + z_1 - z_2} (1-p_4)^{z_2}, \quad (\mathbf{x}, \mathbf{y}, \mathbf{z}) \in \mathbf{X}. \quad (1)$$

Вероятность успешного установления соединения $P_{\text{усп}}$ находится по формуле (2).

$$P_{\text{усп}} = \frac{(1-p_4^{N_4})(1-p_3^{N_3})}{1-p_3} \left[\frac{1 - (p_1 + (1-p_1)(p_3^{N_3} + p_4^{N_4}(1-p_3)))^{N_1}}{(1-p_1)(1-p_3^{N_3} - p_4^{N_4}(1-p_3))} - \right. \\ \left. - p_3 \frac{1 - (p_1 + p_3^{N_3}(1-p_1)(p_3 + p_4^{N_4}(1-p_3)))^{N_1}}{(1-p_1)(1-p_3^{N_3}(p_3 + p_4^{N_4}(1-p_3)))} \right]. \quad (2)$$

Численный эксперимент

На рис. 1 показаны вероятности успешного установления соединения $P_{\text{усп}}$ для двух наборов значений повторных передач сообщений, а именно $N_1 = 4, N_3 = N_4 = 2$ и $N_1 = 10, N_3 = N_4 = 5$.

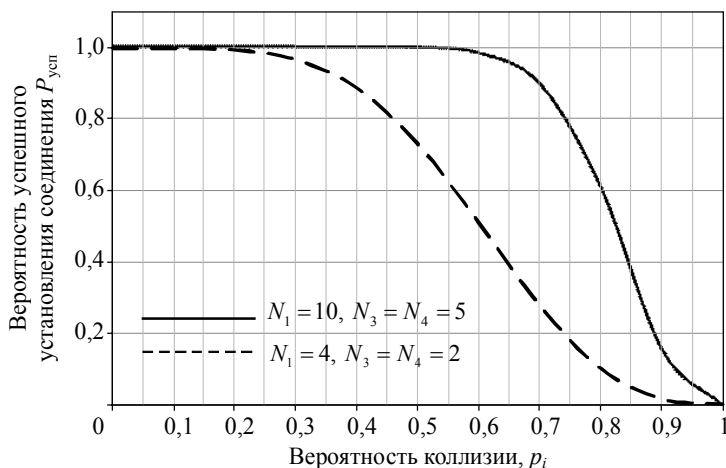


Рис. 1. Вероятность успешного установления соединения

При увеличении вероятности коллизии p_i , $i = \{1, 3, 4\}$, уменьшается вероятность успешного установления соединения. При меньшем числе возможных повторных передач поведение графика вероятности успешного соединения имеет более пологий характер.

Дальнейшие исследования вероятностно-временных характеристик для процедуры установления соединения по каналу радиодоступа планируется расширить для нахождения аналитической формулы среднего времени установления соединения.

ЛИТЕРАТУРА

1. GPP TR 37.869 – Study on Enhancements to Machine-Type Communications and other Mobile Data Applications; RAN Aspects. Release 12. Sept. 2013.
2. Cheng R.G., Wei C.H., Tsao S.L., Ren F.C. RACH collision probability for Machine-Type Communications // Proc. IEEE 75th Veh. Technol. Conf., May 2012. – P. 1–5.
3. Condoluci M., Araniti G., Dohler M., et al. Virtual code resource allocation for energy-aware MTC access over 5G systems // Ad Hoc Networks. – 2016. – V. 43. – P. 3–15.
4. Gerasimenko M., Petrov V., Galinina O., et al. Impact of M2M on energy and delay performance of random-access channel in LTE-Advanced // Transactions on Emerging Telecommunications Technologies. – 2013. – V. 24. – No. 4. – P. 366–377.
5. Borodakiy V.Y., Samouylov K.E., Gaidamaka Y.V., et al. Modelling a random access channel with collisions for M2M traffic in LTE networks // Lecture Notes in Computer Science. – 2014. – V. 8638. – P. 301–310.

6. *Samouylov K.E., Gaidamaka Y.V., Gudkova I.A., et al.* Baseline analytical model for machine-type communications over 3GPP RACH in LTE-advanced networks // 31st International Symposium on Computer and Information Sciences (ISCIS), October 27–28th, 2016, Krakow, Poland. T. Czachórski et al. (Eds.): ISCIS 2016, CCIS, 2016. – V. 659. – P. 203–213.
7. *Wei C.H., Bianchi G., Cheng R.G.* Modelling and analysis of random access channels with bursty arrivals in OFDMA wireless networks // IEEE Trans. on Wireless Commun. – 2015. – V. 14. – No. 4. – P. 1940–1953.
8. *Зарипова Э.Р., Ардила П.А.* Метод оценки времени установления соединения по радиоканалу случайного доступа // Вестник РУДН. Математика. Информатика. Физика. – 2017. – Т. 25. – № 1. – С. 9–18.

ИМИТАЦИОННОЕ МОДЕЛИРОВАНИЕ ЦИКЛИЧЕСКИХ СИСТЕМ МАССОВОГО ОБСЛУЖИВАНИЯ

А.Н. Моисеев, Д.А. Буртовая

Томский государственный университет, г. Томск, Россия

Циклические СМО – это класс систем массового обслуживания, в которых обслуживающие приборы в определенном (циклическом) порядке подключаются к различным накопителям для обслуживания заявок, находящихся в этих накопителях. Заявки попадают в накопители из нескольких входящих в систему потоков заявок. Аналитическим исследованиям циклических систем обслуживания посвящено достаточно много научных трудов (см., например, обзор [1]), но во многих случаях не удастся получить аналитические результаты. В этом случае применяют другие методы исследования, в том числе и имитационное моделирование [2]. Имитационный подход позволяет воспроизвести поведение системы и таким образом достичь требуемых результатов практически для любой конфигурации исследуемой системы. В настоящей работе представлено достаточно общее описание модели циклических систем и предложены механизмы для построения объектно-ориентированного приложения имитационного моделирования таких систем.

Описание модели циклической системы массового обслуживания

Рассматривается класс циклических систем массового обслуживания следующего вида (рис. 1). На вход системы поступает $K \geq 1$ входящих потоков A_1, \dots, A_K , каждый из которых описывается заданной для него моделью случайного потока событий с определенными параметрами. Все заявки k -го входящего потока ($k = 1, \dots, K$) поступают в k -й накопитель (буфер), который может иметь заданный максимальный размер или может быть неограничен. Имеется некоторое число $N \geq 1$ обслуживающих приборов (серверов) B_1, \dots, B_N , которые подключаются к накопителям, забирают из них заявки в порядке, соответствующем некоторой дисциплине, и обслуживают эти заявки в соответствии с некоторым заданным для каждого прибора законом. После этого заявки покидают систему. Функция распределения времени обслуживания может определяться для каждого прибора независимо от накопителя, из которого

взята заявка, но также для каждого прибора может быть задан индивидуальный набор из K функций распределения, определяющих продолжительность обслуживания заявок, поступивших из разных входящих потоков. Таким образом, при K входящих потоках (накопителях) и N обслуживающих приборах можно задать $N \times K$ функций распределения $B_{nk}(x)$, $n = \overline{1, N}$, $k = \overline{1, K}$, определяющих продолжительность обслуживания разными приборами заявок из разных входящих потоков.

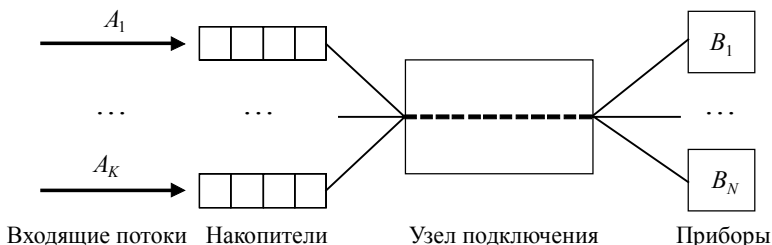


Рис. 1. Модель циклической системы

Порядок, способ и длительность подключения приборов к накопителям определяется заданными параметрами, которые в дальнейшем будем называть параметрами подключения. Переключение приборов в циклических СМО всегда производится циклически – от предыдущего k -го накопителя к следующему $(k + 1)$ -му и от последнего K -го к первому. В случае, когда количество серверов $N > 1$, возможны различные варианты их «группового» поведения. В качестве основных групповых дисциплин выделим три:

1. *Жесткая группа.* Все обслуживающие приборы системы образуют одну группу обслуживания: все приборы в один момент времени работают только с одним накопителем, переключение между накопителями и пребывание в состоянии переключения (см. далее) для всех приборов производится синхронно.

2. *Одиночные приборы.* Обслуживающие приборы подключаются к накопителям независимо друг от друга, причем с одним накопителем в один момент времени может работать только один прибор. В случае, если по окончании сеанса подключения прибора к накопителю следующий накопитель уже обслуживается другим сервером, то прибор ищет следующий по циклическому порядку свободный (не находящийся в сеансе подключения с другим сервером) накопитель и подключается

ся к нему. Эта дисциплина подключения доступна только при условии, что количество серверов не превышает количество накопителей ($N \leq K$).

3. *Гибкие группы.* Обслуживающие приборы подключаются к накопителям независимо друг от друга, в один момент времени к накопителю может быть подключено любое число приборов. По окончании сеанса подключения прибор подключается к следующему накопителю независимо от того, обслуживается ли он другими серверами или нет.

Кроме групповых дисциплин также важна дисциплина подключения, определяющая продолжительность одного сеанса подключения сервера или группы серверов к накопителю. В данной работе рассматривается 2 типа дисциплин подключения:

1. *С разделением времени.* Продолжительность каждого подключения определяется заданной функцией распределения длительности сеанса. По окончании этого периода обслуживание всех заявок, производившееся в рамках данного подключения, прерывается, а сами заявки возвращаются в накопитель. Продолжительность сеанса подключения может определяться одной функцией распределения для всех накопителей либо задаваться индивидуально для каждого накопителя.

2. *До полного исчерпания.* Подключение сохраняется до тех пор, пока не будет закончено обслуживание последней заявки в накопителе. При этом в случае жесткой группы все незадействованные приборы простаивают, пока последний из них не закончит обслуживание.

Когда прибор или группа приборов завершают обслуживание одного накопителя и должны переключиться к следующему, то параметрами задачи может быть задано, что это переключение требует некоторого времени. В случае применения дисциплины подключения «до полного исчерпания» это время должно быть обязательно больше нуля, так как в противном случае, если все накопители оказались временно пусты, модель перейдет в состояние «зацикливания», когда приборы мгновенно переключаются от очереди к очереди, а модельное время не изменяется. Если параметрами задачи указано, что переключение требует времени, то задается функция распределения, определяющая длительность состояния переключения, в которое прибор или группа приборов переходят по окончании сеанса подключения к одному накопителю. В этом состоянии приборы не выполняют никаких действий. Функция распределения длительности времени переключения может определяться одной функцией распределения для всех накопителей либо задаваться индивидуально для каждого накопителя.

Основные объекты имитационного моделирования циклической системы обслуживания

В настоящей работе в качестве платформы для реализации приложения имитационного моделирования циклических систем будет использована платформа программного комплекса ODIS [3, 4], предназначенного для имитационного моделирования сетей массового обслуживания. Данная платформа использует дискретно-событийный подход моделирования. В ней имеется базовый класс *SimulationModel*, который реализует основной цикл дискретно-событийного моделирования систем массового обслуживания. Исполнение цикла моделирования базируется на взаимодействии объектов событий, заявок и элементов системы.

Для нужд разрабатываемой системы моделирования дополним объектную модель следующими классами (рис. 2). Классы *Element*, *Source*, *Source*, *PassiveBuffer* реализованы каркасом ODIS и предназначены для моделирования элементов СМО общего вида. Классы *InputSource*, *Server* и *ConnectionManager* реализуют объекты, специфичные для циклических СМО.

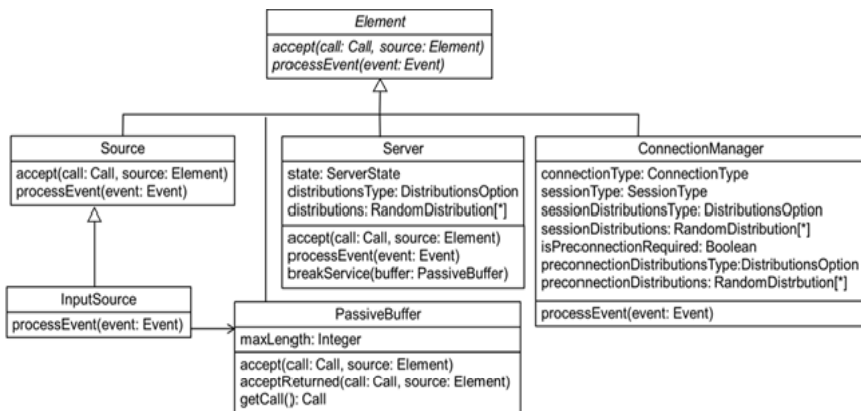


Рис. 2. Объектная модель циклической системы обслуживания

Класс *Element* представляет собой абстрактный класс, потомками которого будут любые элементы системы (например, входящий поток, накопитель, прибор обслуживания), которые могут принимать заявки с

помощью операции `accept(...)`, а также генерировать и/или обрабатывать события. Обработка событий производится с помощью операции `processEvent(...)`.

Класс `Source` представляет входящий поток заявок. В методе `accept(...)` генерируется исключение, так как этот элемент не может принимать заявки. А в методе `processEvent(...)` генерируются новые заявки.

Класс `InputSource`, в отличие от `Source`, включает в себя накопитель заявок (объект стандартного класса `PassiveBuffer`). Метод `processEvent(...)` вызывается автоматически каркасом платформы, если обрабатываемое системой событие сгенерировано данным объектом `InputSource`. В дальнейшем объекты класса `InputSource` будем называть источниками.

Класс `PassiveBuffer` представляет собой реализацию объекта-накопителя заявок. Он имеет атрибут `maxLength`, который определяет максимальный объем накопителя (максимальное число заявок, которые можно в нем разместить). Значение `maxLength = -1` означает, что накопитель не ограничен в размерах. Методы `accept(...)` и `acceptReturned(...)` используются для помещения заявки в накопитель. Если при вызове метода `accept(...)` оказалось, что накопитель заполнен до предела, то заявка получает отказ и немедленно покидает систему. В случае неограниченного объема накопителя или при вызове метода `acceptReturned(...)` заявка будет помещена в накопитель в обязательном порядке. В этом случае считается, что накопитель всегда имеет небольшой резерв для размещения заявок, возвращенных в него в экстренных ситуациях (прекращение обслуживания). Операция `getCall(...)` извлекает очередную заявку из накопителя и возвращает ее в качестве результата.

Объект `Server` реализует обслуживающий прибор и имеет два существенных отличия по сравнению с соответствующим объектом платформы, а именно:

- 1) может пребывать в трех состояниях (поле `state`) – «Свободен», «Занят обслуживанием», «В режиме переключения»;

- 2) реализует возможность задания отдельных функций распределения для длительности обслуживания заявок из разных накопителей (входящих потоков). Для этого используется опция `distributionsType` со значениями `Single` (одна функция распределения для всех накопителей) и `ForEachBuffer` (функции распределения вероятностей заданы для каждого накопителя в отдельности), а также массив `distributions`.

Методы `accept(...)` и `processEvent(...)` используются соответственно для помещения заявки на обслуживание в данном приборе и для обработки события окончания обслуживания. Метод `breakService(...)` немедленно прерывает обслуживание заявки и возвращает ее в соответствующий накопитель с помощью метода `PassiveBuffer.acceptReturned(...)`.

Класс `ConnectionManager` (диспетчер подключений) – искусственно введенный объект, который хранит информацию и реализует логику подключения приборов к очередям. Диспетчер подключений содержит такие атрибуты, как тип соединения `connectionType`, дисциплина обслуживания очередей `sessionType`, способ задания функций распределения длительности подключения `sessionDistributionsType`, функции распределения длительности подключения приборов к накопителям `sessionDistributions`, флаг требуемого времени на переключение приборов `isPreconnectionRequired`, способ задания функций распределения длительности времени переключения `preconnectionDistributionsType`, функции распределения длительности переключения приборов `preconnectionDistributions`. Метод `processEvent(...)` класса `ConnectionManager` обрабатывает событие окончания времени переключения прибора. Другие методы данного класса будут описаны ниже.

Механизм сеансов подключения

Для организации моделирования подключений приборов к накопителям предлагается использовать механизм сеансов подключения. Сеанс подключения (объект `Session`) – это объект, создаваемый на время подключения пула приборов к одному накопителю (рис. 3). Данный объект отвечает за продолжительность подключения, управляет движением поступающих заявок, а также заявок, находящихся в накопителе и на обслуживании.

Объект `Session` реализует операцию `processEvent(...)` интерфейса `Element`, в ней обрабатывается событие окончания сеанса. Операции `start(...)` и `finish()` используются соответственно для старта и окончания сеанса. Метод `serviceEnded(...)` вызывается, когда обслуживающий прибор, находящийся в сеансе, закончил обслуживание и готов принять на обслуживание новую заявку. Операция `acceptCall(...)` используется объектами-источниками `InputSource` во время поступления заявки в систему для поиска сеанса со свободным прибором и передачи ему этой заявки. Атрибут `buffer` ссылается на объект-накопитель источника `source`.

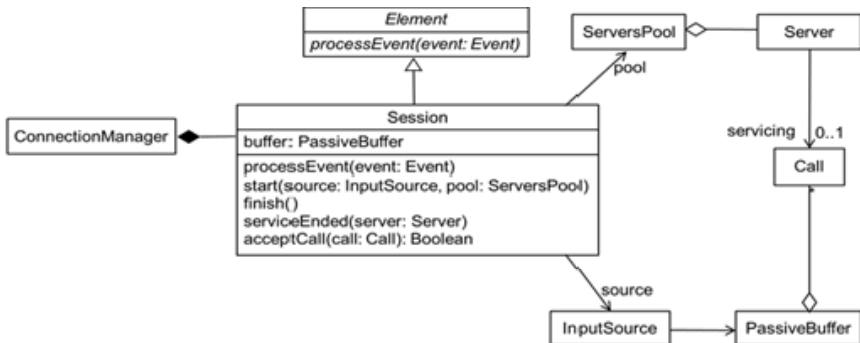


Рис. 3. Структурная схема механизма сеансов подключения

В объектной модели все серверы системы собраны в одну коллекцию – пул серверов (объект класса ServersPool). Все входящие потоки с накопителями также образуют коллекцию. У модели имеется единственный объект класса ConnectionManager, которому предоставлен полный доступ к обеим коллекциям. Управление сеансами полностью осуществляет объект ConnectionManager.

Основные сценарии дискретно-событийного моделирования

Так как вся система построена на дискретно-событийной модели управления, то для выполнения полного цикла имитационного моделирования достаточно выделить типы событий системы, реализовать механизмы их генерации, записи в журнал событий, извлечения из него и процедуры обработки этих событий. Ниже перечислены основные типы выделенных событий:

1. Начало моделирования. Происходит однократно на старте всего процесса моделирования. Не требует записи в журнал.
2. Поступление заявки в систему в определенном входящем потоке.
3. Окончание времени переключения прибора или группы приборов (если задано, что переключение требует времени).
4. Окончание обслуживания заявки на приборе.
5. Завершение сеанса (для дисциплины подключения «С разделением времени»).

Подробные сценарии и реализация обработчиков указанных событий будут выполнены в дальнейших работах.

Заключение

Представлено достаточно общее описание модели циклических систем обслуживания. Разработаны основные классы программного приложения для имитационного моделирования циклических систем. Проектируемое приложение будет иметь возможность моделировать широкий класс циклических систем обслуживания за счет вариативности задаваемых параметров конфигурации: различные типы входящих потоков и их параметры для каждого источника, индивидуальные параметры соответствующих им накопителей, индивидуальные параметры для каждого обслуживающего прибора, в том числе и в зависимости от обслуживаемого источника, учет возможных затрат времени на переключение приборов от одного источника к другому и так далее. В ближайшее время планируется разработка подробных сценариев основных событий моделирования и реализация их обработчиков.

ЛИТЕРАТУРА

1. Вишневецкий В.М., Семенова О.В. Математические методы исследования систем поллинга // Автоматика и телемеханика. – 2006. – № 2. – С. 3–56.
2. Лоу А., Кельтон В. Имитационное моделирование. – 3-е изд. – СПб.: Питер, 2004. – 848 с.
3. Моисеев А.Н. Программная система имитационного моделирования сетей массового обслуживания // Хроники объединенного фонда электронных ресурсов «Наука и образование». – 2015. – № 11 (78). – С. 34.
4. Мещеряков Р.В., Моисеев А.Н., Демин А.Ю. и др. Применение параллельных вычислений в имитационном моделировании сетей массового обслуживания // Известия Томского политехнического университета. – 2014. – Т. 325. – № 5. – С. 99–109.

ИССЛЕДОВАНИЕ ПОТОКОВ В ДВУХФАЗНОЙ БЕСКОНЕЧНОЛИНЕЙНОЙ СМО С ПОВТОРНЫМИ ОБРАЩЕНИЯМИ МЕТОДОМ ПРЕДЕЛЬНОЙ ДЕКОМПОЗИЦИИ

С.П. Моисеева, М.А. Шкленник, О.О. Набокова

Томский государственный университет, г. Томск, Россия

Задачи исследования потоков в системах массового обслуживания (СМО) с неограниченным числом обслуживающих приборов и произвольным временем обслуживания рассматривались в работах [1, 2], в которых предложен метод предельной декомпозиции. Данный метод работает только для систем с простейшим входящим потоком, но, как показывает статистический анализ реальных экономических систем [3], такие потоки встречаются довольно часто. В настоящей работе предлагается метод предельной декомпозиции для анализа двухфазной СМО с неограниченным числом обслуживающих приборов и возможностью повторного обращения заявки на каждой фазе.

Математическая модель

Рассмотрим двухфазную систему $M|GI^{(2)}|_{\infty}$ (рис. 1), на вход которой поступает простейший поток с параметром λ .

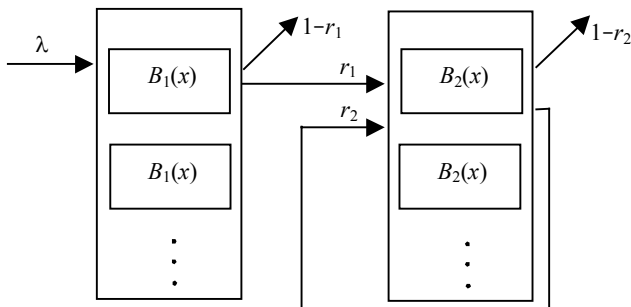


Рис. 1. Неоднородная двухфазная СМО $M|GI^{(2)}|_{\infty}$
с повторными обращениями на второй фазе

Время обслуживания заявок на первой фазе задано произвольной функцией распределения $B_1(x)$, одинаковой для каждого прибора фазы. По истечении этого времени заявка переходит на вторую фазу с вероятностью r_1 или покидает систему с вероятностью $(1 - r_1)$. Время обслуживания заявки на каждом из приборов второй фазы имеет также произвольную функцию распределения $B_2(x)$. После обслуживания на второй фазе заявка может вернуться для повторного обслуживания на вторую фазу с вероятностью r_2 или же может покинуть систему с вероятностью $(1 - r_2)$.

Ставится задача исследования суммарного потока обращений к системе.

Метод предельной декомпозиции

Для исследования потоков в системах массового обслуживания с входящим пуассоновским потоком, неограниченным числом обслуживающих устройств и произвольно заданным временем обслуживания заявок воспользуемся методом предельной декомпозиции [1].

Входящий поток разделим по полиномиальной схеме с равными вероятностями на N независимых пуассоновских потоков с интенсивностями λ/N [4]. Для заявок каждого из этих потоков определяется единственная линия обслуживания. Таким образом, получаем совокупность N независимых однолинейных двухфазных СМО с отказами (рис. 2). Линия считается занятой, если занята одна из ее фаз обслуживания.

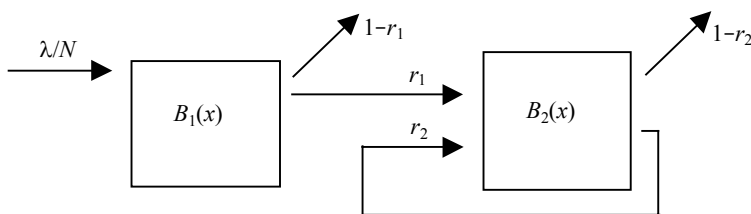


Рис. 2. Однолинейная двухфазная СМО с отказами

При $N \rightarrow \infty$ вероятностью потерь заявок можно пренебречь, и тогда суммарные характеристики совокупности N однолинейных двухфазных СМО сходятся к характеристикам исходной модели.

Исследование однолинейной двухфазной СМО с отказами и повторными обращениями

В указанной однолинейной системе будем рассматривать соответствующий случайный процесс $n(t, N)$ – суммарное число обращений, реализованных за время t . Поскольку процесс $n(t, N)$ не является марковским, введем в рассмотрение дополнительные переменные: процесс $k(t)$ – состояние линии обслуживания, то есть если k -я фаза линии занята, то $k(t) = k$, $k = 1, 2$, и если линия свободна, то $k(t) = 0$; $z(t)$ – длина интервала времени от момента t до момента окончания текущего обслуживания, если линия занята. Трехмерный процесс $\{k(t), n(t, N), z(t)\}$ является марковским.

Обозначим распределение вероятностей следующим образом:

$P_0(n, t, N) = P\{k(t) = 0, n(t, N) = n\}$ – вероятность того, что в момент времени t линия свободна и за это время к системе обратилось n заявок;

$P_k(n, t, N) = P\{k(t) = k, n(t, N) = n, z(t) < z\}$ – вероятность того, что в момент времени t в системе суммарное число обращений равно n , занята k -я фаза и до конца обслуживания остается времени меньше z .

Для распределения вероятностей трехмерного марковского процесса $\{k(t), n(t, N), z(t)\}$ составим прямую систему дифференциальных уравнений Колмогорова:

$$\frac{\partial P_0(n, t, N)}{\partial t} = \sum_{k=1}^2 (1 - r_k) \frac{\partial P_k(n, 0, t, N)}{\partial z} - \frac{\lambda}{N} P_0(n, t, N); \quad (1)$$

$$\frac{\partial P_1(n, z, t, N)}{\partial t} = \frac{\partial P_1(n, z, t, N)}{\partial z} - \frac{\partial P_1(n, 0, t, N)}{\partial z} + \frac{\lambda}{N} P_0(n - 1, t, N) B_1(z); \quad (2)$$

$$\begin{aligned} \frac{\partial P_2(n, z, t, N)}{\partial t} &= \frac{\partial P_2(n, z, t, N)}{\partial z} - \frac{\partial P_2(n, 0, t, N)}{\partial z} + \\ &+ r_1 B_2(z) \frac{\partial P_1(n - 1, 0, t, N)}{\partial z} + r_2 B_2(z) \frac{\partial P_2(n - 1, 0, t, N)}{\partial z}; \end{aligned} \quad (3)$$

с начальными условиями

$$\begin{aligned} P_0(n, 0, N) &= \begin{cases} 0, & n > 0, \\ R_0(N), & n = 0, \end{cases} \\ P_k(n, z, 0, N) &= \begin{cases} 0, & n > 0, \\ R_k(z, N), & n = 0. \end{cases} \\ R_0(N) + R_1(\infty, N) + R_2(\infty, N) &= 1. \end{aligned}$$

Здесь $R_0(N)$, $R_1(z, N)$, $R_2(z, N)$ – стационарное распределение вероятностей занятости фаз в однолинейной СМО.

Рассмотрим производящие функции:

$$\sum_{n=0}^{\infty} x^n P_0(n, t, N) = H_0(x, t, N); \quad (4)$$

$$\sum_{n=0}^{\infty} x^n P_k(n, z, t, N) = H_k(x, z, t, N), \quad k = 1, 2. \quad (5)$$

Тогда система (1) – (3) удовлетворяет системе дифференциальных уравнений в частных производных первого порядка

$$\begin{aligned} \frac{\partial H_0(x, t, N)}{\partial t} &= \sum_{k=1}^2 (1 - r_k) \frac{\partial H_k(x, 0, t, N)}{\partial z} - \frac{\lambda}{N} H_0(x, t, N), \\ \frac{\partial H_1(x, z, t, N)}{\partial t} - \frac{\partial H_1(x, z, t, N)}{\partial z} &= - \frac{\partial H_1(x, 0, t, N)}{\partial z} + H_0(x, t, N) x \frac{\lambda}{N} B_1(z), \\ \frac{\partial H_2(x, z, t, N)}{\partial t} - \frac{\partial H_2(x, z, t, N)}{\partial z} &= (r_2 x B_2(z) - 1) \frac{\partial H_2(x, 0, t, N)}{\partial z} + \\ &+ r_1 x B_2(z) \frac{\partial H_1(x, 0, t, N)}{\partial z}, \end{aligned}$$

решение которой будем искать в виде

$$H_0(x, t, N) = 1 - \frac{1}{N} F_0(x, t) + o(N^{-2}); \quad (6)$$

$$H_k(x, z, t, N) = \frac{1}{N} F_k(x, z, t) + o(N^{-2}), \quad k = 1, 2. \quad (7)$$

Тогда уравнения для $F_0(x, t)$, $F_1(x, z, t)$ и $F_2(x, z, t)$ имеют вид

$$\frac{\partial F_0(x, t)}{\partial t} = \lambda - \sum_{k=1}^2 (1 - r_k) f_k(x, t); \quad (8)$$

$$\frac{\partial F_1(x, z, t)}{\partial t} = \frac{\partial F_1(x, z, t)}{\partial z} + \lambda x B_1(z) - f_1(x, t); \quad (9)$$

$$\frac{\partial F_2(x, z, t)}{\partial t} = \frac{\partial F_2(x, z, t)}{\partial z} + r_1 x B_2(z) f_1(x, t) + (r_2 x B_2(z) - 1) f_2(x, t), \quad (10)$$

где $f_k(x, t) = \partial F_k(x, 0, t) / \partial z$, $k = 1, 2$.

Вероятности $R_0, R_k(z)$, $k = 1, 2$, можно найти как решение системы (8) – (10) при $x = 1$ и условии нормировки, то есть

$$R_0 = \frac{\lambda(r_1 b_2 + (1 - r_2)b_1)}{(1 - r_2)},$$

$$R_1(z) = \lambda \int_0^z (1 - B_1(y)) dy, \quad R_2(z) = \frac{r_1 \lambda}{(1 - r_2)} \int_0^z (1 - B_2(y)) dy.$$

где b_1 и b_2 – математические ожидания времени обслуживания на первой и второй фазах соответственно.

Решение дифференциального уравнения (9) с учетом начальных условий имеет вид

$$F_1(x, z, t) = \lambda \int_0^{z+t} (1 - B_1(y)) dy + \int_0^t [\lambda x B_1(z + t - s) - f_1(x, s)] ds. \quad (11)$$

Дифференцируя это тождество по z в нуле, получаем выражение для нахождения неизвестной функции $f_1(x, t)$:

$$f_1(x, t) = \lambda(1 - B_1(t)) + \lambda x B_1(t). \quad (12)$$

Аналогично решению уравнения (9), найдем решение уравнения (10):

$$F_2(x, z, t) = \frac{r_1 \lambda}{(1 - r_2)} \int_0^{z+t} (1 - B_2(y)) dy + \int_0^t [r_1 x B_2(z + t - s) f_1(x, s) + (r_2 x B_2(z + t - s) - 1) f_2(x, s)] ds. \quad (13)$$

Дифференцируя полученное тождество (13) по z в нуле, получаем интегральное уравнение для нахождения неизвестной функции $f_2(x, t)$:

$$f_2(x, t) = \frac{r_1 \lambda}{(1 - r_2)} (1 - B_2(t)) + x r_1 \int_0^t b_2(t - s) f_1(x, s) ds + x r_2 \int_0^t b_2(t - s) f_2(x, s) ds, \quad (14)$$

где $f_1(x, t)$ определяется выражением (12), а $b_2(t)$ – плотность распределения времени обслуживания на второй фазе.

Решение интегрального уравнения относительно $f_2(x, t)$ можно получить, используя преобразование Фурье – Стильтьеса вида

$$\int_0^{\infty} \frac{\partial f_2(x, t)}{\partial t} e^{j\alpha t} dt = \int_0^{\infty} e^{j\alpha t} d_t f_2(x, t) = \varphi_2(\alpha, x), \quad \int_0^{\infty} e^{j\alpha t} b_2(t) dt = B_2^*(\alpha)$$

Нетрудно показать, что

$$\varphi_2(\alpha, x) = \left[\frac{\lambda r_1 (x-1)}{(1-r_2)} + x r_1 \varphi_1(\alpha, x) \right] \frac{B_2^*(\alpha)}{1 - x r_2 B_2^*(\alpha)}.$$

Следовательно, решение уравнения (8) имеет вид

$$F_0(x, t) = \frac{\lambda(r_1 b_2 + (1-r_2)b_1)}{(1-r_2)} + \lambda t + (r_1 - 1) \int_0^t f_1(x, s) ds + (r_2 - 1) \int_0^t f_2(x, s) ds, \quad (15)$$

где $f_1(x, t)$ и $f_2(x, t)$ уже известные нам функции.

При $z \rightarrow \infty$ выражения (11) и (13) приобретают вид

$$F_1(x, t) = \lambda b_1 + \lambda x t - \int_0^t f_1(x, s) ds, \quad (16)$$

$$F_2(x, t) = \frac{r_1 \lambda b_2}{(1-r_2)} + r_1 x \int_0^t f_1(x, s) ds + (r_2 x - 1) \int_0^t f_2(x, s) ds. \quad (17)$$

Тогда, суммируя (15) – (17), получаем

$$F(x, t) = (x-1) \left\{ \lambda t + r_1 \int_0^t f_1(x, s) ds + r_2 \int_0^t f_2(x, s) ds \right\}.$$

Исследование потока в двухфазной неоднородной СМО с неограниченным числом обслуживающих приборов

Обозначим $n_i(t, N)$ – суммарное число обращений заявок в i -й однолинейной системе. Эти величины стохастически независимы и одинаково распределены. Их производящая функция $H(x, t, N)$ имеет вид

$$H(x, t, N) = H_0(x, t, N) + H_1(x, \infty, t, N) + H_2(x, \infty, t, N).$$

Следовательно, для производящей функции $G(x, t)$ суммарного числа $n(t)$ обращений в исходной системе с неограниченным числом линий можно записать

$$G(x, t) = \lim_{N \rightarrow \infty} Mx^{\sum_{i=1}^N n_i(t, N)} = \lim_{N \rightarrow \infty} (H(x, t, N))^N.$$

Учитывая (6), (7), получаем

$$G(x, t) = \lim_{N \rightarrow \infty} \left(1 + \frac{1}{N} F(x, t) + o\left(\frac{1}{N}\right) \right)^N = \exp\{F(x, t)\}.$$

Подставляя (23), окончательно получаем выражение для нахождения производящей функции $G(x, t)$:

$$G(x, t) = \exp \left\{ (x-1) \left\{ \lambda t + r_1 \int_0^t f_1(x, s) ds + r_2 \int_0^t f_2(x, s) ds \right\} \right\}.$$

Заключение

Таким образом, в работе построена математическая модель суммарного потока обращений в неоднородной двухфазной системе массового обслуживания с неограниченным числом обслуживающих приборов и повторным обращением на второй фазе. Методом предельной декомпозиции исследован суммарный поток обращений в рассматриваемой системе массового обслуживания. Получено выражения для производящей функции исследуемого потока.

Полученные результаты могут быть использованы при проведении анализа потоков различных социально-экономических систем, где наблюдаются повторные обращения заявок к системе, например в страховых и торговых компаниях.

ЛИТЕРАТУРА

1. Назаров А.А., Моисеева С.П., Морозова А.С. Исследование СМО с повторным обращением и неограниченным числом обслуживающих приборов методом предельной декомпозиции // Вычислительные технологии. – 2008. – Т. 13. – № 55. – С. 88 – 92.
2. Моисеева С.П., Ананина И.А., Назаров А.А. Исследование потоков в системе M/GI/∞ с повторными обращениями // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. – 2009. – № 3(8). – С. 56 – 66.
3. Brown L., Gans N., Mandelbaum A., et al. Statistical Analysis of a Telephone Call Center // A Queueing-Science Perspective. J. American Statistical Association. – 2005. – V. 100. – P. 36–50.
4. Хинчин А.Я. Работы по математической теории массового обслуживания. – М.: Физмат-лит. – 1963. – 236 с

ИССЛЕДОВАНИЕ НЕОДНОРОДНОЙ БЕСКОНЕЧНОЛИНЕЙНОЙ СМО $SM|M^{(n)}|_{\infty}$

С.П. Моисеева¹, Е.В. Панкратова², М.П. Фархадов²

¹*Томский государственный университет, г. Томск, Россия*

²*Институт проблем управления им. В.А. Трапезникова РАН,
г. Москва, Россия*

В условиях постоянного роста требований к эффективности устройств, применяемых в системах передачи и обработки информации, к сокращению сроков исследования и разработки новых телекоммуникационных систем и сетей актуально их исследование с помощью построения математических моделей. Особенностью выбранной для исследования математической модели является то, что сигналы в систему поступают поодиночке, но они имеют абсолютно разную природу и соответственно требуют совершенно разного обслуживания. Такие ситуации моделируются с помощью неоднородных СМО с ординарными входящими потоками разнотипных заявок.

Постановка задачи

Рассмотрим СМО $SM|M^{(n)}|_{\infty}$ с n типами неоднородных (в смысле скорости обслуживания) обслуживающих приборов, на вход которой поступает SM-поток разнотипных заявок. Входящий SM-поток (Semi Markovian Process) задан полумарковской матрицей $A(x)$, состоящей из элементов $A_{k_1 k_2}(x)$ ($k_1 = 0, \dots, N, k_2 = 0, \dots, N$),

$$A_{k_1 k_2}(x) = F_2(k_2, x; k_1) = P(\xi(k+1) = k_2, \tau(k+1) < x | \xi(k) = k_1), \quad (1)$$

где $\xi(k)$ эргодическая цепь Маркова с дискретным временем и матрицей $P = [p_{vk}]$ вероятностей перехода за один шаг, процесс $\tau(k)$ принимает неотрицательные значения из непрерывного множества и определяет длины интервалов в SM-потоке. Дисциплину обслуживания определим следующим образом. Заявки, поступившие в систему, с вероят-

ностью p_i ($i = 1, \dots, n, \sum_{i=1}^n p_i = 1$) обслуживаются в течение случайного времени, распределенного по экспоненциальному закону с параметром μ_i ($i = 1, \dots, n$), соответствующим типу заявки.

Поставим задачу исследования n -мерного случайного процесса $\{l_1(t), \dots, l_n(t)\}$, где $l_i(t)$ – число занятых приборов i -го типа в системе в момент времени t . Очевидно, что исследуемый случайный процесс не является марковским. Воспользуемся методом «внешнего» марковизирования [1, 2], который заключается во введении дополнительных переменных таким образом, чтобы исследуемый многомерный случайный процесс стал марковским. Рассмотрим $(n+2)$ -мерный марковский случайный процесс $\{s(t), z(t), l_1(t), \dots, l_n(t)\}$, где $z(t)$ – длина интервала от момента времени t до момента наступления очередного события в СМ-потоке, а дискретный процесс $s(t)$ определяется следующим образом: $s(t) = \xi(k+1)$, если $t_k < t \leq t_{k+1}$, где моменты восстановления t_k определяются равенством $t_k = \sum_{i=1}^k \tau(i)$.

Для совместного распределения вероятностей

$$P(s, z, l_1, \dots, l_n, t) = P\{s(t) = s, z(t) < z, l_1(t) = l_1, \dots, l_n(t) = l_n\}$$

можно записать систему дифференциальных уравнений Колмогорова

$$\begin{aligned} \frac{\partial P(s, z, l_1, \dots, l_n, t)}{\partial t} = & \frac{\partial P(s, z, l_1, \dots, l_n, t)}{\partial z} - \frac{\partial P(s, 0, l_1, \dots, l_n, t)}{\partial z} - \\ & - \sum_{i=1}^n l_i \mu_i P(s, z, l_1, \dots, l_n, t) + \\ & + p_1 \sum_{v=1}^K \frac{\partial P(v, 0, l_1 - 1, \dots, l_n, t)}{\partial z} A_{vs} + \dots + p_n \sum_{v=1}^K \frac{\partial P(v, 0, l_1, \dots, l_n - 1, t)}{\partial z} A_{vs} + \\ & + \mu_1 (l_1 + 1) P(s, z, l_1 + 1, \dots, l_n, t) + \dots + \mu_n (l_n + 1) P(s, z, l_1, \dots, l_n + 1, t), s = 1, \dots, K. \end{aligned} \quad (2)$$

Решение системы (2) будем искать при стационарном режиме функционирования рассматриваемой системы. Обозначим

$$\lim_{t \rightarrow \infty} P(s, z, l_1, \dots, l_n, t) = \Pi(s, z, l_1, \dots, l_n), s = 1, \dots, K.$$

Введем частичные характеристические функции вида

$$H(s, z, u_1, \dots, u_n) = \sum_{l_1=0}^{\infty} \dots \sum_{l_n=0}^{\infty} e^{ju_1 l_1} \times \dots \times e^{ju_n l_n} \Pi(s, z, l_1, \dots, l_n), \quad s = 1, \dots, K,$$

$j = \sqrt{-1}$ – мнимая единица.

Для функций $H(s, z, u_1, \dots, u_n)$ получим систему уравнений

$$\begin{aligned} \frac{\partial H(s, z, u_1, \dots, u_n)}{\partial z} - \frac{\partial H(s, 0, u_1, \dots, u_n)}{\partial z} + j \sum_{i=1}^n \mu_i (1 - e^{-ju_i}) \frac{\partial H(s, z, u_1, \dots, u_n)}{\partial u_i} + \\ + \sum_{i=1}^n p_i e^{ju_i} \sum_v \frac{\partial H(v, 0, u_1, \dots, u_n)}{\partial z} A_{vs}(z) = 0, \quad s = 1, \dots, K, \end{aligned} \quad (3)$$

которую перепишем в виде векторно-матричного уравнения

$$\begin{aligned} \frac{\partial \mathbf{H}(z, u_1, \dots, u_n)}{\partial z} + j \sum_{i=1}^n \mu_i (1 - e^{-ju_i}) \frac{\partial \mathbf{H}(z, u_1, \dots, u_n)}{\partial u_i} + \\ + \frac{\partial \mathbf{H}(0, u_1, \dots, u_n)}{\partial z} \left\{ \sum_{i=1}^n p_i e^{ju_i} \mathbf{A}(z) - \mathbf{I} \right\} = 0, \end{aligned} \quad (4)$$

решение $\mathbf{H}(z, u_1, \dots, u_n)$ которого, удовлетворяющее условию $\mathbf{H}(z, 0, \dots, 0) = \mathbf{r}(z)$, определяет характеристическую функцию числа занятых приборов в системе $SM | M^{(n)} | \infty$, функционирующей в стационарном режиме, равенством $Me^{j \sum_{i=1}^n u_i l_i(t)} = \mathbf{H}(\infty, u_1, \dots, u_n) \mathbf{e}$.

Здесь $\mathbf{r}(z) = \kappa_1 \mathbf{r} \int_0^z (\mathbf{P} - \mathbf{A}(x)) dx$ – стационарное распределение вероятностей значений двумерного случайного процесса $\{s(t), z(t)\}$, \mathbf{r} – стационарное распределение вероятностей значений вложенной цепи Маркова, величина κ_1 определяется равенством $\kappa_1 = \frac{1}{\mathbf{rAe}}$, где матрица

\mathbf{A} определяется равенством $\mathbf{A} = \int_0^{\infty} (\mathbf{P} - \mathbf{A}(x)) dx$.

Векторно-матричное уравнение (4) является основным для дальнейших исследований.

Вероятностные характеристики числа занятых приборов в системе $SM|M^{(n)}|\infty$

Теорема 1. Для начальных моментов числа занятых приборов каждого типа при стационарном функционировании неоднородной системы $SM|M^{(n)}|\infty$ верны следующие утверждения:

Утверждение 1. Среднее значение числа занятых приборов i -го типа $fm_i (i = 1, \dots, n)$ в неоднородной системе $SM|M^{(n)}|\infty$ имеет вид

$$fm_i = \frac{p_i}{\mu_i} \lambda, \quad (5)$$

где $\lambda = \mathbf{r}'(0)\mathbf{e}$, $\mathbf{e} = [1, \dots, 1]^T$ – единичный вектор-столбец.

Утверждение 2. Начальный момент второго порядка числа занятых приборов i -ого типа $sm_i (i = 1, \dots, n)$ в неоднородной системе $SM|M^{(n)}|\infty$ имеет вид

$$sm_i = \frac{p_i}{\mu_i} \left[\lambda + p_i \mathbf{r}'(0) \mathbf{A}^*(\mu_i) (\mathbf{I} - \mathbf{A}^*(\mu_i))^{-1} \mathbf{e} \right], \quad (6)$$

где $\mathbf{A}^*(\alpha) = \int_0^\infty e^{-\alpha z} d\mathbf{A}(z)$, \mathbf{I} – единичная матрица.

Утверждение 3. Корреляционный момент $cm_{ig} (i = 1, \dots, n, g = 1, \dots, n, i \neq g)$ числа занятых приборов типа i и g в неоднородной системе $SM|M^{(n)}|\infty$ имеет вид

$$cm_{ig} = \frac{p_i p_g}{\mu_i + \mu_g} \mathbf{r}'(0) \left[\mathbf{A}^*(\mu_i) (\mathbf{I} - \mathbf{A}^*(\mu_i))^{-1} + \mathbf{A}^*(\mu_g) (\mathbf{I} - \mathbf{A}^*(\mu_g))^{-1} \right] \mathbf{e}. \quad (7)$$

Доказательство. Из свойств характеристической функции

$$\begin{aligned} \frac{\partial \mathbf{H}(z, u_1, \dots, u_n)}{\partial u_i} \bigg|_{u_1=0, \dots, u_n=0} &= j \mathbf{f} \mathbf{m}_i(z), \\ \frac{\partial^2 \mathbf{H}(z, u_1, \dots, u_n)}{\partial u_i^2} \bigg|_{u_1=0, \dots, u_n=0} &= j^2 \mathbf{s} \mathbf{m}_i(z), \end{aligned} \quad (8)$$

$$\left. \frac{\partial^2 \mathbf{H}(z, u_1, \dots, u_n)}{\partial u_i \partial u_g} \right|_{u_1=0, \dots, u_n=0} = j^2 \mathbf{cm}_{ig}(z),$$

$$i = 1, \dots, n, \quad g = 1, \dots, n, \quad i \neq g.$$

Начальные моменты первого порядка

Продифференцируем уравнение (4) по переменной $u_i, i = 1, \dots, n$, положим $u_1 = 0, \dots, u_n = 0$ и получим систему дифференциальных уравнений для нахождения вектор-строки $\mathbf{fm}_i(z)$

$$\mathbf{fm}'_i(z) - \mu_i \mathbf{fm}_i(z) + \mathbf{fm}'_i(0)(\mathbf{A}(z) - \mathbf{I}) + p_i \mathbf{r}'(0)\mathbf{A}(z) = 0, \quad i = 1, \dots, n. \quad (9)$$

Это уравнение будем решать методом преобразования Лапласа – Стильтеса, обозначив

$$\Phi_i(\alpha) = \int_0^\infty e^{-\alpha z} d\mathbf{fm}_i(z), \quad i = 1, \dots, n, \quad \mathbf{A}^*(\alpha) = \int_0^\infty e^{-\alpha z} d\mathbf{A}(z). \quad (10)$$

Выполнив в (9) преобразование Лапласа – Стильтеса, получим равенство

$$(\mu_i - \alpha)\Phi_i(\alpha) = \mathbf{fm}'_i(0)(\mathbf{A}^*(\alpha) - \mathbf{I}) + p_i \mathbf{r}'(0)\mathbf{A}^*(\alpha) = 0, \quad i = 1, \dots, n, \quad (11)$$

положив в котором $\alpha = \mu_i, i = 1, \dots, n$, найдём вид вектора $\mathbf{fm}'_i(0)$

$$\mathbf{fm}'_i(0) = p_i \mathbf{r}'(0)\mathbf{A}^*(\mu_i) \left(\mathbf{I} - \mathbf{A}^*(\mu_i) \right)^{-1}. \quad (12)$$

Подставив (12) в (11), получим равенство

$$\Phi_i(\alpha) = \frac{1}{\mu_i - \alpha} \left\{ \mathbf{fm}'_i(0)(\mathbf{A}^*(\alpha) - \mathbf{I}) + p_i \mathbf{r}'(0)\mathbf{A}^*(\alpha) \right\}, \quad i = 1, \dots, n, \quad (13)$$

определяющее преобразование Лапласа – Стильтеса вектор-функции $\mathbf{fm}_i(z)$.

Так как $\mathbf{fm}'_i(\infty) = \Phi_i(0)$ и $\mathbf{A}^*(\infty) = \mathbf{P}$, то для момента первого порядка числа занятых приборов разного типа в системе $\text{SM} | M^{(n)} | \infty$ можно записать

$$fm_i = \mathbf{fm}_i(\infty)\mathbf{e} = \frac{p_i}{\mu_i} \mathbf{r}'(0)\mathbf{e}, \quad i = 1, \dots, n. \quad (14)$$

Начальные моменты второго порядка

Для нахождения момента второго порядка числа занятых приборов, продифференцируем дважды по $u_i, i = 1, \dots, n$, равенство (4) и с учетом (8) получим систему дифференциальных уравнений

$$\mathbf{sm}'_i(z) - \mu_i \mathbf{fm}_i(z) - 2\mu_i \mathbf{sm}_i(z) + \mathbf{sm}'_i(0)(\mathbf{A}(z) - \mathbf{I}) + p_i \{\mathbf{fm}'_i(0) + \mathbf{r}'(0)\} \mathbf{A}(z) = 0,$$

выполнив в которой преобразование Лапласа – Стильбеса, имеем равенство

$$\begin{aligned} \Psi_i(\alpha) &= \frac{1}{2\mu_i - \alpha} \left\{ \mu_i \Phi_i(\alpha) + \mathbf{sm}'_i(0)(\mathbf{A}^*(\alpha) - \mathbf{I}) + p_i [2\mathbf{fm}'_i(0) + \mathbf{r}'(0)] \mathbf{A}^*(\alpha) \right\}, \\ \Psi_i(\alpha) &= \int_0^\infty e^{-\alpha z} d\mathbf{sm}_i(z), \quad i = 1, \dots, n, \quad \mathbf{A}^*(\alpha) = \int_0^\infty e^{-\alpha z} d\mathbf{A}(z), \end{aligned} \quad (15)$$

откуда можно записать равенство, определяющее второй момент числа занятых приборов разного типа в системе $\text{SM} | M^{(n)} | \infty$

$$sm_i = \mathbf{sm}_i(\infty)\mathbf{e} = \Psi_i(0)\mathbf{e} = \frac{p_i}{\mu_i} \left[\lambda + p_i \mathbf{r}'(0) \mathbf{A}^*(\mu_i)(\mathbf{I} - \mathbf{A}^*(\mu_i))^{-1} \mathbf{e} \right], \quad i = 1, \dots, n.$$

Корреляционный момент

Продифференцировав (4) по u_i , и по $u_g, i = 1, \dots, n, g = 1, \dots, n, i \neq g$, получим систему уравнений для вектор-функции $\mathbf{cm}_{ig}(z)$:

$$\begin{aligned} \mathbf{cm}'_{ig}(z) - (\mu_i + \mu_g) \mathbf{cm}_{ig}(z) + \mathbf{cm}'_{ig}(0)(\mathbf{A}(z) - \mathbf{I}) + \\ + \{p_g \mathbf{fm}'_i(0) + p_i \mathbf{fm}'_g(0)\} \mathbf{A}(z) = 0, \end{aligned}$$

применив к полученному уравнению преобразование Лапласа – Стильбеса

$$\text{еса вида } \Theta_{ig} = \int_0^\infty e^{-\alpha z} d\mathbf{cm}_{ig}(z) \text{ и положив } \alpha = \mu_i + \mu_g, \quad i, g = 1, \dots, n, \quad i \neq g,$$

получим выражение для корреляционного момента

$$cm_{ig} = \frac{p_i p_g}{\mu_i + \mu_g} \mathbf{r}'(0) [\mathbf{A}^*(\mu_i)(\mathbf{I} - \mathbf{A}^*(\mu_i))^{-1} + \mathbf{A}^*(\mu_g)(\mathbf{I} - \mathbf{A}^*(\mu_g))^{-1}] \mathbf{e}.$$

Теорема доказана.

Для анализа зависимости между компонентами вектора $\{l_1(t), \dots, l_n(t)\}$ запишем выражение для коэффициента корреляции

$$r_{ig} = \frac{\text{cov}_{ig}}{\sqrt{\text{Var}_i \text{Var}_g}} = \frac{cm_{ig} - fm_i fm_g}{\sqrt{\text{Var}_i \text{Var}_g}}, \quad i = 1, \dots, n, \quad g = 1, \dots, n, \quad i \neq g,$$

где Var_i – дисперсия числа занятых приборов типа $i, i = 1, \dots, n$.

ЛИТЕРАТУРА

1. Мусеев А.Н., Назаров А.А. Бесконечнолинейные системы и сети массового обслуживания. – Томск: Изд-во НТЛ, 2015. – 240 с.
2. Рыков В.В. Управляемые системы массового обслуживания // Итоги науки и техн. Сер. Теор. вероятн. Мат. стат. Теор. кибернет. – М: ВИНТИ, 1975. – Вып. 12. – С. 43–153.

УСКОРЕННОЕ ОЦЕНИВАНИЕ ХАРАКТЕРИСТИК СИСТЕМЫ С ЗАЯВКАМИ СЛУЧАЙНОГО ОБЪЕМА*

Е.В. Морозов^{1,2}, Л.В. Потахина¹

¹Институт прикладных математических исследований

Карельского научного центра РАН, г. Петрозаводск, Россия

²Петрозаводский государственный университет, г. Петрозаводск, Россия

Современные задачи, связанные с высокотехнологичными телекоммуникационными системами, могут быть описаны моделями, в которых заявки описываются двумя характеристиками: *случайным временем обслуживания* и *случайным объемом*, при этом буфер для суммарного объема ожидающих заявок является ограниченным [1, 2]. В работе [3] была предложена аналогия между объемом теряемой заявки и накрывающим интервалом соответствующего процесса восстановления, позволяющая получить полезную аппроксимацию моментов теряемого объема на основании *парадокса времени восстановления*. Но, в случае, когда загрузка системы невелика или буфер для накопленного объема достаточно велик, моделирование методом Монте-Карло требует значительного времени для получения оценок, поскольку потеря заявки является редким событием.

Для решения этой проблемы в данной работе мы применяем *метод расщепления*, позволяющий участить наблюдение редкого события. Данный метод основан на идее создания нескольких стохастических копий процесса, лежащего в основе динамики системы, при пересечении заданных порогов. Мы моделируем динамику системы для разных значений загрузки, различных распределений объемов заявок и величины буфера для накопленного объема.

Описание модели и связь с теорией восстановления

Рассмотрим систему типа $GI/G/1$, в которой каждая заявка имеет случайное время обслуживания и случайный объем. Предполагается, что времена обслуживания $\{S_n\}$, $n \geq 0$, и объемы $\{v_n\}$, $n \geq 0$, являются

* Исследование выполнено при финансовой поддержке РФФИ в рамках научных проектов № 15-07-02341, 15-07-02354, 15-07-02360.

независимыми одинаково распределенными (н.о.р.) случайными величинами (с.в.) с типичными элементами S и ν соответственно. Будем считать, что пары $\{S_n, \nu_n\}$ являются также н.о.р., но для фиксированного n величины S_n и ν_n могут быть зависимы. Обозначим через t_n момент поступления в систему заявки n . Интервалы между поступлениями заявок в систему $\tau_n = t_{n+1} - t_n$ являются н.о.р. с типичным элементом τ . Определим величину загрузки системы как $\rho = ES / E\tau$.

Определим через $V(t)$ величину накопленного объема в системе в момент времени t , который равен сумме объемов всех ожидающих заявок. Будем считать, что буфер для числа заявок, находящихся в очереди в системе, является бесконечным, в то время как буфер для суммарного объема ожидающих заявок ограничен конечной константой M . Таким образом, заявка с номером n будет потеряна, если $V(t_n^-) + \nu_n > M$. Обозначим через $R(t)$ множество номеров заявок, потерянных в интервале $(0, t]$.

Рассмотрим следующую важную характеристику системы: средний теряемый объем заявки ν_{loss} , определяемую как предел

$$\nu_{\text{loss}} = \lim_{t \rightarrow \infty} \frac{\sum_{i \in R(t)} \nu_i}{|R(t)|}, \quad (1)$$

где $|R(t)|$ – мощность множества $R(t)$. С другой стороны, теряемый объем может быть проинтерпретирован как интервал, накрывающий уровень M , в процессе восстановления, порожденном н.о.р. величинами $\{\nu_k\}$ [3]. Определим случайные суммы $Z_k = \nu_1 + \dots + \nu_k$, $k \geq 1$ такие, что $0 \leq Z_1 < Z_2 < \dots$. Пусть F является функцией распределения объема заявки, $F(x) = P(\nu \leq x)$, $x \geq 0$ [4]. Обозначим через ν_t интервал, накрывающий момент t . Тогда из формулы полной вероятности для ν_t верно следующее:

$$\begin{aligned} P(\nu_t \leq x) &= \sum_{k=1}^{\infty} \int_{t-x}^x P(t-x < \nu \leq x) P(Z_k \in du) = \\ &= \int_0^t I(x > t-u) [F(x) - F(t-u)] dH(u), \end{aligned} \quad (2)$$

где $H(t)$ – функция восстановления, определенная для $t \geq 0$ как

$$H(t) := EN_t, \quad N_t = \sum_{n=1}^{\infty} I(Z_n \leq t). \quad (3)$$

Из вышеупомянутой аналогии теряемый объем в рассматриваемой системе совпадает с v_t при $t = M$, что влечет $v_{\text{loss}} \equiv Ev_M$. Это дает возможность применить асимптотические результаты теории восстановления для оценки Ev_M при достаточно большом (но конечном) буфере M .

Из ключевой теоремы восстановления [4] следует, что если g вещественнозначная ограниченная функция, функция восстановления F нерешетчатая и $Ev < \infty$, то при $t \rightarrow \infty$ верно

$$\int_0^t g(t-u)dH(u) \rightarrow \frac{1}{Ev} \int_0^\infty g(u)du. \quad (4)$$

Применяя это к равенству (2), получаем

$$\lim_{t \rightarrow \infty} P(v_t \leq x) = \frac{1}{Ev} \int_0^x [F(x) - F(u)]du = \frac{1}{Ev} \int_0^x uF(du), \quad (5)$$

где на последнем шаге было применено интегрирование по частям. Откуда получаем следующую аппроксимацию для момента порядка $p > 0$ теряемого объема v_M при достаточно большом M и $\rho < 1$:

$$E(v_M^p) \approx \frac{E(v^{p+1})}{Ev}. \quad (6)$$

В частности, для первых двух моментов теряемого объема верно следующее:

$$Ev_M \approx \frac{Ev^2}{Ev}, \quad Ev_M^2 \approx \frac{Ev^3}{Ev}. \quad (7)$$

Метод расщепления и моделирование

В данной работе мы иллюстрируем при помощи имитационного моделирования точность аппроксимации (7) в зависимости от величины буфера M , загрузки системы ρ для различных распределений объема заявок. Но, как уже было отмечено ранее, при большом значении буфера M потеря заявки становится редким событием, что делает сложной задачу получения значимой выборки в рамках приемлемого времени моделирования.

Для ускорения времени моделирования мы применяем метод расщепления [5, 6], в основе которого лежит идея копирования базового процесса, описывающего состояние системы, по достижении некоторых заданных порогов, что позволяет сделать редкое событие более частым.

В рамках постановки задачи рассмотрим равноотстоящие пороги $0 < x_1 < x_2 < \dots < x_K < M$. Когда процесс $V(t)$ пересекает порог x_i в первый раз, мы формируем R_i независимых копий базового процесса. Необходимо отметить, что согласно методу расщепления, базовый процесс должен быть марковским, иначе нельзя утверждать, что новая система будет стохастической копией оригинальной. Однако в процессе моделирования расщепление основывается только на величине основной компоненты $V(t)$.

Основным вопросом при реализации метода расщепления является выбор порогов x_i и множителей R_i оптимальным способом. Применим следующую схему, предложенную для ускорения моделирования редких событий процесса величины очереди системы М/М/1 с ограниченным буфером и дающую минимальную дисперсию величины оценки [7, 8]:

$$K = -\log P_{\text{loss}} / 2; R_i = e^2, i > 0; R_0 = 1, \quad (14)$$

где P_{loss} – вероятность потери заявки в системе, K – количество порогов, R_i – количество копий для порога x_i .

Моделирование проводилось с использованием среды для статистических вычислений “R” [9] с использованием высокопроизводительного кластера Карельского научного центра РАН [10].

Была рассмотрена система с экспоненциально распределенными временами обслуживания с параметром $\mu = 1$ и экспоненциально распределенными временами между поступлениями заявок с параметрами $\lambda \in [0,7; 2]$ с шагом 0,05, что позволило рассмотреть случаи сильно и слабо загруженной системы. В процессе моделирования построены оценки среднего теряемого и среднего принимаемого объема, а также оценки дисперсии теряемого объема. Эти данные позволили построить доверительный интервал для оценки среднего теряемого объема.

Приведем полученные результаты. Рисунки 1 и 2 соответствуют случаю объемов заявок, распределенных по закону Вейбулла с легким хвостом с функцией распределения

$$F(x) = 1 - e^{-x^2}, \quad x > 0.$$

Результаты для объемов заявок, распределенных по закону Парето с тяжелым хвостом с функцией распределения

$$F(x) = 1 - \left(\frac{1}{x}\right)^4, \quad x \geq 1,$$

($F(x) = 0$ для $x < 1$) показаны на рис. 3 и 4.

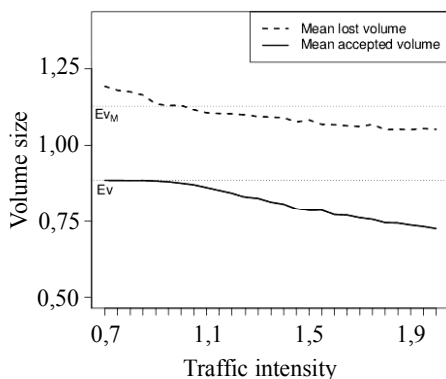


Рис. 1. Средний теряемый и средний принимаемый объем

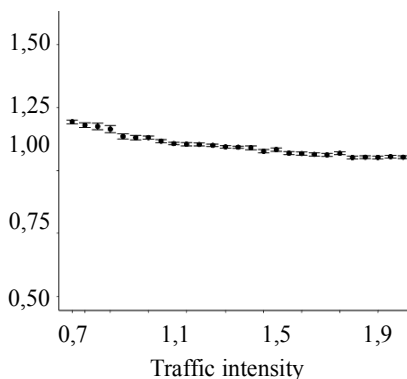


Рис. 2. Доверительный интервал для оценки среднего теряемого объема

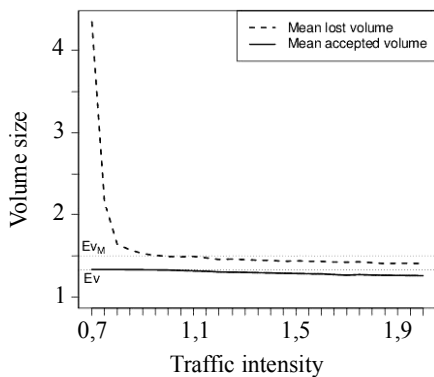


Рис. 3. Средний теряемый и средний принимаемый объем

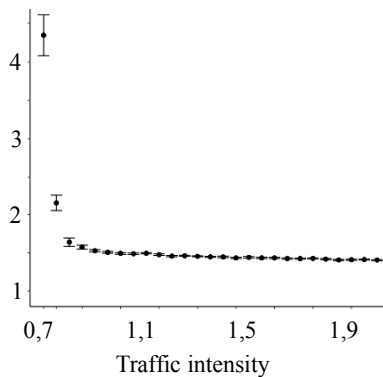


Рис. 4. Доверительный интервал для оценки среднего теряемого объема

Объемы заявок с распределением Вейбулла

Объемы заявок с распределением Парето

В частности, рис. 1 и 3 показывают оценки для среднего теряемого и среднего принимаемого объема в зависимости от значения параметра ρ . Теоретическое значение среднего объема $E\nu$ и значение аппроксимации $E\nu_M$ также приведены на рисунках. Очевидно, что, когда $\rho \in (1; 2]$, значительная часть заявок теряется, что ведет к сближению оценки $E\nu_M$ со средним объемом $E\nu$. Факт, что величина среднего принимаемого объема меньше $E\nu$, можно объяснить тем, что в основном заявки с малым объемом принимаются в систему. С другой стороны, когда $\rho \in [0, 7; 1]$, оценка принимаемого объема близка к $E\nu$, поскольку редкие потери почти не влияют на результирующее значение. Рост оценки теряемого объема можно объяснить тем, что теряются заявки с нетипично большим объемом. Особенно хорошо это выражено на рис. 3 в силу наличия тяжелого хвоста распределения объема [11]: если с.в. $\{X_i\}$ являются н.о.р. и субэкспоненциальными (этому классу принадлежат, например, распределения Парето и Вейбулла с тяжелым хвостом), то

$$P(X_1 + X_2 + \dots + X_n > x) \sim P(\max(X_1, X_2, \dots, X_n) > x), \quad x \rightarrow \infty.$$

То есть сумма с.в. превышает сколь угодно большое фиксированное число за счет одного нетипично большого слагаемого.

На рис. 2 и 4 изображены доверительные интервалы для оценок среднего теряемого объема (в зависимости от значения ρ) для уровня значимости 0,9. То, что доверительные интервалы при $\rho \in [0, 7; 1]$ больше, чем при $\rho \in (1; 2]$ на рис. 4, можно объяснить ростом дисперсии потеряннного объема при (редких) потерях заявок с тяжелохвостым распределением.

Как показывают эксперименты, аппроксимация (11) хорошо согласуется с результатами имитационного моделирования при значениях ρ близких к 1. Отклонения полученных оценок от аппроксимации в остальных случаях могут быть объяснены тем, что параметры метода расщепления не слишком подходят для рассмотренной постановки задачи. Аналогию с теорией восстановления следует применять с осторожностью в случае низкой загрузки (особенно в случае тяжелохвостых распределений), поскольку в общем случае процесс накопленного объема не является классическим процессом восстановления. В случае большой загрузки заявки с нетипично малым объемом имеют большую вероятность быть принятыми за счет того, что значения процесса $V(t)$ близки к M .

Заключение

Мы рассмотрели односерверную систему, в которой каждая заявка описывается не только временем обслуживания, но и случайным объемом. При этом суммарный объем заявок, находящихся в системе, ограничен конечной константой. Предложенная ранее аппроксимация для моментов потерянного объема была проиллюстрирована при помощи имитационного моделирования для систем с различными параметрами. В случае, когда потеря заявки является редким событием, для сокращения времени моделирования было использовано ускоренное моделирование, в частности метод расщепления. Численные результаты показывают, что аппроксимацию величины потерянного и принятого объема основе теории восстановления можно применять, однако это можно делать с осторожностью, принимая во внимание распределение объемов заявок и загруженность системы.

ЛИТЕРАТУРА

1. *Morozov E., Potakhina L., Tikhonenko O.* Regenerative analysis of a system with a random volume of customers // *Communications in Computer and Information Science*. – 2016. – V. 638. – P. 261–272.
2. *Tikhonenko O.M.* Queuing system with processor sharing and limited resources // *Automation and Remote Control*. – 2010. – V. 71. – No. 5. – P. 803–815.
3. *Morozov E., Nekrasova R., Potakhina L., Tikhonenko O.* Asymptotic analysis of queueing systems with finite buffer space // *Communications in Computer and Information Science*. – 2014. – V. 431. – P. 223–232.
4. *Asmussen S.* Applied Probability and Queues. 2nd Ed. – New York: Springer-Verlag, 2003.
5. *Botev Z.I., Kroese D.P.* Efficient Monte Carlo Simulation via The Generalized Splitting Method // *Statistics and Computing*. – 2012. – V. 22. – No. 1. – P. 1–16.
6. *Rubinstein R.Y., Ridder A., Vaisman R.* Fast Sequential Monte Carlo Methods for Counting and Optimization. – Wiley, 2014.
7. *Borodina A.V.* PhD Thesis “Regenerative modification of the splitting method for overload probability estimation in queueing systems”. – The Petrozavodsk State University, 2008. (in Russian)
8. *Garvels M.* PhD Thesis “The splitting method in rare event simulation. – The University of Twente, The Netherlands, 2000.
9. *R Foundation* for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. – URL: <http://www.R-project.org/>.
10. *Center* for collective use of Karelian Research Centre of Russian Academy of Science. – URL <http://cluster.krc.karelia.ru/>.
11. *Asmussen S., Klppelberg C., Sigman C.* Sampling at subexponential times, with queueing applications. *Stochastic Processes and their Applications*. – 1999. – V. 79. – No. 2. – P. 265–286.

АСИМПТОТИЧЕСКИЙ АНАЛИЗ ДВУХФАЗНОЙ RQ-СИСТЕМЫ С ДВУМЯ ОРБИТАМИ В УСЛОВИИ БОЛЬШОЙ ЗАДЕРЖКИ ЗАЯВОК НА ОРБИТАХ

А.А. Назаров, А.А. Анисимова

Томский государственный университет, г. Томск, Россия

С середины XX века все большую роль стали играть телекоммуникационные системы: компьютерные и телефонные сети, системы передачи данных, радио, телевидение, мобильная связь и другие. В отличие от классических систем массового обслуживания [1], для телекоммуникационных систем характерна ситуация, при которой заявка, заставшая обслуживающий прибор занятым, не встает в очередь, а уходит на орбиту, откуда через некоторые промежутки времени предпринимает попытки вновь обратиться за обслуживанием. Такие модели описываются в виде систем массового обслуживания с повторными вызовами или RQ-систем (Retrial queueing system), которые были подробно изучены в работах Дж.Р. Арталехо [2] и Г.И. Фалина [3].

В данной работе исследуется RQ-система $M|M|1$ с двумя фазами, каждая из которых содержит один обслуживающий прибор и орбиту. Впервые двухфазные RQ-системы рассмотрели С.М. Krishna и Y.H. Lee [4]. Также их анализом занимались В.Т. Doshi [5], В. Krishnakumar [6], G. Choudhury [7] и другие.

Математическая модель

Рассмотрим RQ-систему, на вход которой поступает простейший входящий поток интенсивности λ . Каждая заявка проходит последовательно 2 фазы. Продолжительность обслуживания на обеих фазах имеет экспоненциальное распределение с параметрами μ_1 и μ_2 соответственно. Если на очередной фазе заявка застает прибор занятым, она отправляется на соответствующую орбиту, откуда через случайные моменты времени, имеющие экспоненциальное распределение с параметрами σ_1 и σ_2 , делает попытки вновь обратиться за обслуживанием. После обслуживания на второй фазе заявка покидает систему (рис. 1).

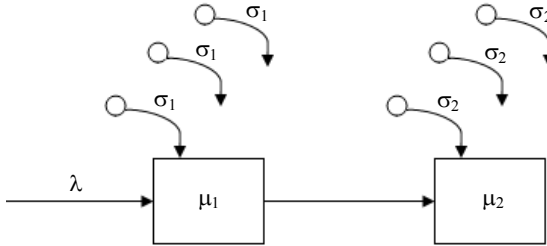


Рис. 1. Модель двухфазной RQ-системы

Обозначим через $\{i_j(t), j = \overline{1, 2}\}$ число заявок на орбите, соответствующей j -й фазе, в момент времени t . Процессы $\{i_j(t), j = \overline{1, 2}\}$ не являются марковскими, поэтому введем дополнительную переменную:

$$\{k_j(t), j = \overline{1, 2}\} = \begin{cases} 0, & \text{прибор на } j\text{-й фазе свободен,} \\ 1, & \text{прибор на } j\text{-й фазе занят.} \end{cases}$$

Пусть $P_{k_1 k_2}(i_1, i_2, t) = P\{k_1(t) = k_1, k_2(t) = k_2, i_1(t) = i_1, i_2(t) = i_2\}$ – вероятность того, что в системе в момент времени t , в то время как приборы на обеих фазах находятся в состояниях k_1 и k_2 соответственно, на орбитах на первой и второй фазе содержится i_1 и i_2 заявок. Будем рассматривать систему в стационарном режиме, при котором $\lim_{t \rightarrow \infty} P_{k_1 k_2}(i_1, i_2, t) = P_{k_1 k_2}(i_1, i_2)$.

Очевидно, что процесс $\{i_j(t), j = \overline{1, 2}, k_j(t), j = \overline{1, 2}\}$ является марковским. Для получения распределений вероятностей $P_{k_1 k_2}(i_1, i_2)$ составим систему уравнений Колмогорова:

$$\begin{cases} -(\lambda + i_1 \sigma_1 + i_2 \sigma_2) P_{00}(i_1, i_2) + \mu_2 P_{01}(i_1, i_2) = 0, \\ \lambda P_{10}(i_1 - 1, i_2) + \lambda P_{00}(i_1, i_2) + \sigma_1 (i_1 + 1) P_{00}(i_1 + 1, i_2) - \\ - (\lambda + \mu_1 + i_2 \sigma_2) P_{10}(i_1, i_2) + \mu_2 P_{11}(i_1, i_2) = 0, \\ \mu_1 P_{10}(i_1, i_2) + \mu_1 P_{11}(i_1, i_2 - 1) + \sigma_2 (i_2 + 1) P_{00}(i_1, i_2 + 1) - \\ - (\lambda + i_1 \sigma_1 + \mu_2) P_{01}(i_1, i_2) = 0, \\ \lambda P_{11}(i_1 - 1, i_2) + \lambda P_{01}(i_1, i_2) + \sigma_1 (i_1 + 1) P_{01}(i_1 + 1, i_2) + \\ + \sigma_2 (i_2 + 1) P_{10}(i_1, i_2 + 1) - (\lambda + \mu_1 + \mu_2) P_{11}(i_1, i_2) = 0. \end{cases} \quad (1)$$

Асимптотический анализ в условии большой задержки заявок на орбитах

В системе (1) перейдем к частичным характеристическим функциям:

$$H_{k_1 k_2}(u_1, u_2) = \sum_{i_1} \sum_{i_2} e^{j(u_1 i_1 + u_2 i_2)} P_{k_1 k_2}(i_1, i_2),$$

где $j = \sqrt{-1}$ – мнимая единица.

Перепишем систему уравнений (1) для характеристических функций:

$$\left\{ \begin{array}{l} -\lambda H_{00}(u_1, u_2) + j\sigma_1 \frac{\partial H_{00}(u_1, u_2)}{\partial u_1} + j\sigma_2 \frac{\partial H_{00}(u_1, u_2)}{\partial u_2} + \mu_2 H_{01}(u_1, u_2) = 0, \\ \lambda e^{ju_1} H_{10}(u_1, u_2) + \lambda H_{00}(u_1, u_2) - j\sigma_1 e^{-ju_1} \frac{\partial H_{00}(u_1, u_2)}{\partial u_1} - \\ - (\lambda + \mu_1) H_{10}(u_1, u_2) + j\sigma_2 \frac{\partial H_{10}(u_1, u_2)}{\partial u_2} + \mu_2 H_{11}(u_1, u_2) = 0, \\ \mu_1 H_{10}(u_1, u_2) + \mu_1 e^{ju_2} H_{11}(u_1, u_2) - j\sigma_2 e^{-ju_2} \frac{\partial H_{00}(u_1, u_2)}{\partial u_2} - \\ - (\lambda + \mu_2) H_{01}(u_1, u_2) + j\sigma_1 \frac{\partial H_{01}(u_1, u_2)}{\partial u_1} = 0, \\ \lambda e^{ju_1} H_{11}(u_1, u_2) + \lambda H_{01}(u_1, u_2) - j\sigma_1 e^{-ju_1} \frac{\partial H_{01}(u_1, u_2)}{\partial u_1} - \\ - j\sigma_2 e^{-ju_2} \frac{\partial H_{10}(u_1, u_2)}{\partial u_2} - (\lambda + \mu_1 + \mu_2) H_{11}(u_1, u_2) = 0. \end{array} \right. \quad (2)$$

Систему (2) будем решать методом асимптотического анализа в условии большой задержки заявок на орбитах, характеризующимся тем, что $\sigma_1, \sigma_2 \rightarrow 0$.

Асимптотика первого порядка

В соответствии с условием большой задержки заявок на орбитах обозначим $\sigma_k = \gamma_k \sigma$, $\sigma \rightarrow 0$. В системе (2) выполним замены:

$$\sigma = \varepsilon, u_k = \varepsilon w_k, H_{k_1 k_2}(u_1, u_2) = F_{k_1 k_2}(w_1, w_2, \varepsilon).$$

Теорема 1. Пусть $i_1(t)$ и $i_2(t)$ – число заявок на орбитах на первой и второй фазе соответственно в момент времени t . Тогда выполняется равенство

$$\lim_{\sigma \rightarrow 0} M \exp \{ jw_1 \sigma i_1(t) + jw_2 \sigma i_2(t) \} = \exp \{ jw_1 a_1 + jw_2 a_2 \}, \quad (3)$$

где параметры a_1 и a_2 определяются совместным решением систем уравнений

$$R = (v \overline{M}^T) \left(\overline{M M}^T \right)^{-1},$$

$$\begin{cases} \lambda R_{10} + \lambda R_{11} - \gamma_1 a_1 R_{00} - \gamma_1 a_1 R_{01} = 0, \\ \mu_1 R_{11} - \gamma_2 a_2 R_{00} - \gamma_2 a_2 R_{10} = 0. \end{cases}$$

Здесь матрица \overline{M} и вектор v имеют вид

$$\overline{M} = (M | E), \text{ где}$$

$$M = \begin{pmatrix} -(\lambda + \gamma_1 a_1 + \gamma_2 a_2) & \lambda + \gamma_1 a_1 & \gamma_2 a_2 & 0 \\ \mu_2 & 0 & -(\lambda + \mu_2 + \gamma_1 a_1) & \lambda + \gamma_1 a_1 \\ 0 & -(\mu_1 + \gamma_2 a_2) & \mu_1 & \gamma_2 a_2 \\ 0 & \mu_2 & \mu_1 & -(\mu_1 + \mu_2) \end{pmatrix},$$

$$v = (E^T \ 0 | 1).$$

Асимптотическая характеристическая функция имеет вид

$$H(u_1, u_2) = \exp \left\{ ju_1 \frac{a_1}{\sigma} + ju_2 \frac{a_2}{\sigma} \right\}.$$

Асимптотика второго порядка

Для более детального исследования характеристической функции $Me^{j(u_1 i_1 + u_2 i_2)}$ решение $H_{k_1 k_2}(u_1, u_2)$ системы (2) будем искать в виде

$$H_{k_1 k_2}(u_1, u_2) = e^{\frac{j}{\sigma}(u_1 a_1 + u_2 a_2)} H_{k_1 k_2}^{(2)}(u_1, u_2).$$

Введем обозначения:

$$\sigma_k = \gamma_k \sigma, \quad \sigma = \varepsilon^2, \quad u_k = \varepsilon w_k, \quad H_{k_1 k_2}^{(2)}(u_1, u_2) = F_{k_1 k_2}^{(2)}(w_1, w_2, \varepsilon).$$

Теорема 2. Пусть $i_1(t)$ и $i_2(t)$ – число заявок на орбитах на первой и второй фазе соответственно в момент времени t . Тогда выполняется равенство

$$\begin{aligned} \lim_{\sigma \rightarrow 0} M \exp \left\{ jw_1 \sqrt{\sigma} \left(i_1(t) - \frac{a_1}{\sigma} \right) + jw_2 \sqrt{\sigma} \left(i_2(t) - \frac{a_2}{\sigma} \right) \right\} = \\ = \exp \left\{ \frac{(jw_1)^2}{2} K_{11} + jw_1 jw_2 K_{12} + \frac{(jw_2)^2}{2} K_{22} \right\}, \end{aligned} \quad (4)$$

где параметры K_{11} , K_{12} и K_{22} определяется решением неоднородных систем линейных уравнений:

$$\begin{cases} \frac{1}{2} R_{00} \gamma_1 a_1 + \frac{1}{2} R_{01} \gamma_1 a_1 + \frac{1}{2} R_{10} \lambda + \frac{1}{2} R_{11} \lambda - R_{00} \gamma_1 K_{11} - R_{01} \gamma_1 K_{11} - \gamma_1 a_1 f_{00} - \\ - \gamma_1 a_1 f_{01} + \lambda f_{10} + \lambda f_{11} = 0, \\ \frac{1}{2} R_{00} \gamma_2 a_2 + \frac{1}{2} R_{10} \gamma_2 a_2 + \frac{1}{2} R_{11} \mu_1 - R_{00} \gamma_2 K_{22} - R_{10} \gamma_2 K_{22} - \gamma_2 a_2 g_{00} - \gamma_2 a_2 g_{10} + \\ + \mu_1 g_{11} = 0, \\ -R_{00} \gamma_1 K_{12} - R_{00} \gamma_2 K_{12} - R_{10} \gamma_2 K_{12} - R_{01} \gamma_1 K_{12} - \gamma_1 a_1 g_{00} - \gamma_2 a_2 f_{00} - \gamma_1 a_1 g_{01} + \\ + \lambda g_{10} - \gamma_2 a_2 f_{10} + \lambda g_{11} + \mu_1 f_{11} = 0, \\ -\lambda f_{00} - \gamma_1 a_1 f_{00} - \gamma_1 R_{00} K_{11} - \gamma_2 a_2 f_{00} - \gamma_2 R_{00} K_{12} + \mu_2 f_{01} = 0, \\ \lambda f_{10} + \lambda f_{00} + \lambda R_{10} + \gamma_1 a_1 f_{00} - \gamma_1 a_1 R_{00} + \gamma_1 R_{00} K_{11} - (\lambda + \mu_1) f_{10} - \gamma_2 a_2 f_{10} - \\ - \gamma_2 R_{10} K_{12} + \mu_2 f_{11} = 0, \\ \mu_1 f_{10} + \mu_1 f_{11} + \gamma_2 a_2 f_{00} + \gamma_2 R_{00} K_{12} - (\lambda + \mu_2) f_{01} - \gamma_1 a_1 f_{01} - \gamma_1 R_{01} K_{11} = 0, \\ -\gamma_1 R_{01} K_{11} = 0, \\ \lambda f_{11} + \lambda R_{11} + \lambda f_{01} + \gamma_1 a_1 f_{01} - \gamma_1 a_1 R_{01} + \gamma_1 R_{01} K_{11} + \gamma_2 a_2 f_{10} + \gamma_2 R_{10} K_{12} - \\ - (\lambda + \mu_1 + \mu_2) f_{11} = 0, \\ -\lambda g_{00} - \gamma_1 a_1 g_{00} - \gamma_1 R_{00} K_{12} - \gamma_2 a_2 g_{00} - \gamma_2 R_{00} K_{22} + \mu_2 g_{01} = 0, \\ \lambda g_{10} + \lambda g_{00} + \gamma_1 a_1 g_{00} - \gamma_1 R_{00} K_{12} - (\lambda + \mu_1) g_{10} - \gamma_2 a_2 g_{10} - \gamma_2 R_{10} K_{22} + \mu_2 g_{11} = 0, \\ + \mu_2 g_{11} = 0, \\ \mu_1 g_{10} + \mu_1 g_{11} + \gamma_2 a_2 g_{00} + \mu_1 R_{11} - \gamma_2 a_2 R_{00} + \gamma_2 R_{00} K_{22} - (\lambda + \mu_2) g_{01} - \\ - \gamma_1 a_1 g_{01} - \gamma_1 R_{01} K_{12} = 0, \\ \lambda g_{11} + \lambda g_{01} + \gamma_1 a_1 g_{01} + \gamma_1 R_{01} K_{12} + \gamma_2 a_2 g_{10} - \gamma_2 a_2 R_{10} + \gamma_2 R_{10} K_{22} - \\ - (\lambda + \mu_1 + \mu_2) g_{11} = 0. \end{cases}$$

Из (3) и (4) следует, что характеристическая функция имеет вид

$$H(u_1, u_2) = \exp \left\{ ju_1 \frac{a_1}{\sigma} + ju_2 \frac{a_2}{\sigma} + \frac{(ju_1)^2}{2} \frac{K_{11}}{\sigma} + j^2 u_1 u_2 \frac{K_{12}}{\sigma} + \frac{(ju_2)^2}{2} \frac{K_{22}}{\sigma} \right\}.$$

Отсюда плотность нормального распределения числа заявок на орбитах

$$P_1(i_1, i_2) = \frac{1}{2\pi\sqrt{K_{11}K_{22}(1-\rho^2)}} \exp \left\{ -\frac{1}{2(1-\rho^2)} \left(\frac{(i_1 - a_1)^2}{K_{11}} - \rho \frac{2(i_1 - a_1)(i_2 - a_2)}{\sqrt{K_{11}K_{22}}} + \frac{(i_2 - a_2)^2}{K_{22}} \right) \right\},$$

где

$$\rho = \frac{K_{12}}{\sqrt{K_{11}K_{22}}}.$$

Результаты имитационного моделирования

В данном разделе проводится сравнение асимптотических результатов с результатами имитационного моделирования. Описание имитационной модели приводится в [8, 9].

В качестве критерия точности оценивания будем использовать расстояние Колмогорова между асимптотическим распределением вероятностей числа заявок на орбите и распределением, полученным в результате работы модели. Пусть Δ – двумерное расстояние для обеих орбит, а Δ_1 и Δ_2 – расстояния для первой и второй орбиты соответственно. В качестве критерия точности оценивания было выбрано расстояние Колмогорова меньше 0,05. В таблице представлены результаты сравнения в зависимости от изменения параметра σ и заданных значениях параметров: $\lambda = 0,5$, $\mu_1 = \mu_2 = 1$, $\gamma_1 = 1$, $\gamma_2 = 2$.

Область применимости асимптотических результатов

σ	Δ	Δ_1	Δ_2
0,1	0,172	0,038	0,069
0,05	0,045	0,031	0,039
0,03	0,051	0,028	0,029
0,01	0,077	0,015	0,021

Заключение

Из таблицы видно, что при уменьшении σ повышается точность аппроксимации для маргинальных распределений, но не для двумерного. Это связано со слабой корреляцией распределений вероятностей числа заявок на первой и второй орбитах.

ЛИТЕРАТУРА

1. Гнеденко Б.В., Коваленко И.Н. Введение в теорию массового обслуживания. – 2-е изд., перераб. и доп. – М.: Наука, 1987. – 336 с.
2. Artalejo J.R., Gomez-Corral A. Retrial Queueing Systems: A Computational Approach. – Berlin: Springer, 2008. – 267 p.
3. Falin G.I., Templeton J.G.C. Retrial queues. London: Chapman & Hall, 1997. – 320 p.
4. Krishna C.M., Lee Y.H. A study of two-phase service // Operations Research Letters. – 1990. – V. 9. – P. 91–97.
5. Doshi B.T. Analysis of a two phase queueing system with general service times // Operations Research Letters. – 1991. – V. 10. – P. 265–272.
6. Krishnakumar B., Arivudainambi D. An M/G/1 retrial queueing system with two phases of services and preemptive resume // Annals of Operations Research. – 2002. – V. 113. – P. 61–79.
7. Choudhury G. A single server queueing system with two-phases of service and vacations // QTQM. – 2008. – V. 5. – No. 1. – P. 33–49.
8. Nazarov A.A., Sudyko E.A. Method of asymptotic semiinvariants for studying a mathematical model of a random access network // Problems of information transmission. – 2010. – V. 46. – No. 1. – P. 86–102.
9. Анисимова А.А. Имитационное моделирование двухфазной RQ-системы // Научное творчество молодежи. Математика. Информатика: материалы XX Всероссийской научно-практической конференции (28–29 апреля 2016 г.) / сост. Ю.А. Намкина. – Томск: Изд-во Том. ун-та, 2016. – Ч. 1. – С. 81–85.

ИССЛЕДОВАНИЕ RQ-СИСТЕМЫ M/GI/1 С R-НАСТОЙЧИВЫМ ВЫТЕСНЕНИЕМ ЗАЯВОК*

А.А. Назаров, Я.Е. Измайлова

Томский государственный университет, г. Томск, Россия

В современном мире, где активно развиваются информационные, транспортные системы, исследование RQ-систем (систем с повторными вызовами) является актуальной научной задачей. Системы с повторами характеризуются тем, что прибывшая в систему заявка, обнаружив прибор занятым, уходит в зону ожидания и через некоторое случайное время повторяет попытку обслуживания. Между повторами заявки (клиенты) находятся в «источнике повторных вызовов» (ИПВ или орбита). Обзор работ по этой тематике приведен в работах J.R. Artalejo [1]. Системы с повторными вызовами и вытеснением заявок могут являться математическими моделями реальных телекоммуникационных систем, а также операционных компьютерных систем.

В последние годы акцент направлен на исследование приоритетных систем с повторами. К таким работам можно отнести исследования В. Kim [2], С. D'Arice [3], С. Kim, В.И. Клименок [4], А.Н. Дудина [5, 6]. Также можно встретить немало работ, касающихся изучения RQ-систем с дискретным временем, например работы I.M. Atencia [7]. В [7] рассматривается система с повторными вызовами, в которой прибывшая заявка, обнаружившая прибор занятым, может решить: начать обслуживание или присоединиться к орбите и повторить попытку позже согласно дисциплине FCFS.

Но в приоритетных RQ-системах не учитывается эффект наблюдаемого в реальных системах вытеснения требований.

Математическая модель

Рассмотрим RQ-систему с вытеснением заявок. На вход системы поступает простейший поток заявок с интенсивностью λ . Требование, заставшее прибор свободным, занимает его для обслуживания в течение

* Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 16-31-00292 мол_а.

случайного времени с функцией распределения $B(x)$. Если прибор занят, то с вероятностью r поступившая заявка вытесняет обслуживаемую и сама встает на прибор, а заявка, которая обслуживалась, переходит на орбиту, где осуществляет случайную задержку, продолжительность которой имеет экспоненциальное распределение с параметром σ . Иначе с вероятностью $(1-r)$ пришедшая заявка уходит на орбиту сама, не влияя на обслуживание стоящей заявки на приборе. С орбиты, после случайной задержки, заявка вновь обращается к прибору с повторной попыткой его захвата. Дисциплина обращений заявок с орбиты аналогична дисциплине обращения заявок, которые впервые прибыли в систему (рис. 1).

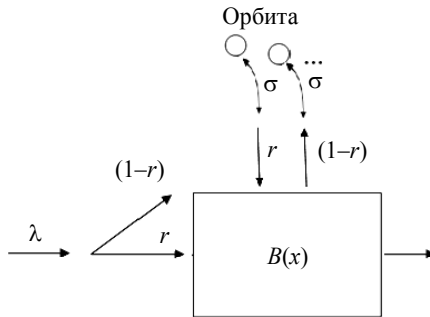


Рис. 1. RQ-система M/GI/1
с r -настойчивым вытеснением заявок

Обозначим $i(t)$ – число заявок на орбите, $k(t)$ определяет состояние прибора следующим образом:

$$k(t) = \begin{cases} 0, & \text{если прибор свободен,} \\ 1, & \text{если прибор занят.} \end{cases}$$

Ставится задача нахождения стационарного распределения вероятностей числа заявок на орбите и состояний прибора.

Так как процесс $\{k(t), i(t)\}$ не является марковским, то для его марковизации рассмотрим процесс с переменным числом компонент.

Если $k(t) = 0$, то рассматриваем процесс $\{k(t), i(t)\}$. Если $k(t) = 1$, то рассматриваем процесс $\{k(t), i(t), z(t)\}$, где $z(t)$ остаточное время от момента t до момента окончания обслуживания.

Обозначим $P\{k(t)=0, i(t)=i\} = P_0(i, t)$, $i \geq 0$, вероятность того, что прибор в момент времени t находится в состоянии 0 и в источнике повторных вызовов находится i заявок; $P\{k(t)=1, i(t)=i, z(t) < z\} = P_1(i, z, t)$, $i \geq 0$ вероятность того, что прибор в момент времени t находится в состоянии 1, остаточное время обслуживания меньше z и в источнике повторных вызовов находится i заявок.

Для этого распределения вероятностей нетрудно получить систему уравнений Колмогорова:

$$\begin{aligned} -\frac{\partial P_1(i, z)}{\partial z} + \frac{\partial P_1(i, 0)}{\partial z} &= \lambda B(z)P_0(i) - (\lambda + i\sigma)P_1(i, z) + (1-r)\lambda P_1(i-1, z) + \\ &+ (1-r)i\sigma P_1(i, z) + i\sigma rB(z)P_1(i, \infty) + \lambda rB(z)P_1(i-1, \infty) + (i+1)\sigma B(z)P_0(i+1), \\ -\frac{\partial P_1(i, 0)}{\partial z} &= -(\lambda + i\sigma)P_0(i), \end{aligned}$$

в которой применяется обозначение

$$\frac{\partial P_1(i, 0)}{\partial z} = \frac{\partial P_1(i, z)}{\partial z} \Big|_{z=0}.$$

От системы для распределения вероятностей перейдем к системе уравнений для частичных характеристических функций вида

$$H_0(u) = \sum_{i=0}^{\infty} e^{ju i} P_0(i), \quad H_1(u, z) = \sum_{i=0}^{\infty} e^{ju i} P_1(i, z),$$

где $j = \sqrt{-1}$ – мнимая единица. Обозначим

$$H_1(u, \infty) = H_1(u), \quad \frac{\partial H_1(u, 0)}{\partial z} = \frac{\partial H_1(u, z)}{\partial z} \Big|_{z=0}.$$

Запишем систему уравнений для частичных характеристических функций:

$$\begin{aligned} \frac{\partial H_1(u, z)}{\partial z} + \frac{\partial H_1(u, 0)}{\partial z} + j\sigma \frac{\partial H_1(u, z)}{\partial u} - j\sigma rB(z) \frac{\partial H_1(u)}{\partial u} - \\ - j\sigma(1-r)B(z) \frac{\partial H_1(u, z)}{\partial u} - j\sigma B(z)e^{-ju} \frac{\partial H_0(u)}{\partial u} + \lambda B(z)H_0(u) - \\ - \lambda H_1(u, z) + \lambda rB(z)e^{ju} H_1(u) + \lambda(1-r)e^{ju} H_1(u, z) = 0, \\ \frac{\partial H_1(u, 0)}{\partial z} - \lambda H_0(u) + j\sigma \frac{\partial H_0(u)}{\partial u} = 0. \end{aligned}$$

Аналитически данную систему решить затруднительно. Будем решать ее методом асимптотического анализа [8, 9] в условии большой задержки ($\sigma \rightarrow 0$).

Асимптотический анализ

Обозначим κ_1 – асимптотическое среднее значение числа заявок на орбите, а R_k , $k = 0, 1$ – распределение вероятностей состояний прибора. Сформулируем следующее утверждение.

Теорема 1. Пусть $i(t)$ число заявок на орбите RQ-системы с r -настойчивым вытеснением заявок. Тогда выполняется равенство

$$\lim_{\sigma \rightarrow 0} M \exp \{ju \sigma i(t)\} = \exp \{j\kappa_1\},$$

где κ_1 является решением уравнения

$$\lambda = (\lambda + \kappa_1)R_0, \quad (1)$$

а R_0 , R_1 определяются равенствами

$$R_0 = \frac{rB^*((\lambda + \kappa_1)r)}{1 - (1-r)B^*((\lambda + \kappa_1)r)}, \quad R_1 = \frac{1 - B^*((\lambda + \kappa_1)r)}{1 - (1-r)B^*((\lambda + \kappa_1)r)}. \quad (2)$$

Из-за трудоемких вычислений доказательство теоремы не приводится. Из полученных в теореме 1 уравнения (1) и равенств (2) определяются: распределение вероятностей состояний прибора и среднее число заявок в источнике повторных вызовов на фазах.

Для более детального исследования RQ-системы построим гауссовскую аппроксимацию числа заявок на орбите.

Теорема 2. Пусть $i(t)$ число заявок на орбите RQ-системы с r -настойчивым вытеснением заявок. Тогда выполняется равенство

$$\lim_{\sigma \rightarrow 0} M \exp \left\{ ju \frac{\sigma i(t) - \kappa_1}{\sqrt{\sigma}} \right\} = \exp \left\{ \frac{(ju)^2}{2} \kappa_2 \right\},$$

где

$$\kappa_2 = \frac{\lambda R_1 + (1-r)\lambda R_1^*(r(\lambda + \kappa_1))}{r[B^*(r(\lambda + \kappa_1)) - R_1^*(r(\lambda + \kappa_1))]},$$

κ_1 является решением уравнения (1), величины R_0 , R_1 удовлетворяют равенствам (2),

$$R_1(z) = (\lambda + \kappa_1) e^{r(\lambda + \kappa_1)z} \int_0^z e^{-r(\lambda + \kappa_1)x} (R_0 - R_0 B(x) - r R_1 B(x)) dx,$$

$R_1^*(r(\lambda + \kappa_1))$ – преобразование Лапласа – Стильесса от функции $R_1(z)$ в точке $r(\lambda + \kappa_1)$.

Таким образом, стационарное распределение вероятностей числа заявок на орбите RQ-системы с простейшим входящим потоком, произвольной функцией распределения времени обслуживания и г-настойчивым вытеснением заявок можно аппроксимировать гауссовским распределением с параметрами $\frac{\kappa_1}{\sigma}, \frac{\kappa_2}{\sigma}$.

Заключение

Предложенная новая математическая модель RQ-системы с г-настойчивым вытеснением заявок расширяет класс RQ-систем с вытеснением заявок [10, 11], что позволит решить ряд практических задач в области телекоммуникаций и транспортных системах.

ЛИТЕРАТУРА

1. *Artalejo J.R., Gomez-Corral A.* Retrial Queueing Systems: A Computational Approach. – Springer, 2008.
2. *Kim B., Kim J.* Stability of a two-class two-server retrial queueing system // Performance Evaluation. – 2015. – V. 88–89. – P. 1–17.
3. *D'Apice C., De Simone T., Manzo R., Rizelian G.* Priority service of primary customers in the M/G/1/r retrial queueing system with server searching for customers // J. Information Theory and Information Processing. – 2004. – V. 4. – No. 1. – P. 13–23.
4. *Kim C., Klimenok V.I., Dudin A.N.* Priority tandem queueing system with retrials and reservation of channels as a model of call center // Computers & Industrial Engineering. – 2016. – V. 96. – P. 61–71.
5. *Dudin A.N., Krishnamoorthy A., Joshua V.C., Tsarenkov G.V.* Analysis of the BMAP/G/1 retrial system with search of customers from the orbit // European J. Operational Research. – 2004. – V. 157. – No. 1. – P. 169–179.
6. *Dudin A.N., Kim C., Dudin S., Dudina O.* Priority retrial queueing model operating in random environment with varying number and reservation of servers // Applied Mathematics and Computation. – 2015. – V. 269. – P. 674–690.
7. *Atencia I.M.* A Geo/G/1 retrial queueing system with priority services // European Journal of Operational Research. – V. 256. – Issue 1. – 2017. – 1 January. – P. 178–186.
8. *Боровков А.А.* Асимптотические методы в теории массового обслуживания. – М.: Наука, 1980. – 384 с.

-
9. Анисимов В.В. Асимптотические методы анализа стохастических систем – Киев; Тбилиси: Мецниереба, 1984. – 178 с.
 10. Nazarov A., Izmailova Ya. Asymptotic analysis retrial queueing system M/GI/1 with hyper exponential distribution the delay time in the orbit and exclusion of alternative customers // Information Technologies and Mathematical Modelling. Queueing Theory and Applications: Proceeding of the 15th International Scientific Conference, ITMM 2016. Katun, September 12–16, 2016. – Springer International Publishing Switzerland, 2016. – V. 638: Communications in Computer and Information Science. – P. 292–302. – DOI: 10.1007/978-3-319-44615-8.
 11. Nazarov A., Chernikova Ya. Gaussian approximation of distribution of states of the retrial queueing system with r-persistent exclusion of alternative customers // Information Technologies and Mathematical Modelling. Queueing Theory and Applications: Proceeding of the 14th International Scientific Conference, ITMM 2015. Anzhero-Sudzhensk, November 18–22, 2015. – Springer International Publishing Switzerland, 2015. – V. 564: Communications in Computer and Information Science. – P. 200–208. – DOI: 10.1007/978-3-319-25861-4_19

СРАВНЕНИЕ МЕТОДОВ ОСТАТОЧНОГО И ИСТЕКШЕГО ВРЕМЕНИ ОБСЛУЖИВАНИЯ ДЛЯ ИССЛЕДОВАНИЯ ЗАМКНУТОЙ RQ-СИСТЕМЫ M/GI/1/N С КОНФЛИКТАМИ ЗАЯВОК И НЕНАДЕЖНЫМ ПРИБОРОМ

А.А. Назаров, А.С. Квач

Томский государственный университет, г. Томск, Россия

RQ-системы [1–3] широко используются для исследования телекоммуникационных и компьютерных систем, при проектировании мобильных сотовых радиосетей, телефонных сетей и во многих других областях. На практике часто необходимо учитывать тот факт, что с увеличением числа заявок в системе интенсивность генерирования новой заявки уменьшается. В данной ситуации используются модели с конечным числом источников или замкнутые RQ-системы.

Очень часто при исследовании различных систем массового обслуживания обслуживающие приборы предполагаются абсолютно надежными. Но на практике необходимо учитывать возможность выхода прибора из строя и его восстановления. Например, системы с конечным числом источников и ненадежным прибором ранее были исследованы в работах [4–7].

Во многих практических ситуациях, при передаче данных из разнообразных источников возможно возникновение конфликтной ситуации. Нескоординированные попытки захвата единственного сервера несколькими источниками могут привести к столкновениям, приводящим к потере передаваемых данных и, следовательно, необходимости в их повторной передаче. В работах [8–11] были рассмотрены различные RQ-системы с конфликтами заявок.

Целью данной работы является исследование RQ- системы, которая содержит все вышеперечисленные свойства, а именно: конечное число источников, ненадежный прибор и возможность возникновения конфликтных ситуаций.

Математическая модель

Рассматривается немарковская замкнутая RQ-система с конфликтами заявок и ненадежным прибором (рис. 1). Каждый из N источников генерирует заявку и отправляет ее на прибор с интенсивностью λ/N . Источник, который отправил заявку на обслуживание, находится в режиме ожидания и не генерирует новую до тех пор, пока заявка не завершит свое успешное обслуживание. Заявка, заставшая прибор свободным, занимает его для обслуживания. Время обслуживания каждой заявки является случайной величиной с произвольной функцией распределения $B(x)$. Если прибор занят, то поступившая заявка вступает в конфликт с обслуживаемой заявкой и они обе переходят на орбиту. С орбиты после случайной задержки, продолжительность которой имеет экспоненциальное распределение с параметром σ/N , заявка вновь обращается к прибору с повторной попыткой его захвата.

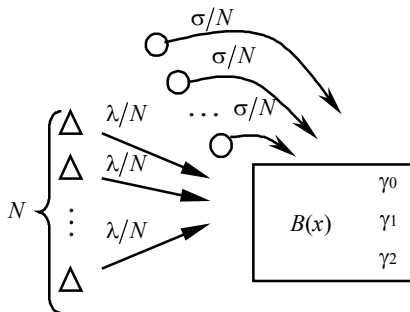


Рис. 1. Замкнутая RQ-система с конфликтами заявок и ненадежным прибором

В рассматриваемой системе прибор является ненадежным. Если прибор свободен, то он выходит из строя с интенсивностью γ_0 , если прибор занят – с интенсивностью γ_1 . В случае поломки прибор немедленно переходит в состояние восстановления, продолжительность которого имеет экспоненциальное распределение с параметром γ_2 . Если в момент выхода из строя прибор был занят, то обслуживаемая заявка покидает прибор и переходит на орбиту. Отметим, что во время периода восстановления (ремонта) генерация заявок продолжается и заявки, обратившиеся к прибору во время ремонта, мгновенно переходят на орбиту.

Пусть $i(t)$ – число заявок в системе, то есть общее число заявок на приборе и на орбите, а $k(t)$ определяет состояние прибора следующим образом:

$$k(t) = \begin{cases} 0, & \text{прибор свободен,} \\ 1, & \text{прибор занят,} \\ 2, & \text{прибор на ремонте.} \end{cases}$$

Для рассматриваемой системы процесс $\{k(t), i(t)\}$ не является марковским. Для его марковизации воспользуемся методом дополнительной переменной, а именно, рассмотрим два варианта: метод остаточного времени и метод истекшего времени, и впоследствии сравним их.

Метод остаточного времени для замкнутой RQ-системы с конфликтами заявок и ненадежным прибором

Введем случайный процесс $z(t)$, имеющий смысл длины интервала от момента t до момента окончания успешного обслуживания заявки. Таким образом, исследуется марковский процесс $\{k(t), z(t), i(t)\}$, который имеет переменное число компонент в зависимости от состояния прибора, так как компонента $z(t)$ определена только в те моменты, когда $k(t) = 1$.

Определим следующие вероятности:

$$P_0(i, t) = P\{k(t) = 0, i(t) = i\}, \quad P_1(i, z, t) = P\{k(t) = 1, i(t) = i, z(t) < z\},$$

$$P_2(i, t) = P\{k(t) = 2, i(t) = i\}.$$

Для распределения вероятностей $P_0(i, t)$, $P_1(i, z, t)$, $P_2(i, t)$ состояний системы составим систему дифференциальных уравнений Колмогорова и запишем ее для стационарного распределения

$$\begin{aligned} -\left[\lambda \frac{N-i}{N} + i \frac{\sigma}{N} + \gamma_0\right] P_0(i) + \frac{\partial P_1(i+1, 0)}{\partial z} + \lambda \frac{N-i+1}{N} P_1(i-1) + \\ + \sigma \frac{i-1}{N} P_1(i) + \gamma_2 P_2(i) = 0, \\ \frac{\partial P_1(i, z)}{\partial z} - \frac{\partial P_1(i, 0)}{\partial z} - \left(\lambda \frac{N-i}{N} + \sigma \frac{i-1}{N} + \gamma_1\right) P_1(i, z) + \\ + \lambda \frac{N-i+1}{N} P_0(i-1) B(z) + \sigma \frac{i}{N} P_0(i) B(z) = 0, \end{aligned} \quad (1)$$

$$-\left[\lambda \frac{N-i}{N} + \gamma_2\right] P_2(i) + \gamma_0 P_0(i) + \gamma_1 P_1(i) + \lambda \frac{N-i+1}{N} P_2(i-1) = 0.$$

Введем частичные характеристические функции вида

$$H_k(u) = \sum_{i=0}^N e^{ju i} P_k(i), \quad k=0,2; \quad H_1(u,z) = \sum_{i=1}^N e^{ju i} P_1(i,z),$$

тогда систему (1) перепишем в виде

$$\begin{aligned} \frac{j}{N}(\sigma - \lambda) \frac{\partial H_0(u)}{\partial u} + \frac{j}{N}(\lambda e^{ju} - \sigma) \frac{\partial H_1(u)}{\partial u} - (\lambda + \gamma_0) H_0(u) + \\ + \left(\lambda e^{ju} - \frac{\sigma}{N} \right) H_1(u) + \gamma_2 H_2(u) + e^{-ju} \frac{\partial H_1(u,0)}{\partial z} = 0, \\ \frac{\partial H_1(u,z)}{\partial z} - \frac{\partial H_1(u,0)}{\partial z} + \frac{j}{N}(\lambda e^{ju} - \sigma) B(z) \frac{\partial H_0(u)}{\partial u} + \frac{j}{N}(\sigma - \lambda) \frac{\partial H_1(u,z)}{\partial u} + \\ + \lambda e^{ju} B(z) H_0(u) + \left(\frac{\sigma}{N} - \lambda - \gamma_1 \right) H_1(u,z) = 0, \end{aligned}$$

$$\frac{j}{N} \lambda (e^{ju} - 1) \frac{\partial H_2(u)}{\partial u} + [\lambda (e^{ju} - 1) - \gamma_2] H_2(u) + \gamma_0 H_0(u) + \gamma_1 H_1(u) = 0.$$

Систему (2) будем решать методом асимптотического анализа [12] в предельном условии неограниченно растущего числа источников, то есть при $N \rightarrow \infty$.

Теорема 1. Пусть $i(t)$ – число заявок в замкнутой RQ-системе M/GI/1//N с конфликтами заявок и ненадежным прибором, тогда выполняется равенство

$$\lim_{N \rightarrow \infty} M \exp \left(jw \frac{i(t)}{N} \right) = \exp(jw\kappa),$$

где значением параметра κ является положительное решение уравнения

$$\lambda(1 - \kappa) - \delta(\kappa)[R_0(\kappa) - R_1(\kappa)] + \gamma_1 R_1(\kappa) = 0,$$

а стационарное распределение вероятностей $R_k(\kappa)$ состояний k прибора определяются равенствами

$$R_0(\kappa) = \gamma_2 \left\{ \gamma_0 + \gamma_2 + (\gamma_1 + \gamma_2) \cdot \frac{\delta(\kappa)}{\delta(\kappa) + \gamma_1} \left[1 - B^*(\delta(\kappa) + \gamma_1) \right] \right\}^{-1},$$

$$R_1(\kappa) = \frac{\delta(\kappa)}{\delta(\kappa) + \gamma_1} \left[1 - B^*(\delta(\kappa) + \gamma_1) \right] R_0(\kappa),$$

$$R_2(\kappa) = \frac{1}{\gamma_2} [\gamma_0 R_0(\kappa) + \gamma_1 R_1(\kappa)].$$

Здесь $\delta(\kappa) = \lambda(1 - \kappa) + \sigma\kappa$.

Метод истекшего времени для замкнутой RQ-системы с конфликтами заявок и ненадежным прибором

Введем дополнительную переменную $y(t)$ и рассмотрим трехмерный процесс $\{k(t), y(t), i(t)\}$, где $y(t)$ – длина интервала от момента начала обслуживания до момента t , то есть компонента $y(t)$ определяется только в те моменты, когда $k(t) = 1$. Пусть

$$p_k(i, t) = P\{k(t) = k, i(t) = i\}, k = 0, 2;$$

$$p_1(i, y, t) = \frac{P\{k(t) = 1, i(t) = i, y(t) < y\}}{\partial y},$$

тогда для распределения вероятностей $p_0(i, t)$, $p_1(i, y, t)$, $p_2(i, t)$ состояний системы составим систему дифференциальных уравнений Колмогорова и запишем ее для стационарного распределения

$$\begin{aligned} - \left[\lambda \frac{N-i}{N} + i \frac{\sigma}{N} + \gamma_0 \right] p_0(i) + \int_0^\infty p_1(i+1, y) \mu(y) dy + \lambda \frac{N-i+1}{N} p_1(i-1) + \\ + \sigma \frac{i-1}{N} p_1(i) + \gamma_2 p_2(i) = 0, \end{aligned} \quad (3)$$

$$\frac{\partial p_1(i, y)}{\partial y} = - \left[\lambda \frac{N-i}{N} + \sigma \frac{i-1}{N} + \mu(y) + \gamma_1 \right] p_1(i, y),$$

$$- \left[\lambda \frac{N-i}{N} + \gamma_2 \right] p_2(i) + \gamma_0 p_0(i) + \gamma_1 p_1(i) + \lambda \frac{N-i+1}{N} p_2(i-1) = 0,$$

с краевым условием

$$p_1(i, 0) = \lambda \frac{N-i+1}{N} p_0(i-1) + i \frac{\sigma}{N} p_0(i). \quad (4)$$

Введем частичные характеристические функции вида

$$H_k(u) = \sum_{i=0}^N e^{ju i} p_k(i), \quad k=0, 2; \quad H_1(u, y) = \sum_{i=1}^N e^{ju i} p_1(i, y),$$

тогда систему (3) и равенство (4) перепишем в виде

$$\begin{aligned} & \frac{j}{N}(\sigma - \lambda) \frac{\partial H_0(u)}{\partial u} + \frac{j}{N}(\lambda e^{ju} - \sigma) \frac{\partial H_1(u)}{\partial u} - (\lambda + \gamma_0) H_0(u) + \\ & + \left(\lambda e^{ju} - \frac{\sigma}{N} \right) H_1(u) + \gamma_2 H_2(u) + e^{-ju} \int_0^\infty H_1(u, y) \mu(y) dy = 0, \\ & \frac{\partial H_1(u, y)}{\partial y} = \frac{j}{N}(\sigma - \lambda) \frac{\partial H_1(u, y)}{\partial u} + \left[\frac{\sigma}{N} - \lambda - \mu(y) - \gamma_1 \right] H_1(u, y), \quad (5) \\ & \frac{j}{N} \lambda (e^{ju} - 1) \frac{\partial H_2(u)}{\partial u} + \left[\lambda (e^{ju} - 1) - \gamma_2 \right] H_2(u) + \gamma_0 H_0(u) + \gamma_1 H_1(u) = 0, \\ & H_1(u, 0) = \frac{j}{N}(\lambda e^{ju} - \sigma) \frac{\partial H_0(u)}{\partial u} + \lambda e^{ju} H_0(u). \end{aligned}$$

Полученную систему (5) будем решать методом асимптотического анализа в том же предельном условии неограниченно растущего числа источников.

Теорема 2. Пусть $i(t)$ – число заявок в замкнутой RQ-системе M/GI/1//N с конфликтами заявок и ненадежным прибором, тогда выполняется равенство

$$\lim_{N \rightarrow \infty} M \exp \left(jw \frac{i(t)}{N} \right) = \exp(jw\kappa),$$

где значением параметра κ является положительное решение уравнения

$$\lambda(1 - \kappa) - \delta(\kappa)[R_0(\kappa) - R_1(\kappa)] + \gamma_1 R_1(\kappa) = 0,$$

а стационарное распределение вероятностей $R_k(\kappa)$ состояний k прибора определяются равенствами

$$R_0(\kappa) = \gamma_2 \left\{ \gamma_0 + \gamma_2 + (\gamma_1 + \gamma_2) \cdot \frac{\delta(\kappa)}{\delta(\kappa) + \gamma_1} \left[1 - B^*(\delta(\kappa) + \gamma_1) \right] \right\}^{-1},$$

$$R_1(\kappa) = \frac{\delta(\kappa)}{\delta(\kappa) + \gamma_1} \left[1 - B^*(\delta(\kappa) + \gamma_1) \right] R_0(\kappa), \quad R_2(\kappa) = \frac{1}{\gamma_2} [\gamma_0 R_0(\kappa) + \gamma_1 R_1(\kappa)].$$

Здесь $\delta(\kappa) = \lambda(1 - \kappa) + \sigma\kappa$.

Сравнение методов остаточного и истекшего времени

При исследовании замкнутой RQ-системы M/GI/1//N с конфликтами заявок и ненадежным прибором методом асимптотического анализа для марковизации процесса $\{k(t), i(t)\}$ в работе были рассмотрены два метода: метод истекшего времени и метод остаточного времени. В результате были сформулированы теоремы 1 и 2, из которых следует, что основные характеристики системы, такие, как стационарное распределение вероятностей R_k состояний k прибора и асимптотическое среднее нормированного числа заявок в системе κ имеют одинаковый вид и не зависят от метода исследования.

Использование метода истекшего времени необходимо для исследования числа переходов заявки на орбиту, а также для дальнейшего исследования времени пребывания заявки на орбите.

Метод остаточного времени используется для нахождения распределения вероятностей числа заявок в системе. Также метод остаточного времени необходим при исследовании времени пребывания заявки на приборе.

Заключение

Рассмотрена замкнутая RQ-система M/GI/1//N с конфликтами заявок и ненадежным прибором. Исследование системы выполнено методом асимптотического анализа в предельном условии неограниченно растущего числа источников. При исследовании системы были использованы методы истекшего и остаточного времени. В результате было показано, что стационарное распределение вероятностей R_k состояний k прибора и асимптотическое среднее нормированного числа заявок в системе κ имеют одинаковый вид и не зависят от применяемого метода.

ЛИТЕРАТУРА

1. Назаров А.А., Терпугов А.Ф. Теория массового обслуживания: учебное пособие. – 2-е изд., испр. – Томск: Изд-во НТЛ, 2010. – 228 с.
2. Artalejo J.R., Gomez-Corral A. Retrial queueing systems // A Computational Approach. – Springer, 2008. – 309 p.
3. Falin G.I., Templeton J.G.C. Retrial queues. – London: Chapman & Hall, 1997. – 328 p.
4. Sztrik J., Almási B., Roszik J. Heterogeneous finite-source retrial queues with server subject to breakdowns and repairs // J. Mathematical Sciences. – 2006. – V. 132. – P. 677–685.
5. Dragieva V.I. System State Distributions in One Finite Source Unreliable Retrial Queue. – URL: <http://elib.bsu.by/handle/123456789/35903>.
6. Gharbi N., Dutheillet C. An algorithmic approach for analysis of finite-source retrial systems with unreliable servers // Computers & Mathematics with Applications. – 2011. – V. 62. – No. 6. – P. 2535–2546. DOI 10.1016/j.camwa.2011.03.109
7. Wang J., Zhao L., Zhang F. Analysis of the finite source retrial queues with server breakdowns and repairs // J. Industrial and Management Optimization. – 2011. – V. 7. – No. 3. P. 655–676. DOI 10.3934/jimo.2011.7.655.
8. Choi B.D., Shin Y.W., Ahn W.C. Retrial queues with collision arising from unslotted CSMA/CD protocol // Queueing Syst. – 1992. – V. 11. – No. 4. – P. 335–356. DOI 10.1007/BF01163860.
9. Nazarov A., Kvach A., Yampolsky V. Asymptotic analysis of closed Markov Retrial Queueing System with collision // Communications in Computer and Information Science. – 2014. – V. 487: Information Technologies and Mathematical Modelling. – P. 334–341.
10. Квач А.С. Назаров А.А. Исследование замкнутой RQ-системы $M|GI|1|N$ с конфликтами заявок в условии неограниченно растущего числа источников // Теория вероятностей, случайные процессы, математическая статистика и приложения: материалы Международной научной конференции, г. Минск, 2015. – С. 65–70. URL: <http://vital.lib.tsu.ru/vital/access/manager/Repository/vtls:000535452>
11. 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.
12. Назаров А.А., Моисеева С.П. Метод асимптотического анализа в теории массового обслуживания. – Томск: Изд-во НТЛ, 2006. – 112 с.

ИССЛЕДОВАНИЕ ВЫХОДЯЩЕГО ПОТОКА В RQ-СИСТЕМЕ М/М/1/1 С ВЫЗЫВАЕМЫМИ ЗАЯВКАМИ

А.А. Назаров, И.Л. Лапатын, С.В. Пауль

Томский государственный университет, г. Томск, Россия

В системах обслуживания, таких, как call-центры, время простоя операторов должно быть минимизировано для увеличения продуктивности систем. С этой целью операторы могут не только принимать входящие вызовы, но и делать исходящие, например для продвижения услуг центра. Эти ситуации моделируются RQ-системами с вызываемыми заявками, где сервер имеет возможность принимать вызовы и совершать исходящие.[1]

В работе проведено исследование потока обслуженных заявок в такой RQ-системе [2, 3]. Хотя информация о их характеристиках представляет большой практический интерес, исследованию выходящих потоков в настоящее время уделяется недостаточно внимания, так как не существует общих подходов к их исследованию. Основные результаты по аналитическому исследованию выходящих потоков в рамках классической теории были сделаны во второй половине XX в. такими учеными, как П. Берк [4], Е. Рейч [5], П. Финч [6]. Поэтому достаточно актуальной является проблема разработки новых и модификация имеющихся методов исследования выходящих потоков.

В предложенной работе основным методом исследования является метод асимптотического анализа, который позволяет в RQ-системе М/М/1/1 с двумя классами заявок [7] найти вид предельного распределения числа событий в выходящем потоке, наступивших за время t в асимптотическом условии большой задержки на орбите. На основе найденного распределения построено дискретное распределение (будем называть его гауссовской аппроксимацией), которое аппроксимирует дискретное распределение числа событий в выходящем потоке, наступивших за время t .

Математическая модель RQ-системы M/M/1/1 с вызываемыми заявками и постановка задачи

Рассмотрим RQ-систему, на вход которой поступает простейший поток заявок с интенсивностью λ . Заявка входящего потока, поступая в систему и обнаруживая прибор свободным, занимает его, а прибор начинает обслуживание в течение экспоненциально-распределенного времени с параметром μ_1 . Если же заявка, поступая в систему, обнаруживает прибор занятым, она мгновенно уходит на орбиту и осуществляет там случайную задержку в течение экспоненциально-распределенного времени с параметром σ .

Когда прибор свободен, он вызывает для обслуживания дополнительные заявки из внешней среды с интенсивностью α . Вызываемая заявка занимает прибор, на котором обслуживается в течение случайного времени, распределенного по экспоненциальному закону с параметром μ_2 . Обозначим:

$i(t)$ – число заявок в системе в момент времени t ;

$k(t)$ – состояние прибора: 0 – прибор свободен, 1 – прибор занят обслуживанием заявки входящего потока, 2 – прибор обслуживает вызываемую заявку;

$m(t)$ – число событий в выходящем потоке, наступивших за время t .

Рассмотрим трехмерный марковский процесс $\{k(t), i(t), m(t)\}$. Для распределения вероятностей

$$P\{k(t) = k, i(t) = i, m(t) = m\} = P_k(i, m, t)$$

составим прямую систему Колмогорова

$$\frac{\partial P_0(i, m, t)}{\partial t} = -(\lambda + i\sigma + \alpha)P_0(i, m, t) + \mu_1 P_1(i + 1, m - 1, t) + \mu_2 P_2(i + 1, m - 1, t),$$

$$\frac{\partial P_1(i, m, t)}{\partial t} = -(\lambda + \mu_1)P_1(i, m, t) + \lambda P_0(i - 1, m, t) + \lambda P_1(i - 1, m, t) + i\sigma P_0(i, m, t),$$

$$\frac{\partial P_2(i, m, t)}{\partial t} = -(\lambda + \mu_2)P_2(i, m, t) + \alpha P_0(i - 1, m, t) + \lambda P_2(i - 1, m, t). \quad (1)$$

Введем частичные характеристические функции, обозначив $j = \sqrt{-1}$,

$$H_k(u_1, u_2, t) = \sum_{i=0}^{\infty} \sum_{m=0}^{\infty} e^{ju_1 i} e^{ju_2 m} P_k(i, m, t).$$

Тогда систему (1) перепишем в виде

$$\begin{aligned}\frac{\partial H_0(u_1, u_2, t)}{\partial t} &= -(\lambda + \alpha)H_0(u_1, u_2, t) + j\sigma \frac{\partial H_0(u_1, u_2, t)}{\partial u_1} + \\ &+ \mu_1 e^{-ju_1} e^{ju_2} H_1(u_1, u_2, t) + \mu_2 e^{-ju_1} e^{ju_2} H_2(u_1, u_2, t), \\ \frac{\partial H_1(u_1, u_2, t)}{\partial t} &= \left(\lambda(e^{ju_1} - 1) - \mu_1 \right) H_1(u_1, u_2, t) + \\ &+ \lambda e^{ju_1} H_0(u_1, u_2, t) - j\sigma \frac{\partial H_0(u_1, u_2, t)}{\partial u_1}, \\ \frac{\partial H_2(u_1, u_2, t)}{\partial t} &= \left(\lambda(e^{ju_1} - 1) - \mu_2 \right) H_2(u_1, u_2, t) + \alpha e^{ju_1} H_0(u_1, u_2, t).\end{aligned}\quad (2)$$

Просуммируем уравнения системы (2), получим

$$\begin{aligned}\frac{\partial H(u_1, u_2, t)}{\partial t} &= \lambda(e^{ju_1} - 1)H(u_1, u_2, t) + \alpha(e^{ju_1} - 1)H_0(u_1, u_2, t) + \\ &+ \mu_1(e^{-ju_1} e^{ju_2} - 1)H_1(u_1, u_2, t) + \mu_2(e^{-ju_1} e^{ju_2} - 1)H_2(u_1, u_2, t).\end{aligned}$$

Таким образом, получим систему из четырех уравнений

$$\begin{aligned}\frac{\partial H_0(u_1, u_2, t)}{\partial t} &= -(\lambda + \alpha)H_0(u_1, u_2, t) + j\sigma \frac{\partial H_0(u_1, u_2, t)}{\partial u_1} + \\ &+ \mu_1 e^{-ju_1} e^{ju_2} H_1(u_1, u_2, t) + \mu_2 e^{-ju_1} e^{ju_2} H_2(u_1, u_2, t), \\ \frac{\partial H_1(u_1, u_2, t)}{\partial t} &= \left(\lambda(e^{ju_1} - 1) - \mu_1 \right) H_1(u_1, u_2, t) + \\ &+ \lambda e^{ju_1} H_0(u_1, u_2, t) - j\sigma \frac{\partial H_0(u_1, u_2, t)}{\partial u_1}, \\ \frac{\partial H_2(u_1, u_2, t)}{\partial t} &= \left(\lambda(e^{ju_1} - 1) - \mu_2 \right) H_2(u_1, u_2, t) + \alpha e^{ju_1} H_0(u_1, u_2, t), \\ \frac{\partial H(u_1, u_2, t)}{\partial t} &= \lambda(e^{ju_1} - 1)H(u_1, u_2, t) + \alpha(e^{ju_1} - 1)H_0(u_1, u_2, t) + \\ &+ \mu_1(e^{-ju_1} e^{ju_2} - 1)H_1(u_1, u_2, t) + \mu_2(e^{-ju_1} e^{ju_2} - 1)H_2(u_1, u_2, t).\end{aligned}\quad (3)$$

Асимптотика первого порядка

Основным содержанием работы является решение системы (3) методом асимптотического анализа в предельном условии большой задержки на орбите ($\sigma \rightarrow 0$) и растущего времени наблюдения за потоком ($t \rightarrow \infty$).

Теорема 1. Пусть $m(t)$ – число событий в выходящем потоке, наступивших за время t в RQ-системе M/M/1/1 с вызываемыми заявками, тогда

$$\lim_{\sigma \rightarrow 0} E \exp(ju_1 i(t) + ju_2 m(t)) = \exp\{ju_1 \kappa_1 + ju_2 at\}, \quad (4)$$

где среднее число заявок в системе κ определяется равенством

$$\kappa_1 = \frac{\lambda(\lambda\mu_2 + \alpha\mu_1)}{\mu_2(\mu_1 - \lambda)}, \quad (5)$$

среднее число событий в выходящем потоке

$$a = \lambda + \mu_2 \frac{\alpha(\mu_1 - \lambda)}{\mu_1(\mu_2 + \alpha)}. \quad (6)$$

Для более детального исследования процесса $m(t)$ рассмотрим асимптотику второго порядка.

Асимптотика второго порядка

Теорема 2. В условиях теоремы 1 выполняется равенство

$$\begin{aligned} \lim_{\sigma \rightarrow 0} E \exp\left(ju_1 \frac{\sigma i(t) - \kappa_1}{\sqrt{\sigma}} + ju_2 \frac{\sigma m(t) - at}{\sqrt{\sigma}}\right) = \\ = \exp\left\{\frac{(ju_1)^2}{2} \kappa_2 + \frac{(ju_2)^2}{2} a_2 t + ju_1 ju_2 a_{12}\right\}, \end{aligned} \quad (7)$$

где параметры κ_2 , a_{12} и a_2 определяются решением уравнений

$$\begin{aligned} 2\lambda f + 2\alpha f_0 - 2\mu_1 f_1 - 2\mu_2 f_2 + 2a \\ \lambda(g_0 + g_1 + g_2) - a(f_0 + f_1 + f_2) + \alpha g_0 + \mu_1 f_1 + \mu_2 f_2 - \mu_1 g_1 - \mu_2 g_2 - a = 0, \\ -2a(g_0 + g_1 + g_2) + a + 2\mu_1 g_1 + 2\mu_2 g_2 = a_2. \end{aligned} \quad (8)$$

Здесь $f_0, f_1, f_2, g_0, g_1, g_2$ определяются системами уравнений

$$\begin{aligned}
 -(\lambda + \alpha + \kappa_1)f_0 + \mu_1f_1 + \mu_2f_2 &= \kappa_2r_0 + a, \\
 (\lambda + \kappa_1)f_0 - \mu_1f_1 &= -\lambda(r_1 - r_0) - \kappa_2r_0, \\
 \alpha f_0 - \mu_2f_2 &= -\alpha r_0 - \lambda r_2.
 \end{aligned} \tag{9}$$

$$\begin{aligned}
 -(\lambda + \alpha + \kappa_1)g_0 + \mu_1g_1 + \mu_2g_2 &= ar_0 + a_{12}r_0 - a, \\
 (\lambda + \kappa_1)g_0 - \mu_1g_1 &= -a_{12}r_0 + ar_1 \\
 \alpha g_0 - \mu_2g_2 &= ar_2.
 \end{aligned} \tag{10}$$

Заключение

Исследование выходящего потока рассматриваемой RQ-системы проводится методом асимптотического анализа в условии большой задержки на орбите. В результате проведенных исследований было получено, что распределение вероятностей числа заявок, закончивших обслуживание в системе за некоторое время t , в этом асимптотическом условии является асимптотически нормальным.

ЛИТЕРАТУРА

1. *Artalejo J.R., Gomez-Corral A.* Retrial queueing systems: a computational approach. – Springer, 2008.
2. *Falin G.I., Templeton J.G.C.* Retrial Queues. – London: Chapman and Hall, 1997.
3. *Deslauriers A., L'Ecuyer P., Pichitlamken J., Ingolfsson A., Avramidis A.N.* Markov chain models of a telephone call center with call blending // *Computers and Operations Research*. – 2007. – V. 34. – P. 1616–1645.
4. *Burke P.J.* The output of queueing systems // *Operations Research*. – 1956. – V. 4. – P. 699–704.
5. *Reich E.* Waiting times when queues are in tandem // *Ann. Math. Statist.* – 1957. – V. 28. – No. 3. – P. 768.
6. *Finch P.D.* The output process of the queueing system M[G]1 // *J. Roy. Statist. Soc.* – 1959. – V. 21. – No. 2. – P. 375–380.
7. *Nazarov A., Paul S., Gudkova I.* Asymptotic analysis of Markovian retrial queue with two-way communication under low rate of retrials condition // *Proceedings 31st European Conference on Modelling and Simulation, ECMS, Budapest, 2017.* – P. 687–693.

МОДАЛЬНЫЙ СИНТЕЗ СИСТЕМЫ УПРАВЛЕНИЯ ДЛИНОЙ ОЧЕРЕДИ ПАКЕТОВ В БУФЕРЕ МАРШРУТИЗАТОРА

Е.А. Перепелкин

*Алтайский государственный технический университет
им. И.И. Ползунова, г. Барнаул, Россия*

Состояние маршрутизатора сети передачи данных, работающего по протоколу ТСР и алгоритму управления перегрузкой ТСР NewReno, определяется размером ТСР-окна $W(t)$ и длиной очереди пакетов $q(t)$ в буфере маршрутизатора [1]. Управление потоком пакетов осуществляется отбрасыванием пакетов с вероятностью $p(t)$. Задача управления потоком пакетов заключается в поддержании заданной длины очереди пакетов в буфере маршрутизатора.

Нелинейная модель маршрутизатора с учетом запаздываний по состоянию и управлению имеет следующий вид [1]:

$$\dot{W}(t) = \frac{1}{R(t)} - \frac{W(t)W(t - R(t))}{2R(t - R(t))} p(t - R(t)); \quad (1)$$

$$\dot{q}(t) = \frac{W(t)}{R(t)} N - C, \quad R(t) = T_p + \frac{q(t)}{C}, \quad (2)$$

где $R(t)$ – время между отправкой пакета и подтверждением о получении пакета; T_p – время прохождения сигнала по линии связи; N – количество потоков пакетов; C – пропускная способность маршрутизатора.

В положении равновесия должны выполняться соотношения

$$\frac{1}{\bar{R}} - \frac{\bar{W}^2}{2\bar{R}} \bar{p} = 0, \quad \frac{\bar{W}}{\bar{R}} N - C = 0, \quad \bar{R} = T_p + \frac{\bar{q}}{C}.$$

Здесь \bar{q} – заданное значение длины очереди пакетов, \bar{W} и \bar{p} – соответствующие значения ТСР-окна и сигнала управления. Следовательно, номинальное значение сигнала управления, обеспечивающего заданную длину очереди пакетов, равно

$$\bar{p} = \frac{2N^2}{\bar{R}^2 C^2}.$$

Рассмотрим возможность применения классического ПИД-регулятора для решения задачи управления длиной очереди пакетов. Линеаризованная модель объекта управления описывается передаточной функцией [1]

$$g_1(s) = \frac{b}{a(s)} e^{-sd},$$

где

$$b = -\frac{C^2}{2N}; \quad a(s) = s^2 + a_1 s + a_2, \quad a_1 = \frac{1}{\bar{R}} + \frac{2N}{C\bar{R}^2}, \quad a_2 = \frac{2N}{C\bar{R}^3}.$$

Передаточная функция ПИД-регулятора

$$g_2(s) = \frac{k(s)}{s}, \quad k(s) = k_i + k_p s + k_d s^2$$

содержит три коэффициента k_p , k_i , k_d соответственно при пропорциональной, интегральной и дифференцирующей составляющей управления. Значения этих коэффициентов должны быть найдены из условия асимптотической устойчивости замкнутой системы, передаточная функция которой имеет следующий вид:

$$g_3(s) = \frac{b}{a(s)s + bk(s)} e^{-s\bar{R}}.$$

Для асимптотической устойчивости замкнутой системы необходимо и достаточно, чтобы все полюсы замкнутой системы имели отрицательные действительные части [2]. Рассмотрим задачу о размещении полюсов замкнутой системы в левой части комплексной плоскости.

Полюсы замкнутой системы есть корни характеристического квази-полинома

$$c(s) = a(s)s + bk(s)e^{-\bar{R}s}. \quad (3)$$

В общем случае у квазиполинома бесконечно много корней. Выбирая коэффициенты регулятора можно назначить только три полюса. При этом коэффициенты регулятора будут зависеть от величины запаздывания \bar{R} .

Заметим, что коэффициенты квазиполинома (3) также зависят от \bar{R} , что усложняет решение рассматриваемой задачи.

Предлагаемый метод модального синтеза заключается в следующем. Сначала решим задачу расчета коэффициентов регулятора для системы с характеристическим полиномом

$$d(s) = a(s)s + bk(s). \quad (4)$$

Затем рассмотрим квазиполином

$$h(s) = a(s)s + bk(s)e^{-sL}, \quad L > 0. \quad (5)$$

Найдем критическое значение $L = L_{\max}$, при котором корни квазиполинома (5) пересекают мнимую ось [2]. Если $\bar{R} < L_{\max}$, то это означает, что корни квазиполинома (3) находятся в левой части комплексной плоскости и система является асимптотически устойчивой.

Решим первую задачу для полинома (4). Зададим параметр $r > 0$ и рассмотрим равенство полиномов

$$a(s)s + bk(s) = (s + r)^3.$$

Приравнявая коэффициенты при одинаковых степенях s , получим

$$k_p = \frac{3r^2 - a_2}{b}, \quad k_i = \frac{r^3}{b}, \quad k_d = \frac{3r - a_1}{b}.$$

Зафиксируем $r > 0$ и рассмотрим квазиполином (5) при $L > 0$. Предположим, что при некотором значении $L > 0$ корни $h(s)$ пересекают мнимую ось. Это означает, что существует действительное число $w > 0$, такое, что

$$h(jw) = 0, \quad h(-jw) = 0. \quad (6)$$

Из равенств (6) следует, что

$$a(jw)a(-jw)w^2 - b^2k(jw)k(-jw) = 0$$

или

$$v^3 + (a_1^2 - 2a_2 - (3r - a_1)^2)v^2 + (a_2^2 + 2r^3(3r - a_1) - (3r^2 - a_2)^2)v - r^6 = 0, \quad (7)$$

$$v = w^2.$$

Предположим, что уравнение (7) имеет одно положительное действительное решение и два отрицательных действительных решения, или одно положительное действительное решение и два комплексно сопряженных. Тогда значение $w > 0$, при котором справедливы равенства (6), равно $\bar{w} = \sqrt{v}$, где v – положительное решение уравнения (7).

Из равенства

$$e^{-j\bar{w}L} = -\frac{a(j\bar{w})j\bar{w}}{bk(j\bar{w})}$$

получим критическое значение L , при котором корни полинома (5) пересекают мнимую ось,

$$L_{\max} = \frac{1}{\bar{w}} \arccos \operatorname{Re} \left(-\frac{a(j\bar{w})j\bar{w}}{bk(j\bar{w})} \right).$$

Сравниваем \bar{R} и L_{\max} . Замкнутая система будет асимптотически устойчивой, если $\bar{R} < L_{\max}$.

Рассмотрим пример. Пусть параметры маршрутизатора равны: $C = 3750$; $N = 100$; $T_p = 0,24$. Заданное значение длины очереди пакетов $\bar{q} = 200$. При этом $\bar{W} = 11$, $\bar{R} = 0,29$.

Расчеты и моделирование выполнялись в системе компьютерной математики Scilab с использованием подсистемы компьютерного моделирования Xcos. На рис. 1 показана зависимость L_{\max} от r . При всех указанных значениях $1 \leq r \leq 5$ уравнение (7) имеет только одно действительное положительное решение. Прямая линия на графике обозначает значение \bar{R} . Таким образом, при $1 \leq r \leq 2,8$ система с регулятором асимптотически устойчива.

На рис. 2 показаны результаты моделирования нелинейной системы (1), (2) с регулятором в контуре управления. Коэффициенты регулятора, равные

$$k_p = -6,59 \cdot 10^{-5},$$

$$k_i = -4,8 \cdot 10^{-5},$$

$$k_d = 6,7 \cdot 10^{-6},$$

были получены при $r = 1,5$.

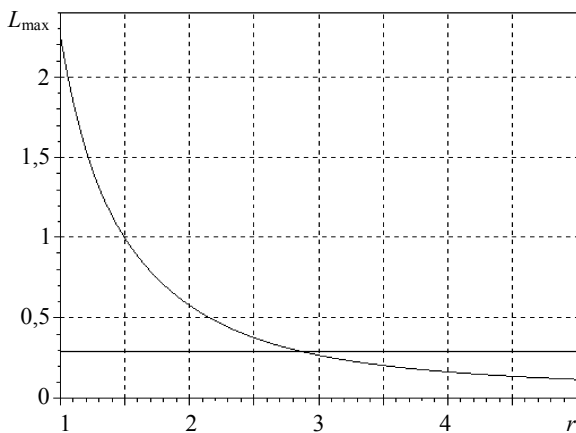


Рис. 1. Зависимость L_{\max} от r

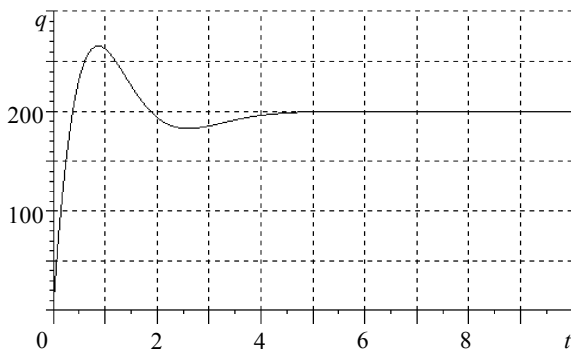


Рис. 2. Переходный процесс в системе с регулятором

ЛИТЕРАТУРА

1. *Alvarez T., Heras H., Reguera J.* Controller design for congestion control: some comparative studies // *Proceedings of the World Congress on Engineering*. – London, U.K., 2014. – V. 2. – P. 756–772.
2. *Silva G., Datta A., Bhattacharyya S.* PID Controllers for Time-Delay Systems. – Birkhauser Boston, 2005.

ИССЛЕДОВАНИЕ НЕМАРКОВСКОЙ НЕОДНОРОДНОЙ БЕСКОНЕЧНОЛИНЕЙНОЙ СИСТЕМЫ МАССОВОГО ОБСЛУЖИВАНИЯ В СЛУЧАЙНОЙ СРЕДЕ

Е.П. Полин, Е.А. Павлова

Томский государственный университет, г. Томск, Россия

Большинство современных технических систем, в том числе систем передачи информации и телекоммуникационных систем, функционируют в условиях изменяющейся внешней среды, которые носят как регулярный (смены времен года), так и случайный характер.

Возникновение в последние несколько десятилетий новых практических задач, связанных с появлением систем гибкого автоматического производства, в которых возможно отключение, переподключение и переналадка оборудования, систем управления запасами и экономических систем, информационно-вычислительных сетей и сетей связи, дало существенный толчок к развитию исследований систем с изменяемыми параметрами. Особенно актуальным представляется исследование таких систем при оценке ситуации в современных и перспективных инфокоммуникационных системах и сетях связи.

Такие задачи в теории массового обслуживания называются системами массового обслуживания (СМО) в случайной среде. Такие СМО более адекватно по сравнению с классическими марковскими системами отображают реальные процессы, связанные с изменяющейся во времени внешней случайной средой и реакцией самой системы на эти изменения.

Постановка задачи

Рассматривается бесконечнолинейная система массового обслуживания, функционирующая в случайной среде. Входящий МММР-поток задается матрицей инфинитезимальных характеристик \mathbf{Q} для цепи Маркова $I(t)$ и матрицей интенсивностей $\mathbf{\Lambda} = [\lambda_1, \dots, \lambda_L]$ [1].

Вектор стационарного распределения вероятностей \mathbf{R} удовлетворяет системе уравнений [2]:

$$\begin{cases} \mathbf{RQ} = \mathbf{0}, \\ \mathbf{Re} = 1. \end{cases}$$

Время обслуживания есть случайная величина, которая определяется одной из функций распределения $A_1(x), \dots, A_L(x)$. Выбор функции распределения $A_l(x)$ производится по значению процесса $l(t) = l$ в момент поступления сообщения и не меняется до конца обслуживания.

Уравнения Колмогорова

Обозначим $i_l(t)$ – число сообщений l -го типа, находящихся на обслуживании в системе в момент времени t , $l = 1, 2, \dots, L$.

Обозначим $\mathbf{i}(t) = [i_1(t), i_2(t), \dots, i_L(t)]$.

Многомерный процесс $\{\mathbf{i}(t), l(t)\}$ не является марковским. Поэтому воспользуемся методом многомерного динамического просеивания [3], который для данной задачи можно сформулировать следующим образом. Будем считать, что сообщение входящего потока, пришедшее в момент времени t , генерирует событие в l -м просеянном потоке с вероятностью $S_l(t)$, либо с вероятностью $1 - S_l(t)$ не генерирует события.

Вероятности просеивания $S_l(t)$ определяются выражениями

$$S_l(t) = 1 - A_l(T - t),$$

где T – некоторый фиксированный момент времени.

Обозначим через $\mathbf{n}(t) = [n_1(t), n_2(t), \dots, n_L(t)]$ число событий, наступивших соответственно в просеянном потоке до момента t .

Случайный процесс $\{l(t), n_1(t)\}$ является марковским, и для его распределения вероятностей $P(l, n_1, t) = P\{l(t) = l, n_1(t) = n_1\}$ можно записать следующую систему дифференциальных уравнений Колмогорова:

$$\frac{\partial P(l, \mathbf{n}, t)}{\partial t} = \lambda_l ((P(l, \mathbf{n}, t) - P(l, \mathbf{n} - \mathbf{e}_l, t)) S_l(t) + \sum_v q_{vl} P(v, \mathbf{n}, t)), \quad (1)$$

$$l = 1, 2, \dots, L.$$

Введем характеристические функции

$$H(\mathbf{u}, t) = \sum_{n_1=0}^{\infty} e^{ju_1 n_1} \dots \sum_{n_L=0}^{\infty} e^{ju_L n_L} P(l, \mathbf{n}, t), \quad j = \sqrt{-1}.$$

Применяя векторное обозначение $\mathbf{H}(\mathbf{u}, t) = [H(1, \mathbf{u}, t), \dots, H(L, \mathbf{u}, t)]$, запишем уравнение в матричном виде

$$\frac{\partial \mathbf{H}(\mathbf{u}, t)}{\partial t} = \mathbf{H}(\mathbf{u}, t) \{ \mathbf{Q} + \Lambda [(e^{j\mathbf{u}} - 1) \mathbf{S}(t)] \} \quad (2)$$

с начальным условием

$$\mathbf{H}(\mathbf{u}, t_0) = \mathbf{R}. \quad (3)$$

Здесь $\mathbf{S}(t)$ – диагональные матрицы с элементами $S_i(t)$.

Асимптотический анализ

Решение системы (2) будем искать в асимптотическом условии высокой интенсивности входящего потока [4]. Для достижения этого результата ММРР зададим представлением $(N\mathbf{Q}, N\Lambda)$, где числовой параметр $N > 0$ будет определять высокую интенсивность входящего потока. Асимптотический анализ будем проводить в условии $N \rightarrow \infty$.

Подставим вместо матриц \mathbf{Q}, Λ матрицы $N\mathbf{Q}, N\Lambda$, тогда уравнение (3) переписется в виде

$$\frac{1}{N} \frac{\partial \mathbf{H}(\mathbf{u}, t)}{\partial t} = \mathbf{H}(\mathbf{u}) \{ \mathbf{Q} + \Lambda [(e^{j\mathbf{u}} - 1) \mathbf{S}(t)] \}. \quad (4)$$

Асимптотический анализ первого порядка

В уравнении (5) выполним замены

$$\frac{1}{N} = \varepsilon, \quad \mathbf{u} = \varepsilon \mathbf{w}, \quad \mathbf{H}(\mathbf{u}, t) = \mathbf{F}_1(\mathbf{w}, t, \varepsilon). \quad (5)$$

Тогда уравнение (4) переписется в виде

$$\varepsilon \frac{\partial \mathbf{F}_1(\mathbf{w}, t, \varepsilon)}{\partial t} = \mathbf{F}_1(\mathbf{w}, t, \varepsilon) \{ \mathbf{Q} + \Lambda [(e^{j\varepsilon \mathbf{w}} - 1) \mathbf{S}(t)] \}, \quad (6)$$

$$\mathbf{F}(\mathbf{w}, t_0, \varepsilon) = \mathbf{R}. \quad (7)$$

Найдем асимптотическое решение $\mathbf{F}_1(\mathbf{w}, t) = \lim_{\varepsilon \rightarrow 0} \mathbf{F}_1(\mathbf{w}, t, \varepsilon)$ этой задачи.

Положим в выражении (7) $\varepsilon \rightarrow 0$. Получим

$$\mathbf{F}_1(\mathbf{w}, t) \mathbf{Q} = \mathbf{0}.$$

Сравнивая это уравнение с (1), представим функцию $F_1(\mathbf{w}, t)$ в виде

$$F_1(\mathbf{w}, t) = \mathbf{R}\Phi_1(\mathbf{w}, t), \quad (8)$$

Домножим обе части матричного уравнения (6) справа на единичный вектор \mathbf{e} , поделим на ε , подставим (8) и выполним предельный переход $\varepsilon \rightarrow 0$. В результате получим следующее дифференциальное уравнение:

$$\frac{\partial \Phi_1(\mathbf{w}, t)}{\partial t} = \Phi_1(\mathbf{w}, t) \mathbf{R} \Lambda \mathbf{j} \mathbf{w} \mathbf{S}(t) \mathbf{e}.$$

Обозначив $s(t) = \mathbf{R} \Lambda \left[\int_0^t \mathbf{S}(\tau) d\tau \right] \mathbf{e}$, с учетом начального условия (10)

получим

$$\Phi_1(\mathbf{w}, t) = e^{j\mathbf{w} s(t)}.$$

Тогда асимптотическое решение первого порядка будет иметь вид

$$F_1(\mathbf{w}, t) = \mathbf{R} e^{j\mathbf{w} s(t)}. \quad (11)$$

Для получения более точных характеристик проведен асимптотический анализ второго порядка. Показано, что асимптотическая характеристическая функция числа занятых приборов различного типа имеет вид многомерной гауссовской.

ЛИТЕРАТУРА

1. Назаров А.А., Моисеева С.П. Метод асимптотического анализа в теории массового обслуживания. – Томск: Изд-во НТЛ, 2006. – 112 с.
2. Назаров А.А., Терпугов А.Ф. Теория массового обслуживания: учебное пособие. – Томск: Изд-во НТЛ, 2004. – 228 с.
3. Моисеев А.Н., Назаров А.А. Бесконечнолинейные системы и сети массового обслуживания. – Томск: Изд-во НТЛ, 2015. 240 с.

КВАЗИОПТИМАЛЬНОЕ УПРАВЛЕНИЕ НЕОРДИНАРНЫМИ ПУАССОНОВСКИМИ ПОТОКАМИ

М.А. Рачинская, М.А. Федоткин

*Нижегородский государственный университет им. Н.И. Лобачевского,
г. Нижний Новгород, Россия*

Изучается система массового обслуживания с переменной структурой. Обслуживающее устройство (ОУ) может находиться в одном из нескольких возможных состояний, каждое из которых связано с особым режимом функционирования системы. Стратегия управления подобной системой подразумевает выбор длительностей пребывания ОУ в каждом из состояний (управляющих параметров), а также выбор алгоритма переключения состояний. Стратегия считается оптимальной, если она доставляет минимум среднему времени ожидания начала обслуживания произвольной заявкой системы. Работа посвящена поиску квазиоптимальных значений управляющих параметров при выбранном алгоритме управления с обратной связью путем имитационного моделирования.

Математическая модель

Рассмотрим (рис. 1) систему, на вход которой поступает два независимых конфликтных потока случайных заявок Π_1 и Π_2 . Поток Π_1 обладает большим приоритетом, а поток Π_2 – большой интенсивностью поступления заявок. Каждый из потоков Π_j (здесь и далее $j = 1, 2$) можно аппроксимировать неординарным пуассоновским потоком с параметрами λ_j – интенсивность пачек, p_j, q_j, s_j ($p_j + q_j + s_j = 1$) – вероятности поступления пачки из одной, двух, трех заявок [1]. Заявки потока Π_j ожидают начала обслуживания в бункере O_j , организованном согласно дисциплине очереди. Из бункера-накопителя заявки поступают на обслуживание, следуя экстремальной стратегии δ_j . Содержательно экстремальная стратегия означает отсутствие немотивированных простоев ОУ: заявки на обслуживание выбираются как можно в большем количестве, но не превышающем пропускной способности. В ОУ выделены внутренние состояния $\Gamma^{(1)}, \Gamma^{(2)}, \dots, \Gamma^{(5)}$ и задан алгоритм $s(\Gamma)$ их смены. В состоянии $\Gamma^{(1)}$ происходит обслуживание с интенсивностью μ_1 заявок потока Π_1 , в состояниях $\Gamma^{(3)}$ и $\Gamma^{(4)}$ с интенсивностью μ_2 обслуживаются

заявки потока Π_2 . В целях безопасности при переходе от обслуживания одного из конфликтных потоков к другому выделяются состояния для переналадки ОУ: $\Gamma^{(2)}$ и $\Gamma^{(5)}$. Обслуживание новых заявок в них не совершается, но допускается дообслуживание заявок, обслуживание которых было начато в предшествующем состоянии. В состоянии вида $\Gamma^{(k)}$, где $k \in M = \{1, 2, \dots, 5\}$, ОУ находится в течение промежутка времени длительностью T_k . При равной интенсивности обслуживания в состояниях $\Gamma^{(3)}$ и $\Gamma^{(4)}$ длительность пребывания ОУ в них различна: $T_3 > T_4$. По завершении такого промежутка происходит смена или продление текущего состояния. Граф управляющего алгоритма $s(\Gamma)$ смены состояний представлен на рис. 1. В силу неоднородности входных потоков предлагается организация обратной связи по числу ожидающих заявок в очереди потока с большим приоритетом. В случае, если это количество достигло пороговой величины $h_1 \geq 0$, то совершаются переходы типа 1 (рис. 1), которые имеют цель приблизить во времени начало обслуживания потока Π_1 . В противном случае осуществляются переходы типа 2, направленные на увеличение суммарного времени обслуживания потока Π_2 . Обслуженные заявки потока Π_j образуют выходной поток Π'_j .

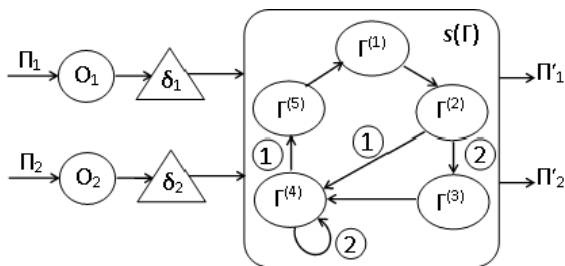


Рис. 1. Система управления конфликтными потоками и обслуживания их заявок

Пусть τ_i при $i = 0, 1, \dots$ есть случайные последовательные моменты, в которые происходит принятие решение о смене или продлении текущего состояния ОУ ($0 \leq \tau_0 < \tau_1 < \dots$). Вводятся следующие случайные величины: $\kappa_{j,i}$ – количество заявок в очереди O_j в момент τ_i , $\xi'_{j,i}$ – количество заявок потока Π_j , покинувших систему на промежутке $[\tau_i, \tau_{i+1})$, $\xi'_{j,-1}$ – количество заявок потока Π_j , покинувших систему на промежутке $[0, \tau_0)$. Пусть также случайный элемент Γ_i есть случайное состояние

ОУ на промежутке $[\tau_i, \tau_{i+1})$. В работе [2] была представлена математическая модель рассматриваемой системы обслуживания.

Лемма 1. Векторная последовательность

$$\{(\Gamma_i, \kappa_{1,i}, \kappa_{2,i}, \xi'_{1,i-1}, \xi'_{2,i-1}), i = 0, 1, \dots\} \quad (1)$$

с заданным начальным распределением вектора $(\Gamma_0, \kappa_{1,0}, \kappa_{2,0}, \xi'_{1,-1}, \xi'_{2,-1})$ является многомерной однородной цепью Маркова.

Благодаря этому результату, удастся аналитически найти условия существования в системе стационарного режима.

Введем для любых $k \in M, x_j \in \{0, 1, \dots\}, y_j \in \{0, 1, \dots, [\mu_j T_{2j-1}]\}, j = 1, 2$, и при $i = 0, 1, \dots$ обозначения для одномерных распределений цепи (1):

$$Q_i(\Gamma^{(k)}, x_1, x_2, y_1, y_2) = P(\Gamma_i = \Gamma^{(k)}, \kappa_{1,i} = x_1, \kappa_{2,i} = x_2, \xi'_{1,i-1} = y_1, \xi'_{2,i-1} = y_2).$$

Справедливы следующие утверждения.

Лемма 2. При любом начальном распределении многомерной цепи Маркова (1) либо для любых допустимых $k, x_j, y_j, j = 1, 2$,

$$\lim_{i \rightarrow \infty} Q_i(\Gamma^{(k)}, x_1, x_2, y_1, y_2) = 0$$

и стационарного распределения не существует либо существуют пределы

$$\lim_{i \rightarrow \infty} Q_i(\Gamma^{(k)}, x_1, x_2, y_1, y_2) = Q(\Gamma^{(k)}, x_1, x_2, y_1, y_2),$$

причем $Q(\Gamma^{(k)}, x_1, x_2, y_1, y_2) = 0$ для несущественных и $Q(\Gamma^{(k)}, x_1, x_2, y_1, y_2) > 0$ для существенных состояний вида $(\Gamma^{(k)}, x_1, x_2, y_1, y_2)$, и стационарное распределение существует и единственно.

Теорема 1. Для существования в системе стационарного режима по потоку Π_1 необходимо и достаточно выполнение неравенства

$$\lambda_1(3s_1 + 2q_1 + p_1)(\sum_{k=1}^5 T_k - T_3) - [\mu_1 T_1] < 0.$$

Теорема 2. При одновременном выполнении неравенств

$$\lambda_2(3s_2 + 2q_2 + p_2)(\sum_{k=1}^5 T_k - T_4) - [\mu_2 T_3] < 0,$$

$$\lambda_2(3s_2 + 2q_2 + p_2)\sum_{k=1}^5 T_k - [\mu_2 T_3] - [\mu_2 T_4] > 0$$

стационарного режима по потоку Π_2 в системе не существует.

Имитационная модель системы

Имитационная модель изучаемой системы построена на основе метода дискретных событий [3]. Состоянием системы в некоторый момент является набор следующих величин: количество заявок в очереди по каждому потоку, состояние ОУ. События следующих типов могут происходить в системе: 1) смена или решение о продлении текущего состояния ОУ; 2) приход пачки заявок по некоторому потоку; 3) завершение обслуживания некоторой заявки. Для каждого из указанных событий реализован обработчик, который меняет текущее состояние системы и обновляет собираемые в процессе имитации статистические данные. После фиксации значений входных параметров функционирование системы имитируется $N \geq 1$ раз с различными независимыми реализациями входных потоков. Таким образом получается N независимых реализаций случайного процесса имитации при неизменных входных параметрах. Будем считать, что новый цикл имитации начинается при очередной смене состояния ОУ на $\Gamma^{(1)}$. Процесс имитации заканчивается по завершении обслуживания хотя бы одного из заранее сгенерированных входных потоков с установленным максимальным количеством пачек или по завершении установленного количества циклов имитации. Пусть далее индекс $n = 0, 1, \dots$ указывает на номер цикла имитации, индекс $j = 1, 2, \dots$ – на номер потока, индекс $r \in \{1, 2, \dots, N\}$ – на номер реализации. В каждой реализации отслеживаются значения следующих характеристик системы: 1) $\gamma_{j,v}^r$ – время ожидания начала обслуживания заявкой с номером v потока P_j ; 2) $\alpha_{j,n}^r$ – количество заявок потока P_j , обслуженных за цикл с номером n ; 3) $\beta_{j,n}^r$ – количество заявок в очереди O_j в момент начала обслуживания потока P_j на цикле с номером n . Применяемый алгоритм определения момента окончания переходных процессов и достижения квазистационарного режима представлен в работе [4]. Так, пусть в реализации с номером r квазистационарный режим был достигнут на цикле с номером n_r^* по завершении обслуживания заявки с номером v_r^* потока P_{j^*} . Определяются величины $n^* = \max_{r \in \{1, 2, \dots, N\}} n_r^* + 1$ и $v^* = \max_{r \in \{1, 2, \dots, N\}} v_r^* + 1$ – номер первого цикла в квазистационарном режиме и номер первой заявки, обслуженной в квазистационарном режиме. С применением статистических критериев было по-

казано, что гипотезу о независимости и одинаковом распределении величин $\gamma_{j,v^*}^r, \gamma_{j,v^*+1}^r, \gamma_{j,v^*+2}^r, \dots$ при фиксированных r и j следует отвергнуть, в то время как для величин $\gamma_{j,v^*}^1, \gamma_{j,v^*}^2, \dots, \gamma_{j,v^*}^N$ при фиксированном j аналогичную гипотезу на уровне значимости 0,05 следует принять. В силу случайности момента достижения системой квазистационарного режима можно считать заявку с номером v^* в каждой реализации произвольной заявкой потока P_j в системе. Предлагается следующая оценка $\hat{M}\gamma$ для среднего времени ожидания начала обслуживания произвольной заявкой системы:

$$\hat{M}\gamma_j = \frac{1}{N} \sum_{r=1}^N \gamma_{j,v^*}^r, \quad \hat{M}\gamma = \frac{\sum_{j=1}^2 \lambda_j (3s_j + 2q_j + p_j) \hat{M}\gamma_j}{\sum_{j=1}^2 \lambda_j (3s_j + 2q_j + p_j)}.$$

Аналогичные рассуждения справедливы и для оценки $\hat{M}\beta$ среднего количества заявок в очереди в момент начала обслуживания произвольного потока системы:

$$\hat{M}\beta_j = \frac{1}{N} \sum_{r=1}^N \beta_{j,n^*}^r, \quad \hat{M}\beta = \frac{\sum_{j=1}^2 \lambda_j (3s_j + 2q_j + p_j) \hat{M}\beta_j}{\sum_{j=1}^2 \lambda_j (3s_j + 2q_j + p_j)}.$$

Квазиоптимальное управление

Величины $\hat{M}\gamma$ и $\hat{M}\beta$ позволяют оценить качество работы системы. Так, целью квазиоптимального управления системой является минимизация среднего времени ожидания $\hat{M}\gamma$. В качестве управляемых параметров могут выступать длительности пребывания ОУ в различных состояниях, величина порога h_1 . Приведем пример исследования системы при фиксированных длительностях $T_2 = 2$ и $T_5 = 2$ для состояний переналадки при значениях параметров $\mu_1 = 0,9$, $\mu_2 = 0,95$, $N = 10$. Поиск квазиоптимальных значений длительностей T_1 и T_3 происходит перебором в области, где, согласно теоремам 1 и 2, стационарный режим мо-

жет существовать. Перед началом работы алгоритма фиксируется также величина порога h_1 и длительность T_4 . Процесс имитации запускается в точках (T_1, T_3) , для которых сумма $\sum_{k=1}^5 T_k$ изменяется от $T_{\min} = 18$ до

$T_{\max} = 100$ с шагом a , при этом при фиксированной сумме $\sum_{k=1}^5 T_k$ длительность T_3 меняется с шагом b .

Вычисляется $\hat{M}\gamma$ для каждой из указанных точек, затем определяется точка (T_1^*, T_3^*) с минимальным значением $\hat{M}\gamma$ – она и назначается квазиоптимальной.

На рис. 2 и в таблице представлены результаты поиска квазиоптимальных значений для различных входных потоков. Отметим, что в случае, если потоки Π_1 и Π_2 становятся близки по интенсивности, переходы типа 2 алгоритма $s(\Gamma)$ совершаются редко (см. рис. 1), управляющий алгоритм приближается к циклическому, среднее время ожидания растет (см. рис. 2 и записи 6–8 в таблице). Если же входные потоки удовлетворяют изначальному предположению о неоднородности, то минимально достижимое значение величины $\hat{M}\gamma$ также существенно зависит от

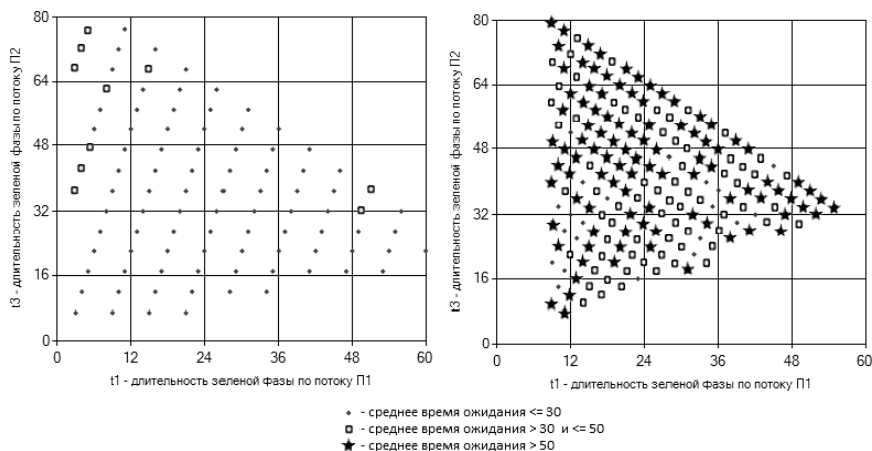


Рис. 2. Точки (T_1, T_3) подсчета для алгоритма поиска квазиоптимального управления: слева – входные потоки неоднородны по интенсивности, справа – входные потоки имеют равную интенсивность заявок

Результаты поиска квазиоптимальных значений

№	j	$\lambda_j(3s_j + 2q_j + p_j)$	h_1	T_4	a, b	(T_1^*, T_3^*)	\hat{M}_γ
1	1	0,144	4	3	5, 4	(15, 6)	2,116
	2	0,234					
2	1	0,144	4	8	5, 4	(19, 62)	3,624
	2	0,234					
3	1	0,144	10	3	5, 4	(20, 46)	2,821
	2	0,234					
4	1	0,144	7	3	5, 4	(27, 54)	3,86
	2	0,234					
5	1	0,144	10	4	6, 5	(15, 37)	8,432
	2	0,234					
6	1	0,216	10	5	6, 4	(21, 6)	8,889
	2	0,234					
7	1	0,216	10	10	6, 4	(42, 22)	9,932
	2	0,234					
8	1	0,216	10	10	6, 4	(44, 26)	10,398
	2	0,234					

порога h_1 и длительности T_4 . Чем меньше интенсивность потока Π_1 , тем меньше следует назначать порог h_1 . В свою очередь, чем больше интенсивность потока Π_1 , тем больше нужно назначать T_4 , поскольку в состоянии $\Gamma^{(3)}$ ОУ будет попадать редко и единственным состоянием обслуживания потока с большой интенсивностью останется $\Gamma^{(4)}$.

Заключение

Представлена имитационная модель неклассической системы обслуживания и управления независимыми конфликтными неординарными пуассоновскими потоками. Предложен алгоритм поиска квазиоптимальных значений управляющих параметров системы. Дальнейшее развитие имитационной модели связано с поиском алгоритма, позволяющего эффективно отыскивать квазиоптимальные значения в четырехмерном пространстве параметров h_1 , T_1 , T_3 и T_4 . Кроме того, по результатам исследований было замечено, что оценка \hat{M}_β имеет тенденцию к нарастанию и для некоторых точек (T_1, T_3) в области существования стационарного режима по теоремам 1 и 2. Это указывает на необходимость аналитического поиска дополнительных условий стационарности и дальнейшего сужения области изменения управляемых параметров.

ЛИТЕРАТУРА

1. *Fedotkin M., Rachinskaya M.* Parameters estimator of the probabilistic model of moving batches traffic flow // Distributed Computer and Communication Networks. Ser. Communications in Computer and Information Science. – 2014. – V. 279. – P.154–168.
2. *Федоткин М.А., Рачинская М.А.* Модель функционирования системы управления и обслуживания потоков разной интенсивности и приоритетности // Вестник Волжской государственной академии водного транспорта. – 2016. – № 48. – С. 62–69.
3. *Averill M. Law, W. David Kelton.* Simulation modeling and analysis. – McGraw-Hill, 2000. – 760 p.
4. *Федоткин М.А., Рачинская М.А.* Имитационная модель циклического управления конфликтными неординарными пуассоновскими потоками // Вестник Волжской государственной академии водного транспорта. – 2016. – № 47. – С. 43–51.

МЕТОД ПРОГОНКИ ДЛЯ МНОГОКАНАЛЬНЫХ СИСТЕМ ОБСЛУЖИВАНИЯ

Ю.И. Рыжиков

*Военно-космическая академия им. А.Ф. Можайского,
г. Санкт-Петербург, Россия*

*Санкт-Петербургский Институт информатики
и автоматизации РАН, г. Санкт-Петербург, Россия*

Фазовая аппроксимация многоканальных систем

Эффективные численные метода расчета актуальных для многих применений многоканальных систем массового обслуживания (СМО) стали возможными лишь после представления немарковских распределений в виде системы фаз с экспоненциально распределенной задержкой в каждой [1, 2].

Общепринята «фольклорная» рекомендация: при коэффициенте вариации $v > 1$ применять H_2 -аппроксимацию, в противном случае – эрлангову. Последняя при малых v приводит к чрезвычайно быстрому росту числа микросостояний по числу каналов. К тому же, она позволяет выравнять лишь первый и приближенно – второй момент исходного распределения.

H_2 -распределение, выравнивающее *три* момента, порождает диаграмму переходов ширины всего $n + 1$. При $v < 1/\sqrt{2}$ его параметры становятся комплексными, но конечные результаты имеют традиционный вероятностный смысл и хорошо согласуются с полученными иными методами. По этим причинам далее рассматривается модель $M/H_2/n$.

Работа такой системы может быть интерпретирована как обслуживание неоднородного потока заявок. «Ключ» микросостояния указывает количество находящихся в каналах обслуживания заявок каждого типа. Завершение обслуживания с вероятностями $\{y_i\}$ в зависимости от типа выбранной из очереди заявки приводит в одно из микросостояний вышележащего яруса.

Приведем общую для сопоставляемых методов постановку задачи. Обозначим через S_j множество всех возможных микросостояний системы, при которых на обслуживании находится ровно j заявок, а через σ_j – количество элементов в S_j . Далее в соответствии с диаграммой переходов построим матрицы интенсивностей переходов: $A_j[\sigma_j \times \sigma_{j+1}]$ – в S_{j+1} (прибытие заявки), $B_j[\sigma_j \times \sigma_{j-1}]$ – в S_{j-1} (полное завершение обслуживания заявки), $D_j[\sigma_j \times \sigma_j]$ – ухода из состояний яруса j (в квадратных скобках указан размер матриц).

Введем векторы-строки $\gamma_j = \{\gamma_{j,1}, \gamma_{j,2}, \dots, \gamma_{j,\sigma_j}\}$ нахождения СМО в состоянии (j, i) , $j = 0, 1, \dots$. Теперь можно записать векторно-матричные уравнения баланса переходов между состояниями

$$\begin{aligned} \gamma_0 D_0 &= \gamma_1 B_1, \\ \gamma_j D_j &= \gamma_{j-1} A_{j-1} + \gamma_{j+1} B_{j+1}, \\ j &= 1, 2, \dots \end{aligned} \quad (1)$$

Итерационный метод

Такахаси и Таками [3] предложили алгоритм итерационного расчета подобных систем, центральной идеей которого является переход к расчету условных (нормированных к единице) вероятностей микросостояний системы $\{\bar{\gamma}_{j,i}^{(m)}\}$ для фиксированного числа заявок в системе (j -й ярус диаграммы) и параллельно – вычисление отношений $\{x_j\}$ суммарных безусловных вероятностей на смежных ярусах, $x_j = p_{j+1}/p_j$, $j = 0, 1, \dots, N$. В итерации номер m вектор условных вероятностей $\bar{\gamma}_j^{(m)}$ для каждого яруса при прогонке сверху вниз выражается через $\bar{\gamma}_{j-1}^{(m)}$ и $\bar{\gamma}_{j+1}^{(m-1)}$. При обсчете N -го яруса применяется замыкающее систему уравнений приближенное равенство

$$\bar{\gamma}_{N+1}^{(m-1)} \approx \bar{\gamma}_{N-1}^{(m)}. \quad (2)$$

Расчетная схема [3] была усовершенствована в [4, 5].

Матрично-геометрическая прогрессия

Применительно к расчету *разомкнутых* систем обслуживания применяется развиваемый М. Ньютоном и его последователями [6–8] метод *матрично-геометрической прогрессии* (МГП). Здесь вектора вероятностей микросостояний полностью занятой системы представляются как

$$\gamma_j = \gamma_n R^{j-n}, \quad j = n, n+1, \dots, \quad (3)$$

где R – матричный знаменатель прогрессии. Легко показать, что он должен удовлетворять матричному квадратному уравнению

$$R^2 B - R D + A = 0. \quad (4)$$

Найдя этот знаменатель и имея вектор p_n вероятностей микросостояний n -го яруса, можно согласно (3) вычислить вероятности микросостояний для $j > n$.

Реализации МГП различаются методом расчета знаменателя прогрессии (варианты простой итерации и метода Ньютона) и способом расчета векторов вероятностей микросостояний для ярусов $j = \overline{0, n-1}$. Подробное их сопоставление проведено опять же в [5]. Для дальнейшего усовершенствования выбран вариант МГП с расчетом знаменателя согласно

$$R = A(D - RB)^{-1}. \quad (5)$$

Метод прогонки

Сходимость обоих рассмотренных методов по числу каналов быстро ухудшается. Обнаружилось, что в методе МГП необходимость решения системы линейных уравнений относительно неизвестных, различающихся на много порядков, приводит к появлению отрицательных начальных вероятностей. Соответственно размерность решаемых задач ограничивается. Эти ограничения снимает описываемый ниже *метод прогонки*.

Метод прогонки давно и успешно используется при решении систем линейных трехдиагональных алгебраических уравнений с *числовыми* коэффициентами (в [9] приведены соответствующие примеры). В зада-

чах теории очередей для процессов с переходами между микросостояниями только соседних ярусов (QBD) глобальная матрица содержит ненулевые *матрицы* интенсивностей перехода тоже на трех диагоналях, что дает основание для использования соответствующих аналогий.

Введем матрицы *обратного* пересчета векторов вероятностей γ_j согласно

$$\gamma_j = \gamma_{j+1} F_j, \quad j = \overline{0, N-1}. \quad (6)$$

Из первого уравнения системы (1) следует, что

$$F_0 = B_1 D_0^{-1} = B_1 / \lambda. \quad (7)$$

Для последующих ярусов из той же системы имеем

$$\gamma_j D_j = \gamma_{j-1} A_{j-1} + \gamma_{j+1} B_{j+1}.$$

Воспользовавшись (6), его можно переписать в виде

$$\gamma_j (D_j - F_{j-1} A_{j-1}) = \gamma_{j+1} B_{j+1},$$

откуда следует рекуррентный пересчет матриц на этапе прямой прогонки

$$F_j = B_{j+1} (D_j - F_{j-1} A_{j-1})^{-1}, \quad j = \overline{1, n}. \quad (8)$$

С учетом правил формирования матриц интенсивностей переходов между ярусами для последующих ярусов имеем

$$F_j = B_{n+1} (D_n - \lambda F_{j-1})^{-1}, \quad j = \overline{n+1, N-1}. \quad (9)$$

Здесь N – предельный индекс обсчитываемых ярусов, $N > n$.

Будем обозначать предельные при $j \rightarrow \infty$ матрицы, векторы условных (нормированных к единице в пределах яруса) вероятностей и отношения смежных вероятностей ярусов прежними символами, но без индексов. Если существует $\gamma_\infty = \gamma$, то существуют и предельные значения отношений смежных вероятностей x и $z = 1/x$, причем из (1) следует

$$\gamma = (x^{-1} \lambda \gamma + x \gamma B) D^{-1} = \gamma (x^{-1} \lambda I + x B) D^{-1} = \gamma Q.$$

Обозначим $(Q - I)_1$ матрицу, полученную из $Q - I$ заменой ее первой строки на единичную, и положим $\delta_1 = \{1, 0, 0, \dots, 0\}^T$. Тогда заведомо $\det(Q - I) \neq 0$ и искомый вектор получается как решение системы линейных алгебраических уравнений

$$\gamma(Q - I)_1 = \delta_1. \quad (10)$$

Выполняя обратную прогонку для $j = N - 1, N - 2, \dots, 0$ согласно (7), можно получить масштабированные векторы вероятностей микросостояний и для каждого j – кумулянтные вероятности ярусов. Затем выполняется их перенормировка к единице.

Для предельного отношения смежных вероятностей в моделях $A/B/n$ Такахаси [10] предложил весьма сложный алгоритм. Довольно близкие к нему результаты дает формула из [11]

$$x_\infty = \rho^{2/(v_A^2 + v_B^2)}. \quad (11)$$

Численный эксперимент

Приведем результаты расчета системы $M/H_2/5$ (табл. 1). Допуск $\varepsilon = 10^{-8}$ для итерационного метода Iter определял максимальный модуль уточнения отношений смежных вероятностей $\{x_j\}$, а для метода матрично-геометрической прогрессии MGP – максимальную из сумм модулей столбцов поправки к знаменателю прогрессии. Столбец SWP соответствует методу прогонки. Согласие результатов следует признать вполне удовлетворительным, что подтверждает корректность как расчетных зависимостей, так и их программных реализаций.

Теперь сопоставим трудоемкости тестируемых методов (табл. 2). Здесь количество обсчитываемых ярусов назначалось как $n + 20$. В этой таблице через слэш записаны число итераций и время счета в секундах. Нулевое время счета указывает трудоемкость ниже порога системных часов (0,01 с). Прочерками отмечены случаи расходимости итераций, а звездочками – появление в МГП отрицательных вероятностей начальных состояний.

Т а б л и ц а 1

Вероятности состояний системы $M/H_2/5$

j	$\beta = 3,0$			$\beta = 0,25$		
	Iter	MGP	SWP	Iter	MGP	SWP
0	1,2440e-2	1,2440e-2	1,2440e-2	1,3799e-2	1,3790e-2	1,3786e-2
1	5,0313e-2	5,0312e-2	5,0312e-2	5,4703e-2	5,4670e-2	5,4643e-2
2	1,0235e-1	1,0235e-1	1,0235e-1	1,0695e-1	1,0693e-1	1,0689e-1
3	1,4049e-1	1,4049e-1	1,4049e-1	1,3527e-1	1,3533e-1	1,3529e-1
4	1,4851e-1	1,4851e-1	1,4851e-1	1,2082e-1	1,2094e-1	1,2091e-1
5	1,3374e-1	1,3375e-1	1,3375e-1	7,7329e-2	7,7447e-2	7,7424e-2
6	1,0767e-1	1,0767e-1	1,0767e-1	5,6147e-2	5,6264e-2	5,6247e-2
7	8,2209e-2	8,2201e-2	8,2201e-2	4,4571e-2	4,4689e-2	4,4676e-2
8	6,1082e-2	6,1064e-2	6,1064e-2	3,7444e-2	3,7564e-2	3,7553e-2
9	4,4732e-2	4,4705e-2	4,4705e-2	3,2543e-2	3,2665e-2	3,2655e-2
10	3,2507e-2	3,2471e-2	3,2471e-2	2,8850e-2	2,8973e-2	2,8964e-2
11	2,3525e-2	2,3485e-2	2,3485e-2	2,5874e-2	2,5997e-2	2,5989e-2
12	1,6989e-2	1,6946e-2	1,6946e-2	2,3363e-2	2,3484e-2	2,3477e-2
13	1,2255e-2	1,2213e-2	1,2213e-2	2,1179e-2	2,1299e-2	2,1293e-2
14	8,8348e-3	8,7962e-3	8,7962e-3	1,9246e-2	1,9363e-2	1,9357e-2
15	6,3675e-3	6,3330e-3	6,3330e-3	1,7515e-2	1,7628e-2	1,7623e-2
16	4,5886e-3	4,5587e-3	4,5587e-3	1,5953e-2	1,6062e-2	1,6057e-2
17	3,3065e-3	3,2812e-3	3,2812e-3	1,4539e-2	1,4643e-2	1,4638e-2
18	2,3825e-3	2,3616e-3	2,3616e-3	1,3255e-2	1,3353e-2	1,3349e-2
19	1,7167e-3	1,6997e-3	1,6997e-3	1,2087e-2	1,2179e-2	1,2175e-2
20	1,2370e-3	1,2233e-3	1,2233e-3	1,1023e-2	1,1110e-2	1,1106e-2

Т а б л и ц а 2

Трудоемкость обсе́та модели $M/H_2/n$

n	$\beta = 3,0$			$\beta = 0,25$		
	Iter	MGP	SWP	Iter	MGP	SWP
5	48/0	11/0	0/0	120/0	42/0	0/0
10	94/0,016	20/0	0/0	202/0,031	55/0/016	0/0
20	170/0,078	34*/0,047	0/0,031	320/0,172	70/0,062	0/0,031
30	232/0,0312	-/-	0/0,094	500/0,672	80/0,219	0/0,078
50	-/-	-/-	0/0,469	778/4,859	-/-	0/0,485
70	-/-	-/-	0/1,578*	782/14,532	-/-	0/1,562
100	-/-	-/-	0/5,515*	/78,500	-/-	0/5,500

Заключение

Из анализа вычислительных схем и сопоставления результатов расчета (в том числе не представленных в данной статье из-за ограниченности места) вытекают следующие *выводы*:

1. Метод прогонки (SWP) применим только для QBD-процессов (с переходами между соседними ярусами диаграммы) и неограниченной длине очереди, но в этих условиях существенно превосходит упомянутые альтернативы, особенно при больших n . В отличие от других подходов, SWP *не требует итерационного счета*. Для H_2 -обслуживания с вещественными параметрами он применим по крайней мере для числа каналов $n = 100$, с комплексными при коэффициенте вариации обслуживания $1/\sqrt{3} = 0,577$ – до 50.

2. Итерационный метод легко модифицируется применительно к системам с интенсивностью входящего потока, зависящей от состояния системы, и к системам с ограниченной очередью, легко обобщается на системы с потоком групповых заявок. Благодаря работе с векторами *относительных* вероятностей и наличию этапа агрегации на каждом слое его точность практически не зависит от числа N обсчитываемых ярусов.

3. Метод матрично-геометрической прогрессии (МГП) предполагает итерационный расчет знаменателя прогрессии, сходимость которого от количества ярусов не зависит. Он принципиально применим только для QBD-процессов. Метод удобен при работе с бесконечными суммами вероятностей и при числе каналов $n < 30$. Серьезной проблемой является расчет вероятностей состояний для $j \leq n$ (здесь можно воспользоваться итерациями).

4. Некоторые из этих выводов могут быть пересмотрены при увеличении разрядности вычислений (они выполнялись в *удвоенной* разрядной сетке).

5. Все перечисленные методы в принципе могут быть обобщены применительно к рекуррентному потоку, распределение интервалов между заявками которого опять же аппроксимируется гиперэкспонентой H_2 .

ЛИТЕРАТУРА

1. *Osogami T., Harchol-Balter M.* A Closed-form solution for mapping general distribution to minimal PH distributions // Computer Performance Evaluation. Modeling Technique and Tools: Proc. of the 13th Internat. Conf., Urbana, IL, USA, Sept. 2003. – P. 200–217.
2. *Brandwajn A., Begin T.* Preliminary results on simple approach to G/G/c-Like queues // Analytical and Stochastic Modeling Techniques and Applications: Proc. of the 16th Internat. Conf. ASMTA, Madrid, Spain, June 2009. – P. 159–173.
3. *Takahashi Y., Takami Y.* A numerical method for the steady-state probabilities of a GI/G/c queueing system in a general class // J. Operat. Res. Soc. of Japan. – 1976. – V. 19. – No. 2. – P. 147–157.
4. *Рыжиков Ю.И.* Алгоритм расчета многоканальной системы с эрланговским обслуживанием // Автоматика и телемеханика. – 1980. – Т. 5. – С. 30–37.
5. *Рыжиков Ю.И.* Развитие и сопоставление методов расчета многоканальных систем обслуживания // Труды XII Всероссийского совещания по проблемам управления. – М.: ИПУ РАН, 2014. – С. 5208–5219.
6. *Бочаров П.П., Печинкин А.В.* Теория массового обслуживания: учебник. – М.: Изд-во РУДН, 1995. – 529 с.
7. *Вишневский В.М.* Теоретические основы проектирования компьютерных сетей. – М.: Техносфера, 2003. – 512 с.
8. *Daigle J.N.* Queuing Theory with Applications to Packet Telecommunication. – Boston: Springer, 2005. – 326 p.
9. *Демидович Б.П., Марон И.А., Шувалова Э.З.* Численные методы анализа. – М.: Физматгиз, 1963. – 400 с.
10. *Takahashi Y.* Asymptotic exponentiality of the tail of the waiting time distribution in a Ph/Ph/c queue // Adv. in Applied Probability. – 1981. – V. 13. – P. 619–630.
11. *Башарин Г.П., Бордукова В.Т.* Некоторые точные и приближенные результаты для многолинейных систем массового обслуживания с ограниченным накопителем // Теория телетрафика и информационные сети: сб. – М.: Наука, 1977. – С. 6–15.

МЕТОД МОДЕЛИРОВАНИЯ ИНФОРМАЦИОННО-ВЫЧИСЛИТЕЛЬНОЙ СЕТИ СЕТЬЮ МАССОВОГО ОБСЛУЖИВАНИЯ С НЕСТАЦИОНАРНОЙ СТРУКТУРОЙ

И.Е. Тананко, Н.П. Фокина

*Саратовский государственный университет
им. Н.Г. Чернышевского, г. Саратов, Россия*

Сети массового обслуживания с ненадежными элементами [1–3], и поэтому имеющие нестационарную структуру, широко используются в качестве математических моделей дискретных сетевых стохастических систем, в которых один или несколько элементов в процессе функционирования могут выходить из строя и восстанавливаться. К таким системам относятся системы обработки и передачи информации. Интерес представляет построение моделей таких систем, получение характеристик и изучение свойств этих моделей.

Например, в работе [4] представлена модель сети передачи данных с ненадежными каналами и исследуется эффективность различных алгоритмов маршрутизации пакетов. Методу анализа открытой сети массового обслуживания, в которой в момент отказа прибора уничтожаются все требования, находящиеся в системе обслуживания, посвящена работа [5]. Предлагается использовать эту сеть обслуживания в качестве математической модели компьютерных систем и сетей, подверженных вирусным атакам.

В данной работе рассматривается модель информационно-вычислительной сети (ИВС) с ненадежными каналами передачи данных в виде открытой экспоненциальной сети массового обслуживания с нестационарной структурой и управлением. Разработан метод получения стационарных характеристик и приведен пример анализа этой сети обслуживания.

Модель информационно-вычислительной сети

Рассмотрим информационно-вычислительную сеть, состоящую из периферийного оборудования и сети передачи данных (СПД). Периферийное оборудование является источником и приемником передаваем-

мых данных. СПД представляет собой множество маршрутизаторов, соединенных каналами передачи данных (КПД).

Любой из КПД имеет по одному входу и по одному выходу, которые соединены со входами и выходами маршрутизаторов. Каждый маршрутизатор имеет два или более входов и два или более выходов. Некоторые маршрутизаторы соединены с периферийным оборудованием.

Информационно-вычислительные сети подвержены динамическому изменению структуры. Одной из таких причин является негативное воздействие окружающей среды на каналы передачи данных, которое приводит к потере передаваемых по сети пакетов и к необходимости повторной передачи [6, 7]. Будем называть такие каналы ненадежными.

Один из методов надежной доставки пакетов от отправителя к получателю реализован на сетевом уровне эталонной модели взаимодействия открытых систем. Механизм его состоит в следующем. Маршрутизатор передает пакет в соответствии со своей таблицей маршрутизации следующему маршрутизатору. Если через определенный интервал времени маршрутизатор-отправитель не получает подтверждение о полученном пакете от соседнего маршрутизатора, то производится повторная передача пакета посредством протокола ICMP.

Число k повторных передач пакетов за единицу времени Δ от маршрутизатора-отправителя к маршрутизатору-получателю по каналу связи является одним из параметров формирования таблицы маршрутизации и используется маршрутизатором-отправителем в течение следующего интервала времени Δ . Если $k \geq K$, где K – заданный параметр, то таблица маршрутизации изменяется так, что в следующий интервал Δ пакеты не передаются по ненадежному каналу [7].

Передача пакетов по ненадежному каналу передачи данных возобновляется через случайный интервал времени кратный Δ и определяется информационно-вычислительной сетью как появление нового устройства сети [7].

Информационно-вычислительная сеть является достаточно надежной системой, поэтому полагаем, что все маршрутизаторы являются надежными, а к выходу каждого маршрутизатора подключено не более одного ненадежного КПД. Таким образом, отключение любого из КПД сети не приводит к нарушению связности СПД.

В качестве модели ИВС будем использовать открытую экспоненциальную сеть массового обслуживания. Передаваемым по ИВС пакетам в сети массового обслуживания поставим в соответствие требования. Все

периферийное оборудование, которое является источником и приемником пакетов, отобразим в модели пуассоновским источником требований одного класса с интенсивностью потока λ_0 .

Каждому маршрутизатору, а также каналу передачи данных СПД в модельной сети обслуживания поставим в соответствие систему массового обслуживания типа $M/M/1$. Для каждого ненадежного канала введем дополнительную систему массового обслуживания типа $M/M/1$ для отображения задержек, связанных с обработкой повторно передаваемых пакетов. Эта система образует обратную связь в цепи двух последовательно соединенных систем обслуживания, отображающих маршрутизатор и ненадежный канал передачи данных.

Пусть модельная сеть массового обслуживания состоит из источника требований S_0 и L систем массового обслуживания S_i с интенсивностями обслуживания μ_i , $i = 1, \dots, L$. Введем обозначения: $\mathbf{L} = \{1, \dots, L\}$, $\mathbf{M} \subset \mathbf{L}$ – подмножество номеров систем обслуживания, отображающих ненадежные КПД.

Рассмотрим два произвольных маршрутизатора с номерами l и r , соединенных ненадежным КПД с номером m (рис. 1). Обозначим через p_m – вероятность потери пакета в канале m . Ненадежному КПД с номером m и маршрутизаторам с номерами l и r поставим в соответствие системы массового обслуживания S_m , S_l и S_r , $m, l, r \in \mathbf{L}$. Дополнительную систему массового обслуживания, отображающую задержки, связанные с обработкой повторно передаваемых пакетов, обозначим S_n , $n \in \mathbf{L}$.

Переходы требований между системами обслуживания сети определяются маршрутной матрицей $\Theta = (\theta_{ij})$, $i, j = 0, \dots, L$, в которой $\theta_{mn} = p_m$, $\theta_{mr} = 1 - p_m$, $\theta_{nl} = 1$.

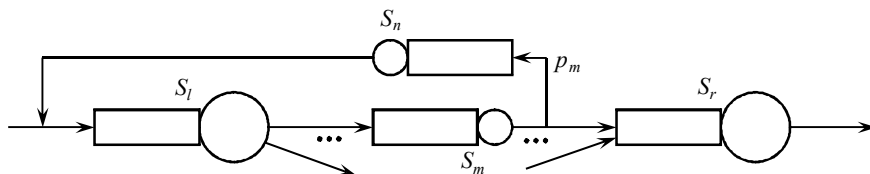


Рис. 1. Фрагмент модельной сети

Пусть k_m – число требований, возвращенных за интервал времени Δ из системы обслуживания S_m в систему обслуживания S_l для повторной передачи.

Введем управление маршрутизацией в сети массового обслуживания, которое заключается в выключении и включении потока требований в систему S_m . Будем называть эту систему управляемой. Если $k_m \geq K_m$, где параметр управления маршрутизацией K_m обозначает предельное число возвращенных требований за интервал времени Δ , то поток требований в систему S_m перекрывается. Последующее включение потока требований в систему S_m производится через кратный Δ интервал времени, являющийся геометрически распределенной случайной величиной с параметром β_m . Будем считать, что вероятность p_m не изменяется после подключения системы S_m .

Обозначим вектор состояния управляемых систем обслуживания $w = (w_m)$, $m \in \mathbf{L}$, где $w_m = 1$, если в систему S_m поступает поток требований, и $w_m = 0$, если поток требований в систему S_m перекрыт. Множество всех векторов w обозначим W .

Пусть $\Theta(w)$ – маршрутная матрица сети при условии, что состояние управляемых систем определяется вектором w . Рассмотрим два состояния структуры сети $w, w' \in W$, отличающиеся только тем, что $w_m = 1$, $w'_m = 0$. В момент выключения потока из системы S_l , $l \in \mathbf{L} \setminus \mathbf{M}$, в систему S_m , $m \in \mathbf{M}$, вектор структуры сети w преобразуется в w' , а маршрутная матрица $\Theta(w)$ преобразуется в матрицу $\Theta(w')$ так, что требования не поступают в систему S_m до включения потока в эту систему. Элементы матрицы $\Theta(w')$ равны соответствующим элементам матрицы $\Theta(w)$ за исключением элементов l -й строки, которые определяются по формулам

$$\theta_{li}(w') = \frac{\theta_{li}(w)}{1 - \theta_{lm}(w)}, \quad i = 0, \dots, L, \quad i \neq m,$$

$$\theta_{lm}(w') = 0.$$

Требуется найти вероятностно-временные характеристики сети обслуживания.

Относительные интенсивности потоков требований $\omega(w)$ при условии w определим из решения системы уравнений $\omega(w)\Theta(w) = \omega(w)$ с условием нормировки $\sum_{i=0}^L \omega_i(w) = 1$, для всех $w \in W$. Тогда интенсивности потоков требований в системы S_i

$$\lambda_i(w) = \omega_i(w) \frac{\lambda_0}{\omega_0(w)}, \quad i = 1, \dots, L, \quad w \in W.$$

Вероятность того, что за интервал Δ из системы S_m в систему S_l поступит s требований, при условии w

$$P(k_m(w) = s) = \frac{(p_m \lambda_m(w) \Delta)^s}{s!} e^{-p_m \lambda_m(w) \Delta}, \quad s = 0, 1, 2, \dots$$

Тогда

$$P(k_m(w) \geq K_m) = 1 - \sum_{s=0}^{K_m-1} \frac{(p_m \lambda_m(w) \Delta)^s}{s!} e^{-p_m \lambda_m(w) \Delta}.$$

Обозначим $\alpha_m(w) = P(k_m(w) \geq K_m)$. Очевидно, что $\alpha_m(w) > 0$ при $w_m = 1$ и $\alpha_m(w) = 0$ при $w_m = 0$.

Будем считать, что длительность переходного режима в сети массового обслуживания, вызванного включением или выключением потока требований в систему S_m , много меньше длительности Δ . Поэтому в дальнейшем будем пренебрегать переходным режимом, полагая, что стационарный режим в сети обслуживания наступает сразу после включения или выключения потока требований в систему S_m , $m \in \bar{L}$.

Рассмотрим два состояния $w, w' \in W$, отличающиеся только тем, что $w_m = 1$, $w'_m = 0$. Тогда стационарные вероятности состояний управления системы S_m , $m \in \bar{L}$, при условии, что структура сети задана векторами w и w' , определяются следующим образом:

$$\pi_m(w) = \beta_m / (\alpha_m(w) + \beta_m), \quad \pi_m(w') = \alpha_m(w) / (\alpha_m(w) + \beta_m).$$

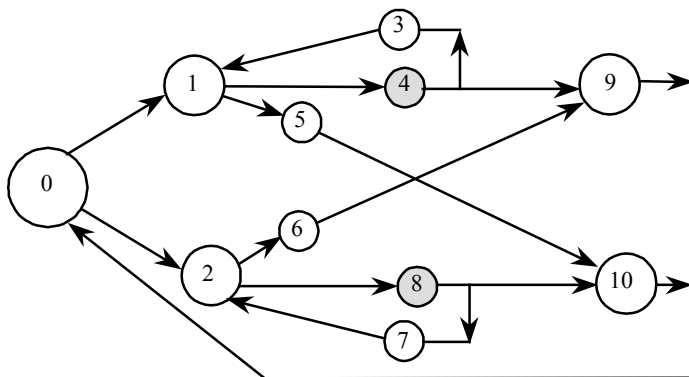


Рис. 2. Структура модельной сети обслуживания

Остальные маршрутные матрицы $\Theta(0,1)$, $\Theta(1,0)$, $\Theta(0,0)$ отличаются от Θ только вероятностями перехода в соответствующие ненадежные системы, которые согласно топологии отключены:

$$\Theta_{16}(0,1) = 1, \quad \Theta_{1m}(0,1) = 0, \quad \text{для всех } m \neq 6;$$

$$\Theta_{27}(1,0) = 1, \quad \Theta_{2m}(1,0) = 0, \quad \text{для всех } m \neq 7;$$

$$\Theta_{16}(0,0) = 1, \quad \Theta_{1m}(0,0) = 0, \quad \text{для всех } m \neq 6;$$

$$\Theta_{27}(0,0) = 1, \quad \Theta_{2m}(0,0) = 0, \quad \text{для всех } m \neq 7.$$

Время реакции сети обслуживания с заданными параметрами без отключения систем обслуживания S_4 и S_8 равно 1,039. В сети обслуживания с периодическим отключением систем обслуживания S_4 и S_8 , время реакции равно 0,976.

Используя выражение (1), математические ожидания (м.о.) числа требований в системах S_1 и S_2 соответственно равны 0,130 и 0,204. В сети обслуживания без управления потоком (с неизменной структурой) м.о. числа требований в S_1 и S_2 несколько больше и равны 0,133 и 0,231.

Предложенная в работе модель сети с управлением и нестационарной структурой и метод ее анализа позволяют проводить исследование ИВС. В частности, предполагается провести исследование влияния вероятностей потерь пакетов в каналах на характеристики сети.

ЛИТЕРАТУРА

1. *Chakka R., Mitrani I.* Approximate solutions for open networks with breakdowns and repairs // *Stochastic Networks, Theory and Applications. Royal Statistical Society Lecture Notes Series*, 1996. – V. 4. – P. 267–280.
2. *Vinod B., Altiok T.* Approximating unreliable queueing networks under the assumption of exponentiality // *J. Opl. Res. Soc.*, 1986. – V. 37. – No. 3. – P. 309–316.
3. *Цицианивили Г.Ш., Осипова М.А.* Предельные распределения в сетях массового обслуживания с ненадежными элементами // *Проблемы передачи информации*, 2008. – Т. 44, № 4. – С. 109–119.
4. *Economides A.A., Silvester J.A.* Optimal routing in a network with unreliable links // *IEEE INFOCOM'88, August 1988.* – P. 288–297.
5. *Chao X.* A queueing network model with catastrophes and product form solution // *Operations Research Letters.* – 1995. – No. 18. – P. 75–79.
6. *Таненбаум Э., Уэзеролл Д.* Компьютерные сети. 5-е изд. – СПб.: Питер, 2012. – 960 с.
7. *Siyon K.S., Parker T.* TCP/IP Unleashed. – Sams. Publishing, 2002.
8. *Митрофанов Ю.И.* Анализ сетей массового обслуживания: учеб. пособие. – Саратов: Научная книга, 2005. – 175 с.

ДИСКРЕТНАЯ СТОХАСТИЧЕСКАЯ ОПТИМИЗАЦИЯ СИСТЕМ МАССОВОГО ОБСЛУЖИВАНИЯ НА ИМИТАЦИОННЫХ МОДЕЛЯХ*

***В.М. Трегубов¹, А.П. Кирпичников²,
И.М. Якимов¹, Р.М. Шакирзянов¹***

*¹Казанский национальный исследовательский
технический университет им. А.Н.Туполева – КАИ, г. Казань, Россия*

*²Казанский национальный исследовательский
технологический университет, г. Казань, Россия*

В настоящее время методы имитационного моделирования (ИМ) стали одним из мощных инструментов исследования и проектирования различных классов систем массового обслуживания (СМО). Методы ИМ применяются для исследования экономических, производственных, транспортных, вычислительных и других систем [1, 2]. Важным этапом моделирования СМО в общей методологии проектирования является поиск оптимальных параметров функционирования системы. Существенной особенностью возникающих задач является стохастичность функционирования, чрезвычайно большая размерность, а также, во многих случаях, дискретность варьируемых переменных задачи [2, 3]. В том случае, когда для описания моделируемой системы используется имитационная модель, дополнительно возникают сложности, обусловленные отсутствием аналитических выражений для целевой функции и ограничений, случайным характером получаемых в процессе моделирования величин, чрезвычайно высокими затратами на проведение имитационных экспериментов и решение собственно оптимизационной задачи.

Методы оптимизации систем, описываемых имитационными моделями, в настоящее время интенсивно развиваются, обзор соответствующих результатов приведен в [2, 4]. Одним из перспективных подходов для решения таких задач является применение методов стохастического программирования [5], а в случае оптимизации СМО на имитационной модели – методов дискретного стохастического программирования [4, 6–8, 10].

* Исследование выполнено в рамках государственного задания № 2.1724.2017/4.6.

В общей постановке указанная задача является чрезвычайно сложной, и в настоящее время нет эффективных методов ее решения. В работах [7, 8] предложен алгоритм, основанный на разбиении множества X на симплексы с последующей кусочно-линейной интерполяцией оптимизируемой функции на этих симплексах. Структура симплексов позволяет далее легко вычислить стохастический градиент (или его аналог) и применить для решения задачи методы типа обобщенного стохастического градиента [9]. Метод показал достаточную эффективность в численных экспериментах [8], однако без дополнительных предположений о структуре множества оптимизируемых параметров и свойствах целевой функции невозможно обосновать качество получаемого решения. В [8] показано, что в случае, когда оптимизируется выпуклая сепарабельная функция, предложенный алгоритм находит решение задачи на всем пространстве E^n .

В данной работе рассматривается случай, когда целевая функция имеет специальную структуру и не сепарабельна. Показывается, что возникающий класс задач является достаточно широким и охватывает обширную проблематику задач оптимизации СМО на имитационных моделях. Для этого класса задач строится алгоритм, основанный на разбиении исходного допустимого множества на симплексы с последующей кусочно-линейной интерполяцией функции на построенных симплексах и применением аналогов метода обобщенного стохастического градиента [9]. За счет учета структуры оптимизируемой функции удастся доказать сходимость построенного алгоритма к решению исходной задачи.

Постановка задачи

Пусть $X_0 = \{x_i, : 0 \leq x_i \leq 1, i = \overline{1, n}\}$ – единичный куб в E^n с вершиной в нуле. Обозначим Z_0 – множество целочисленных вершин куба X_0 .

Рассмотрим следующую задачу целочисленного стохастического программирования: требуется найти x^* :

$$F(x^*) = \min_{x \in Z_0} M \sum_{j=1}^m f_j(A_j x, \omega). \quad (1)$$

Здесь M – символ математического ожидания, ω – элемент вероятностного пространства (Ω, F, P) . В общем случае, ω является вектором

случайных параметров, влияющих на функционирования СМО. Будем предполагать, что аналитические выражения для $f_j(\mathbf{A}_j x, \omega)$ неизвестны, и для $x \in Z_0$ невозможно вычислить точное значение $F(x)$, однако существует возможность наблюдать в ходе имитационного эксперимента с моделью СМО значение случайных показателей $f_j(\mathbf{A}_j x, \omega)$ для $\forall x \in Z_0$. Функции $f_j(\mathbf{A}_j x, \omega)$ предполагаются измеримыми и ограниченными для $\forall \omega$. Кроме того, предполагается, что ω не зависит от параметров x .

Будем считать, что $f_j(\mathbf{A}_j x, \omega)$, $j = \overline{1, m}$, – выпуклые функции при каждом ω , $\mathbf{x} = (x_1, x_2, \dots, x_n)$ – вектор-столбец оптимизируемых параметров, \mathbf{A}_j – строка матрицы $\mathbf{A} = (a_{j,i})$, $j = \overline{1, m}$, $i = \overline{1, n}$. Будем также считать, что матрица \mathbf{A} обладает следующими свойствами:

A1. $a_{j,i} \in \{0, 1\}$ $j = \overline{1, m}$, $i = \overline{1, n}$.

A2. В каждой строке матрицы \mathbf{A} содержится не более двух единиц.

A3. Если J_m – множество номеров строк матрицы \mathbf{A} , в которых содержится ровно 2 единицы, то система уравнений

$$\mathbf{A}_j \mathbf{v} = 1, \quad j \in J_m, \quad (2)$$

имеет решение $\mathbf{v} = (v_1, v_2, \dots, v_n)$, $v_i \in \{0, 1\}$, $i = \overline{1, n}$.

В виде (1) может быть сформулировано большое число практически важных задач оптимизации СМО на имитационных моделях.

Алгоритм решения задачи

Пусть $x \in X_0$, вектор \mathbf{v} удовлетворяет (2). Обозначим

$$t_i(x) = x_i(1 - v_i) + v_i(1 - x_i), \quad i = \overline{1, n}. \quad (3)$$

Очевидно, что $0 \leq t_i(x) \leq 1$, $i = \overline{1, n}$. (4)

Для координат вектора $t_i(x)$, $i = \overline{1, n}$, справедливо упорядочивание вида

$$0 \leq t_{i_1}(x) \leq t_{i_2}(x) \leq \dots \leq t_{i_n}(x) \leq 1. \quad (5)$$

Упорядочиванию (5) соответствует перестановка $\pi(t(x)) = (i_1, i_2, \dots, i_n)$ координат вектора $t(x)$. Как показано в [7, 8], множество (5) является симплексом $C(x)$ с вершинами $\mathbf{H}^0 = (h_1^0, h_2^0, \dots, h_n^0)$, ..., $\mathbf{H}^n = (h_1^n, h_2^n, \dots, h_n^n)$.

Координаты вершин этого симплекса вычисляются следующим образом [7]:

$$h_{ij}^0 = v_{ij}, j = \overline{1, n},$$

$$h_{ij}^s = h_{ij}^{s-1} + r_{ij}^{s-1}, j = \overline{1, n}, s = \overline{1, n}, \quad (6)$$

где
$$r_{ij}^s = \begin{cases} 0 & \text{при } j \neq n - s, \\ 1 - 2v_{ij}^{s-1} & \text{при } j = n - s. \end{cases} \quad (7)$$

Здесь $i_l, l = \overline{1, n}$, – элемент перестановки $\pi(t(x))$.

Далее, обозначим $\varphi(x, \omega) = \sum_{j=1}^m f_j(\mathbf{A}_j x, \omega)$ и задачу (1) запишем в виде

$$F(x^*) = \min_{x \in Z_0} M_\omega \varphi(x, \omega). \quad (8)$$

Следуя [7, 8], построим кусочно-линейную интерполяцию функции $\varphi(x, \omega)$ на кубе X_0 . Заметим, что существует $n!$ различных симплексов C_k , $k = 1, \dots, n!$, образующих разбиение куба X_0 , при этом, если $x \in C(x)$, то существует k^* , такой, что $x \in C_{k^*}$.

Обозначим $\mathbf{H}_k^s, s = \overline{0, n}, k = \overline{1, n!}$, – s -я вершина симплекса C_k .

Построим семейство определенных на X_0 линейных функций $L_k(x, \omega)$ так, чтобы в вершинах симплекса C_k выполнялось условие

$$L_k(\mathbf{H}_k^s, \omega) = \varphi(\mathbf{H}_k^s, \omega), s = \overline{0, n}. \quad (9)$$

Учитывая, что точка \mathbf{H}^0 принадлежит каждому симплексу C_k выражение для $L_k(x, \omega)$ запишем в виде

$$L_k(x, \omega) = \varphi(\mathbf{H}^0, \omega) + \sum_{i=1}^n b_i^k(\omega)(x_i - v_i).$$

Коэффициенты $b_i^k(\omega)$ выберем из условия (9), тогда из (6), (7) получаем

$$b_{i_j}^k(\omega) = \left(\varphi(\mathbf{H}_k^{n-j-1}, \omega) - \varphi(\mathbf{H}_k^{n-j}, \omega) \right) / (1 - 2v_j), \quad j = \overline{1, n}. \quad (10)$$

Положим теперь $L(x, \omega) = L_k(x, \omega)$, если $x \in C_k$.

Наряду с задачей (8) рассмотрим задачу

$$\Phi(x^*) = \min_{x \in X} M_\omega L(x, \omega) \text{ при } X = X_0.$$

Из свойств функции $L(x, \omega)$ [7] следует, что $\Phi(x) = F(x)$ для $x \in Z_0$.

Обозначим $b_i(x, \omega) = b_i^k(\omega)$, $i = \overline{1, n}$, при $x \in C_k = C(x)$, где $b_i^k(\omega)$ определяются из (10). Заметим, что при $0 < t_{i_1}(x) < t_{i_2}(x) < \dots < t_{i_n}(x) < 1$ $\partial L(x, \omega) / \partial x_i = b_i(x, \omega)$, так что градиент функции $L(x, \omega)$ при $0 < t_{i_1}(x) < t_{i_2}(x) < \dots < t_{i_n}(x) < 1$ равен $\nabla L(x, \omega) = (b_1(x, \omega), b_2(x, \omega), \dots, b_n(x, \omega))$, где $b_i(x, \omega)$ определяется из (10), а переход от точки \mathbf{H}^s к точке \mathbf{H}^{s+1} на симплексе $C(x)$ осуществляется в соответствии с формулами (6), (7).

Обобщенный градиент $\mathbf{g}(x, \omega)$ функции $L(x, \omega)$ [7, 9] вычисляется при $0 \leq t_{i_1}(x) < t_{i_2}(x) < \dots < t_{i_n}(x) \leq 1$ по формуле $\mathbf{g}(x, \omega) = \nabla L(x, \omega)$. Если же $t_{i_l}(x) = t_{i_{l+1}}(x)$, то будем считать, что $t_{i_l}(x) < t_{i_{l+1}}(x)$, если $i_l < i_{l+1}$, при этом полностью определяется способ вычисления обобщенного градиента $\mathbf{g}(x, \omega)$.

Применим для решения данной задачи метод проекции стохастических обобщенных градиентов [7, 9]:

$$x^{k+1} = \text{Pr}_{X_0}(x^k - \rho_k \mathbf{g}(y^k, \omega)), \quad (11)$$

где $\text{Pr}_{X_0}(u)$ – оператор проектирования вектора u на множество X_0 [5], $\mathbf{g}(y^k, \omega)$ – стохастический обобщенный градиент функции $L(x, \omega)$ в точке $y^k = (y_1^k, y_2^k, \dots, y_n^k)$, y_l^k – $l = \overline{1, n}$, независимые случайные величины, равномерно распределенные на отрезках $[x_l^k - \beta_k, x_l^k + \beta_k]$.

Если

$$\sum_{k=1}^{\infty} \rho_k = \infty, \quad \sum_{k=1}^{\infty} \rho_k^2 < \infty, \quad \rho_k / \beta_k \xrightarrow{k \rightarrow \infty} \infty, \quad \beta_k \xrightarrow{k \rightarrow \infty} 0, \\ \rho_k > 0, \quad |\beta_k - \beta_{k+1}| / \rho_k \xrightarrow{k \rightarrow \infty} 0, \quad \beta_k > 0,$$

то последовательность x_k , определенная формулой (11), сходится почти наверное (п.н.) к множеству локальных минимумов функции $\Phi(x)$ на множестве X_0 [7, 9].

С учетом выпуклости $L(x, \omega)$ на X_0 и свойств функции $L(x, \omega)$ получаем сходимость алгоритма (11) к множеству решений задачи (1), при этом

$$\Phi(x^*) = \min_{x \in Z_0} M \sum_{j=1}^m f_j(A_j x, \omega).$$

Оптимизация системы обслуживания сети коммуникаций.

Рассмотрим задачу оптимизации СМО на имитационной модели следующего вида.

Пусть имеется транспортная сеть поставки продукции от поставщиков к потребителям. Пусть n – общее количество узлов сети, m – количество коммуникаций сети. Сеть задается графом G , который является двудольным, и пусть $A = a_{ij}$, $i = \overline{1, n}$, $j = \overline{1, m}$, – матрица реберной инцидентности графа G . Пусть граф G не имеет петель, тогда $\sum_{i=1}^n a_{ji} = 2$,

$j = \overline{1, m}$, при этом система уравнений $A_j v = 1$, $j = \overline{1, m}$ разрешима для $v_i \in \{0, 1\}$ и матрица A удовлетворяет условиям A1 – A3.

Процесс перевозки продукции от поставщиков к потребителям рассматривается как процесс обслуживания транспортных средств в узлах и на коммуникациях сети. Спрос на продукцию предполагается случайным, случайными являются также время перевозки продукции от поставщиков к потребителям и другие параметры системы.

Важной особенностью рассматриваемой СМО является то, что в процессе перевозок на коммуникациях сети могут возникать отказы (например, поломки транспортных средств), которые носят случайный характер, для ликвидации которых в узлах сети размещаются некоторые ресурсы, причем ресурсы, размещаемые в узле, могут быть направлены

на ликвидацию отказов в любой коммуникации сети, связанной с данным узлом. Для моделирования функционирования сети используется имитационная модель, которая позволяет в имитационном эксперименте ω моделировать процесс перевозок и отказы на коммуникации j сети, а также вычислять ресурсы $R_j(\omega)$, необходимые для восстановления работоспособности коммуникации. Пусть

$$x_i = \begin{cases} 1, & \text{если в } i\text{-м узле планируется размещать ресурсы,} \\ 0, & \text{в противном случае.} \end{cases}$$

Пусть $d_{ji} > 0$ – планируемый объем ресурса, выделяемого в узле i на восстановление коммуникации j , связанной с узлом i . Пусть также q_j^- – стоимость потерь (в расчете на единицу ресурса), которые возникают в связи с тем, что выделенных ресурсов не хватает на ликвидацию отказа на коммуникации j , а q_j^+ – стоимость потерь (в расчете на единицу ресурса), которые возникают в связи с тем, что количество выделенных на ликвидацию отказа на коммуникации j ресурсов превышает необходимое.

Выражение для средних потерь, связанных с размещением ресурсов в узлах сети с учетом потерь как на недостаток, так и на избыток ресурсов, запишется в виде

$$F(x) = M \left(\sum_{j=1}^m \max \left\{ q_j^+ \left(\sum_{i=1}^n d_{ji} a_{ji} x_i - R_j(\omega) \right), q_j^- \left(R_j(\omega) - \sum_{i=1}^n d_{ji} a_{ji} x_i \right) \right\} \right),$$

$$x_i \in \{0, 1\}, \quad i = \overline{1, n}. \quad (12)$$

Задача заключается в нахождении такого распределения ресурсов по узлам сети, которое минимизирует средние потери (12).

На рис. 1 приведены результаты вычислительного эксперимента для сети, содержащей 5 поставщиков, 5 потребителей и 12 коммуникаций.

Анализ эффективности разработанного алгоритма проводился в сравнении с эффективностью работы для данной задачи «жадного» (greedy) алгоритма [11], в котором вычисление значения функции $F(x)$ в каждой точке x проводилось путем усреднения по 100 реализациям величин $R_j(\omega)$ для каждой коммуникации j . Эффективность работы алгоритмов оценивалась по точности найденного решения в зависимости от количества обращений к имитационной модели ω . Такой анализ был выбран в связи с тем, что основная трудоемкость алгоритмов в рассматриваемом классе задач связана с высокой трудоемкостью имитационных экспериментов.

На рис. 1. показан график поиска решения разработанным алгоритмом в сравнении с «жадным» алгоритмом. Предложенный алгоритм нашел приближенное решение $x_1 = (0; 0,0601; 1; 0,01695; 0; 0; 0,0299; 0,9980; 0; 0,0258)$, $F(x_1) = 7,2710$ за 2000 обращений к имитационной модели. Точка x_1 уже может рассматриваться как окрестность оптимального решения $x_{\text{opt}} = (0; 0; 1; 0; 0; 0; 0; 1; 0; 0)$.

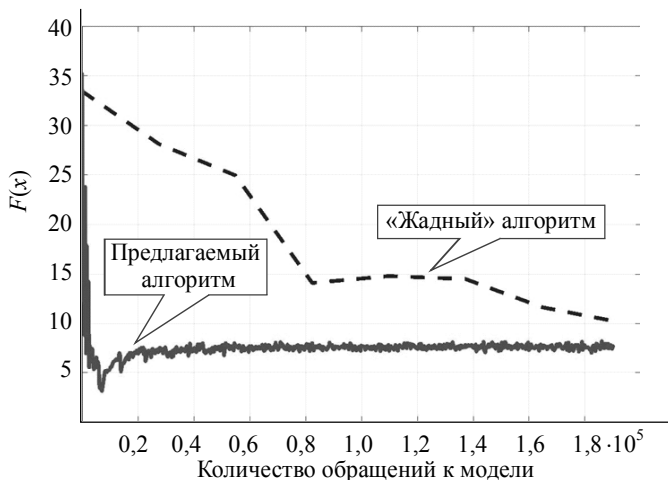


Рис. 1. Сравнение алгоритмов

«Жадный» алгоритм нашел решение задачи за 200 000 обращений к имитационной модели. Алгоритм нашел решение $x_{\text{opt}} = (0; 0; 1; 0; 0; 0; 0; 1; 0; 0)$ со значением $F(x_{\text{opt}}) = 10,1748$.

Заключение

Предложенный алгоритм существенно расширяет класс задач дискретного стохастического программирования, для которого, за счет учета специфики задачи, удастся обосновать качество получаемого решения.

Предложенный алгоритм может быть эффективно применен для решения широкого круга практически важных задач оптимизации стохастических систем на основе имитационных моделей.

Разработанный алгоритм превосходит существующие методы по времени решения задачи, в первую очередь, на начальных этапах поиска.

ЛИТЕРАТУРА

1. Сидоренко В.Н., Красносельский А.В. Имитационное моделирование в науке и бизнесе: подходы, инструменты, применение // Бизнес-информатика. – 2008. – № 2. – С. 52–57.
2. Fu M.C., Glover F.W., April J. Simulation optimization: a review, new developments, and applications // Proceedings of the 2005 Winter Simulation Conference / Kuhl M.E., Steiger N.M., Armstrong F.B., Joines J.A. (eds.). – 2005. – P. 83–95.
3. Лифшиц Ф.Л., Мальц Э.М. Статистическое моделирование систем массового обслуживания. – М.: Советское радио, 1978. – 248 с.
4. Barry L. Nelson. Optimization via simulation over discrete decision variables // Tutorials in Operations Research, 2010 INFORMS. – P. 193–207.
5. Ермолов Ю.М. Методы стохастического программирования. – М.: Наука, 1976. – 235 с.
6. Abspoel S.J., Etman L.F.P., Vervoort J. Simulation optimization of stochastic systems with integer variables by sequential linearization // Proceedings of the 2000 Winter Simulation Conference. – P. 715–723.
7. Трегубов В.М. Алгоритм целочисленной стохастической оптимизации имитационных моделей сложных систем // Имитационное моделирование. Теория и практика (ИММОД-2013): сб. докл. Шестой Всероссийской научно-практической конференции по имитационному моделированию и его применению в науке и промышленности. – Казань: Академия наук РТ, 2013. – С. 271–277.
8. Kagan Gokbayrak, Christos G. Cassandras. A generalized “surrogate problem” methodology for on-line stochastic discrete optimization // Submitted to JOTA, January 2001. – P. 3–29.
9. Гунал А.М. Стохастические методы решения негладких экстремальных задач. – Киев: Наукова думка, 1979. – 152 с.
10. Трегубов В.М. Применение методов стохастического программирования в задачах оптимизации СМО на имитационных моделях // Исследование операций и аналитическое проектирование в технике: межвузовский сборник. – Казань: Казан. авиац. ин-т, 1988. – С. 39–44.
11. Глебов Н.И. Об условиях разрешимости оптимизационных задач «жадным» алгоритмом // Дискретный анализ и исследование операций. – Июль – декабрь 2002. – Сер. 2. – Т. 9. – № 2. – С. 3–12.

АНАЛИТИЧЕСКИЕ И ЧИСЛЕННЫЕ МЕТОДЫ ИССЛЕДОВАНИЯ СИСТЕМЫ АДАПТИВНОГО УПРАВЛЕНИЯ КОНФЛИКТНЫМИ ПОТОКАМИ

М.А. Федоткин, Е.В. Кудрявцев

*Нижегородский государственный университет им. Н.И. Лобачевского,
г. Нижний Новгород, Россия*

Рассматривается система адаптивного управления конфликтными потоками неоднородных требований. В качестве математического описания такой системы выбирается состояние обслуживающего устройства и длины очередей по конфликтным входным потокам. Доказано марковское свойство последовательности состояний системы и проведена их классификация. Получены как необходимые, так и достаточные условия существования стационарного распределения. Проведено имитационное моделирование системы. Численное исследование имитационной модели позволило найти квазиоптимальные параметры адаптивного управления.

Математическая модель

Данная работа связана с важной проблемой создания алгоритмов в интеллектуальных транспортных системах, которые управляют конфликтными потоками на пересечениях магистралей в крупных городах. Предлагается простой класс нециклических управлений транспортными потоками. Алгоритмы управления потоками из этого класса зависят не только от длин очередей, но и от очередности поступления автомобилей к стоп-линии перекрестка. Построена и изучена математическая модель такой системы управления потоками.

Транспортный поток состоит из разнотипных автомобилей, различающихся скоростью движения. Таким образом, рассматриваются входные потоки с неоднородными требованиями. В работе [1] предложен механизм образования таких потоков. Показана возможность аппроксимации потоков такого вида неординарными пуассоновскими по-

токами. Это дает возможность рассматривать в качестве входных потоков два конфликтных неординарных пуассоновских потока Π_1 и Π_2 . В каждый вызывающий момент по потоку Π_j , где $j = 1, 2$, приходит k заявок с вероятностями $P_j(k)$, $k = 1, 2, \dots$. Эти вероятности имеют следующий вид:

$$P_j(1) = (1 + \alpha_j + \alpha_j \beta_j / (1 - \gamma_j))^{-1} = p_j,$$

$$P_j(1) = \alpha_j (1 + \alpha_j + \alpha_j \beta_j / (1 - \gamma_j))^{-1},$$

$$P_j(1) = \alpha_j \beta_j \gamma_j^{k-3} (1 + \alpha_j + \alpha_j \beta_j / (1 - \gamma_j))^{-1}, \quad k \geq 3,$$

где α_j , β_j и γ_j – некоторые параметры распределения, физический смысл которых был определен в [1]. Интенсивность поступления вызывающих моментов по потоку Π_j равна λ_j . Свойства таких потоков с неоднородными требованиями изучены в [1, 2].

Управление конфликтными потоками при их обслуживании производится с помощью адаптивного нециклического алгоритма, подробное описание которого приведено в работах [3, 4].

В системе обслуживающим устройством является светофор, а требованиями – разнотипные автомобили, подъезжающие к светофору. Множество состояний светофора обозначим через $\Gamma = \{\Gamma^{(1)}, \Gamma^{(2)}, \Gamma^{(3)}, \Gamma^{(4)}, \Gamma^{(5)}, \Gamma^{(6)}, \Gamma^{(7)}, \Gamma^{(8)}\}$.

Приведем описание каждого из состояний.

Состояние $\Gamma^{(3j-2)}$ (зеленый сигнал светофора для j -го потока) соответствует первому этапу периода обслуживания потока Π_j . Длительность обслуживания одной заявки, поступившей из накопителя, равна постоянной величине $\mu_{j,1}^{-1}$. Длительность пребывания в $\Gamma^{(3j-2)}$ равна T_{3j-2} .

Состояние $\Gamma^{(3j-1)}$ (зеленый сигнал светофора для j -го потока) соответствует второму этапу периода обслуживания потока Π_j . Длительность обслуживания одной заявки равна величине $\mu_{j,2}^{-1} < \mu_{j,1}^{-1}$. Длительность пребывания в этом состоянии – случайная величина, принимающая значения kT_{3j-1} , $k = \overline{1, n_j}$, где n_j – максимальное число продлений и T_{3j-1} – длительность одного продления. Продление происходит в 2 случаях: 1) длина очереди по j -му потоку не меньше параметра K_j , 2) на предыдущем этапе продлений поступили требования, которые необходимо обслужить.

Состояние $\Gamma^{(3j)}$ (желтый сигнал светофора для j -го потока) соответствует режиму переналадки для потока Π_j . Длительность пребывания в этом состоянии равна T_{3j} .

Состояние $\Gamma^{(6+j)}$ (зеленый сигнал светофора для j -го потока) соответствует первому этапу периода обслуживания потока Π_j в случае, когда возможен мгновенный переход в состояние $\Gamma^{(3j)}$. Длительность пребывания в $\Gamma^{(6+j)}$ является случайной величиной. Максимальное время пребывания в этом состоянии равно $T_{3j} - 2$.

Далее систему будем рассматривать в моменты τ_i , $i \geq 0$, и на промежутках $[\tau_i, \tau_{i+1})$. Здесь τ_0 – начальный момент времени, а τ_i , $i \geq 1$, – моменты смены состояний обслуживающего устройства. Обозначим $y_0 = (0, 0)$, $y_1 = (1, 0)$, $y_2 = (0, 1)$ и пусть X – целочисленная одномерная неотрицательная решетка. Для нелокального описания системы при $i = 0, 1, \dots$ введем следующие случайные величины и элементы:

$\Gamma_i \in \Gamma$ – состояние обслуживающего устройства на временном интервале $[\tau_i, \tau_{i+1})$ (на i -м такте);

$\eta_{j,i} \in X$ – число заявок j -го потока, поступивших в систему за промежутки $[\tau_i, \tau_{i+1})$, $\eta_i = (\eta_{1,i}, \eta_{2,i})$;

$\eta'_{j,i}$ – случайный вектор, принимающий значение y_0 , если на i -м такте $[\tau_i, \tau_{i+1})$ в систему не поступило ни одной заявки, значения y_j , если на i -м такте первой пришла заявка (или заявки) j -го потока;

$\kappa_{j,i} \in X$ – число заявок j -го потока, которые находятся в системе в момент τ_i , $\kappa_i = (\kappa_{1,i}, \kappa_{2,i})$;

$\xi_{j,i}$ – максимально возможное число заявок j -го потока, которые система может обслужить на интервале $[\tau_i, \tau_{i+1})$, $\xi_i = (\xi_{1,i}, \xi_{2,i})$.

Примем следующие соотношения:

$$T_{3j-2} = \mu_{j,1}^{-1} + l_{3j-2} \theta_j \mu_{j,1}^{-1},$$

$$T_{3j-1} = l_{3j-1} \theta_j \mu_{j,2}^{-1},$$

$$T_{3j} = l_{3j} \theta_j \mu_{j,2}^{-1},$$

где $l_{3j-2} \in X$, $l_{3j-1}, l_{3j} \in \mathbb{N}$. Параметр l_{3j} выбирается так, чтобы выполнялось неравенство $T_{3j} \geq \mu_{j,1}^{-1}$. Величина $0 < \theta_j \leq 1$ обозначает часть обслуживания, которую необходимо пройти требованию, чтобы можно было начать обслуживать следующую заявку. В случае $\theta_j < 1$ одновременно может обслуживаться несколько требований.

Адаптивный алгоритм смены состояний обслуживающего устройства из множества Γ задается с помощью рекуррентного соотношения:

$$\Gamma_{i+1} = \begin{cases} \Gamma^{(3j-2)}, & \left\{ \left[\Gamma_i = \Gamma^{(3s)} \right] \cap \left[\left(\kappa_{j,i} > 0 \right) \cup \left(\kappa_{s,i} \geq K_s \right) \cup \left(\eta'_i = y_j \right) \right] \right\} \cup \\ & \left\{ \left[\Gamma_i = \Gamma^{(3j)} \right] \cap \left[\kappa_{s,i} = 0 \right] \cap \left[\kappa_{j,i} \leq K_j \right] \cap \left[\eta'_i = y_j \right] \right\}, \\ \Gamma^{(3j-1)}, & \left\{ \Gamma_i = \Gamma^{(3j-2)} \right\} \cup \left\{ \left[\Gamma_i = \Gamma^{(6+j)} \right] \cap \left[\eta'_i = y_j \right] \right\}, \\ \Gamma^{(3j)}, & \left\{ \Gamma_i = \Gamma^{(3j-1)} \right\} \cup \left\{ \left[\Gamma_i = \Gamma^{(6+j)} \right] \cap \left[\eta'_i \neq y_j \right] \right\}, \\ \Gamma^{(6+j)}, & \left[\Gamma_i = \Gamma^{(3s)} \right] \cap \left[\kappa_{j,i} = 0 \right] \cap \left[\kappa_{s,i} < K_s \right] \cap \left[\eta'_i = y_0 \right], \end{cases}$$

где $j, s = 1, 2; j \neq s$.

Как видно из приведенного соотношения, состояние обслуживающего устройства на следующем шаге зависит от состояния на предыдущем шаге, длины очередей и очередности прихода заявок. При этом динамика длины очереди задается следующими рекуррентными соотношениями:

$$\kappa_{j,i+1} = \begin{cases} \max(0, \kappa_{j,i} + \eta_{j,i} - \xi_{j,i}), & \text{если } \Gamma_i \in \Gamma \setminus \{ \Gamma^{(3)}, \Gamma^{(6)} \}, \\ \eta_{j,i} + \max(0, \kappa_{j,i} - \xi_{j,i}), & \text{если } \Gamma_i \in \{ \Gamma^{(3)}, \Gamma^{(6)} \}. \end{cases}$$

Теоремы существования стационарного распределения

В работе [3] приведена теорема марковости векторной последовательности $\{(\Gamma_i, \kappa_i); i \geq 0\}$ и проведена классификация ее состояний.

Для любого $i \geq 0$, $r = \overline{1, 8}$; $x \in X^2$ введем обозначение:

$$Q_i^{(r)}(x) = P(\Gamma_i = \Gamma^{(r)}, \kappa_i = x).$$

В работе [4] приведены рекуррентные соотношения для одномерных распределений $\{Q_i^{(r)}(x) : r = \overline{1, 8}, x \in X^2\}$, $i \geq 0$, марковской последовательности $\{(\Gamma_i, \kappa_i); i \geq 0\}$.

Пусть $z = (z_1, z_2)$, где компоненты z_1, z_2 — действительные или комплексные переменные, удовлетворяющие условиям $|z_1| \leq 1$, $|z_2| \leq 1$. Положим $z^x = z_1^{x_1} z_2^{x_2}$, где $x = (x_1, x_2) \in X^2$. Рассмотрим производящие функции

$$W_i^{(r)}(z) = \sum_{x \in X^2} Q_i^{(r)}(x) z^x, \quad r = \overline{1, 8},$$

$$W_i(z) = \sum_{r=1}^8 W_i^{(r)}(z).$$

Используя рекуррентные соотношения для одномерных распределений $\{Q_i^{(r)}(x) : r = \overline{1, 8}, x \in X^2\}$, $i \geq 0$, векторной марковской последовательности $\{(\Gamma_i, \kappa_i); i \geq 0\}$, стандартным образом были получены рекуррентные соотношения для производящих функций $W_i^{(r)}(z)$, $r = \overline{1, 8}, i > 0$.

Используя рекуррентные соотношения для производящих функций, получено следующее достаточное условие существования предельного распределения изучаемой векторной марковской последовательности.

Теорема 1. Для существования предельного распределения векторной последовательности $\{(\Gamma_i, \kappa_i); i \geq 0\}$ достаточно выполнения неравенств

$$\lambda_j M_j T - L_j < 0, \quad j = 1, 2,$$

где

$$\begin{aligned} T &= T_1 + T_3 + T_4 + T_6 + n_1 T_2 + n_2 T_5, \\ L_j &= l_{3j-2} + n_j l_{3j-1} + l_{3j}, \\ M_j &= \left(1 + 2\alpha_j + \alpha_j \beta_j \left(2/(1-\gamma_j) + 1/(1-\gamma_j)^2\right)\right) p_j. \end{aligned}$$

Обозначим

$$\tilde{T}_j = T_1 + T_3 + T_4 + T_6 + \tilde{n}_j T_{3j-1} + n_s T_{3s-1},$$

где \tilde{n}_j – нижняя оценка числа продлений для состояния $\Gamma^{(3j-1)}$ в случае, когда продления происходят только при условии прихода новых заявок на предыдущем такте продлений. Тогда получаем следующую теорему.

Теорема 2. Для существования предельного распределения последовательности $\{(\Gamma_i, \kappa_i); i \geq 0\}$ необходимо выполнение неравенств

$$\lambda_j M_j \tilde{T}_s - L_j < 0, \quad j, s = 1, 2; j \neq s.$$

Также были получены следующие необходимые условия существования предельного распределения.

Теорема 3. Если существует предельное распределение марковской последовательности $\{(\Gamma_i, \kappa_i); i \geq 0\}$, то

$$\frac{\theta_1 \lambda_1 M_1}{\mu_{1,2}} + \frac{\theta_2 \lambda_2 M_2}{\mu_{2,2}} < 1.$$

Из условия теоремы 3 легко вытекает следующее следствие.

Следствие 1. *Предельное распределение $\{(\Gamma_i, \kappa_i); i \geq 0\}$ существует только тогда, когда выполняются условия*

$$\theta_j \lambda_j M_j < \mu_{j,2}, \quad j = 1, 2.$$

Теорема 4. *Для существования предельного распределения последовательности $\{(\Gamma_i, \kappa_i); i \geq 0\}$ необходимо выполнение неравенства*

$$\lambda_1 M_1 T - L_1 + \frac{\lambda_1 M_1 T_5}{l_5 - \lambda_2 M_2 T_5} (\lambda_2 M_2 T - L_2) < 0.$$

Теорема 5. *Для существования предельного распределения последовательности $\{(\Gamma_i, \kappa_i); i \geq 0\}$ необходимо выполнение неравенства*

$$\lambda_1 M_1 T - L_1 + \frac{l_2 - \lambda_1 M_1 T_2}{\lambda_2 M_2 T_2} (\lambda_2 M_2 T - L_2) < 0.$$

Следствие 2. *Для существования предельного распределения последовательности $\{(\Gamma_i, \kappa_i); i \geq 0\}$ необходимо, чтобы хотя бы для одного $j = 1, 2$ выполнялось неравенство*

$$\lambda_j M_j T - L_j < 0.$$

Исследование системы с помощью имитационного моделирования

Изученный выше алгоритм ориентирован, прежде всего, на уменьшение среднего времени пребывания произвольного требования в системе (в дальнейшем мы будем называть эту характеристику просто средним временем пребывания). Однако получить аналитическое выражение для этой характеристики практически невозможно. Поэтому для проверки эффективности данного алгоритма необходимо применять численные методы. Основным способом численного исследования сложных систем массового обслуживания является метод имитационного моделирования. Для этого была построена имитационная модель, с помощью которой было проведено численное исследование.

Основной задачей численного исследования является определение квазиоптимального набора параметров (\bar{T}, \bar{n}) , где $\bar{T} = (T_1, T_2, T_3, T_4, T_5, T_6)$ и $\bar{n} = (n_1, n_2)$, при котором среднее время пребывания минимально. Проведенные аналитические исследования и имитационное моделирование позволяют не только дать практические рекомендации адаптив-

ного управления потоками, но и определить квазиоптимальные параметры \bar{T}, \bar{n} .

Следующая задача исследования имитационной модели связана с оценкой $\rho(\bar{T}, \bar{n})$ для загрузки системы. Показано, что квазиоптимальному набору параметров соответствует близкое к минимальному значение этой оценки. Изучено влияние параметров n_1 и n_2 на среднее время пребывания.

Наконец, выявлены и рассмотрены адаптивные свойства предложенного алгоритма, благодаря которым этот алгоритм позволяет значительно уменьшить среднее время пребывания. Для изучения адаптивных свойств рассмотрены конкретные реализации со специально выбранными начальными условиями.

Заключение

Была построена и изучена система массового обслуживания с адаптивным нециклическим управлением. Основным методом исследования системы был аппарат производящих функций. Для нахождения условий существования предельного распределения в системе использовался итеративно-мажорантный метод. С помощью имитационного моделирования проведено численное исследование системы. Оказалось, что при большой загрузке алгоритм начинает работать как циклический.

ЛИТЕРАТУРА

1. Fedotkin M.A., Fedotkin A.M., Kudryavtsev E.V. Construction and analysis of a mathematical model of spatial and temporal characteristics of traffic flows // Automatic Control and Computer Sciences. – Allerton Press, Inc., 2014. – V. 48. – No 6. – P. 358–367.
2. Fedotkin M.A., Fedotkin A.M., Kudryavtsev E.V. Nonlocal description of the time characteristic for input flows by means of observations // Automatic Control and Computer Sciences. – Allerton Press, Inc., 2015. – V. 49. – No 1. – P. 29–36.
3. Федоткин М.А., Кудрявцев Е.В. Предельные свойства системы адаптивного управления конфликтными потоками неоднородных требований // Материалы Международной научной конференции «Распределенные компьютерные и телекоммуникационные сети: управление, вычисление, связь (DCCN-15)». – М., 2015. – С. 233–240.
4. Федоткин М.А., Кудрявцев Е.В. Построение математической модели адаптивного управления неординарными потоками // Материалы Международной научной конференции «Теория вероятностей, случайные процессы, математическая статистика и приложения». – Минск: БГУ, 2015. – С. 106–111.

СРАВНЕНИЕ СИСТЕМ СТРУКТУРНОГО И ИМИТАЦИОННОГО МОДЕЛИРОВАНИЯ ВЕРОЯТНОСТНЫХ ОБЪЕКТОВ

И.М. Якимов¹, А.П. Кирпичников², В.В. Мокшин¹, З.Т. Яхина²

*¹Казанский национальный исследовательский
технический университет им. А.Н.Туполева – КАИ, г. Казань, Россия*

*²Казанский национальный исследовательский
технологический университет, г. Казань, Россия*

В последнее время в сфере имитационного моделирования (ИМ) произошли значительные изменения, связанные в первую очередь с переходом от применения специализированных языков ИМ к системам структурного и имитационного моделирования (ССИМ). ССИМ позволяют пользователям отказаться от программирования моделей и перейти к процессу их «рисования». Процесс программирования остаётся только на долю программистов, создающих подпрограммы имитации функционирования элементов моделируемых объектов, совокупность которых представляется специализированными библиотеками.

В монографии [1] достаточно подробно рассмотрены системы ИМ на конец 1985 года, но за более чем 30 лет, прошедших после её выхода в свет, появилось большое количество новых перспективных ССИМ, которые, естественно, не могли попасть в эту монографию. В статье [2] рассмотрено более 10 перспективных ССИМ с данными по удельным весам их применения. Авторами данной статьи проведено сравнение 13 ССИМ с их кратким описанием, приведены структурные модели системы массового обслуживания (СМО) типа М/М/5 в них и сравнение достоверности результатов ИМ, полученных в этих системах, с результатами аналитического моделирования (АМ) СМО М/М/5. Такой подход является дальнейшим развитием работ [3, 4]. И, по мнению авторов, в полной мере будет содействовать квалифицированному выбору средств моделирования по предметной области, к которому относится объект моделирования, по достоверности результатов моделирования и по личным предпочтениям пользователя.

Из совокупности ССИМ, приведённых в [2], выбраны 13 систем, в которых отсутствует процесс программирования для массовых пользо-

вателей и имеется возможность получения бесплатной версии из интернета: AnyLogic, Arena, Bizagi Modeler, Business Studio, Enterprise Dynamics, ExtendSim, Flexsim, GPSS W, Plant Simulation, Process Simulator, Rand Model Designer, Simio и Simul8.

1. Краткое описание ССИМ

1.1. AnyLogic

ССИМ AnyLogic [5] разработана российской компанией «Экс Джей Текнолджис» (англ. XJ Technologies). Первая версия системы AnyLogic 4.0 разработана в 2003 г. В 2014 г. разработана AnyLogic 7.0. ССИМ AnyLogic включает в себя графический язык моделирования и позволяет пользователю расширять создаваемые модели с помощью языка Java.

Понятия, принятые в ССИМ AnyLogic по отношению к понятиям, принятым в теории массового обслуживания, следующие: заявки – сущности, очереди – очереди, ОА – задачи. Имеется русифицированная версия ССИМ AnyLogic и литература на русском языке, в том числе [5].

1.2. Arena

ССИМ Arena [6] разработана компанией Systems Modeling Corporation. Первая версия этой системы разработана в 1993 г. В 2014 г. разработана ССИМ Arena 3.0. Основу системы Arena составляют: транслятор языка моделирования Siman и система анимации Cinema Animation.

Понятия, принятые в системе, по отношению к понятиям, принятым в теории массового обслуживания следующие: заявки – сущности, очереди – очереди, ОА – задачи. Имеется литература на русском языке в том числе [6]. Несомненным достоинством ССИМ Arena является возможность автоматического перехода с диаграммы IDEF3, широко распространённой системы BPwin [6], к структурной модели в ССИМ Arena.

1.3. Bizagi Modeler

ССИМ Bizagi Modeler [7] разработана группой компаний Object Management Group, созданной в 1989 г. Год разработки системы ССИМ Bizagi Modeler – 2007. В 2016 г. разработана ССИМ Bizagi BPM Suite

11-я версия. Для обозначения элементов моделируемых систем использует нотацию моделирования бизнес-процессов (BPMN 2.0).

Понятия, принятые в ССИМ Bizagi Modeler по отношению к понятиям, принятым в теории массового обслуживания следующие: заявки – сообщения, очереди – очереди, ОА – действия. Для указания условий выбора маршрута движения сущностей используются шлюзы.

Имеется русифицированная версия Bizagi Modeler. Литературы на русском языке нет.

1.4. Business Studio

ССИМ Business Studio [8] разработана группой компаний «Современные технологии управления», основанной в 1991 г. Первая версия системы Business Studio разработана в 2004 г., а в 2013 г. разработана версия 4.0. Для обозначения элементов моделируемых систем использует нотацию моделирования бизнес-процессов (BPMN 2.0).

Понятия, принятые в ССИМ Business Studio, те же что и в ССИМ Bizagi Modeler [7].

Русифицированной версии Business Studio и литературы на русском языке по ней нет.

1.5. Enterprise Dynamics

ССИМ Enterprise Dynamics разработана компанией InControl Simulation, организованной в 1998 г. Год разработки системы Enterprise Dynamics – 2004 [9].

Понятия, принятые в ССИМ Enterprise Dynamics: заявки – продукты, очереди – очереди, ОА – серверы.

Русифицированной версии Enterprise Dynamics и литературы на русском языке по ней нет.

1.6. ElxtendSim

Первая версия ССИМ ElxtendSim [10] разработана компанией Imagine That Inc. в начале 1987 г. Она была одним из первых программных продуктов на рынке в области ИМ с реализацией создания имитационной модели по вводимой в ССИМ структурной схеме. Последняя версия ССИМ поступила в продажу в 2015 г. [10].

Понятия, принятые в ССИМ ElxtendSim, по отношению к понятиям, принятым в теории массового обслуживания, следующие: заявки – элементы, очереди – очереди, ОА – задачи (activity).

Русифицированной версии ElxtendSim и литературы на русском языке по ней нет.

1.7. Flexim

ССИМ Flexim [11] разработана компанией Flexim Software Products Inc (ESP), организованной в 1993 г. Год разработки системы Flexim – 2003.

Понятия, принятые в ССИМ Flexim по отношению к понятиям, принятым в теории массового обслуживания следующие: заявки – заявки, очереди – очереди, ОА – процессоры, памяти [11].

Недостаток – отсутствие русифицированной версии и описания на русском языке.

1.8. GPSS W с расширенным редактором

ОАО «Элина-Компьютер» в 2014 г. для системы GPSS W создан расширенный редактор [12], который для массового пользователя позволяет отказаться от программирования и перейти к рисованию моделей. Программистам в то же время предоставляется возможность создания программ моделирования новых элементов и включения их в структурные и имитационные модели. Кроме того, расширенный редактор позволяет проводить планирование имитационных экспериментов обработки результатов моделирования.

Понятия, принятые в ССИМ GPSS W, по отношению к понятиям, принятым в теории массового обслуживания, следующие: заявки – транзакты, очереди – очереди, ОА – устройства, памяти.

Имеется русифицированная версия ССИМ GPSS W и документация на русском языке.

1.9. Plant Simulation

ССИМ Plant Simulation [13] с 2007 г. поставляется компанией Siemens PLM Software. В 2016 году начались поставки 13 версии ССИМ Plant Simulation.

Понятия, принятые в системе GPSS W, по отношению к понятиям, принятым в теории массового обслуживания, следующие: заявки – де-

тали, очереди – накопители, ОА – единичные операции, параллельные операции. Имеется возможность моделирования производственных процессов с продуктами в жидкообразном состоянии.

Недостаток – отсутствие русифицированной версии и описания на русском языке.

1.10. Process Simulator

ССИМ Process Simulator разработана корпорацией ProModel Corporation и впервые поступила в продажу в 2001 г. [14]. Последней версией ССИМ Process Simulator является версия 9.3.0.2701, разработанная в 2016 г.

Понятия, принятые в ССИМ Process Simulator, по отношению к понятиям, принятым в теории массового обслуживания, следующие: заявки – заявки, очереди – очереди, ОА – действия.

Недостаток – отсутствие русифицированной версии и описания на русском языке.

1.11. Rand Model Designer

Известная ССИМ Model Vision Studium сменила своё название и с 2011 г. стала называться Rand Model Designer [15]. Она разработана Компанией MVSTUDIUM Group, основанной в 1997 г. Последняя версия ССИМ Rand Model Designer (RMD) разработана в 2016 г. на базе языка моделирования Modelica [15].

Понятия, принятые в ССИМ Rand Model Designer, по отношению к понятиям, принятым в теории массового обслуживания, следующие: заявки – транзакты, очереди – очереди, ОА – сервисы.

Недостаток – отсутствие русифицированной версии и описания на русском языке.

1.12. Simio

ССИМ Simio [16] разработана в 2007 г. компанией Simio LLC, которая была организована в 2005 г.

Понятия, принятые в ССИМ Simio, следующие: заявки – агенты (инициаторы), очереди – очереди, ОА – серверы.

Недостаток – отсутствие русифицированной версии и описания на русском языке.

1.13. Simul8

Первая версия ССИМ Simul8 [17] была разработана одноименной корпорацией в начале 2003 г. Последняя версия ССИМ Simul8 2017 разработана в 2017 г. В ССИМ Simul8 указываются маршруты движения заявок и процессы их обслуживания [17]. Понятия, принятые в системе Bizagi Modeler, следующие: заявки – сущности, очереди – очереди, обслуживающие аппараты – рабочие центры.

Демовersion ССИМ Simul8 можно бесплатно скачать из Интернета и пользоваться ей до 14 дней. Недостаток – отсутствие русифицированной версии и описания на русском языке.

2. Аналитическое моделирование СМО М/М/5

Аналитическое моделирование (АМ) СМО М/М/5 проведено по формулам, приведённым в [18], для М/М/м и сведено к расчёту показателей функционирования СМО по ним. Для расчёта принято: количество обслуживающих аппаратов $m = 5$, среднее время между поступлением заявок $\bar{t}_{\text{post}} = 10$ единиц времени, среднее время обслуживания заявок $\bar{t}_{\text{obsl}} = 30$ единиц времени. Результаты аналитического моделирования СМО М/М/5 приведены в табл. 3.

3. Структурные модели СМО М/М/5

Структурные модели СМО М/М/5 в ССИМ: AnyLogic, Arena, Bizagi Modeler, Business Studio, Enterprise Dynamics, ExtendSim, Flexsim, GPSS W, Plant Simulation, Process Simulator, Rand Model Designer, Simio и Simul8 приведены на рис. 1 – 12.

4. Сравнение результатов имитационного и аналитического моделирования

Оценка достоверности результатов ИМ СМО М/М/5 в анализируемых ССИМ произведена по результатам их сравнения с результатами АМ по разнице средних значений по формуле

$$\Delta_{ij} = \frac{y_{ij}^* - y_i}{y_i} \cdot 100,$$

где y_{ij}^* – оценка i -го параметра, определённая по результатам ИМ в j -й ССИМ; y_i – значение i -го параметра, вычисленное по результатам АМ.

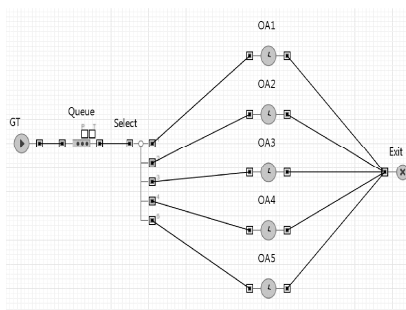


Рис. 1. Структурная модель СМО М/М/5 в ССИМ AnyLogic

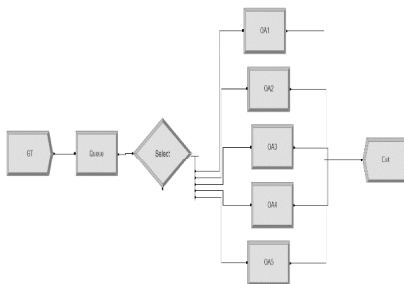


Рис. 2. Структурная модель СМО М/М/5 в ССИМ Arena

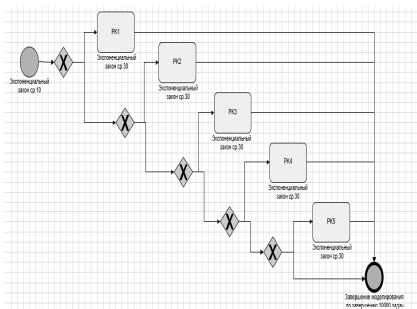


Рис. 3. Структурная модель СМО М/М/5 в ССИМ Bizagi Modeler и Business Studio

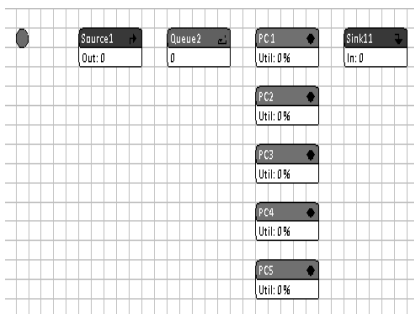


Рис. 4. Структурная модель СМО М/М/5 в ССИМ Enterprise Dynamics

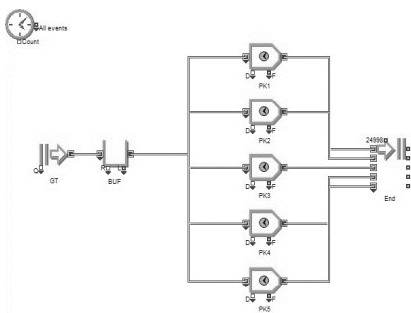


Рис. 5. Структурная модель СМО М/М/5 в ССИМ ExtendSim

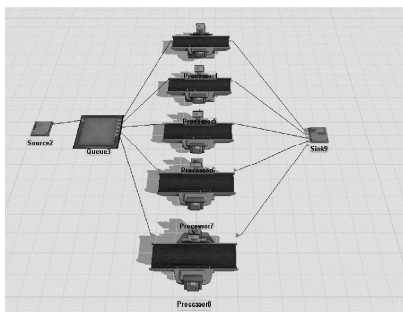


Рис. 6. Структурная модель СМО М/М/5 в ССИМ Flexsim

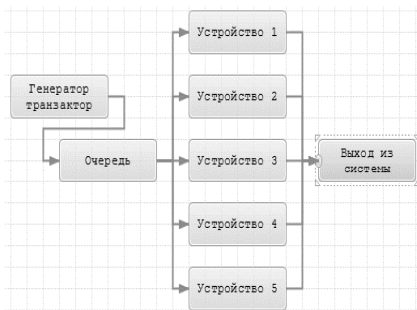


Рис. 7. Структурная модель СМО М/М/5 в ССИМ GPSS W с расширенным редактором

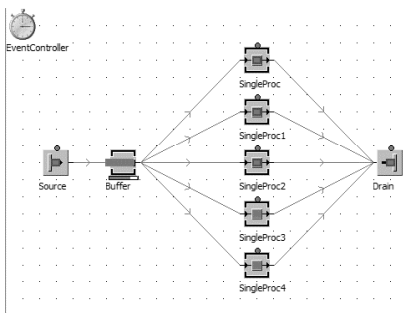


Рис. 8. Структурная модель СМО М/М/5 в ССИМ Plant Simulation

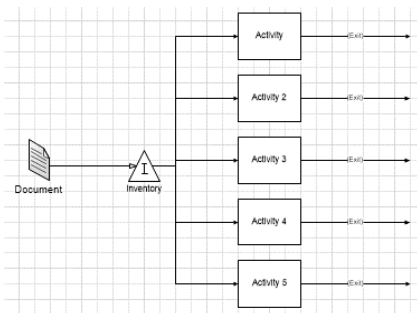


Рис. 9. Структурная схема СМО М/М/5 в ССИМ ProcessSimulator

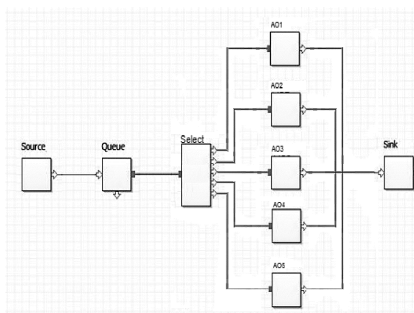


Рис. 10. Структурная модель СМО М/М/5 в ССИМ Rand Model Designer

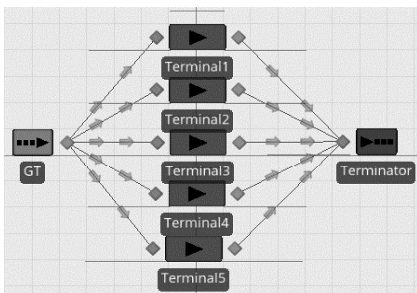


Рис. 11. Структурная модель СМО М/М/5 в ССИМ Simio

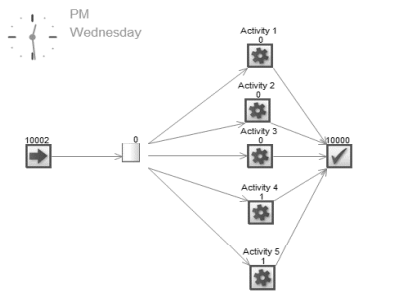


Рис. 12. Структурная модель СМО М/М/5 в ССИМ Simul8

Для ИМ принято: количество обслуживающих аппаратов $m = 5$, среднее время между поступлением заявок $\bar{t}_{\text{post}} = 10$ единиц времени, среднее время обслуживания заявок $\bar{t}_{\text{obsl}} = 30$ единиц времени. Количество обслуженных заявок 25000. Результаты АМ и ИМ в системах AnyLogic, Arena, Bizagi Modeler и Business Studio и оценка их разности в процентах по формуле (16) приведены в табл. 1.

Средняя разница в процентах по шести тестам остальных 9 ССИМ приведены в табл. 2.

Заключение

В работе проведено сравнение 13 ССИМ, которые позволяют основным пользователям – не программистам – создавать структурные и имитационные модели отказавшись от программирования и перейдя к изображению структурных схем объектов моделирования. Профессиональным программистам остаётся деятельность по программированию функционирования элементов моделируемых объектов и новых методов обработки результатов моделирования. Для сравнения ССИМ использованы структурные, имитационные и аналитические модели СМО, в частности СМО типа М/М/5. По результатам проведённой работы сделаем следующие выводы:

1. Все 13 ССИМ, структурные модели СМО М/М/5 которых приведены на рис. 1 – 12, с помощью стрелок достаточно наглядно отражают маршруты движения заявок. На основании этих рисунков можно выбрать ССИМ, наиболее подходящую по предметной области, к которой относится моделируемый объект, и на основании предпочтений пользователя.

2. Наиболее подробно условия изменения маршрутов движения заявок без отображения процессов их обработки можно указать в ССИМ Bizagi Modeler и Business Studio, которые используют нотацию моделирования бизнес-процессов BPMN 2.0.

3. Наиболее детально моделируемые процессы отражаются в ССИМ AnyLogic.

4. Средняя разница результатов ИМ и АМ СМО М/М/5 в процентах по 6 тестам для всех 13 ССИМ не превышает 5 %, что является удовлетворительным результатом.

5. По достоверности результатов по средней разнице между ИМ и АМ в процентах ССИМ можно выбрать по проранжированному перечню,

Таблица 1

Результаты оценки основных показателей функционирования СМО М/М/5 в ССИМ: AnyLogic, Arena, Bizagi Modeler и Business Studio по 6 параметрам

№	Код	Наименование	АМ	ИМ в AnyLogic	Разница в %	ИМ в Arena	Разница в %	ИМ в Bizagi Modeler	Разница в %	ИМ в Business Studio	Разница в %
1	\bar{l}	Среднее количество заявок в очереди	0,354	0,340	3,954	0,355	0,282	0,352	0,565	0,369	4,237
2	\bar{m}	Среднее количество заявок в ОА	3,000	3,090	3,000	3,152	5,067	3,075	2,500	3,125	4,167
3	\bar{k}	Среднее количество заявок в системе	3,354	3,43	2,265	3,507	4,562	3,427	2,177	3,494	4,174
4	$\bar{t}_{ожид}$	Среднее время ожидания заявок в очереди	3,542	3,500	1,185	3,681	1,920	3,520	0,621	3,664	3,444
5	$\bar{t}_{обсл}$	Среднее время задержки заявок в ОА	30,00	30,900	3,000	30,850	2,827	30,790	2,633	31,020	3,400
6	$\bar{t}_{преб}$	Среднее время пребывания заявок в системе	33,542	34,400	2,557	34,531	2,325	34,310	2,289	34,684	3,405
Средняя разница в % по 6 тестам				2,660		2,830		1,798		3,805	

Таблица 2

Средняя разница в процентах по шести тестам в ССИМ: Enterprise Dynamics, ExtendSim, Flexsim, GPSS W, Plant Simulation, ProcessSimulator, Rand Model Designer, Simio и Simul8

Наименование	ИМ в Enterprise Dynamics	ИМ в ExtendSim	ИМ в Flexsim	ИМ в GPSS W	ИМ в Plant Simul	ИМ в Process Simulator	ИМ в Rand Model	ИМ в Simio	ИМ в Simul8
Средняя разница в % по 6 тестам	1,534	1,700	3,673	1,020	2,326	1,004	2,211	1,978	0,633

в котором для ССИМ в скобках приведена средняя разница в процентах: Simul8 (0,633), Proccess Simulator (1,004), GPSS W (1,020), Enterprise Dynamics (1,534), ExtendSim (1,700), Bizagi Modeler (1,798), Simio (1,978), Rand Model Designer (2,211), Plant Simulation (2,326), AnyLogic (2,660), Arena (2,830), Flexsim (3,673), Business Studio (3,805).

ЛИТЕРАТУРА

1. Киндлер Е. Языки моделирования. – М.: Энергоатомиздат, 1985. – 288 с.
2. Борщев А.В. Имитационное моделирование: Состояние области на 2015 год, тенденции и прогноз // Труды Седьмой Всероссийской научно-практической конференции «Имитационное моделирование. Теория и практика» (ИММОД_2015). Том пленарных докладов. – С. 14–22.
3. Якимов И.М., Кирпичников А.П., Зайнуллина Г.Р., Яхина З.Т. Оценка достоверности результатов имитационного моделирования по результатам аналитического моделирования // Вестник Казанского технологического ун-та. – 2015. – Т. 18. – № 6. – С. 173–178.
4. Якимов И.М., Кирпичников А.П., Зайнуллина Г.Р. и др. Информационная система имитационного и аналитического моделирования систем массового обслуживания // Вестник Казанского технологического ун-та. – 2016. – Т. 19. – № 5. – С. 141–145.
5. Мокшин В.В., Якимов И.М. Моделирование систем в среде AnyLogic: практикум. Рекомендации к выполнению лабораторных и практических работ. – Казань: Ред.-изд. центр «Школа», 2014. – 252 с.
6. Щербаков С.М. Имитационное моделирование экономических процессов в системе Agena: учеб. пособие. – Ростов н/Д: Рост. гос. эконом. ун-т, 2012. – 128 с.
7. Якимов И.М., Кирпичников А.П., Мокшин В.В. и др. Имитационное моделирование бизнес-процессов в системе Bizagi Modeler // Вестник Казанского технологического ун-та, 2015. – Т. 18. – № 9. – С. 236–239.
8. Якимов И.М., Кирпичников А.П., Захарова З.Х., Низамиев Ф.А. Имитационное моделирование бизнес-процессов в системе Buisness Studio // Вестник Казанского технологического ун-та. – 2016. – Т. 19. – № 11. – С. 173–177.
9. Якимов И.М., Елисеев И. А, Шуляпова Н.И. Имитационное моделирование в системе Enterprise Dynamics 9.0 // Журнал Интеллект. Инновации. Инвестиции. – 2016. – № 4. – С. 108–113.
10. Якимов И.М., Кирпичников А.П., Павлов А.Д. Моделирование систем массового обслуживания и обучение моделированию в среде EXTENDSIM // Вестник Казанского технологического ун-та. – 2016. – Т. 19. – № 24. – С. 126–129.
11. Якимов И.М., Кирпичников А.П. Имитационное моделирование вероятностных объектов в системе Flexsim // Вестник Казанского технологического ун-та. – 2016. – Т. 19. – № 21. – С.170–173.
12. Девятков В.В. Расширенный редактор GPSS World: основные возможности. – М.: Издание ООО «Принт-сервис», 2013. – 143 с.

-
13. Якимов И.М., Кирпичников А.П., Захарова З.Х., Железнякова Д.Д. Имитационное моделирование в системе Plant Simulation // Вестник Казанского технологического ун-та. – 2017. – Т. 20. – № 2. – С. 107–111.
 14. Интернет-ресурс: <http://www.promodel.com/>.
 15. Якимов И.М., Кирпичников А.П., Халиуллин Р.Ф. и др. Имитационное моделирование в системе Rand Model Designer // Вестник Казанского технологического ун-та. – 2017. – Т. 20. – № 2. – С.116–120.
 16. Якимов И.М., Кирпичников А.П., Родина Р.В., Сафина Д.И. Моделирование бизнес-процессов в системе имитационного моделирования Simio // Вестник Казанского технологического ун-та. – 2016. – Т. 19. – № 11. – С. 149–153.
 17. *Concannon et al.* Simulation Modeling with SIMUL8, USA, 2003-11-01, ISBN 0-9734285-0-3. – Р. 410.
 18. Кирпичников А.П. Методы прикладной теории массового обслуживания. – Казань: Изд-во Казан. ун-та, 2011. – 200 с.

МОДЕЛИРОВАНИЕ ТЕЛЕКОММУНИКАЦИОННЫХ СЕТЕЙ И ТЕЛЕТРАФИКА

USING DOBRUSHIN MEAN-FIELD APPROACH FOR LARGE-SCALE TRANSPORT NETWORKS ANALYSIS*

A.I. Blinov, S.A. Vasilyev, L.A. Sevastianov

Peoples' Friendship University of Russia (RUDN University), Moscow, Russia

In this paper large-scale transport systems was studied using Dobrushin mean-field approach [1–3]. We assume that the transport networks faced with the problem of proving the global convergence of the solutions of certain infinite systems of ordinary differential equations to a time-independent solution. In work [3] the infinite systems of differential equations that modelling of large-scale transport systems are studied and the sufficient conditions of global stability and global asymptotic stability are obtained. Cauchy problems for systems of ordinary differential equations of infinite order was investigated A.N.Tihonov [4] and other researchers. In this paper we apply Dobrushin mean-field approach from [3] for the singular perturbed systems of ordinary differential equations of infinite order of Tikhonov type. It was studied the singular perturbed systems of ordinary differential equations by A.N. Tihonov [5] and other researchers.

In the paper [6] we investigated same questions which relate to systems of ordinary differential equations of infinite order with a small parameter and initial conditions. It was proved same essential theorems such as the local existence of solutions theorem.

In this paper we considered large-scale transport network systems that consists of infinite number of network service nodes with a Poisson input

* The reported study was funded within the Agreement No 02.a03.21.0008 dated 24.11.2016 between the Ministry of Education and Science of the Russian Federation and RUDN University.

flow of requests. Each requests arriving to the system randomly selects two network service nodes and is instantly sent to the one with the shorter queue. In this case a share $u_k(t)$ of the servers nodes that have the queues lengths with not less than k can be described using an infinite system of differential equations. It is possible to investigate Tikhonov type Cauchy problem for this transport system with a small parameter μ and initial conditions using the singular perturbation methods [5]. Using the truncation method it was studied a finite transport system with nodes using system of differential equations with a small parameter μ order N . The evolution analysis of $(k = 1, 2, \dots, N)$ be applied to application in nodes queueing for large-scale transport network systems analysis.

Large-scale transport networks model

Let's consider a large-scale transport networks that consists of nodes in which can be services for the production, storage, sale of goods, or in nodes can be provided certain services. Let all these servers have a Poisson input flow of requests of intensity $N\lambda$. Each request arriving to the system randomly selects two server nodes and is instantly sent to the one with the shorter queue. The service time in a node is distributed exponentially with mean $\bar{t} = 1$. Let $u_k(t)$ be a share server nodes that have the queues lengths with not less than k . The considered system of the servers nodes is described by ergodic Markov chain. There is a stationary probability distribution for the states of the system and if $N \rightarrow \infty$ the evolution of the values $u_k(t)$ becomes deterministic and the Markov chain asymptotically converges to a dynamic system the evolution of which is described by infinite system of differential-difference equations

$$\begin{cases} \dot{u}_0(t) = \lambda u_1(t) - u_0(t), \\ \dot{u}_1(t) = \lambda(u_2(t) - u_1(t)) - u_0(t)(1 - u_1(t)), \\ \dot{u}_k(t) = \lambda(u_{k+1}(t) - u_k(t)) - u_0(t)(u_{k-1}(t) - u_k(t)), \\ \sum_{k=0}^{\infty} u_k(t) = r, \quad r > 0, \quad u_k(0) = g_k \geq 0, \quad k = 0, 1, 2, \dots, \end{cases} \quad (1)$$

where $t \geq 0$, r is a positive parameter and $g = \{g_k\}_{k=1}^{\infty}$ is a numerical sequence $(1 = g_0 \geq g_1 \geq g_2 \geq \dots)$ [3].

We can investigate infinite system of differential-difference equations with a small parameter such form

$$\left\{ \begin{array}{l} \dot{u}_0(t) = \lambda u_1(t) - u_0(t), \\ \dot{u}_1(t) = \lambda (u_2(t) - u_1(t)) - u_0(t)(1 - u_1(t)), \\ \dot{u}_k(t) = \lambda (u_{k+1}(t) - u_k(t)) - u_0(t)(u_{k-1}(t) - u_k(t)), \\ \quad k = 0, 1, \dots, n, \\ \mu \dot{u}_k(t) = \lambda (u_{k+1}(t) - u_k(t)) - u_0(t)(u_{k-1}(t) - u_k(t)), \\ \quad k = n + 1, \\ \sum_{k=0}^{\infty} u_k(t) = r, \quad r > 0, \quad u_k(0) = g_k \geq 0, \quad k = 0, 1, 2, \dots, \end{array} \right. \quad (2)$$

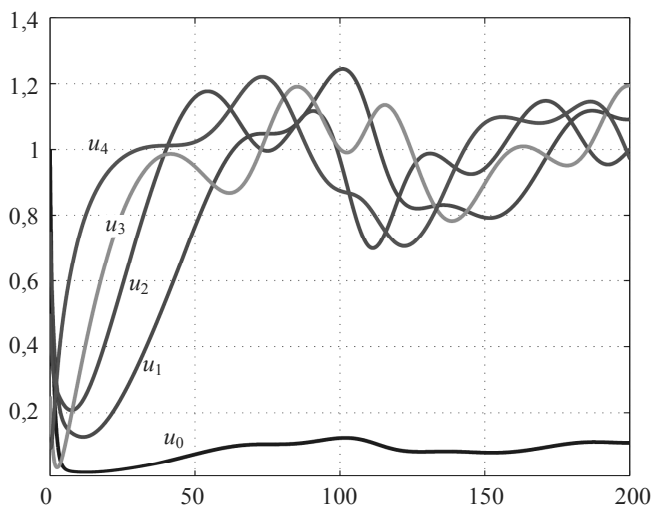
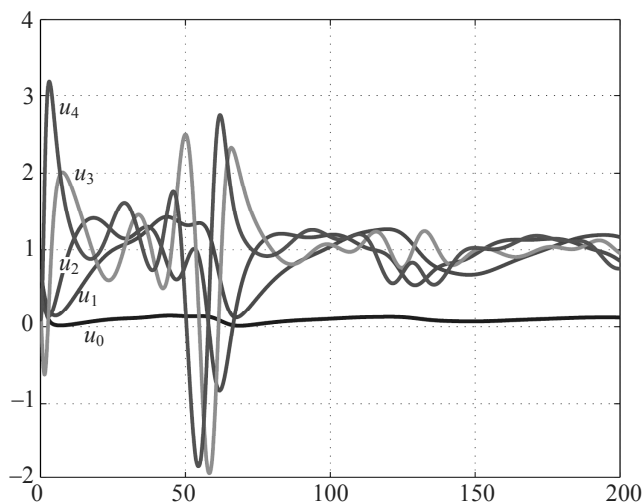
where μ is a small parameter that bring a singular perturbation to the system (1) which allows us to describe the processes of rapid change of the systems.

Truncation large-scale transport networks model and numerical analysis

Using (2) we can write the truncation system of differential-difference equations

$$\left\{ \begin{array}{l} \dot{u}_0(t) = \lambda u_1(t) - u_0(t), \\ \dot{u}_1(t) = \lambda (u_2(t) - u_1(t)) - u_0(t)(1 - u_1(t)), \\ \dot{u}_k(t) = \lambda (u_{k+1}(t) - u_k(t)) - u_0(t)(u_{k-1}(t) - u_k(t)), \\ \quad k = 0, 1, \dots, n, \\ \mu \dot{u}_k(t) = \lambda (u_{k+1}(t) - u_k(t)) - u_0(t)(u_{k-1}(t) - u_k(t)), \\ \quad k = n + 1, \dots, N, \\ \sum_{k=0}^{\infty} u_k(t) = r, \quad r > 0, \quad u_k(0) = g_k \geq 0, \quad k = 0, 1, 2, \dots, N. \end{array} \right. \quad (3)$$

The numerical example is presented in the figure (see Fig. 1, 2) where $n = 2$, $N = 4$, $l = 0.1$, $g_0 = 1$, $g_1 = 0.75$, $g_2 = 0.5$, $g_3 = 0.25$, $g_4 = 0.1$ and a small parameter $\mu = 0.1$ (Fig. 1), $\mu = 0.01$ (Fig. 2). In this numerical example it was shown the existence of steady state conditions for evolutions $u_i(t)$, $i = 0, 1, 2$ and quasi-periodic conditions with boundary layers for evolutions $u_i(t)$, $i = 3, 4$.

Fig. 1. Evolution analysis of u_k ($\mu = 0,1$)Fig. 2. Evolution analysis of u_k ($\mu = 0,01$)

Conclusions

We investigated the large-scale network transport model that consists of infinite number of server nodes with a Poisson input flow of requests of intensity $N\lambda$. Each requests arriving to the system randomly selects two servers and is instantly sent to the one with the shorter queue. In this case a share $u_k(t)$ of the servers that have the queues lengths with not less than k can be described using an infinite system of differential equations. For Tikhonov problem for infinite system of differential equations with a small parameter μ and initial conditions we applied the truncation method and studied a finite system of differential equations with a small parameter μ order N . The numerical analysis of $u_k(t)$ ($k=1,2,\dots,5$) be applied for queueing large-scale networks evolution conditions. It was shown the existence of stady state conditions for evolutions $u_i(t)$ ($i=0,1,2$) and quasi-periodic conditions with boundary layers for evolutions $u_i(t)$ ($i=3,4$).

REFERENCES

1. *Khmelev D.V., Oseledets V.I.* Mean-field approximation for stochastic transportation network and stability of dynamical system // Preprint No. 434 of University of Bremen, 1999.
2. *Khmelev D.V.* Limit theorems for nonsymmetric transportation networks // *Fundamentalnaya i Prikladnaya Matematika*. – 2001. – V. 7. – No. 4. – P. 1259–1266.
3. *Oseledets V.I., Khmelev D.V.* Global stability of infinite systems of nonlinear differential equations, and nonhomogeneous countable Markov chains // *Problemy Peredachi Informatsii* (in Russian). – 2000. – V. 36. – Issue 1. – P. 60–76.
4. *Tihonov A.N.* Uber unendliche Systeme von Differentialgleichungen // *Rec. Math.* – 1934. – V. 41. – Issue 4. – P. 551–555.
5. *Tihonov A.N.* Systems of differential equations containing small parameters in the derivatives // *Mat. Sbornik N. S.* – 1952. – V. 31. – Issue 73. – P. 575–586.
6. *Vasilyev S.A., Bolotova G.O., Udin D.N.* Systems of differential equations of infinite order with small parameter and countable Markov chains // “Distributed computer and communication networks: control, computation, communications (DCCN-2016)”. *Communications in Computer and Information Science* (Springer). – 2016. – V. 678. – P. 565–576 (DOI: 10.1007/978-3-319-51917-3_48).

DOBRUSHIN MEAN-FIELD APPROACH AND LARGE-SCALE NETWORKS WITH A SMALL PARAMETER*

S.A. Vasilyev, G.O. Tsareva

Peoples' Friendship University of Russia (RUDN University), Moscow, Russia

The recent research of service networks with complex routing discipline in transport networks faced with the problem of proving the global convergence of the solutions of certain infinite systems of ordinary differential equations to a time-independent solution. Scattered results of these studies, however, allow a common approach to their just function. This approach will be expounded here. In work the countable systems of differential equations with bounded Jacobi operators are studied and the sufficient conditions of global stability and global asymptotic stability are obtained. In this paper we apply Dobrushin mean-field approaches from [1] for the singular perturbed systems of ordinary differential equations of infinite order of Tikhonov type. Cauchy problems for the systems of ordinary differential equations of infinite order was investigated A.N.Tihonov [2], R. Bellman [3], K.P.Persidsky [4], O.A.Zhautykov [5] and other researchers. It was studied the singular perturbed systems of ordinary differential equations by A.N. Tihonov [6], A.B.Vasil'eva [7] and other researchers.

In the papers [8] we investigated same questions which relate to systems of ordinary differential equations of infinite order with a small parameter and initial conditions. It was proved same essential theorems such as the local existence of solutions theorem. In this paper we considered a system that consists of infinite number of servers with a Poisson input flow of requests of intensity $N\lambda$. Each requests arriving to the system randomly selects two servers and is instantly sent to the one with the shorter queue. In this case a share $u_k(t)$ of the servers that have the queues lengths with not less than k can be described using an infinite system of differential equations. It is possible to investigate Tikhonov type Cauchy problem for this system with a small parameter μ and initial conditions using the singular perturbation

* The reported study was funded within the Agreement No 02.a03.21.0008 dated 24.11.2016 between the Ministry of Education and Science of the Russian Federation and RUDN University.

methods [6, 7]. Using the truncation method it was studied a infinite system of differential equations with a small parameter μ order N . The evolution analysis of $u_k(t)$ ($k=1,2,\dots,N$) be applied to application in queueing large-scale networks analysis

Large-scale network model

Let's consider a system that consists of N servers with a Poisson input flow of requests of intensity $N\lambda$. Each request arriving to the system randomly selects two servers and is instantly sent to the one with the shorter queue. The service time is distributed exponentially with mean $\bar{t}=1$. Let $u_k(t)$ be a share servers that have the queues lengths with not less than k . It is possible to investigate the asymptotic distribution of the queue lengths as $N \rightarrow \infty$ and $\lambda < 1$ [1]. The considered system of the servers is described by ergodic Markov chain. There is a stationary probability distribution for the states of the system and if $N \rightarrow \infty$ the evolution of the values $u_k(t)$ becomes deterministic and the Markov chain asymptotically converges to a dynamic system the evolution of which is described by infinite system of differential-difference equations

$$\begin{cases} \dot{u}_k(t) = u_{k+1}(t) - u_k(t) + \lambda((u_{k-1}(t))^2 - (u_k(t))^2), \\ u_k(0) = g_k \geq 0, \\ k = 0, 1, 2, \dots, \end{cases} \quad (1)$$

where $g = \{g_k\}_{k=1}^{\infty}$ is a numerical sequence ($1 = g_0 \geq g_1 \geq g_2, \dots$) [1].

We can investigate infinite system of differential-difference equations with a small parameter such form

$$\begin{cases} \dot{u}_k(t) = u_{k+1}(t) - u_k(t) + \lambda((u_{k-1}(t))^2 - (u_k(t))^2), \\ k = 0, 1, \dots, n, \\ \mu \dot{u}_k(t) = u_{k+1}(t) - u_k(t) + \lambda((u_{k-1}(t))^2 - (u_k(t))^2), \\ k = n+1, \dots, \\ u_k(0) = g_k \geq 0, \\ k = 0, 1, 2, \dots, \end{cases} \quad (2)$$

where μ is a small parameter that bring a singular perturbation to the system (1) which allows us to describe the processes of rapid change of the systems.

Truncation large-scale network model and numerical analysis

Using (2) we can write the truncation system of differential-difference equations

$$\begin{cases} \dot{u}_k(t) = u_{k+1}(t) - u_k(t) + \lambda \left((u_{k-1}(t))^2 - (u_k(t))^2 \right), \\ k = 0, 1, \dots, n; \\ \mu \dot{u}_k(t) = u_{k+1}(t) - u_k(t) + \lambda \left((u_{k-1}(t))^2 - (u_k(t))^2 \right), \\ k = n+1, \dots, N; \\ u_k(0) = g_k \geq 0, k = 0, 1, 2, \dots, N; \\ u_{N+1} = g_{N+1} \geq 0. \end{cases} \quad (3)$$

The numerical example is presented in the figures (Fig. 1 and 2), where $n=2$, $N=4$, $l=0.2$, $g_k=1$, $k=\overline{0,5}$ and a small parameter $\mu=0,1$ (Fig. 1), $\mu=0,01$ (Fig. 2). In this numerical example it was shown the existence of steady state conditions for evolutions $u_i(t)$ ($i=0,1,2$) and quasi-periodic conditions with boundary layers for evolutions $u_i(t)$ ($i=3,4$).

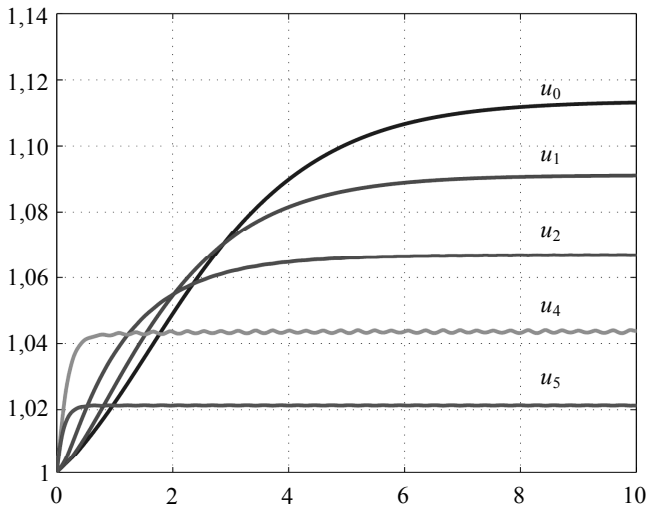


Fig.1. Evolution analysis of u_k ($\mu = 0.1$)

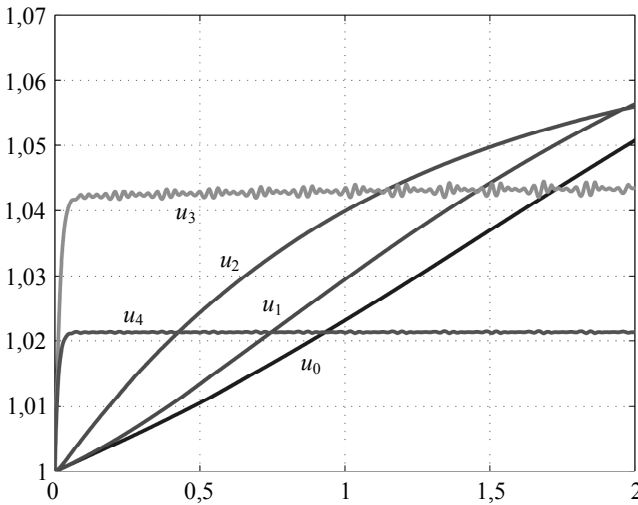


Fig.2. Evolution analysis of u_k ($\mu = 0.01$)

Conclusions

We investigated the large-scale network model that consists of infinite number of servers with a Poisson input flow of requests of intensity $N\lambda$. Each requests arriving to the system randomly selects two servers and is instantly sent to the one with the shorter queue. In this case a share $u_k(t)$ of the servers that have the queues lengths with not less than k can be described using an infinite system of differential equations. For Tikhonov problem for infinite system of differential equations with a small parameter μ and initial conditions we applied the truncation method and studied a finite system of differential equations with a small parameter μ order N . The numerical analysis of $u_k(t)$ ($k = 1, 2, \dots, 5$) be applied for queueing large-scale networks evolution conditions. It was shown the existence of steady state conditions for evolutions $u_i(t)$, $i = 0, 1, 2$ and quasi-periodic conditions with boundary layers for evolutions $u_i(t)$, $i = 3, 4$.

REFERENCES

1. *Vvedenskaya N.D., Dobrushin R.L., Kharpelevich F.I.* Queueing system with a choice of the lesser of two queues – the asymptotic approach // *Probl. inform.* – 1996. – V. 32. – Issue 1. – P.15–27.
2. *Tihonov A.N.* Uber unendliche systeme von differentialgleichungen // *Rec. Math.* – 1934. – V. 41. – Issue 4. – P. 551–555.
3. *Bellman R.* The boundedness of solutions of in_finite systems of linear differential equations // *Duke Mathematical J.* – 1947. – V. 14. – No 3. – P. 695–706.
4. *Persidsky K.P.* // *Izv. AN KazSSR. Ser. Mat. Mach.* – 1948. – Issue 2. – P. 3–34.
5. *Zhautykov O.A.* Extension of the Hamilton-Jacobi theorems to an infinite canonical system of equations // *Mat. Sb. (N.S.).* – 1961. – V. 53. – Issue 95. – P. 313–328.
6. *Tihonov A.N.* Systems of differential equations containing small parameters in the derivatives // *Mat. Sbornik (N. S.).* – 1952. – V. 31. – Issue 73. – P. 575–586.
7. *Vasil'eva A.B.* Asymptotic behaviour of solutions of certain problems for ordinary nonlinear differential equations with a small parameter multiplying the highest derivatives // *Uspehi Mat. Nauk.* – 1963. – V. 18. – Issue 111. – No. 3. – P. 15–86.
8. *Vasilyev S.A., Bolotova G.O., Udin D.N.* Systems of differential equations of infinite order with small parameter and countable Markov chains // “Distributed computer and communication networks: control, computation, communications (DCCN-2016)”. *Communications in Computer and Information Science (Springer).* – 2016. – V. 678. – P. 565–576 (DOI: 10.1007/978-3-319-51917-3_48).
9. *Oseledets V.I., Khmelev D.V.* Global stability of infinite systems of nonlinear differential equations, and nonhomogeneous countable Markov chains // *Problemy Peredachi Informatsii (in Russian).* – 2000. – V. 36. – Issue 1. – P. 60–76.
10. *Vvedenskaya N.D., Suhov Yu.M.* Dobrushin's mean-field approximation for a queue with dynamic routing // *Markov Processes and Related Fields.* – 1997. – Issue 3. – P. 493–526.
11. *Vvedenskaya N.D.* A large queueing system with message transmission along several routes // *Problemy Peredachi Informatsii.* – 1998. – V. 34. – No. 2. – P. 98–108.

STABILITY ANALYSIS OF THE INVERTED PENDULUM UNDER STOCHASTIC DRIVING FORCES*

D.G. Vasilyeva, S.A. Vasilyev, L.A. Sevastianov

Peoples' Friendship University of Russia (RUDN University), Moscow, Russia

Kapitza pendulum is a rigid pendulum in which the pivot point vibrates in a vertical direction, up and down. The unique feature of the Kapitza pendulum is that the vibrating suspension can cause it to balance stably in an inverted position. A pendulum with vibrating point is a classical problem of perturbation theory. The phenomenon of stabilisation of the upper vertical position of the pendulum by fast vertical vibrations of the suspension point was discovered by A. Stephenson [1, 2]. P.L. Kapitza has developed a method of separation of slow and fast motions for the pendulum [3, 4]. Different aspects of this problem were discussed in many publications [5, 6–8]. In this work generalizations to Kapitza pendulum whose suspension point moves in the vertical and horizontal planes is investigated. Lyapunov stability analysis of the motion for this pendulum subjected to excitation of stochastic driving forces that random act in the vertical and horizontal planes has been studied. The numerical study of the random motion for generalized Kapitza pendulum under stochastic driving forces has made. It is shown the existence of stable quasi-periodic motion this pendulum.

Generalized Kapitza pendulum model

Let l and m be length of the massless rod and mass of the bob for this pendulum. Let $x(t)$ and $y(t)$ be horizontal and vertical Cartesian coordinates of the suspension point. Denote by $\theta(t)$ the angle between the rod of the pendulum and the vertical. In this case we can written the coordinates of bob in the form

$$\begin{cases} x(t) = l \sin \theta(t) + a(t), \\ y(t) = l \cos \theta(t) + b(t), \end{cases} \quad (1)$$

* The reported study was funded within the Agreement No 02.a03.21.0008 dated 24.11.2016 between the Ministry of Education and Science of the Russian Federation and RUDN University.

where $a(t)$ and $b(t)$ are random functions that associated with acting upon the pendulum of stochastic forces that drive in the vertical and horizontal planes oscillations along the axes x and y . If in (1) we take the time derivative of $x(t)$ and $y(t)$ we can write

$$\begin{cases} v_x = \dot{x}(t) = l \cos \theta(t) \dot{\theta}(t) + \dot{a}(t), \\ v_y = \dot{y}(t) = -l \sin \theta(t) \dot{\theta}(t) + \dot{b}(t), \end{cases} \quad (2)$$

where $v_x = \dot{x}$, $v_y = \dot{y}$ are velocities along each of the axes. Then the kinetic and potential energies of the bob are

$$K = \frac{m}{2}(\dot{x}^2 + \dot{y}^2) = \frac{m}{2} \left[(l \cos \theta(t) \dot{\theta}(t) + \dot{a}(t))^2 + (-l \sin \theta(t) \dot{\theta}(t) + \dot{b}(t))^2 \right]; \quad (3)$$

$$V = mg y(t) = mg (l \cos \theta(t) + b(t)), \quad (4)$$

where m is mass of the bob and g is acceleration of free fall.

The total energy of the system is given by the sum of the kinetic and potential energies

$$E = K + V = \frac{m}{2}(\dot{x}^2 + \dot{y}^2) + mg (l \cos \theta(t) + b(t)), \quad (4)$$

and then Lagrangian of the system has the form

$$L = K - V = \frac{m}{2}(\dot{x}^2 + \dot{y}^2) - mg (l \cos \theta(t) + b(t)). \quad (5)$$

The Euler – Lagrange equation for the phase of the pendulum as follows

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{\theta}} - \frac{\partial L}{\partial \theta} = 0,$$

and the equation of the pendulum motion has the form

$$\ddot{\theta}(t) + \frac{\ddot{a}(t)}{l} \cos \theta - \frac{\ddot{b}(t) + g}{l} \sin \theta = 0.$$

We can rewrite this equation in the form of the system differential equations first order

$$\begin{cases} \dot{\theta}(t) = \varphi(t), \\ \dot{\varphi}(t) = -l^{-1} \ddot{a}(t) \cos \theta(t) + l^{-1} (\ddot{b}(t) + g) \sin \theta(t), \end{cases} \quad (6)$$

where $\varphi(t)$ is an auxiliary variable.

Lyapunov stability analysis for generalized Kapitza pendulum

In this section Lyapunov stability analysis of the motion for this pendulum subjected to excitation of stochastic driving forces that random act in the vertical and horizontal planes has been studied.

Let θ_0 and φ_0 be parameters of the bob balance when equilibrium conditions can be obtained. Let $\Delta\theta$ and $\Delta\varphi$ be small deviations from the equilibrium state. In this case, the motion of the pendulum can be described as follows:

$$\begin{cases} \theta = \theta_0 + \Delta\theta(t), \\ \varphi = \varphi_0 + \Delta\varphi(t). \end{cases} \quad (7)$$

Parameters θ_0 and φ_0 can be found from

$$\begin{cases} \varphi_0 = 0, \\ \ddot{a}(t)\cos\theta_0 - (\ddot{b}(t) + g)\sin\theta_0 = 0. \end{cases} \quad (8)$$

If the mathematical expectation $M_a = M\ddot{a}(t)$ and $M_b = M\ddot{b}(t)$ exist we can get the coordinates of the equilibrium position for the pendulum

$$\begin{cases} \varphi_0 = 0, \\ \tan\theta_0 = \frac{M_a}{M_b + g}. \end{cases} \quad (9)$$

Using the substitution variables (7) and linearization we can write the system

$$\begin{cases} \Delta\dot{\theta} = \Delta\varphi(t), \\ \Delta\dot{\varphi} = l^{-1}(-M_a \sin\theta_0 + (M_b + g)\cos\theta_0)\Delta\theta(t). \end{cases} \quad (10)$$

The characteristic polynomial of the linear system (10) has the form

$$D(p) = p^2 - l^{-1}(-M_a \sin\theta_0 + (M_b + g)\cos\theta_0). \quad (11)$$

In this case the stability condition given by $D(p) > 0$ and

$$\tan\theta_0 > \frac{M_b + g}{M_a}. \quad (12)$$

Numerical study of the generalized Kapitza pendulum motion

In this section the numerical study of the random motion for generalized Kapitza pendulum under stochastic driving forces has made. Suppose that random functions $a(t)$ and $b(t)$ can be represented in the form

$a(t) = A(t)\sin v_1 t$, $b(t) = B(t)\sin v_2 t$, where $A(t)$ and $B(t)$ are random functions of amplitudes, v_1 and v_2 are the frequencies of the harmonic vertical and horizontal forced oscillations of the pendulum suspension.

We can rewrite the system differential equations (6) in the form

$$\begin{cases} \dot{\theta} = \varphi(t), \\ \dot{\varphi} = -l^{-1} \left(\ddot{A}(t)\sin v_1 t + 2\dot{A}(t)v_1 \cos v_1 t - A(t)v_1^2 \sin v_1 t \right) \cos \theta(t) + \\ + l^{-1} \left(\ddot{B}(t)\sin v_2 t + 2\dot{B}(t)v_2 \cos v_2 t - B(t)v_2^2 \sin v_2 t + g \right) \sin \theta(t), \end{cases} \quad (13)$$

where \dot{A}, \dot{B} are random velocities and \ddot{A}, \ddot{B} are random accelerations of the pendulum suspension long each of the axes.

The numerical example is presented in the figure (Fig. 1), where $l = 0.5 \text{ m}$ is length of the massless rod, $v_1 = 40 \text{ sec}^{-1}$, $v_2 = 10 \text{ sec}^{-1}$ are the frequencies of the harmonic vertical and horizontal forced oscillations of the pendulum suspension, $g = 10 \text{ m/sec}^2$ is acceleration of free fall.

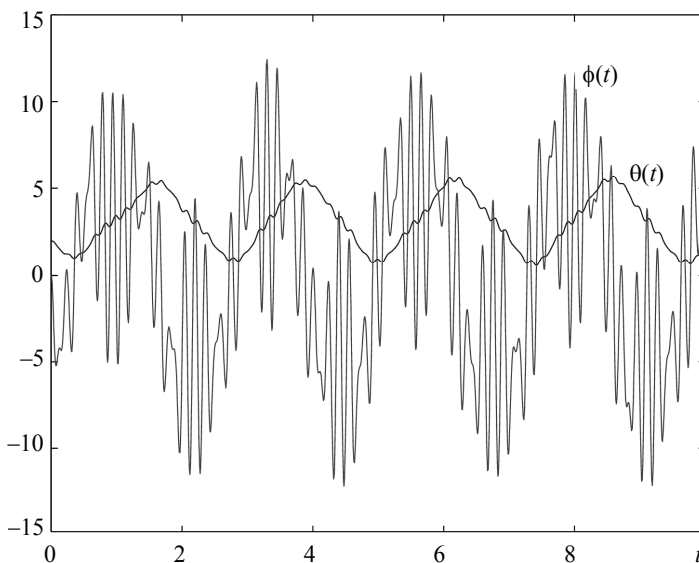


Fig. 1. The motion of generalized Kapitza pendulum

The random functions of amplitudes $A(t)$ and $B(t)$ were taken in the form

$$\begin{cases} A(t) = a_0 + \epsilon_t, \\ B(t) = b_0 + \eta_t, \end{cases} \quad (14)$$

where parameters $a_0 = 0,1$, $b_0 = 0,05$, $\epsilon_t \sim N(0, \sigma_x^2)$ and $\eta_t \sim N(0, \sigma_y^2)$ are independent normally distributed random variables (in the numerical example $\sigma_x = 0,01$ and $\sigma_y = 0,001$). In this numerical example it was shown the existence of stable quasi-periodic motion this pendulum.

Conclusions

In this work generalizations to Kapitza pendulum whose suspension point moves in the vertical and horizontal planes was investigated. Lyapunov stability analysis of the motion for this pendulum subjected to excitation of stochastic driving forces that random act in the vertical and horizontal planes had been studied. The numerical study of the random motion for generalized Kapitza pendulum under stochastic driving forces had made. It was shown the existence of stable quasi-periodic motion this pendulum. The vertical and horizontal motion simulation of the pendulum can be used for analysis of buildings and structure stability during earthquake.

REFERENCES

1. *Stephenson A.* On induced stability // Philosophical Magazine Series. – 1908. – V. 6. – No. 15. – P. 233–236.
2. *Stephenson A.* On a new type of dynamical stability // Mem. Lit Phil Soc. – 1908. – V. 52. – No. 8. – P. 1–10.
3. *Kapitsa P.L.* Dynamic stability of a pendulum with oscillating point of suspension // Sov. Phys. JETP (in Russian). – 1951. – V. 21. – P. 588–597.
4. *Kapitsa P.L.* A pendulum with vibrating point of suspension // Usp. Phys. Nauk. (in Russian). – 1951. – V. 44. – P. 7–20.
5. *Bardin B.S., Markevich A.P.* On the stability of equilibrium of a pendulum with vertical oscillations of its suspension point // J. Appl. Math. Mech. – 1995. – V. 59. – P. 879–886.
6. *Levi M.* Geometry of Kapitza's potential // Nonlinearity. – 1988. – V. 11. – P. 1365–1368.
7. *Neishtadt A.I., Shenga K.* Bifurcations of phase portraits of pendulum with vibrating suspension point. – URL: <https://arxiv.org/abs/1605.09448> (29 Sep. 2016).
8. *Ovseyevich A.I.* The stability of an inverted pendulum when there are rapid random oscillations of the suspension point // J. Appl. Math. Mech. – 2006. – V. 70. – P. 761–768.

HIDDEN INFORMATION TRANSMISSION THROUGH WHATSAPP

G.N. Huseynova¹, E.A. Mustafayeva²

¹Azerbaijan State Oil and Industrial University, Baku, Azerbaijan

²Azerbaijan National Academy of Sciences, Baku, Azerbaijan

The Internet Network gives a wide range of chances to transmit confidential information in terms of creating secret channels. These stenographic channels can be created in internet protocols, social networks, cloud technology.

One of these covert channels founded by Brian Acton and Jan Koum in 2009 and gained a wide audience of users within a short time can be created by the opportunities presented by WhatsApp.

WhatsApp is famous among the most broadly utilized personal messaging mobile applications for free messaging and object sharing (in particular audio, video, pictures, location and contacts), with more than 800 million clients worldwide and was purchased by facebook in 2014 for \$19 Billion. As of late, WhatsApp was procured by world's most youthful business person "Mr. Mark Zuckerberg" after its popularity and being recorded among one of the well known messaging application [1].

Being widespread addition for smartphones, WhatsApp gives users an opportunity to send SMS in a special form of information exchange. Here the information can be sent in forms of short messages, photos, videos and audio. It is possible to create a chance for transmitting the information secretly via WhatsApp, in other words secret information channel can be created.

WhatsApp users can transmit certain information secretly, in other words, can create a secret communication channel by their own accounts. So, using JPEG (Joint Photographic Experts Group) stenographic method, the information can be hidden in the images of JPEG format and thus it can be transmitted to other users.

Among all covert channel, the JPEG is the most prevalent medium for steganography. There are various techniques accessible for steganography. In any case, just a couple of algorithms work with JPEG compressed pictures. JPEG-based steganography calculations work just and just with the DCT (Discrete Cosine Transformation) coefficients, that are whole number values

in the scope of $[-2p-1 \dots 2p-1]$ for a p -bit encoding. Couple of other compression techniques work with DCT quantization values, considering the adjusting error. JPEG techniques do not work with quantization errors. Along these lines, there is no simple approach to control the adjustment bending [2].

Here it should be taken into account that WhatsApp may be exposed to some changes during the placement, transmission and acquisition (recovery) process of the information so that, it can result in loss of some hidden information. For example, it is not appropriate to use the profile photo as stegocontainer. Thus, as the profile photo is exposed to some changes there a loss of hidden information takes place. It has been confirmed in the result of the carried out experience. So, using JPHS and Silent Eye for Windows operating systems, profile photo functions of WhatsApp processed in Stegais and Stegos software for Android operating system, it was attempted to be used as confidential information transmitting channel, but because of the above-mentioned reasons, it resulted in failure.

It is possible to create confidential information channels using WhatsApp service opportunities. So, for example bringing any confidential message or file into the image file one of that digital objects, exchange of confidential information can be carried out between users. In practice, S-Tools, Stegan PEG, Open Stego, Quick Stego, JP Hide and Seek, Image Steganography, DeEgger Embedder, Hide N Send, SilentEye and Invisible Secrets 4 software tools, the Android operating system applied to the Secret Tigings, Stegais, Secret image, PixelKnot, Steganogropia, Ineogaito, Hidden Secrets, Pocket Stego, Photo Hidden, iMessage, VIP Secret, Stegos, Monalisa and Steg APP stenographic programs written for the Windows operating system have been used.

As a result of the experiment, it was clear that, Stegocontainers realized in Stegais and Stegos applications designed for Windows operating systems and SilentEye and JPHS programs for Android operating system have been addressed successfully without any loss.

So, for the Windows operating system that supports the JPEG format ten stenographic means: S-Tools, Stegan PEG, Open Stego, Quick Stego, JP Hide and Seek, Image Steganography, DeEgger Embedder, Hide N Send, SilentEye, Invisible Secrets 4 and for Android OS twelve stenographic means: Secret Tigings, Stegais, Secret image, Steganogropia, Hidden Secrets, Pocket Stego, Photo Hidden, iMessage, VIP Secret, Stegos, Monalisa, Steg APP stegocontainers made on stego programs were tested on the WhatsApp. However, confidential information was able to be recovered from the

stegocontainers processed in JPHS and Silent Eye for the Windows operating system, and Stegais, Stegos for the Android operating system. Hence, it can be concluded that, in Whatsapp service JPHS and Silent Eye and in the Android operating system Stegais and Stegos can be considered favorable stenographic means for exchange of confidential information.

REFERENCES

1. *Amiruzzaman Md.* Steganographic Covert Communication Channels and Their Detection: A thesis submitted to Kent State University in partial fulfillment of the requirements for the degree of Master of Science. August 2011.
2. *Karpisek F., Baggili I., Breitinger F.* WhatsApp network forensics: Decrypting and understanding the WhatsApp call signaling messages // Digital Investigation (2015). – URL: <http://dx.doi.org/10.1016/j.diin.2015.09.002> (16.04.2016).

MODELING END-TO-END BUSINESS PROCESSES OF A TELECOM COMPANY WITH A BCMP QUEUEING NETWORK*

**N. Yarkina¹, N. Popovskaya¹, V. Khalina¹,
A. Gaidamaka¹, K. Samouylov^{1,2}**

¹*Department of Applied Probability and Informatics,
Peoples' Friendship University of Russia (RUDN University), Moscow, Russia*
²*Institute of Informatics Problems, Federal Research Center "Computer Science
and Control" of the Russian Academy of Science, Moscow, Russia*

Today's communication and digital service market is highly competitive. A thorough analysis of business processes allows the service provider to estimate important factors of the telecommunication business success using variety of methods, for instance, queueing theory. Moreover, such an analysis is facilitated by the standards and recommendations, developed by the global telecommunication industry association TM Forum^{**}. TM Forum Business Process Framework, also largely known as eTOM, is one of the core models of the TM Forum Framework standards and best practices suite [1]. Also eTOM standards package includes document [2] offering a set of generic end-to-end business flows applicable to the majority of the companies in the industry. Another TM Forum standard of interest to us [3] contains the definitions of numerous business metrics permitting quantitative estimation of various aspects of telecommunication service provision.

A stand-alone eTOM end-to-end business flow was modelled as an open BCMP [4] network in our previous work [5]. Here, we go further and propose an approach to modelling several end-to-end business flows jointly. Such an approach allows to estimate not only the performance measures related to activity execution time, but also reflects resource sharing among business processes and delays due to such sharing. To illustrate this approach, we model several eTOM business flows as a single open BCMP network and show how the model can be used to evaluate a number of important TM Forum business metrics.

* The publication was financially supported by the Ministry of Education and Science of the Russian Federation (the Agreement number 02.a03.21.0008), RFBR according to the research projects No. 15-07-03608 and No. 16-07-00766.

^{**} <https://www.tmforum.org/>

BCMP Model of a Stand-Alone Business Flow

Consider eTOM end-to-end business flow Complaint-to-Solution [2]. The process deals with customer complaints related to non-technical issues and consists in identifying the source of the issue, initiating resolution and monitoring the progress. We start with the reference diagram using standard eTOM process elements and BPMN (Business Process Model and Notation). In order to model this process with a BCMP network, we first assign the process activities to service stations, or queues, which will serve as network

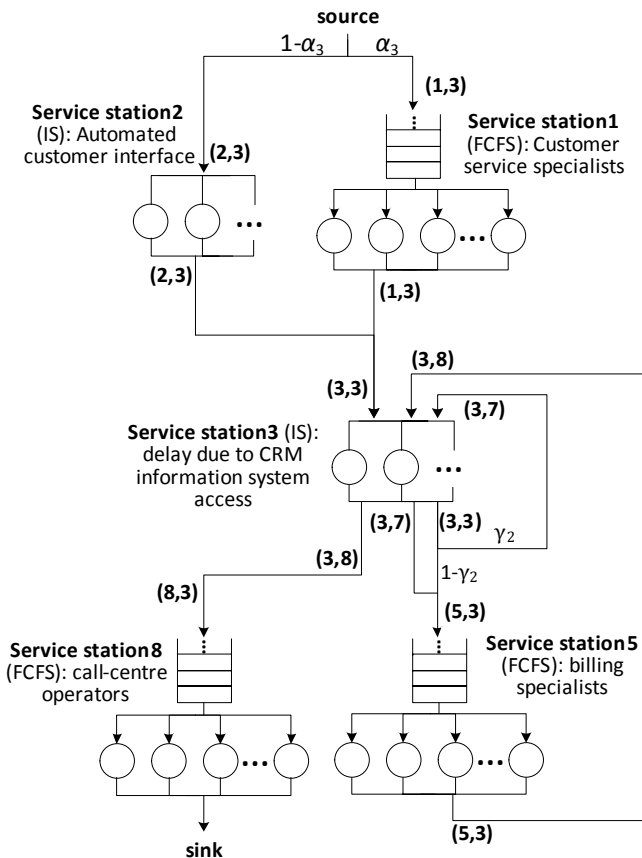


Fig. 1. BCMP model of the Complaint-to-Solution business flow

nodes. Service stations may correspond to the functional units of the company involved in the business flow (in this case, multiple-server FCFS (First Come First Served) queues are used) or to certain random delays (IS (Infinite Server) queues).

The resulting network model is shown in Fig. 1. Here, for each transition we indicate its source and destination service station and class in the form (service station, class); in case of branching, the corresponding probabilities are also indicated.

Joint Model Description and Notation

Let us introduce the formal notation for a joint BCMP model of a set of K business processes sharing certain resources/service stations. Let $M = \{1, \dots, M\}$ be the set of network's service stations, and let $R = \{1, \dots, R\}$ be the set of job classes. Service stations correspond either to the resources shared by the business processes under consideration or to the delays common to some of these processes. All stations have infinite waiting room, however they may be of different types: we shall use $|M|_{c_i}$ – FCFS queues for shared resources (we denote the set of such stations M_{FCFS}), and $|M|_{IS}$ queues for delays (denoted M_{IS}), $M_{FCFS} + M_{IS} = M$. Note that other types of service stations for which the BCMP theorem [4] holds can be used, for example a Processor Sharing (PS) station could correspond to a time-consuming data processing activity. Let $c_i \geq 1$ be the number of servers at FCFS stations, $i \in M_{FCFS}$. Job classes mainly correspond to different customers requests (e.g., information request, complaint, etc.) and serve for routing.

Now, let $\theta_{(i,r)(j,s)}, i, j \in M, r, s \in R$, denote the transition probability, i.e., the probability that a job that completes service in class r at station i will next require service in class s at station j . The routing matrix $\Theta = [\theta_{(i,r)(j,s)}]$ defines a Markov chain (MC) with the state space $L = \{(i, r), i \in M, r \in R\}$. Transition probabilities are chosen so that the MC is decomposable into K ergodic subchains, each corresponding to one business flow. In other words, for the sake of simplicity of the modelling procedure, we assign job classes and transitions between them so that each business flow is modelled with its own ergodic subchain, although this requirement can be relaxed.

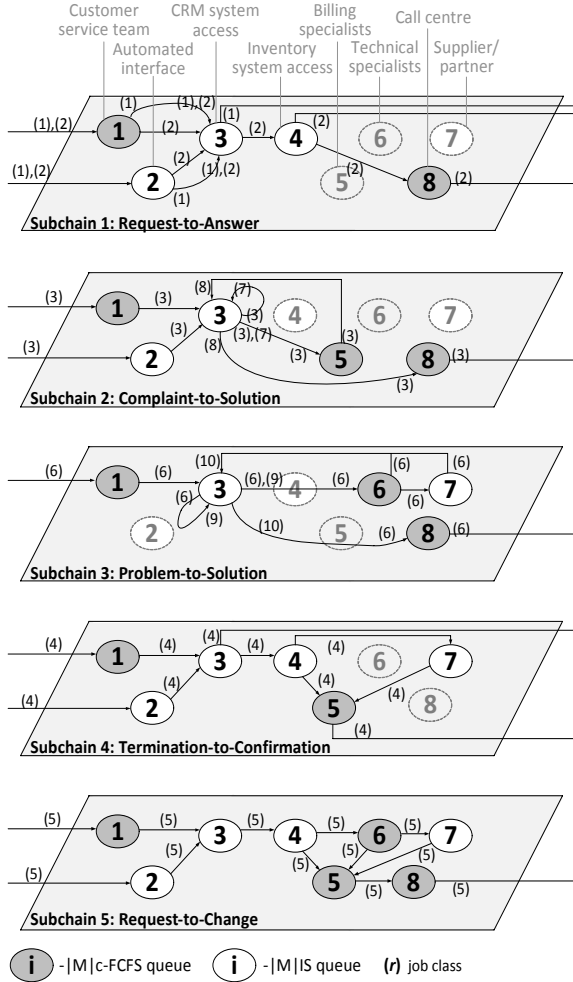


Fig. 2. Joint model transitions

Let L_k denote the set of states in subchain $k=1,\dots,K$, $L = \sum_{k=1}^K L_k$. Each subchain has its Poisson arrival stream of rate $\lambda^{(k)}$, $k=1,\dots,K$. An arrival in stream k will enter service station i in class r with probability

$q_{i,r}, \sum_{(i,r) \in L_k} q_{i,r} = 1$. Finally, a job of class r that completes service at station i

will depart the system with probability $1 - \sum_{(j,s) \in L_k} \theta_{(i,r)(j,s)} =: \theta_{(i,r)(0)}$. De-

pending on the context, (0) denotes the network's source or sink node. The general routing diagram for the joint model is depicted in Fig. 2.

Now, we apply the notation introduced above to describe a joint BCMP model of $K = 5$ end-to-end eTOM business flows: Request-to-Answer, Complaint-to-Solution, Problem-to-Solution, Termination-to-Confirmation and Request-to-Change. Please refer to [2] for the business flows description. The network contains $L = 8$ service stations, $M_{FCFS} = \{1, 5, 6, 8\}$ and $M_{JS} = \{2, 3, 4, 7\}$, and has $R = 10$ job classes. The subchains consist of the following MC states:

$$L_1 = \{(1,1), (1,2), (2,1), (2,2), (3,1), (3,2), (4,2), (8,2)\},$$

$$L_2 = \{(1,3), (2,3), (3,3), (3,7), (3,8), (5,3), (8,3)\},$$

$$L_3 = \{(1,6), (3,6), (3,9), (3,10), (6,6), (7,6), (8,6)\},$$

$$L_4 = \{(1,4), (2,4), (3,4), (4,4), (5,4), (7,4)\},$$

$$L_5 = \{(1,5), (2,5), (3,5), (4,5), (5,5), (6,5), (7,5), (8,5)\}.$$

Performance Measures

We are interested in the mean response time of the service stations, which allows to evaluate certain standard TM Forum business metrics, related to the processes under consideration (see Table 1).

Table 1

Standard TM Forum business metrics that can be estimated with the five-process joint model

Metric ID	Metric name
28 (CM-CE-2a)	<i>Average Hold Time</i>
29 (CM-CE-2b)	<i>Average Handle Time</i>
69 (A-OE-2a)	<i># Minutes per Service Problem Resolution</i>

The mean response time \bar{T}_i of station $i \in M$ can be obtained using Little's law $\bar{T}_i = \bar{n}_i / \lambda_i$, where \bar{n}_i is the mean number of jobs at station i . Thus, we have

$$\bar{T}_i = \begin{cases} \frac{c_i \mu_i}{(c_i \mu_i - \lambda_i)^2} \frac{\rho_i^{c_i}}{c_i!} \left(\sum_{m=0}^{c_i-1} \frac{\rho_i^m}{m!} + \frac{\rho_i^{c_i}}{c_i!} \frac{1}{1 - \frac{\rho_i}{c_i}} \right) + \frac{1}{\mu_i}, & i \in M_{FCFS}, \\ \frac{1}{\mu_i}, & i \in M_{IS}. \end{cases}$$

Now, we can derive expressions for standard business metrics. For example, TM Forum business metric 28 is the average hold time when a customer contacts the company by telephone, which corresponds to the waiting time at service station 1 of the five-process joint model. Hence, the following formula can be used to estimate this measure:

$$I_{28} = \bar{T}_1 - 1/\mu_1.$$

Business metric 29 is the average time needed to handle any request:

$$I_{29} = \sum_{i \in M} \bar{n}_i / \sum_{k=1}^K \lambda^{(k)}.$$

Business metric 69 is related to subchain 3 and corresponds to the mean time between the creation of a trouble ticket and its closure upon the confirmation by the customer that the problem has been resolved. Thus, we can estimate the value of the metric as the sum of the stations' mean response times multiplied by the corresponding visit ratios:

$$I_{69} = \sum_{i \in M \setminus \{1\}} \bar{T}_i \sum_{r: (i,r) \in L_3} e_{i,r}.$$

Fig. 3 shows metrics 29 and 69 as functions of $\lambda^{(3)}$, which is plotted on the X-axis. The metrics are computed for $c_6 = 2$ and $c_6 = 3$ and the curves have vertical asymptotes at $\lambda^{(3)} = 1.5$ and $\lambda^{(3)} = 2.5$ respectively. This is due to a bottleneck at service station 6, since the station utilisation $\frac{\lambda_6}{c_6 \mu_6} = \frac{\lambda^{(3)} + 0.5}{c_6}$ equals 1 at these points. Finally, Fig. 4 shows metric 28 for $c_1 = 2$ and $c_1 = 3$.

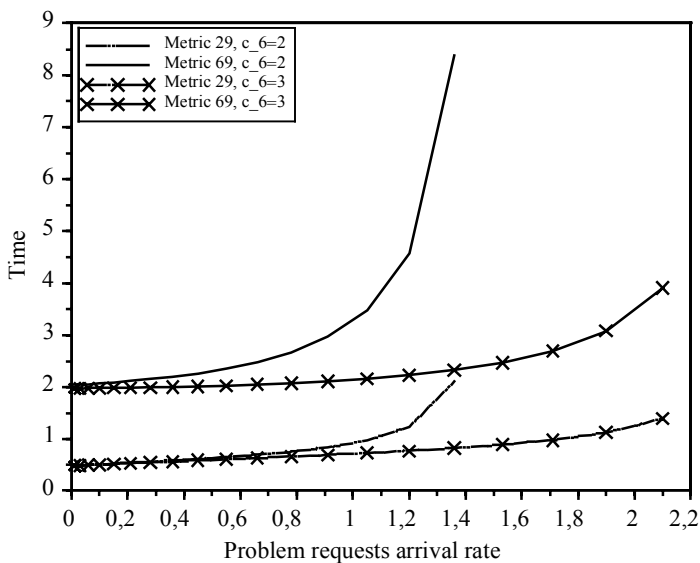


Fig. 3. Metrics 29 and 69 as functions of $\lambda^{(3)}$ for $c_6 = 2$ and $c_6 = 3$

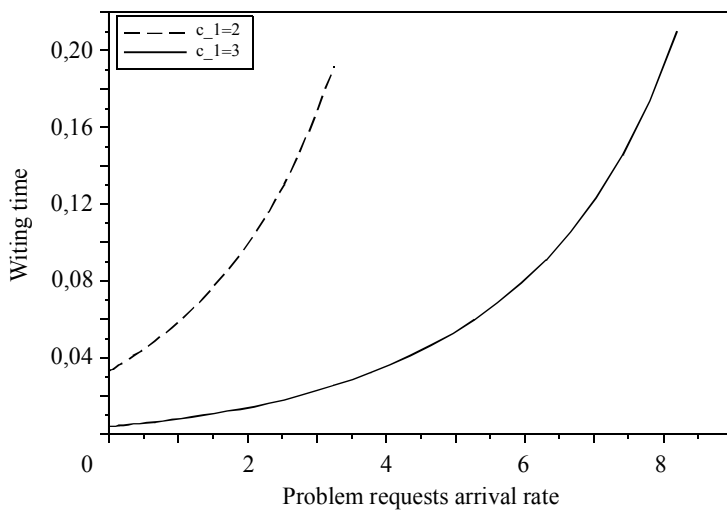


Fig. 4. Metric 28 as a function of $\lambda^{(3)}$, $c_1 = 2, 3$

Conclusion

Combining several business processes in a single queueing network allows to take into consideration resource sharing, occurring at functional units of the company (such as the technical team, the billing team, call-centre operators, etc.). We have proposed an approach to estimating certain important TM Forum business metrics using a BCMP network that combines the models of several standard end-to-end eTOM business flows. The method can be applied to analyze the set of business processes of a particular service provider.

REFERENCES

1. *TM Forum Framework*. <https://www.tmforum.org/tm-forum-frameworkx>. – 2017.
2. *Business process framework (eTOM) End-to-end business flows. Framework Standard GB921 Addendum E*. – 2017 – R16.5.1 – vol 16.5.2.
3. *TM Forum Metrics Definitions. Framework Best Practice GB988*. – 2016 – R16.5.0 – vol 4.0.1.
4. *Baskett F., Chandy K.M., Muntz R.R., Palacios F.G.* Open, closed, and mixed networks of queues with different classes of customers // *Journal of the ACM*. – 1975 – V. 22 – No. 2 – P. 248–260.
5. *Samouylov K., Gaidamaka Y., Zaripova E.* Analysis of Business Process Execution Time with Queueing Theory Models // A. Dudin, A. Gortsev, A. Nazarov, R. Yakupov (Eds.): *Information Technologies and Mathematical Modelling – Queueing Theory and Applications. Communications in Computer and Information Science*. – Springer International Publishing Cham, 2016. – V. 638 – P. 315–326.

MODELING OF SEVEN CLIENT-ORIENTED BUSINESS PROCESSES USING QUEUING NETWORK THEORY*

E.R. Zaripova, N.M. Smirnova, A.S. Alekseeva

Peoples' Friendship University of Russia (RUDN University), Moscow, Russia

Over the last years, customer satisfaction has become one of the most critical factors to success in service environment which is characterized by growing market transparency, customer requirements and reduced customer's loyalty. The key to success is, however, not purely a case of clever customer loyalty programs, but also the consistent orientation of the core processes towards customers' requirements [1].

The paper presents an approach to the analysis of performance indicators of telecommunication company, which represents a symbiosis of simulation modeling in GPSS World and queuing theory. This approach is applied to the business process model consists of seven reference client-oriented processes, taken from Enhanced Telecom Operations Map (eTOM) [2].

Description of client-oriented business processes

The eTOM helps enterprises to atomize and categorize all the business activities that include different spheres such as connections between organizations and customers, search the source of problems and their solutions (technical and non-technical), accounts control. This process is divided into seven processes:

- Requests – to – Answer comprises activities relevant to managing customer requests across all communication channels (customer interfaces);
- Order – to – Payment is linked with all activities which convert the customer request or an accepted offer into a ready for use product;
- Usage – to – Payment deals with all activities related to the handling of the product/service usage;
- Request – to – Change is connected with all activities which convert the customer's change request into a ready for use product;

* The reported study was partially supported by the RFBR, research projects No. 15-07-03051, 16-07-00766, 17-07-00845.

- Termination – to – Confirmation is linked with all activities related to the execution of customer's termination request;
- Problem – to – Solution deals with a technical complaint (problem) initiated by the customer, analyses it to identify the source of the issue, initiates resolution, monitors progress and closes the trouble ticket;
- Complaint – to – Solution deals with customer inquiries in which the customer is not pleased with a product or handling speed of an inquiry etc.

Constructing a generalized process model

The reference client-oriented business processes are aimed at various types of requests, starting from requests for the purchase of services/products and ending with the completion of customer service. Every type of requests will correspond to the classes of requests served in the model (Table 1):

Table 1

Classes of requests

1	Request for change of services/product according to client's wish
2	Request for change of services/product due to the poor quality of services
3	Request for usage and payment for services/product
4	Informational request for the payment for services/product
5	Request for technical failures of service/product
6	Request due to reduced quality of service/ product delivery
7	Request for termination of service
8	Request for preparing last payment and termination of service
9	Complaint about service/ product
10	Complaint about the quality of services/products
11	Informational request for services/ products
12	Informational request for sales
13	Request for receiving the finished services/products
14	Request for ordering services/products

In order to simplify the simulation model, all subprocesses in each of the processes are combined into semantic nodes. They reflect the change in the status of the received requests. After this step, it is noticed that some of the nodes turns out to be common, for example, nodes 1 and 2 are responsible for registration requests received; node 8 generates and provides invoices to the customers; node 9 checks the customer satisfaction.

In order to assess the quality of customer service, a generalized process model is developed in the form of an open heterogeneous exponential queuing network (Fig. 1) with nodes of two types – a single-line node with an infinite queue and an infinity-type node.

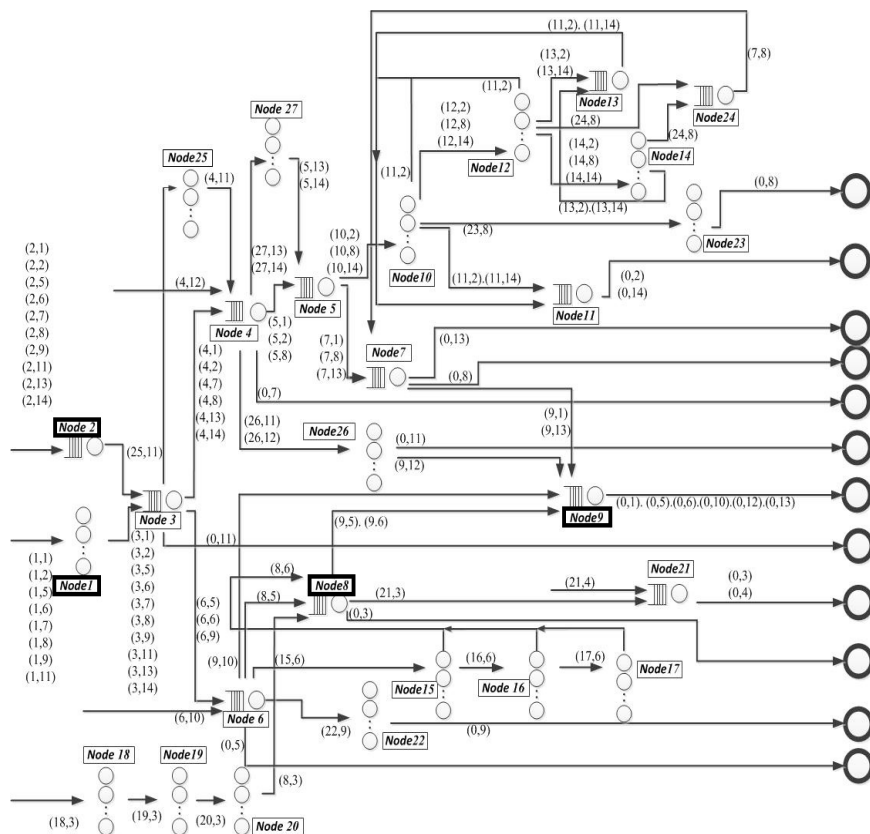


Fig. 1. Queuing network for generalized business process

Constructed queuing network represents the sequence of execution of different types of requests. Moreover it should be noticed that the simulation model built in the GPSS World environment takes into account the impossibility of repeated maintenance of requests in the node during the cyclical processing of the request.

Mathematical model for generalized process

The main performance indicators of constructed model are the average number of requests in the node, as well as the average process execution time in the node and in the whole model. In order to obtain formulas for the basic performance indicators of the model, it is necessary to know the intensities of the flows entering each node [3].

Flow intensity $\lambda_{(i,r)}$ is found from the equilibrium equation system:

$$\lambda_{(j,s)} = \sum_{(i,r) \in L} \lambda_{(i,r)} \theta_{(i,r),(j,s)} + \lambda_0 \theta_{0,(j,s)}, \quad (j,s) \in L \quad (1)$$

The intensity of the total flow λ_i entering the node is determined by the formula

$$\lambda_i = \sum_{r \in R} \lambda_{(i,r)} \quad (2)$$

with a stationary mode $\lambda_i \leq \mu_i$. (3)

Formulas for the average process execution time in the nodes are

$$U_i = \frac{N_i}{\lambda_0} = \frac{1}{\mu_i (1 - \rho_i)}, \quad i \in \{2, 3, 4, 5, 6, 7, 8, 9, 11, 13, 21, 24, 26\}, \quad (4)$$

for a single-line node with an infinite queue and

$$U_j = \frac{N_j}{\lambda_0} = \frac{1}{\mu_j}, \quad j \in \{1, 10, 12, 14, 15, 16, 17, 18, 19, 20, 22, 23, 25, 27\}, \quad (5)$$

for an infinity-type node.

Summarizing the average times over all nodes, we obtain the formula of the average process execution time in the network

$$U = \frac{1}{\lambda_0} \left(\begin{aligned} & \frac{\lambda_1}{\mu_1} + \frac{\lambda_2}{\mu_2 - \lambda_2} + \frac{\lambda_3}{\mu_3 - \lambda_3} + \frac{\lambda_4}{\mu_4 - \lambda_4} + \frac{\lambda_5}{\mu_5 - \lambda_5} + \frac{\lambda_6}{\mu_6 - \lambda_6} + \\ & + \frac{\lambda_7}{\mu_7 - \lambda_7} + \frac{\lambda_8}{\mu_8 - \lambda_8} + \frac{\lambda_9}{\mu_9 - \lambda_9} + \frac{\lambda_{10}}{\mu_{10}} + \frac{\lambda_{11}}{\mu_{11} - \lambda_{11}} + \frac{\lambda_{12}}{\mu_{12}} + \\ & + \frac{\lambda_{13}}{\mu_{13} - \lambda_{13}} + \frac{\lambda_{14}}{\mu_{14}} + \frac{\lambda_{15}}{\mu_{15}} + \frac{\lambda_{16}}{\mu_{16}} + \frac{\lambda_{17}}{\mu_{17}} + \frac{\lambda_{18}}{\mu_{18}} + \frac{\lambda_{19}}{\mu_{19}} + \frac{\lambda_{20}}{\mu_{20}} + \\ & + \frac{\lambda_{21}}{\mu_{21} - \lambda_{21}} + \frac{\lambda_{22}}{\mu_{22}} + \frac{\lambda_{23}}{\mu_{23}} + \frac{\lambda_{24}}{\mu_{24} - \lambda_{24}} + \frac{\lambda_{25}}{\mu_{25}} + \frac{\lambda_{26}}{\mu_{26} - \lambda_{26}} + \frac{\lambda_{27}}{\mu_{27}} \end{aligned} \right). \quad (6)$$

Numerical example

Let 100000 requests enter the constructed simulation model, and we know the average duration of the execution of each subprocesses

$$\begin{aligned}\mu_1^{-1} = \mu_2^{-1} = \mu_{10}^{-1} = \mu_{12}^{-1} = \mu_{16}^{-1} = \mu_{23}^{-1} &= 7 \text{ min}, \quad \mu_3^{-1} = \mu_7^{-1} = 8 \text{ min}, \\ \mu_4^{-1} = \mu_{15}^{-1} = \mu_{24}^{-1} = \mu_{27}^{-1} &= 10 \text{ min}, \quad \mu_5^{-1} = \mu_9^{-1} = \mu_{26}^{-1} = 9 \text{ min}, \\ \mu_6^{-1} = \mu_{21}^{-1} &= 4 \text{ min}, \quad \mu_8^{-1} = \mu_{13}^{-1} = 6 \text{ min}, \quad \mu_{14}^{-1} = 13 \text{ min}, \quad \mu_{17}^{-1} = 15 \text{ min}, \\ \mu_{11}^{-1} = \mu_{18}^{-1} = \mu_{19}^{-1} = \mu_{20}^{-1} = \mu_{22}^{-1} = \mu_{25}^{-1} &= 5 \text{ min}.\end{aligned}$$

As a result, the simulation time of the generalized business process is 19993 minutes, which is approximately equals to 42 days of the company's work, given the eight-hour day.

The graph of the average process execution time, depending on the intensity of the receipt of requests in the queuing network is shown in Fig. 2. According to the graph the system is received a Poisson flow, so the curve of the graph is close to the exponent. With zero intensity of the query flow, the average execution time of the process is very small and tends to zero. And with a single intensity of the flow of incoming requests, the average execution time of the process is approximately 30 minutes. Due to the fact that in each node, which receives the requests, the service is carried out according to

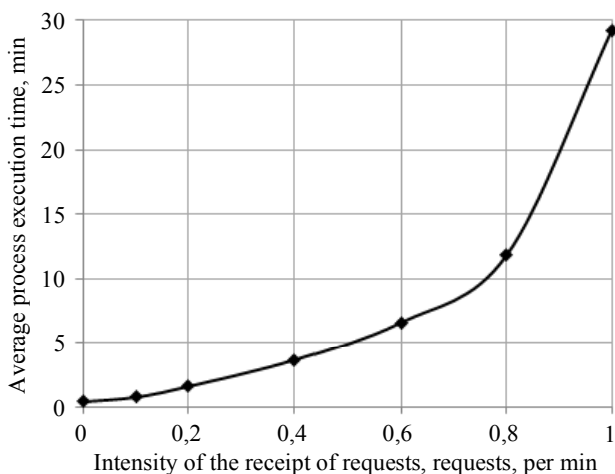


Fig. 2. The graph of the average process execution time

an exponential law with different intensities. In addition, mathematical model takes into account the transfer of requests to certain nodes, corresponding to the types of requests

Conclusions

The advantage of the developed method is scalability. Constructing a common business process model provides significant advantages to evaluate the performance of the company. The proposed technique, shown on the example of analysis of the generalized process, can be used to assess the efficiency of other business processes occurring in the company, for example, responsible for the network operation of the company. In addition, it is proved that the integration of business processes into the overall model significantly influences the evaluation of the company's performance, optimizing the execution time of requests service, which leads to an increase in the flow of incoming requests [4].

REFERENCES

2. *TeleMagement* Forum. Enhanced Telecom Operations Map (eTOM) The Business Process Framework. [Electronic Resource]. – URL: www.tmforum.org (last accessed date: 23.06.2017).
1. *Samouylov K.Y., Chukarin A.V., Yarkina N.V.* Business processes and information technology in the management of telecommunications companies. – M.: Alpina-Publishers, 2016. – 619 p.
4. *Sidnev A., Tuominen J., Krassi B.* Business process modeling and simulation // Helsinki University of Technology. – Industrial Information Technology Laboratory Publications, 2005. – P. 1–116.
3. *Samouylov K., Gaidamaka Yu., Zaripova E.* Analysis of business process execution time with queueing theory models // A. Dudin, A. Gortsev, A. Nazarov, R. Yakupov (Eds.): Information Technologies and Mathematical Modelling. – CCIS 638. – Springer International Publishing Switzerland, 2016 – P. 315–326, DOI: 10.1007/978-3-319-44615-8_28.

АНАЛИЗ ПОКАЗАТЕЛЕЙ КАЧЕСТВА ТРАФИКА РЕАЛЬНОГО ВРЕМЕНИ В БЕСПРОВОДНОЙ СЕТИ С ПОМОЩЬЮ РЕСУРСНЫХ СМО*

К.А. Агеев¹, Э.С. Сопин^{1,2}

¹Российский университет дружбы народов, г. Москва, Россия

²Институт проблем информатики ФИЦ ИУ РАН, г. Москва, Россия

Число пользовательских мобильных устройств неустанно растет, это касается как ассортимента, так и уже существующих моделей. Смартфоны, планшеты, ноутбуки используются различными группами пользователей в совершенно разных сферах жизнедеятельности. В крупных городах люди приобретают и используют сразу по несколько таких устройств: для работы, для дома, для путешествий, для развлечений. В 2017 г. трудно представить такие портативные устройство без доступа к Интернету, который может обеспечиваться различными способами подключения: Bluetooth, Wi-Fi, сотовая связь [1]. С развитием устройств возрастают потребности пользователей, что, в свою очередь, приводит к созданию новых услуг и росту требований качества уже существующих услуг.

В связи с резко растущим объемом трафика, генерируемого в беспроводных сетях (БС), возникает проблема нехватки частотного диапазона. Одним из возможных решений такой проблемы является применение различных планировщиков управления частотно-временными ресурсами, которые позволяют учесть расстояние от устройства до БС [2].

Проведен расчет характеристик трафика реального времени для популярных сервисов голосовых вызовов, видеовызовов, и телеконференций, которые используют передачу данных через сотовую сеть.

Планировщик Full Power

Принцип работы планировщика Full Power (FP) основан на передаче данных с максимальной мощностью. Суть заключается в различных вариациях параметров ресурсного блока – частоты, мощности и времени.

* Исследование выполнено при финансовой поддержке РФФИ в рамках научных проектов № 16-37-60103, 16-07-00766, 17-07-00142.

Базовая станция работает на предельной мощности и обслуживает устройства по мере их обращения к ней. Если же при обращении устройства недостаточно ресурса для обеспечения оказания услуги, то устройство отклоняется до следующей попытки подключения.

Рассмотрим одну пиктосоту беспроводной сети радиуса R . Предположим, что устройства являются стационарными и распределены по территории соты равномерно. С интенсивностью λ они переходят в активное состояние и передают данные с интенсивностью μ в восходящем канале (uplink channel). Обозначим ξ_d – расстояние от устройства до БС, p_{\max} – максимальную мощность передачи сигнала устройством, $\xi_p < p_{\max}$ – текущую мощность передачи. Отметим, что ξ_d и ξ_p являются случайными величинами (СВ). Предположим, что устройства передают данные с гарантированной скоростью r_0 . Достижимая устройством скорость передачи данных $r(\xi_d, \xi_p)$ зависит от ширины полосы частот восходящего канала ω , мощности передачи сигнала ξ_p , удаленности ξ_d от БС и определяется как

$$r(\xi_d, \xi_p) = \omega \ln \left(1 + \frac{G \xi_p}{\xi_d^\kappa N_0} \right).$$

Пусть в соте расположены три устройства (рис. 1) – первое и второе устройства неподвижны и расположены соответственно на расстояниях d_1 и d_2 от БС. Третье устройство может быть расположено либо близко к БС на расстоянии d_3^A , либо далеко – d_3^B . Все устройства работают на максимальной мощности, а временной ресурс делится пропорционально достигаемой скорости передачи $r(\xi_d, \xi_p)$ [2].

В случае А (рис. 2, А) третье устройство получает достаточный для достижения гарантированной скорости временной интервал обслуживания и начинает обслуживаться.

В случае В (рис. 2, В) устройство находится далеко и не может получить достаточный для достижения гарантированной скорости временной интервал, поэтому его запрос на обслуживание блокируется.

$$r_0 = 1 \text{ Мбит/с}, r_1 = 4 \text{ Мбит/с}, r_2 = 2.5 \text{ Мбит/с}, r_3^A = 4 \text{ Мбит/с}, r_3^B = 2 \text{ Мбит/с}$$

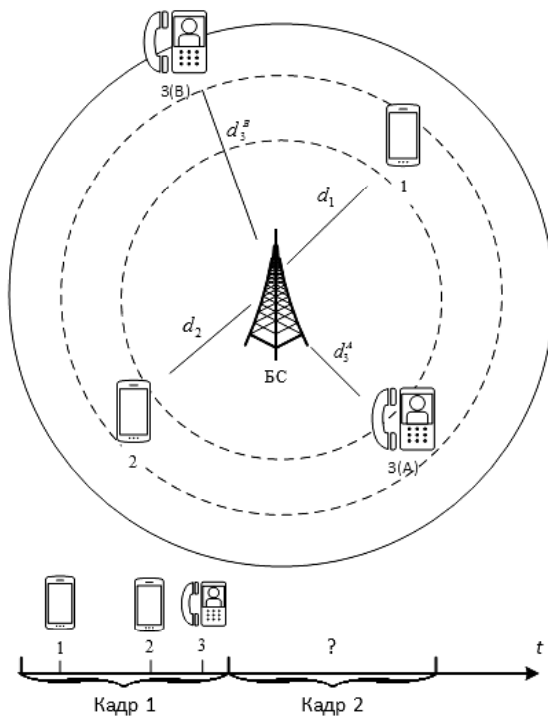


Рис. 1. Пример расположения и активации устройств в сети

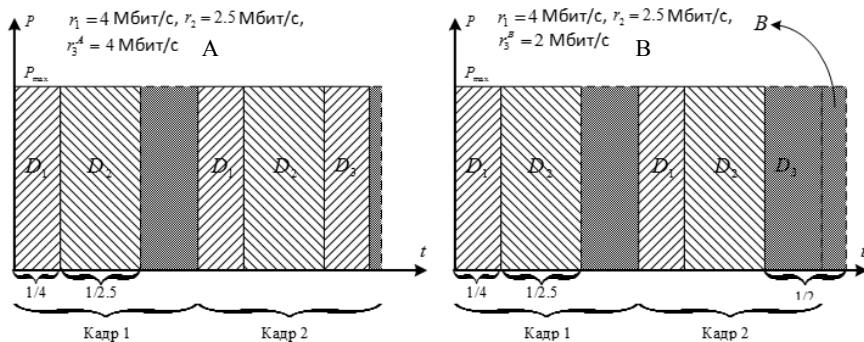


Рис. 2. Схема занятия временного ресурса (кадра) в случае А и В

С учетом описанных выше особенностей планировщика FP, построена функция распределения требований к ресурсам $F_{\xi_\delta}(\delta)$ [2]:

$$F_{\xi_\delta}(\delta) = P\left\{\frac{r_0}{r(\xi_d, p_{\max})} < \delta\right\} = \begin{cases} 0, & \delta \leq 0, \\ \frac{1}{R^2} \left(\frac{Gp_{\max}}{N_0}\right)^{2/\kappa} \left(e^{\frac{r_0}{\delta\omega}} - 1\right)^{-2/\kappa}, & \delta \in (0, \phi], \\ 1, & \delta > \phi, \end{cases} \quad (1)$$

где
$$\phi = \frac{r_0}{\omega \ln\left(\frac{Gp_{\max}}{N_0 R^\kappa} + 1\right)}.$$

Соответствующая плотность $f_{\xi_\delta}(\delta)$ распределения определяется по формуле [2]

$$f_{\xi_\delta}(\delta) = \begin{cases} \frac{2r_0}{R^2 \omega \kappa} \left(\frac{Gp_{\max}}{N_0}\right)^{2/\kappa} \frac{e^{\frac{r_0}{\delta\omega}}}{\delta^2} \left(e^{\frac{r_0}{\delta\omega}} - 1\right)^{-\frac{2+\kappa}{\kappa}}, & \delta \in (0, \phi], \\ 0, & \delta \notin (0, \phi] \end{cases} \quad (2)$$

Математическая модель

Рассмотрим (рис. 3) ресурсную систему, на вход которой поступает поток заявок с параметром λ , время обслуживания каждой заявки распределено по экспоненциальному закону с параметром μ . При поступлении заявка занимает прибор и часть ресурса. Если при поступлении для заявки нет свободного прибора либо не хватает требуемого ресурса, то такая заявка блокируется.

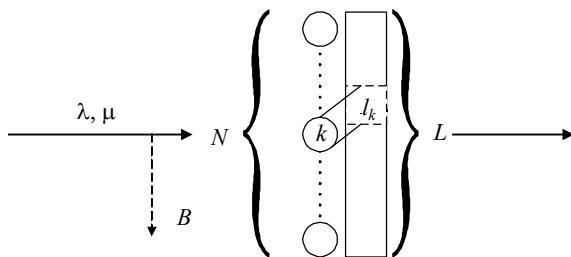


Рис. 3. Схема ресурсной модели

В модели не отслеживается объем занимаемого или освобождаемого ресурса каждой заявкой, рассматривается только суммарный занятый и свободный объем ресурса [3].

Рекуррентный алгоритм

Для расчета вероятностно временных характеристик применен алгоритм дискретизации функции распределения требований к ресурсам [4]. Рекуррентный алгоритм применяется для расчета вероятностно временных характеристик для системы с дискретным ресурсом [5].

$$G(n, l) = G(n-1, l) + \frac{\rho}{n} \sum_{j=0}^l p_j (G(n-1, l-j) - G(n-2, l-j)), \quad (3)$$

где p_j – вероятность того, что j -я заявка занимает l ресурса.

$$G(0, l) = 1, \quad 0 \leq l \leq L; \quad (4)$$

$$G(1, l) = 1 + \rho \sum_{j=0}^l p_j, \quad 0 \leq l \leq L. \quad (5)$$

Вероятность блокировки B может быть посчитана с использованием следующей формулы:

$$B = 1 - G^{-1}(N, L) \sum_{j=0}^L p_j G(N-1, L-j). \quad (6)$$

Средний объем занятого ресурса b определяется по формуле

$$b = L - G^{-1}(N, L) \sum_{m=1}^M e_m \sum_{l=1}^{L_m} G(N, L - l e_m). \quad (7)$$

Расчет показателей для услуг связи

Для проведения расчетов параметры соты были установлены в соответствии с данными в таблице.

Для расчета показателей трафика услуг были взяты требования к скорости, необходимые для осуществления звонков (40 кбит/с) и видео звонков (128 кбит/с) в приложении Skype. Максимальное число устройств в зоне действия соты $N = 1000$. Число устройств, одновременно передающих данные базовой станции, ρ варьируется от 5 до 50.

Исходные значения параметров соты

Параметр	Значение
R	100 м
ω	10 МГц
p_{\max}	0,00398 Вт
N_0	10^{-9} Вт
G	197,43
κ	5
L	1

На рисунках ниже представлено поведение вероятностно-временных характеристик.

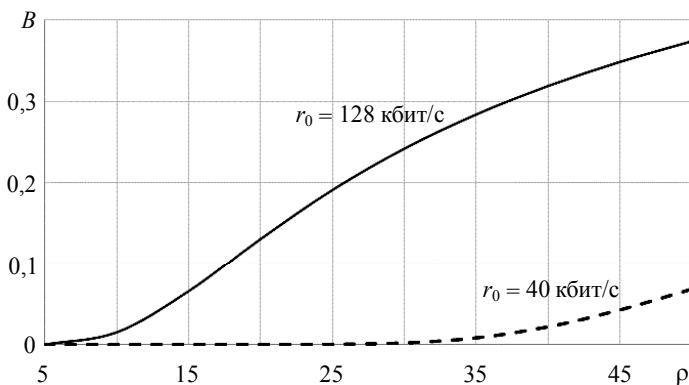


Рис. 4. Вероятность блокировки
в зависимости от предложенной нагрузки

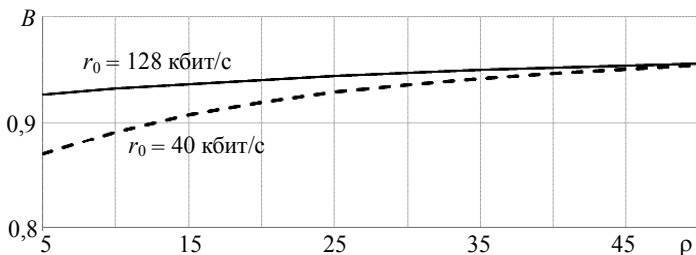


Рис. 5. Объем занятого ресурса
в зависимости от предложенной нагрузки

Логично предположить, что вероятность блокировки для видео звонков больше, чем для голосовых. С увеличением числа одновременно активных пользователей вероятность отказа новому пользователю увеличивается.

Очевидно, что требования к ресурсу для голосовых вызовов ниже, чем для видеозвонков. С увеличением ρ это становится менее заметно, что связано с особенностями работы планировщика FR.

Заключение

В работе проведен расчет показателей характеристик трафика реального времени для услуг, предоставляемых через сотовую сеть. Результаты получены с помощью алгоритма дискретизации функции распределения требований к ресурсам и рекуррентного алгоритма расчета характеристик. Стоит отметить, что область применения данных алгоритмов значительно шире и может применяться для расчетов характеристик на основе других функций распределений.

ЛИТЕРАТУРА

1. 3GPP TS 36.300: Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN); Overall description; Stage 2 (Release 14).
2. Сопин Э., Гудкова И., Маркова Е., Агеев К. Аппроксимация функции требований к ресурсам для анализа характеристик трафика межмашинного взаимодействия // Современные информационные технологии и ИТ-образование. – 2016. – Т. 12. – С. 64–70.
3. Вихрова О., Самуйлов К., Сопин Э., Шоргин С. К анализу показателей качества обслуживания в современных беспроводных сетях // Информатика и ее применение. – 2015. – Т. 9 – С. 48–55.
4. Агеев К.А., Сопин Э.С. Дискретизация функции распределения требований к ресурсу для анализа характеристик M2M трафика // ИТТММ. – 2017. – С. 83–85
5. Сопин Э.С., Самуйлов К.Е. Рекуррентный алгоритм вычисления вероятностных характеристик для СМО с ограниченными ресурсами и случайными требованиями // ИТТММ. – 2017. – С. 52–54.

О ЗАДАЧЕ ОПТИМИЗАЦИИ ЭНЕРГОПОТРЕБЛЕНИЯ ОБЛАЧНОЙ ИНФРАСТРУКТУРЫ*

А.В. Дараселия, Э.С. Сопин

Российский университет дружбы народов, г. Москва, Россия

В целях повышения энергоэффективности облачной системы серверы могут быть переведены в режим ожидания при низкой нагрузке. Перевод в режим ожидания, с одной стороны, позволяет снизить энергопотребление, а с другой – приводит к дополнительным затратам на включение/выключение сервера. Поэтому важно понимать, при каких условиях будет выгодно перевести серверы в режим ожидания, а при каких – оставить их в рабочем состоянии.

В работе [1] мы рассмотрели систему облачных вычислений с учетом разогрева и выключения серверов, причем считалось, что сервер выключается сразу же, как только остается пустым. В данной работе представлена модель, в которой сервер выключается по прошествии некоторого случайного времени после того, как он остался пустым.

Математическая модель

Рассматривается система массового обслуживания, состоящая из N групп приборов (серверов), каждая из которых состоит из V_i приборов. Для краткости, в данной работе приведем результаты анализа для одного сервера.

На систему поступает пуассоновский поток с интенсивностью λ , времена обслуживания заявок, а также периоды разогрева и выключения распределены экспоненциально с параметрами μ , α и β соответственно [1], γ – экспоненциальное время ожидания новой заявки, в течение которого система не уходит в спящий режим, когда заявок не остается. Пространство состояний описывается вектором (s, k) , где k – количество заявок на сервере, s – состояние сервера, где $s = 0$ в состоянии ожидания, 1 – включение сервера, 2 – работа и 3 – выключение соответственно. Система переходит в режим выключения, когда

* Исследование выполнено при частичной финансовой поддержке РФФИ в рамках научных проектов № 15-07-03051, 15-07-03608.

остается пуста. Система переходит в режим выключения, когда остается пуста дольше γ .

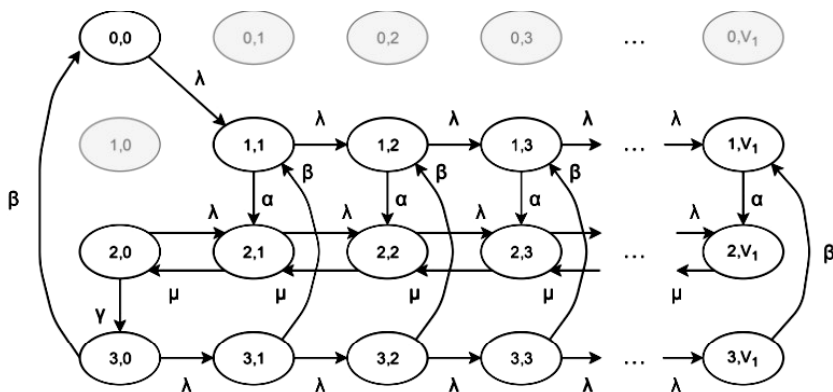


Рис. 1. Диаграмма интенсивностей переходов

На основе диаграммы интенсивностей переходов (рис. 1) выпишем систему уравнений равновесия, которая позволяет получить стационарные вероятности системы:

$$p_{3,0} = \frac{\lambda}{\beta} p_{0,0}; \quad (1)$$

$$p_{3,k} = \frac{\lambda}{(\lambda + \beta)} p_{3,k-1}, \quad 1 \leq k \leq V_1 - 1; \quad (2)$$

$$p_{3,V_1} = \frac{\lambda}{\beta} p_{3,V_1-1}; \quad (3)$$

$$p_{1,1} = \frac{\lambda}{(\lambda + \alpha)} p_{0,0} + \frac{\beta}{(\lambda + \alpha)} p_{3,1}; \quad (4)$$

$$p_{1,k} = \frac{\lambda}{(\lambda + \alpha)} p_{1,k-1} + \frac{\beta}{(\lambda + \alpha)} p_{3,k}, \quad 2 \leq k \leq V_1 - 1; \quad (5)$$

$$p_{1,V_1} = \frac{\lambda}{\alpha} p_{1,V_1-1} + \frac{\beta}{\alpha} p_{3,V_1}; \quad (6)$$

$$p_{2,0} = \frac{\lambda + \beta}{\gamma} p_{3,0}; \quad (7)$$

$$p_{2,1} = \frac{\lambda + \gamma}{\mu} p_{2,0}; \quad (8)$$

$$p_{2,2} = \frac{(\lambda + \mu)}{\mu} p_{2,1} - \frac{\lambda}{\mu} p_{2,0} - \frac{\alpha}{\mu} p_{1,1}; \quad (9)$$

$$p_{2,k+1} = \frac{(\lambda + \mu)}{\mu} p_{2,k} - \frac{\lambda}{\mu} p_{2,k-1} - \frac{\alpha}{\mu} p_{1,k}, \quad 3 \leq k \leq V_1 - 1; \quad (10)$$

$$p_{2,V_1} = \frac{\alpha}{\mu} p_{1,V_1} + \frac{\lambda}{\mu} p_{2,V_1-1}; \quad (11)$$

$$\sum_{i=0}^3 \sum_{j=0}^{V_1} p_{i,j} = 1. \quad (12)$$

Система уравнений равновесия была решена аналитически, ниже представлено ее решение:

$$p_{3,0} = \frac{\lambda}{\beta} p_{0,0}; \quad (13)$$

$$p_{3,k} = \frac{\lambda^{k-1}}{(\lambda + \beta)^k} p_{0,0}, \quad 1 \leq k \leq V_1 - 1; \quad (14)$$

$$p_{3,V_1} = \left(\frac{\lambda}{\lambda + \beta} \right)^{V_1-1} p_{0,0}; \quad (15)$$

$$p_{1,1} = \frac{\lambda(2\lambda + \beta)}{(\lambda + \alpha)(\lambda + \beta)} p_{0,0}; \quad (16)$$

$$p_{1,k} = \frac{\lambda^k (2\lambda + \beta)}{(\lambda + \alpha)^k (\lambda + \beta)} p_{0,0} + \frac{\beta^2 \lambda^{k-1} (\alpha - \beta) [(\lambda + \alpha)^k - (\lambda + \beta)^k]}{(\lambda + \beta)^k (\lambda + \alpha)^k} p_{0,0}, \quad 2 \leq k \leq V_1 - 1; \quad (17)$$

$$p_{1,V_1} = \frac{\lambda}{\alpha} \left[\frac{\lambda^k (2\lambda + \beta)}{(\lambda + \alpha)^k (\lambda + \beta)} + \frac{\beta^2 \lambda^{k-1} (\alpha - \beta) [(\lambda + \alpha)^k - (\lambda + \beta)^k]}{(\lambda + \beta)^k (\lambda + \alpha)^k} \right] p_{0,0} + \frac{\beta}{\alpha} \left(\frac{\lambda}{\lambda + \beta} \right)^{V_1 - 1}; \quad (18)$$

$$p_{2,0} = \frac{(\lambda + \beta)\lambda}{\gamma\beta} p_{0,0}; \quad (19)$$

$$p_{2,1} = \frac{(\lambda + \gamma)(\lambda + \beta)\lambda}{\mu\gamma\beta} p_{0,0}; \quad (20)$$

$$p_{2,2} = \frac{\lambda(\lambda + \beta)(\lambda^2 + \lambda\gamma + \mu\gamma)}{\mu^2\gamma\beta} p_{0,0} - \frac{\alpha\lambda(2\lambda + \beta)}{\mu(\lambda + \alpha)(\lambda + \beta)} p_{0,0}; \quad (21)$$

$$\begin{aligned} p_{2,k} = & \frac{1 - \rho^{k-1}}{1 - \rho} \cdot \left(\frac{\lambda(\lambda + \beta)(\lambda^2 + \lambda\gamma + \mu\gamma)}{\mu^2\gamma\beta} - \frac{\alpha\lambda(2\lambda + \beta)}{\mu(\lambda + \alpha)(\lambda + \beta)} \right) p_{0,0} - \\ & - \rho \frac{1 - \rho^{k-2}}{1 - \rho} \left(\frac{(\lambda + \gamma)(\lambda + \beta)\lambda}{\mu\gamma\beta} \right) p_{0,0} - \\ & - p_{0,0} \left(\frac{\rho}{1 - \rho} \right) \frac{\alpha}{\mu} \left[\frac{\lambda(2\lambda + \beta)}{(\lambda + \beta)} \cdot \left[\frac{\lambda}{\alpha(\lambda + \alpha)} \left(1 - \left(\frac{\lambda}{\lambda + \alpha} \right)^{k-2} \right) - \frac{\lambda^{k-1}}{(\lambda + \beta)\mu} \frac{1 - \left(\frac{\lambda + \alpha}{\mu} \right)^{k-1}}{1 - \frac{\lambda + \alpha}{\mu}} \right] + \right. \\ & + \beta^2 (\alpha - \beta) \left[\left(\frac{\lambda}{\lambda + \beta} \right)^{k-1} \cdot \left(\frac{\left(\frac{\lambda + \beta}{\lambda} \right)^{k-2} - 1}{\beta} - \frac{1 - \left(\frac{\lambda + \beta}{\mu} \right)^{k-2}}{\mu \left(1 - \frac{\lambda + \beta}{\mu} \right)} \right) - \right. \\ & \left. \left. - \left(\frac{\lambda}{\lambda + \alpha} \right)^{k-1} \cdot \left(\frac{\left(\frac{\lambda + \alpha}{\lambda} \right)^{k-2} - 1}{\alpha} - \frac{1 - \left(\frac{\lambda + \alpha}{\mu} \right)^{k-2}}{\mu \left(1 - \frac{\lambda + \alpha}{\mu} \right)} \right) \right] \right], \quad 3 \leq k \leq V_1. \quad (22) \end{aligned}$$

Показатели энергопотребления

Получив стационарное распределение системы, вычислим показатели энергопотребления. Будем считать, что в режиме включения/выключения энергопотребление постоянно и равно среднему значению. В режиме обслуживания заявок потребляемая мощность зависит от загрузки сервера. По аналогии с приведенной в работах [2, 3] формулой, выведем формулу средней потребляемой сервером мощности:

$$P = P_0 \sum_{i=0}^{V_1} p_{0,i} + P_1 \sum_{i=0}^{V_1} p_{1,i} + P_3 \sum_{i=0}^{V_1} p_{3,i} + \sum_{i=0}^{V_1} P_{2,i} p_{2,i},$$

где
$$P_{2,k} = P_{2,\min} + k \frac{P_{2,\max} + P_{2,\min}}{V_1}. \quad (23)$$

Значения P_i были взяты из статьи [4], согласно которой $P_0 = 10$ Вт, $P_1 = 170$ Вт, $P_3 = 120$ Вт, $P_{2,\min} = 105$ Вт и $P_{2,\max} = 268$ Вт.

Результаты численного анализа для значений $V_1 = 7$, $\mu = 20$, $\alpha = 1$, $\beta = 2$ представлены на рис. 2 и 3.

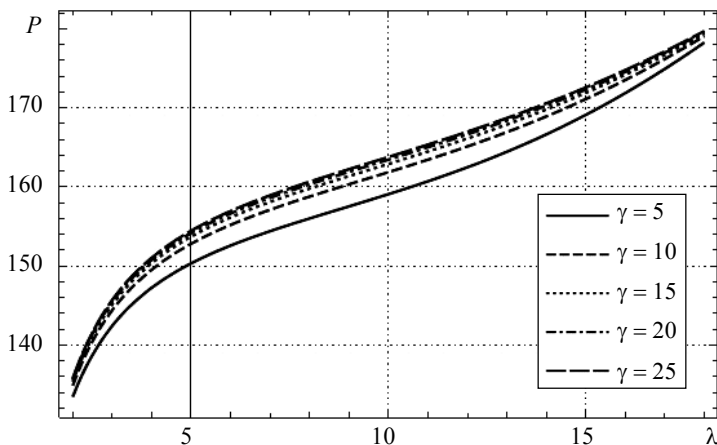


Рис. 2. Графики зависимости мощности P от интенсивности нагрузки λ

График потребляемой сервером мощности (рис. 2) для нашей модели наиболее резко возрастает при небольших значениях интенсивности нагрузки λ , также заметим, что с увеличением экспоненциального време-

ни ожидания новой заявки γ , в течение которого система не уходит в спящий режим, значение потребляемой мощности увеличивается соответственно.

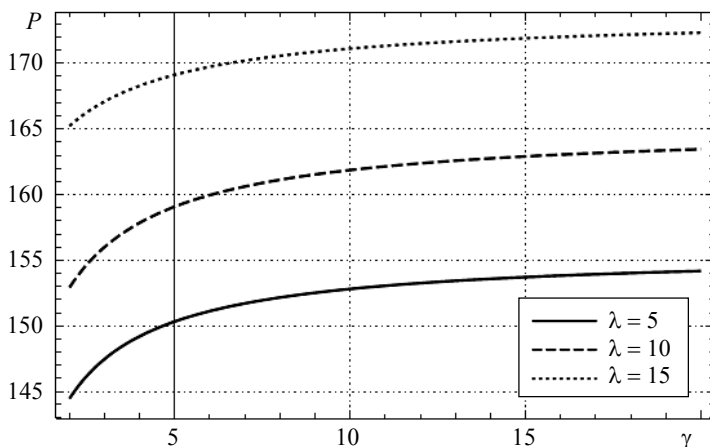


Рис. 3. Графики зависимости мощности P от экспоненциального времени ожидания новой заявки γ , в течение которого система не уходит в спящий режим

На рис. 3 заметим, что наибольший перепад для потребляемой мощности происходит при небольших значениях экспоненциального времени ожидания новой заявки γ и что с увеличением значений интенсивности нагрузки λ также возрастает и энергопотребление.

Заключение

Рассмотрена система облачных вычислений, в которой сервер выключается по прошествии некоторого случайного времени после того, как он остался пустым. В дальнейшем планируется решить задачу оптимизации этого времени для снижения энергопотребления.

ЛИТЕРАТУРА

1. Дараселия А.В., Сопин Э.С. Анализ энергопотребления системы облачных вычислений с учетом разогрева и выключения серверов // Информационно-телекоммуникационные технологии и математическое моделирование высокотехнологичных систем. — 2017. — С. 119–121.

2. Дараселия А.В., Сопин Э.С. К вопросу энергоэффективности облачной системы // Информационно-телекоммуникационные технологии и математическое моделирование высокотехнологичных систем. – 2016. – С. 13–15.
3. Сопин Э.С., Дараселия А.В., Яркина Н.В. Об эффективности миграции виртуальных машин в облачных системах // 19-я Междунар. конф. «Распределенные компьютерные и коммуникационные сети: управление, вычисление, связь». – 2016. – С. 408–411.
4. Conejero J., Rana O., Burnap P., et al. Analysing Hadoop Power Consumption and Impact on Application QoS // Future Generation Computer Systems. – 2016. – V. 55. – Issue C. – P. 213–223.

МОДЕЛИ РАСПРЕДЕЛЕНИЯ СОЕДИНЕНИЙ В БЕСПРОВОДНЫХ ТЕЛЕКОММУНИКАЦИОННЫХ СЕТЯХ*

И.В. Коннов, О.А. Кашина

Казанский (Приволжский) федеральный университет, г. Казань, Россия

Преимущества телекоммуникационных сетей, основанных на использовании беспроводных технологий передачи данных (например, Bluetooth, Wi-Fi или NFC), перед традиционными сетями (использующими провода или оптоволокно) неоспоримо. Это не только мобильность и компактность беспроводных устройств, независимость от кабельной инфраструктуры, простота подключения, удобство применения, но и возможность обеспечения более высокоскоростного доступа к сети Интернет, и финансовая доступность для широких слоёв населения. Названные преимущества обуславливают быстрое развитие беспроводных технологий и расширение сферы применения беспроводных коммуникационных сетей – одной из основных тенденций их современного развития является так называемый «Интернет вещей» (англ.: Internet of Things, IoT) – сеть самых разных объектов, взаимодействующих друг с другом и с внешней средой посредством применения беспроводных технологий. Если изначально «Интернет вещей» рассматривался, в основном, как совокупность «умных» технологий (в первую очередь, «умный дом») [1], то сейчас он включает такие решения, как, например, системы быстрого реагирования, системы интеллектуального управления цепочкой поставок товаров, производство «по запросу», адаптивная торговля, прогнозная клиническая аналитика и т.д. Если в настоящее время к сети Интернет во всём мире подключено 10 млрд устройств, то к 2020 г. это количество, как ожидается, составит 50–60 млрд [2]. Более того, можно ожидать постепенного стирания граней между «Интернетом вещей» и так называемым «Интернетом всего» (или «Всеобщим Интернетом») – интеллектуальной сетью, компонентами которой, наряду с людьми и «вещами», становятся также данные и процессы. Понятно, что для этого необходим универсальный сетевой

* Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований, проект 16-01-00109а, для первого автора также в рамках выполнения государственного задания Минобрнауки России, номер задания 1.460.2016/1.4.

стандарт высокой пропускной способности, и предпосылки для его создания уже есть. Так, новый интернет-протокол IPv6 позволяет использовать практически неограниченное число IP-адресов. Микрочип, разработанный компанией ARM Holdings (Великобритания), за счёт сверхнизкого потребления энергии и сверхмалого размера (1 кв. мм), позволяет подключить к Интернет практически любой объект [2].

В связи с возрастающей востребованностью беспроводных сетей возрастает и актуальность задач, связанных с оптимизацией их производительности. Особенностью этих задач является их большая размерность (ввиду огромного числа абонентов) и частое изменение исходных данных (ввиду мобильности абонентов). Вместе с тем требования к точности решения таких задач обычно не слишком высоки – приоритет отдаётся скорости получения решения. Это позволяет применять к решению задач оптимизации беспроводных сетей приближённые методы нелинейного программирования.

Рассмотрим вначале постановку задачи в фиксированный момент времени (для краткости опускаем его обозначение).

Имеется некоторая территория, входящая в зону покрытия каждого из m провайдеров мобильных услуг. На данной территории находится большое число мобильных абонентов, каждый из которых может выступать и как приёмник, и как передатчик сигнала. Обозначим через n число парных соединений абонентов сети. Естественно считать, что для обеспечения каждого соединения провайдерами расходуется некий сетевой ресурс (например, пропускная способность беспроводного канала связи), «запас» которого (у провайдера i) ограничен заданной величиной γ_i ; через x_{ij} обозначим искомое количество ресурса сети, выделяемого i -м провайдером для обеспечения j -го соединения ($i = 1, \dots, m$; $j = 1, \dots, n$). Далее для краткости будем называть величину x_{ij} потоком (i, j) . Предположим, что существуют нижняя и верхняя границы каждого потока (i, j) – без ограничения общности рассуждений нижнюю границу будем считать равной 0; верхнюю границу потока (i, j) обозначим через α_{ij} . Предположим также, что и суммарный поток для каждого соединения j ограничен сверху некоторой величиной β_j . В [3] предполагается, что для каждой пары j определена величина b_j «цена», которую данная пара готова платить за единицу потока. Аналогично, будем считать заданной величину расходов провайдера i , связанных с обеспечением единицы потока для соединения j – обозначим её через a_{ij} . Тогда чистые суммарные расходы, связан-

ные с обеспечением всех потоков сети, равны

$$\sum_{i=1}^m \sum_{j=1}^n a_{i,j} x_{i,j} - \sum_{j=1}^n b_j \left(\sum_{i=1}^m x_{i,j} \right) \equiv \sum_{i=1}^m \sum_{j=1}^n c_{i,j} x_{i,j},$$

где

$$c_{i,j} \equiv a_{i,j} - b_j.$$

Получаем задачу линейного программирования:

$$\min \rightarrow \sum_{i=1}^m \sum_{j=1}^n c_{i,j} x_{i,j}; \quad (1)$$

$$\sum_{i=1}^m x_{i,j} \geq \beta_j, j = 1, \dots, n; \quad (2)$$

$$\sum_{j=1}^n x_{i,j} \leq \gamma_i, i = 1, \dots, m; \quad (3)$$

$$0 \leq x_{i,j} \leq \alpha_{i,j}, i = 1, \dots, m; j = 1, \dots, n. \quad (4)$$

Таким образом, задача об оптимальном распределении потоков в сети с мобильными абонентами формулируется как открытая транспортная задача с двусторонними ограничениями на переменные. Несмотря на хорошо разработанный инструментарий решения задач линейного программирования и, в частности, транспортной задачи (например, [4]), применять на практике точные методы решения задачи (1) – (4) вряд ли целесообразно. Поскольку в силу быстрого изменения конфигурации беспроводных сетей коэффициенты оптимизационных задач обнаруживают нестационарное поведение, значения их можно рассматривать лишь как приближённые. Поэтому имеет смысл решать задачи вида (1) – (4) приближёнными методами, в частности – методами типа штрафных и/или барьерных функций (например, [5, с. 172 – 182]). Так, в [3] для приближённого решения задачи (1) – (4) в каждый фиксированный момент времени используется метод внешних штрафных функций (МВШ). В [6] для решения той же задачи предложен так называемый метод частичного штрафа (МЧШ). Штрафная функция в этом методе строится не для всех, а лишь для части ограничений:

$$\Phi(X) \equiv \sum_{i=1}^m \left(\sum_{j=1}^n x_{i,j} - \gamma_i \right)_+^2, \quad (5)$$

где знак «+» понимается в смысле функции срезки; $t_+ = \max\{0, t\}$ – для

произвольного вещественного t ; через X обозначена $m \times n$ – матрица с компонентами $x_{i,j}$. Вспомогательная задача, решаемая на каждом шаге k МЧШ, имеет вид

$$\min \rightarrow \Psi(X, \tau_k) \quad (6)$$

при ограничениях (2) и (4). Здесь $\Psi(X, \tau_k) \equiv C, X + \tau_k \Phi(X)$, где C – матрица с компонентами $c_{i,j}$ из формулы (1), τ_k – значение параметра МЧШ. Ограничения (2) и (4) задают множество достаточно простой структуры – этот факт позволил нам предложить конечный алгоритм решения задачи (6), (2), (4) (см. [6], Algorithm A) при каждом фиксированном $k = 0, 1, \dots$.

Экспериментально доказано преимущество МЧШ перед классическим методом «полного» штрафа при $m \leq 20$ и $n \gg m$.

Более общий случай открытой транспортной задачи как модели оптимального распределения соединений в беспроводной сети исследуется в [7], где целевая функция, вообще говоря, является нелинейной (она необязательно является квадратичной или выпуклой), а её параметры подвержены случайным (затухающим со временем) возмущениям. Нелинейность целевой функции связана с тем фактом, что на практике цена на ресурсы сети, как правило, зависит от объёма продаж. Формальная постановка задачи имеет вид

$$\min \rightarrow G(X) \quad (6)$$

при ограничениях (2) – (4), где функция G по предположению является непрерывно дифференцируемой на пространстве вещественных $m \times n$ -матриц X и при всех X и всех $i \in \{1, \dots, m\}$, $j \in \{1, \dots, n\}$ выполнено

условие $\frac{\partial G(X)}{\partial x_{i,j}} < \infty$. Показано, что последовательность приближённых

решений возмущённых задач сходится к решению задачи с нулевым возмущением. Экспериментально подтверждена гипотеза о том, что быстрое действие МЧФ по сравнению с МВШ тем заметнее, чем больше число соединений n , устанавливаемых в сети в единицу времени.

Можно показать, что полученные результаты остаются справедливыми и для более общего случая задачи распределения ресурсов в сети с мобильными абонентами, а именно – когда случайным затухающим возмущениям подвержены не только параметры целевой функции, но и коэффициенты ограничений задачи.

Для простоты изложения в рассмотренных постановках задачи предполагалось, что каждое соединение $j \in \{1, \dots, n\}$ может быть обеспечено любым провайдером $i \in \{1, \dots, m\}$. Вообще говоря, это условие может не выполняться, и для каждого j множество $\{1, \dots, m\}$ следует заменить на некоторое подмножество P_j номеров тех провайдеров, которые могут обеспечить соединение j , положив $x_{i,j} \equiv 0$ для всех $i \notin P_j, j \in \{1, \dots, n\}$. Заметим, что задача оптимального распределения беспроводных соединений может решаться не только в целях расчёта величин всех потоков, но и в целях оценивания производительности сети (т.е. исследователя интересует только значение целевой функции задачи). В этом случае имеет смысл перейти от детерминированной постановки задачи к вероятностной, заменив целевую функцию средним значением случайной величины (математическим ожиданием размера затрат, зависящим от случайного «состава» множеств P_j). Это позволяет снизить размерность задачи (за счёт потери точности решения). Как следствие, возникает вопрос о выборе адекватной модели случайного «поведения» $P_j, j \in \{1, \dots, n\}$ обеспечивающей практически приемлемый «баланс» между точностью решения и временем, затраченным на его получение. Эти вопросы мы планируем изложить в следующих публикациях.

ЛИТЕРАТУРА

1. Kevin Ashton, "That 'Internet of Things' Thing" // RFID Journal. – 22 June 2009.
2. Интернет вещей и Всеобщий Интернет // Интернет в России и в мире [Электронный ресурс]. – URL: http://www.bizhit.ru/index/trend_www_traffic/0-171 (дата обращения: 23.06.2017).
3. Коннов И.В., Кашина О.А. Открытая транспортная модель и её приложения: задача распределения потоков в сети и задача производства-сбыта с учётом транспортных расходов // Социально-экономические и технические системы: исследование, проектирование, оптимизация. – 2016. – № 4. – С. 4–18.
4. Юдин Д.Б., Гольштейн Е.Г. Задачи и методы линейного программирования. Задачи транспортного типа. – М.: Либроком, 2010. – 735 с.
5. Коннов И.В. Нелинейная оптимизация и вариационные неравенства. – Казань: Казан. ун-т, 2013. – 508 с.
6. Konnov I.V., Kashina O.A., Laitinen E. Partial penalty method for flow optimization in wireless networks // WSEAS Transactions on Communications. – 2016. – V. 15. – Art. #40. – P. 363–368.
7. Konnov I.V., Kashina O.A., Laitinen E. Optimization of wireless networks performance: an approach based on a partial penalty method // Int. J. Circuits, Systems and Signal Processing. – 2017. – V. 11. – P. 37–43.

МЕТОДЫ РАСПРЕДЕЛЕНИЯ РЕСУРСОВ В СЕТЯХ ТЕЛЕКОММУНИКАЦИИ*

И.В. Коннов, А.Ю. Кашуба

Казанский федеральный университет, г. Казань, Россия

Несмотря на существенное повышение мощности и производительности современных устройств обработки и передачи данных, значительная нестабильность спроса на ресурсы сетей телекоммуникации, особенно беспроводных, приводит к их неравномерному и поэтому неэффективному использованию. В результате одной из основных задач в управлении сетями телекоммуникации является построение гибких механизмов распределения ресурсов, таких, как спектр радиосвязи (см., например, [1–3]). Большинство работ использует теоретико-игровые модели (см., например, [4, 5]). Также существуют методы, основанные на оптимизации [6, 7, 5, 3].

Рассматривается общая задача распределения ресурсов провайдером (менеджером сети) сети, поделенной на зоны (кластеры), с учетом того, что менеджер сети может привлекать дополнительный внешний ресурс с целью удовлетворения запросов пользователей. Предлагаются методы решения этой задачи.

Описание задачи

Рассмотрим сеть с узлами (пользователями), разбитую на n зон. Для k -й зоны ($k = 1, \dots, n$) I_k обозначает множество индексов узлов, находящихся в этой зоне, b_k – максимальное фиксированное количество ресурса. Менеджер сети с целью удовлетворения запросов на ресурс пользователей в k -й зоне распределяет свой (внутренний) ресурс $x_k \in [0, b_k]$ и привлекаемый внешний ресурс $z_j \in [0, c_j]$, $j \in J_k$ (J_k обозначает множе-

* Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 16-01-00109а; для первого автора также в рамках выполнения государственного задания Минобрнауки России, номер задания 1.460.2016/1.4; для второго автора – за счет средств субсидии, выделенной в рамках государственной поддержки Казанского (Приволжского) федерального университета в целях повышения его конкурентоспособности среди ведущих мировых научно-образовательных центров.

ство индексов внешних провайдеров, находящихся в k -й зоне). Распределяемые ресурсы требуют затрат на обслуживание $f_k(x_k)$ и побочных (сторонних) платежей $h_j(z_j)$, $j \in J_k$ для каждого $k = 1, \dots, n$. Пусть также существует верхняя граница B на общее количество внутреннего ресурса сети. Если i -й пользователь получает y_i ресурса с учетом верхней границы a_i , тогда он выплачивает $\phi_i(y_i)$. Задача менеджера сети заключается в нахождении оптимального распределения ресурсов между зонами и имеет вид

$$\max \rightarrow \sum_{k=1}^n \left[\sum_{i \in I_k} \phi_i(y_i) - f_k(x_k) - \sum_{j \in J_k} h_j(z_j) \right]; \quad (1)$$

$$\sum_{k=1}^n x_k \leq B; \quad (2)$$

$$\sum_{i \in I_k} y_i = x_k + \sum_{j \in J_k} z_j, k = 1, \dots, n; \quad (3)$$

$$0 \leq y_i \leq a_i, i \in I_k, 0 \leq x_k \leq b_k, 0 \leq z_j \leq c_j, j \in J_k, k = 1, \dots, n. \quad (4)$$

Пусть функции $\phi_i(y_i)$, $f_k(x_k)$ и $h_k(z_k)$ являются кубическими, то есть для всех $k = 1, \dots, n$

$$\phi_i(y_i) = \alpha_{3,i} y_i^3 + \alpha_{2,i} y_i^2 + \alpha_{1,i} y_i + \alpha_{0,i}, \alpha_{3,i}, \alpha_{2,i} < 0, \alpha_{1,i} > 0, \alpha_{0,i} \geq 0, i \in I_k; \quad (5)$$

$$f_k(x_k) = \beta_{3,k} x_k^3 + \beta_{2,k} x_k^2 + \beta_{1,k} x_k, \beta_{3,k}, \beta_{2,k} > 0; \quad (6)$$

$$h_j(z_j) = \gamma_{3,j} z_j^3 + \gamma_{2,j} z_j^2 + \gamma_{1,j} z_j, \gamma_{3,j}, \gamma_{2,j} > 0, j \in J_k. \quad (7)$$

Методы решения

Пусть существует хотя бы одна точка, удовлетворяющая условиям (2) – (4). Все функции $f_k(x_k)$, $h_j(z_j)$ и $-\phi_i(y_i)$, согласно (5) – (7), являются выпуклыми. Тем самым (1) – (4) является задачей выпуклой оптимизации. Применим подход, который был предложен в [8]. Для задачи (1) – (4) в качестве функции Лагранжа возьмем следующую функцию:

$$L(x, y, z, \lambda) = \sum_{k=1}^n \left[\sum_{i \in I_k} \phi_i(y_i) - f_k(x_k) - \sum_{j \in J_k} h_j(z_j) \right] - \lambda \left(\sum_{k=1}^n x_k - B \right).$$

Множитель Лагранжа λ был применен только к условию (2). Теперь исходная задача (1)–(4) может быть заменена двойственной:

$$\min_{\lambda \geq 0} \rightarrow \psi(\lambda), \quad (8)$$

$$\begin{aligned} \psi(\lambda) &= \max_{(x, y, z) \in W} L(x, y, z, \lambda) = \\ &= \lambda B + \max_{(x, y, z) \in W} \sum_{k=1}^n \left[\sum_{i \in I_k} \phi_i(y_i) - f_k(x_k) - \lambda x_k - \sum_{j \in J_k} h_j(z_j) \right], \\ W &= \left\{ (x, y, z) \left| \begin{array}{l} \sum_{i \in I_k} y_i = x_k + \sum_{j \in J_k} z_k, 0 \leq y_i \leq a_i, i \in I_k, \\ 0 \leq z_j \leq c_j, j \in J_k, 0 \leq x_k \leq b_k, k = 1, \dots, n. \end{array} \right. \right\} \end{aligned}$$

Благодаря двойственности (см., например, [9, 10]), задачи (1) – (4) и (8) имеют одно и то же оптимальное значение. Но решение задачи (8) может быть найдено с помощью одного из хорошо известных алгоритмов одномерной оптимизации (см., например, [10]). Для подсчета значения $\psi(\lambda)$ необходимо решить внутреннюю задачу:

$$\begin{aligned} \max \rightarrow \sum_{k=1}^n \left[\sum_{i \in I_k} \phi_i(y_i) - f_k(x_k) - \lambda x_k - \sum_{j \in J_k} h_j(z_j) \right], \\ \sum_{i \in I_k} y_i = x_k + \sum_{j \in J_k} z_k, \quad 0 \leq y_i \leq a_i, i \in I_k, \quad 0 \leq z_j \leq c_j, j \in J_k, \\ 0 \leq x_k \leq b_k, k = 1, \dots, n. \end{aligned}$$

Эта задача разбивается на n независимых зональных задач оптимизации

$$\max \rightarrow \left[\sum_{i \in I_k} \phi_i(y_i) - f_k(x_k) - \lambda x_k - \sum_{j \in J_k} h_j(z_j) \right], \quad (9)$$

$$\sum_{i \in I_k} y_i = x_k + \sum_{j \in J_k} z_k, \quad 0 \leq y_i \leq a_i, i \in I_k, \quad 0 \leq z_j \leq c_j, j \in J_k, \quad 0 \leq x_k \leq b_k,$$

для $k = 1, \dots, n$. Каждая k -я независимая зональная задача (9) совпадает с

$$\max_{(x,y,z) \in D} \rightarrow \sum_{i \in I} \phi_i(y_i) - u(x) - \sum_{j \in J_k} h_j(z_j), \quad (10)$$

$$\phi_i(y_i) = \alpha_{3,i} y_i^3 + \alpha_{2,i} y_i^2 + \alpha_{1,i} y_i + \alpha_{0,i}, \alpha_{3,i}, \alpha_{2,i} < 0, \alpha_{1,i} > 0, \alpha_{0,i} \geq 0, i \in I_k,$$

$$u(x) = f_k(x_k) + \lambda x_k = \beta_{3,k} x_k^3 + \beta_{2,k} x_k^2 + \beta_{1,k} x_k + \lambda x_k, \beta_{3,k}, \beta_{2,k} > 0,$$

$$h_j(z_j) = \gamma_{3,j} z_j^3 + \gamma_{2,j} z_j^2 + \gamma_{1,j} z_j, \gamma_{3,i}, \gamma_{2,i} > 0, j \in J_k,$$

$$y = (y_i)_{i \in I}, I = I_k, z = (z_j)_{j \in J}, J = J_k, x = x_k,$$

$$D = \left\{ (x, y, z) \left| \begin{array}{l} \sum_{i \in I} y_i = x + \sum_{j \in J} z_j, \\ 0 \leq y_i \leq a_i, i \in I, 0 \leq z_j \leq c_j, j \in J, 0 \leq x \leq b. \end{array} \right. \right\}$$

Допустим, что множество D не пусто. Тогда (10) является задачей выпуклой оптимизации. К ней можно применить метод условного градиента с линейным поиском (**CGDM**) [11] или **Алгоритм (BS)** [12], адаптировав его к случаю с несколькими продавцами.

Опишем сначала метод **CGDM**. Решаем

$$\min_{\tilde{w} \in \tilde{W}} \rightarrow \tilde{\mu}(\tilde{w}).$$

В качестве начальной точки берем точку $w^0 \in \tilde{W}$. На k -й итерации ($k = 0, 1, \dots$) имеем $w^k \in \tilde{W}$ и ищем $u^k \in \tilde{W}$ как решение

$$\min_{w \in \tilde{W}} \rightarrow \tilde{\mu}(w^k), w > .$$

После этого подсчитываем $p^k = u^k - w^k$. Если $\|p^k\| \leq \delta$, то решение найдено. Иначе ищем $\sigma_k \in (0, 1)$ из условия

$$\tilde{\mu}(w^k + \sigma_k p^k) \leq \tilde{\mu}(w^k) + \alpha \sigma_k < \tilde{\mu}'(w^k), p^k > ,$$

например, $\sigma_k = \gamma^m, m = 0, 1, \dots, \gamma \in (0, 1)$. После этого подсчитываем $w^{k+1} = w^k + \sigma_k p^k$ и переходим на следующую итерацию.

Опишем **Алгоритм (BS)**. Задача выпуклой оптимизации (10) может быть представлена в виде модели двустороннего аукциона с продавцами с нелинейными выпуклыми возрастающими функциями цен и покупателями с нелинейными выпуклыми функциями цен (см. [13, 14]).

Пусть $g(x) = u'(x)$, $v_j(z_j) = h'_j(z_j)$ и $w_i(y_i) = \phi'_i(y_i)$. Необходимое и достаточное условие оптимальности задачи (10) может быть записано в форме вариационного неравенства: найти $(\bar{x}, \bar{y}, \bar{z}) \in D$, такую, что

$$g(\bar{x})(x - \bar{x}) + \sum_{j \in J} v_j(z_j)(z_j - \bar{z}_j) - \sum_{i \in I} w_i(\bar{y}_i)(y_i - \bar{y}_i) \geq 0, \forall (x, y, z) \in D.$$

В этом случае условия оптимальности могут быть записаны как

$$\begin{aligned} (\bar{x}, \bar{y}, \bar{z}) \in D, \exists p^* \quad g(\bar{x}) & \begin{cases} \geq p^* & \text{если } \bar{x} = 0, \\ = p^* & \text{если } \bar{x} \in (0, b), \\ \leq p^* & \text{если } \bar{x} = b; \end{cases} \\ v_j(\bar{z}_j) & \begin{cases} \geq p^* & \text{если } \bar{z}_j = 0, \\ = p^* & \text{если } \bar{z}_j \in (0, c_j), \\ \leq p^* & \text{если } \bar{z}_j = c_j, \end{cases} \quad w_i(\bar{y}_i) & \begin{cases} \leq p^* & \text{если } \bar{y}_i = 0, \\ = p^* & \text{если } \bar{y}_i \in (0, a_i), \\ \geq p^* & \text{если } \bar{y}_i = a_i, \end{cases} \end{aligned} \quad (11)$$

для $j \in J$; для $i \in I$.

Для такого рода задач существует довольно много эффективных методов решения (см., например, [15] и ссылки на нее). Все они в большой степени основаны на теории двойственности. Следуя этому подходу, запишем функцию Лагранжа для задачи (10) с отрицательным знаком:

$$M(x, y, z, p) = u(x) + \sum_{j \in J} h_j(z_j) - \sum_{i \in I} \phi_i(y_i) - p \left(x + \sum_{j \in J} z_j - \sum_{i \in I} y_i \right).$$

С целью нахождения значения двойственной функции затрат

$$\theta(p) = \min_{x \in [0, b], y \in [0, a], z \in [0, c]} M(x, y, z, p),$$

где $a = (a_i)_{i \in I}$ и $c = (c_j)_{j \in J}$, необходимо решить следующие одномерные задачи:

$$\min_{0 \leq x_k \leq b_k} \rightarrow (u(x) - px),$$

$$\min_{0 \leq z_j \leq c_j} \rightarrow (h_j(z_j) - pz_j), \text{ for } j \in J,$$

$$\min_{0 \leq y_i \leq a_i} \rightarrow (-\phi_i(y_i) + py_i), \text{ for } i \in I.$$

Решения этих задач, обозначенные через $x(p)$, $z_j(p)$, $j \in J$, и $y_i(p)$, $i \in I$, соответственно определяются однозначно.

Отсюда следует, что функция $\theta(p)$ выпукла и дифференцируема

$$\theta'(p) = \sum_{i \in I} y_i(p) - x(p) - \sum_{j \in J} z_j(p).$$

Кроме того, одномерная двойственная задача

$$\max_p \rightarrow \theta(p)$$

совпадает с обычным уравнением

$$\theta'(p) = 0, \quad (12)$$

где $\theta'(p)$ не возрастает. Если p^* является решением (12), тогда мы можем найти решение исходной задачи (10) из (11) путем присвоения $p = p^*$.

Пусть $\gamma' = \min_{j \in J} \{g(0), \min v_j(0)\}$. Тогда $\gamma' \leq g(b)$ ($\gamma' < g(b)$ если $\gamma' = g(0)$) и $\gamma' \leq v_j(c_j)$, $\forall j \in J$ ($\gamma' < v_j(c_j)$ если $\exists j' \in J: \gamma' = v_{j'}(0)$, $j = j'$). Пусть $\delta'_i = \beta_i$ и $\delta''_i = \alpha_i a_i + \beta_i$.

Если взять $p'' = \max_{i \in I} \delta'_i$ и $p' = \gamma'$, тогда случай $p'' \leq p'$ сразу дает нулевые решения согласно (11). Тем самым, можно рассматривать только нетривиальные случаи, где $p' < p''$. Тогда из (11) мы должны иметь $\theta'(p') > 0$ и $\theta'(p'') < 0$. Эти свойства позволяют найти решение для (12) с помощью обычного алгоритма деления отрезка пополам, обозначенного как **Алгоритм (BS)**. Для заданной точности $\varepsilon > 0$ и начального отрезка $[p', p'']$, мы берем $\tilde{p} = 0.5(p' + p'')$, подсчитываем $\theta'(\tilde{p})$. Далее устанавливаем $p' = \tilde{p}$, если $\theta'(\tilde{p}) > 0$ и $p' = \tilde{p}$ – иначе, до тех пор, пока $(p'' - p') < \varepsilon$.

Численные эксперименты

Реализованные на C++ методы были протестированы на ПК со следующими характеристиками: Intel(R) Core(TM) i7-4500, CPU 1,80 GHz, RAM 6 Gb.

В качестве начального отрезка для выбора двойственной переменной λ был выбран отрезок $[0, 1000]$. Величины b_k , a_i , c_j были сформированы с помощью тригонометрических функций на отрезках $[1, 51]$, $[1, 2]$ и $[1, 10]$ соответственно. В качестве значения B была взята 1000. В качестве значений γ и α в (CGDM) были взяты значения 0,7 и 0,4 соответственно. Количество внешних провайдеров в каждой зоне было фиксированным и равным 5. Количество зон варьировалось от 5 до 105, количество пользователей – от 210 до 10010. Пользователи по зонам были распределены равномерно либо согласно нормальному закону распределения. Коэффициенты функций $\phi_i(y_i)$, $f_k(x_k)$ и $h_j(z_j)$ выбирались с помощью тригонометрических функций. Для всех методов решения задачи (1)–(4) точность решения верхней задачи варьировалась от 10^{-1} до 10^{-4} . Точность решения задач нижних уровней (внутренних задач) была фиксированной и равной 10^{-2} . Для каждого множества параметров было выполнено 50 тестов.

Во всех случаях предложенные методы смогли найти решение. Более того, для одинаковой точности оба метода давали одинаковое количество верхних итераций, поэтому основное различие было в затраченном процессорном времени, которое показало, что использование **Алгоритма (BS)** для внутренних задач оптимизации является более производительным.

Заключение

Рассмотрена задача оптимального распределения однородного ресурса в телекоммуникационной сети, разбитой на зоны. Применяя двойственный метод Лагранжа к ограничению на общее количество внутреннего ресурса, предлагается уменьшить исходную задачу до одномерной задачи оптимизации, в которой подсчет функции стоимости ведет к поиску решений независимых зональных задач. Решения зональных задач могут быть найдены с помощью простых методов. Результаты проведенных численных экспериментов подтвердили возможность применения предложенных методов.

ЛИТЕРАТУРА

1. *Courcoubetis C., Weber R.* Pricing Communication Networks: Economics, Technology and Modelling. – Chichester: John Wiley & Sons, 2003
2. *Stańczak S., Wiczanowski M., Boche H.* Resource Allocation in Wireless Networks. Theory and Algorithms. – Berlin: Springer, 2006.
3. *Wyglinski A.M., Nekovee M., Hou Y.T.* (eds.) Cognitive Radio Communications and Networks: Principles and Practice. – Amsterdam: Elsevier, 2010.
4. *Leshem A., Zehavi E.* Game theory and the frequency selective interference channel: A practical and theoretic point of view // IEEE Signal Process. – 2009. – V. 26. – P. 28–40
5. *Raoof O., Al-Raweshidy H.* Auction and game-based spectrum sharing in cognitive radio networks // Game Theory / Q. Huang, ed. – Sciyo, Rijeka, 2010. – Ch.2. – P. 13–40
6. *Huang J., Berry R.A., Honig M.L.* Auction-based spectrum sharing // ACM/Springer Mobile Networks and Appl. – 2006. – V. 11. – P. 405–418
7. *Koutsopoulos I., Iosifidis G.* Auction mechanisms for network resource allocation // Proc. of Workshop on Resource Allocation in Wireless Networks, WiOpt 2010. – 2010. – P. 554–563.
8. *Konnov I.V., Kashuba A.Yu., Laitinen E.* A simple dual method for optimal allocation of total network resources // Recent Advances in Mathematics. Proceedings of the International Conference "PMAMCM 2015". Ed. by I.J. Rudas. Zakynthos. – 2015. – P. 19–21.
9. *Поляк Б.Т.* Введение в оптимизацию. – М.: Наука, 1983.
10. *Konnov I.V.* Nonlinear Optimization and Variational Inequalities. – Kazan: Kazan Univ. Press, 2013.
11. *Konnov I.V., Kashuba A.Yu., Laitinen E.* Application of the conditional gradient method to resource allocation in wireless networks // Lobachevskii Journal of Mathematics. – 2016. – V. 37. – Iss. 5. – P. 626–635.
12. *Konnov I.V., Kashuba A.Yu.* Decomposition method for zonal resource allocation problems in telecommunication networks // IOP Conf. Series: Materials Science and Engineering 158, 11th International Conference on "Mesh methods for boundary-value problems and applications". – 2016.
13. *Konnov I.V.* On modeling of auction type markets // Issled. Inform. 10. – 2006. – P. 73–76, Доступно на SSRN: <http://ssrn.com/abstract=2482282> (дата обращения: 30.06.2017).
14. *Konnov I.V.* An alternative economic equilibrium model with different implementation mechanisms // Adv. Model. Optim. – 2015. – V. 17. – P. 245–265
15. *Patriksson M., Strömberg C.* Algorithms for the continuous nonlinear resource allocation problem: New implementations and numerical studies // Eur. J. Oper. Res. – 2015. – V. 243. – P. 703–722.

ПОСТРОЕНИЕ КОНЦЕПТУАЛЬНОЙ МОДЕЛИ ТРАФИКА В БЕСПРОВОДНЫХ СЕТЯХ

Е.А. Печеный, Н.К. Нуриев

*Казанский национальный исследовательский
технологический университет, г. Казань, Россия*

Беспроводные сети передачи данных являются в настоящее время неотъемлемой частью большинства телекоммуникационных информационных систем. Легкость масштабирования, эксплуатационная надежность, широкий набор постоянно обновляющихся технических средств, дает возможность реализации самых разнообразных проектных решений, удовлетворяющих практически любым требованиям заказчиков. Объемы трафика в таких сетях и интенсивность информационных потоков, как правило, весьма значительны. Однако на сегодняшний день единой научно-обоснованной концепции управления трафиком не существует, что вызывает определенные сложности в процессе эксплуатации. На практике управление идет в режиме оперативного реагирования на сложившуюся ситуацию и во многом определяется опытом администратора и особенностями его личности, т.е. факторами часто субъективными. В связи с этим очевидны актуальность и практический интерес формулирования концепции рационального управления трафиком и построение соответствующей математической модели.

Наиболее наглядно явление трафика может быть продемонстрировано с помощью физической модели «дырявое ведро», представленной на рис. 1. Поток пакетированной информации из различных источников поступает в буфер, откуда по специальному каналу доставляется к серверу. При наличии в системе нескольких специализированных серверов может добавляться функция сортировки. Когда наступает момент полной загрузки буфера, система прекращает прием пакетов, что равнозначно их потере.

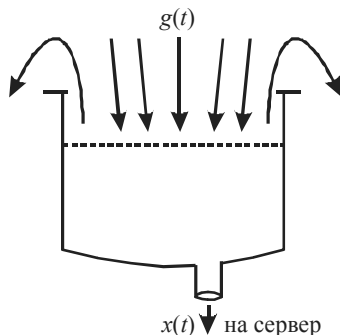


Рис. 1. Модель типа
«дырявое ведро»

В ходе ряда наблюдений, выполненных в режиме реального времени, было замечено, что входной поток информации $g(t)$ представляет собой случайный процесс $\xi(t)$ с наложенной на него периодической функцией $y(t)$, частота которой значительно меньше частоты случайных колебаний (рис. 2). Линия $g(t) = g_0$ соответствует ситуации наполненности буфера.

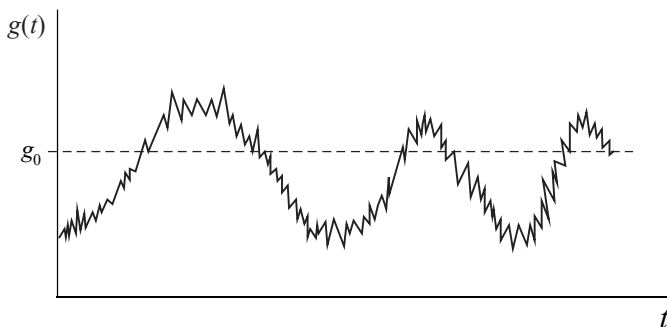


Рис. 2. Изображение фрагмента входного потока

Заметим, что как физический объект буфер выполняет функцию фильтра низких частот. Этот эффект известен и широко применяется в различных областях техники, например для устранения пульсаций жидкости в трубопроводах, в системах стабилизации вакуума, радиотехнике и т.п. Отсюда следует, что при построении модели управления трафиком случайной составляющей $\xi(t)$ можно пренебречь, поскольку ее влияние практически полностью устраняется буфером.

Из рис. 2 видно, что потери информации, обусловленные переполнением буфера, имеют периодический характер. Однако в процессе эксплуатации беспроводных сетей потери информационных пакетов могут происходить по ряду других причин, не связанных непосредственно с наполненностью буфера. Поэтому представляются вполне обоснованным принять в качестве рабочей гипотезы предложение о том, что функция потерь непрерывна и является гармонической. Это позволяет принять в качестве базовой модели уравнение

$$y' + \alpha P(t)y^2 = \beta R^{-1}(1 - P(t)), \quad (1)$$

использованное в работе [1], где y' — скорость передачи информации, $P(t)$ — функция вероятности потерь, предполагаемая гармонической, R —

временная задержка, α и β – настроечные параметры системы, которые в этой работе именуются соответственно мультипликативным уменьшением размера окна передачи данных при потере пакетов и адаптивным увеличением размера окна при отсутствии потери пакетов.

Данное уравнение представляет собой уравнение Риккати, которое в общем случае не разрешимо в квадратурах. Из общей теории обыкновенных дифференциальных уравнений [2] известно, что уравнение Риккати тесно связано с линейными дифференциальными уравнениями второго порядка. В частности, если коэффициент при квадратичном члене есть непрерывная дифференцируемая функция, каждое решение уравнения (1) преобразованием

$$u = \exp\left[\int \alpha P(t) y dt\right] \quad (2)$$

переводится в отличное от нуля решение линейного дифференциального уравнения

$$P(t)u'' - P'(t)u' + \frac{\alpha\beta}{R}P^2(t)(1 - P(t))u = 0. \quad (3)$$

Если корни характеристического многочлена этого уравнения окажутся комплексными сопряженными числами, то решение будет давать описание свободных колебаний некоторой системы. Но поскольку коэффициенты уравнения (3) есть гармонические функции времени, создаются предпосылки для возникновения параметрического резонанса, при котором источником возбуждения становятся изменения параметров внутри системы.

Поведение системы подобного типа практически непредсказуемо, так как решение чрезвычайно чувствительно к изменениям параметров уравнения (3). Ничтожный сдвиг параметров в ту или иную сторону может стать причиной внезапного перехода системы в совершенно иное состояние и кардинально поменять вид интегральных кривых.

Для уравнения (1) была сформулирована задача Коши при естественных начальных условиях $y(0) = 0$ и получено ее решение методом Рунге-Кутты, реализованного с помощью специально созданной компьютерной программы. В процессе вычислений было принято, что функция вероятности потерь изменяется по закону синуса. На рис. 3, а–г приведены интегральные кривые, полученные для некоторых значений параметров. На рис. 3, а наблюдается ярко выраженный пик, свидетельствующий о попадании в область резонанса.

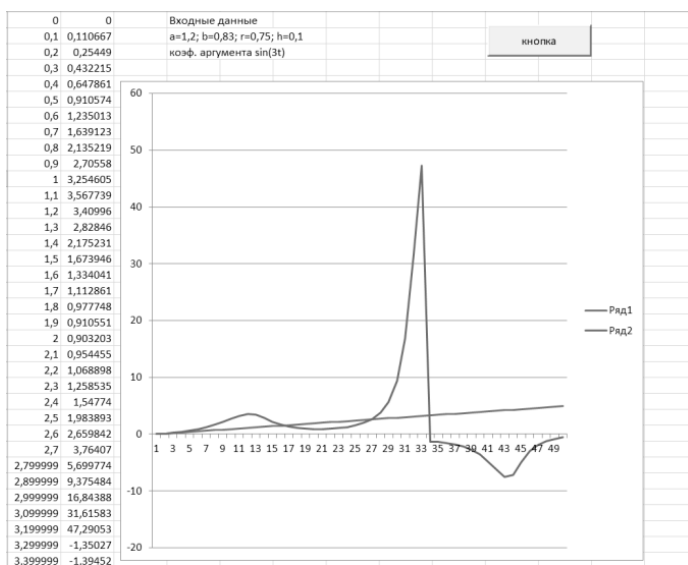


Рис. 3а. Интегральная кривая

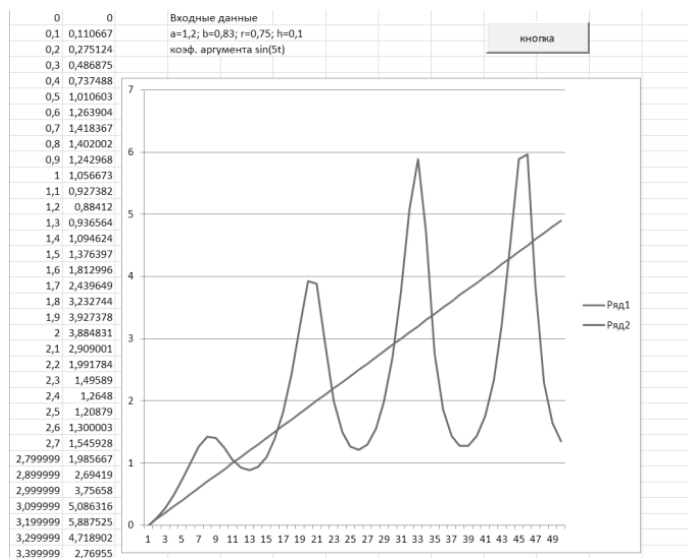


Рис. 3б. Интегральная кривая

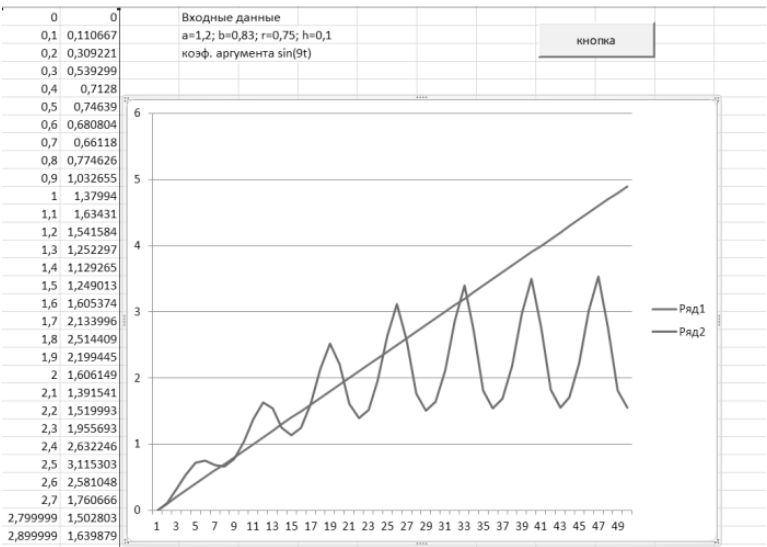


Рис. 3в. Интегральная кривая

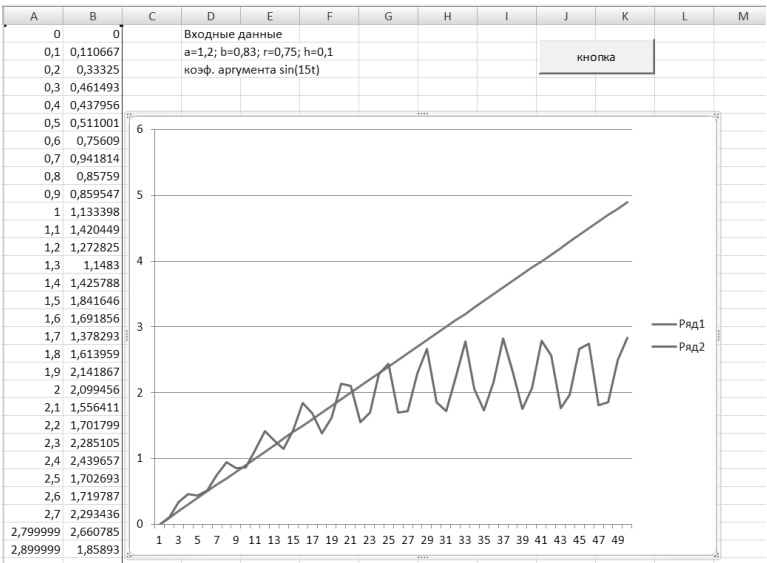


Рис. 3г. Интегральная кривая

По мере увеличения частоты решение сохраняет колебательную природу, однако амплитуда колебаний заметно уменьшается (рис. 3, б–г). Естественно, что при иных значениях параметров α и β вид решения, может претерпеть очень существенные изменения.

Изучение поведения систем, в которых существует возможность возникновения параметрического резонанса, представляет собой задачу чрезвычайной сложности и трудоемкости. В самых общих чертах ее решение предполагает отыскание множества значений параметров, при которых явление параметрического резонанса невозможно, и другого множества, при котором резонанс имеет место. Представляя эти множества в виде точек на плоскости, получают так называемые диаграммы Айнса – Стретта [3], разделяющие «резонансные» и «нерезонансные» области. Основным, и по сути дела единственным, способом эффективного управления такого рода объектами являются отстройка от резонанса, т.е. отыскание значений параметров, принадлежащих «нерезонансной» области и инструментов, позволяющих поддерживать их на желаемом уровне в течение всего времени работы системы.

Заключение

1. Экспериментально показана возможность потока информации в беспроводной сети как суммы случайной (высокочастотной) и квазирегулярной (низкочастотной) составляющих.
2. Обоснована нецелесообразность включения в состав модели случайно составляющей.
3. Обоснован гармонический характер функции вероятности потерь информационных пакетов.
4. Установлен колебательный вид решения уравнения Риккати и доказано наличие предпосылок возникновения параметрического резонанса.
5. Сформулирован общий принцип управления трафиком беспроводной сети.

ЛИТЕРАТУРА

1. Славко О.Г. Імітаційна модель сегменту комп'ютерної мережі з реалізацією методи шейлінгу трафіка // Вісник КрНУ імені Михайла Остроградського. – 2014. – Вип. 6 (89). – С. 44–51.
2. Камке Э. Справочник по обыкновенным дифференциальным уравнениям. – М.: Наука, 1976. – 576 с.
3. Мигулин В.В., Медведев В.И., Мустель Е.Р., Парыгин В.Н. Основы теории колебаний. – М.: Наука, 1978. – 316 с.

ПРИКЛАДНОЙ ВЕРОЯТНОСТНЫЙ АНАЛИЗ

О ПРИМЕНЕНИИ ТАУБЕРОВОЙ ТЕОРЕМЫ ДЛЯ ПРОИЗВОДЯЩИХ ФУНКЦИЙ В ТЕОРИИ ВЕТВЯЩИХСЯ СЛУЧАЙНЫХ ПРОЦЕССОВ

А.Х. Мейлиев, Э.Э. Тухтаев

Каршинский государственный университет, г. Карши, Узбекистан

В настоящем сообщении обсуждаются свойства инвариантных мер для ветвящихся процессов с помощью известных тауберовых теорем. Происхождение тауберовых теорем связано с исследованием асимптотических свойств меры $U(x)$, сосредоточенной на положительной оси \mathbb{R}_+ , с помощью ее преобразование Лапласа

$$\omega(\theta) = \int_{\mathbb{R}_+} e^{-\theta x} U(dx).$$

На самом деле, при весьма общих условиях, поведение $\omega(\theta)$ вблизи нуля однозначно определяет асимптотическое поведение $U(x)$ при $x \rightarrow \infty$. Любое подобное соотношение между $\omega(\theta)$ и $U(x)$ принято называть Тауберовой теоремой. Суть Тауберовой теоремы можно излагать в следующем утверждении, доказанном в [1, с. 508–510].

Теорема Т. Для фиксированного $\rho \in \mathbb{R}_+$ и любого $\lambda \in \mathbb{R}_+$ каждое из соотношений

$$\frac{\omega(\lambda\tau)}{\omega(\tau)} \longrightarrow \frac{1}{\lambda^\rho} \quad \text{при } \tau \downarrow 0 \quad (1)$$

$$\text{и} \quad \frac{U(xt)}{U(t)} \longrightarrow x^\rho \quad \text{при } t \rightarrow \infty \quad (2)$$

влечет другое.

Обратный переход от (2) к (1), как известно, называется Абелевой теоремой.

Помимо Теоремы Т приходится иметь дело с функциями, для которых выполняется условие типа (2). Такие функции называются правильно меняющимися. Автором идеи создания понятия правильного изменения функций является ныне хорошо известный математик И. Карамата, который свои фундаментальные работы по этой тематике опубликовал в начале 30-х годов 20-го века (см. [2, 3]). Из множества положительных и измеримых по Лебегу функций вещественного переменного выделим следующий класс:

$$S_{\infty} := \left\{ l(x) \in \mathbb{R}_+ : \lim_{x \rightarrow \infty} \frac{l(\lambda x)}{l(x)} = 1, \text{ для } \forall \lambda \in \mathbb{R}_+ \right\}.$$

Функция $L(x)$ называется медленно меняющейся функцией в бесконечности (в смысле Карамата), если она принадлежит классу S_{∞} : $L(x) \in S_{\infty}$. Обозначение $L(*)$ для медленно меняющихся (ММ) функций происходило из первой буквы французского слова «Lentement», что в переводе означает «медленно» и, это потому, что основополагающие работы по теории ММ-функций были написаны Караматой на французском языке. Функция $V(x)$ называется правильно меняющейся (ПМ) в бесконечности, если она принадлежит классу

$$R_{\infty}^{\rho} := \left\{ V(x) : V(x) = x^{\rho} l(x), \text{ где } l(x) \in S_{\infty} \right\},$$

где параметр $\rho \in \mathbb{R}_+$ называется показателем правильного изменения. Из сходимости (2) видно, что ММ-функция является ПМ-функцией с показателем $\rho = 0$: $S_{\infty} \equiv R_{\infty}^0$.

Элементы теории функций Караматы в исследовании ветвящихся процессов одним из первых были применены в работе Золотарева [4]. В список ранних работ, где удачно используется ПМ функций в теории ветвящихся процессов, можно внести также работы Слэк [5, 6] и Сенеты [7, 8]. Подробные материалы, связанные с применением ПМ функций в теории ветвящихся процессов, можно найти в монографиях [9] и [10].

Обозначим $\{Z(n), n \in \mathbb{N}_0\}$ численность популяции частиц в ветвящемся процессе дискретного времени с вероятностной производящей функцией (ПФ) $F(s) = E s^{Z(n)}$, где $\mathbb{N}_0 = \{0\} \cup \mathbb{N}$ и $\mathbb{N} = \{1, 2, \dots\}$. Рассматриваемый процесс образует однородную цепь Маркова с пространством состояний $\mathcal{S} \subseteq \mathbb{N}_0$. Введем переходные вероятности $P_{ij}(n) = \mathbb{P}_i \{Z_n = j\}$,

где $\mathbb{P}_i\{*\} := \mathbb{P}\{*|Z_n = i\}$. Эти вероятности определяются функциональными итерациями

$$F_n(s) = F(F_{n-1}(s)) = F_{n-1}(F(s)),$$

где

$$F_n(s) = \sum_{j \in \mathcal{S}} P_{ij}(n) s^j$$

для любых $i, j \in \mathcal{S}$; см. [11].

Эволюция процесса управляется регулирующим параметром $A = F'(1-)$ – средним числом непосредственных потомков одной частицы за одно поколение. Величина $\mathcal{H} := \min\{n \in \mathbb{N} : Z(n) = 0\}$ обозначает момент вырождения процесса. Из теоремы о вырождении известно, что $\mathbb{P}_i\{\mathcal{H} < \infty\} = q^i$, где q есть вероятность вырождения процесса, которая является наименьшим корнем уравнения $x = F(x)$. Этот корень равен 1, если $A \leq 1$ и меньше 1 при $A > 0$. Процесс $\{Z(n)\}$ классифицируется как субкритический, критический и суперкритический, если $A < 1$, $A = 1$ и $A > 1$ соответственно.

Мы изучаем предельные свойства переходных вероятностей $P_{ij}(n)$. Обсуждаем задачу существования инвариантной меры и свойство эргодичности процесса $\{Z(n), n \in \mathbb{N}_0\}$. Докажем теоремы о сходимости к инвариантным мерам. При исследовании свойств инвариантной меры воспользуемся аналогом Теоремы Т для ПФ.

Основные результаты

Для нашей цели существенно используем следующую лемму о монотонной сходимости отношений.

Лемма [11, с. 15]. Для всех $A < \infty$ и $i, j \in \mathcal{S}$

$$\frac{P_{ij}(n)}{P_{11}(n)} \uparrow \pi_j \leq \infty \quad (3)$$

при $n \rightarrow \infty$ и выполняется инвариантное соотношение

$$\beta^i \cdot \pi_j = \sum_{k \in \mathcal{S}} \pi_k P_{kj}(n), \quad (4)$$

при любом $n \in \mathbb{N}$, где $\beta = F'(q)$.

Введем ПФ

$$\mathbf{P}_i(n; s) = \sum_{j \in \mathcal{S}} \frac{P_{ij}(n)}{P_{11}(n)} s^j \quad \text{и} \quad \mathbf{P}(s) = \sum_{j \in \mathcal{S}} \pi_j s^j.$$

В этих обозначениях сходимость (3) эквивалентна тому, что

$$\mathbf{P}_i(n; s) \longrightarrow iq^{i-1} \cdot \mathbf{P}(s)$$

при $n \rightarrow \infty$ равномерно для $0 \leq s < 1$ и, $\mathbf{P}(0) = 0$. А соотношение (4) влечет функциональное уравнение

$$\beta^n \cdot \mathbf{P}(s) = \mathbf{P}(F_n(s)) - \mathbf{P}(F_n(0)).$$

Таким образом, множество неотрицательных чисел $\{\pi_j, j \in \mathcal{S}\}$ представляет собой инвариантную меру для процесса $\{Z(n), n \in \mathbb{N}_0\}$.

Рассмотрим критический процесс в случае, когда распределение числа потомков одной частицы имеет бесконечную дисперсию. Но ПФ $F(s)$ допускает для $0 \leq s < 1$ следующее представление:

$$F(s) = s + (1-s)^{1+\nu} L\left(\frac{1}{1-s}\right), \quad [\mathfrak{R}_\nu] \quad (5)$$

где $0 < \nu \leq 1$ и $L(x) \in S_\infty$. Для случая когда $F''(1-) < \infty$, условие $[\mathfrak{R}_\nu]$ имеет место со значением $\nu = 1$ и

$$L(n) \longrightarrow \frac{F''(1-)}{2}$$

при $n \rightarrow \infty$. Следовательно, рассматриваемый случай содержит процесс с конечной дисперсией. Представление же $[\mathfrak{R}_\nu]$ означает условие правильного изменения «хвоста» распределения числа потомков одной частицы.

Следующие две теоремы доказываются с помощью тауберовой теоремы для ПФ. Они описывают основные свойства инвариантных мер для критических ветвящихся процессов.

Теорема 1. Пусть $A=1$ и выполнено условие $[\mathfrak{R}_\nu]$. Тогда

$$\sum_{j=1}^n \pi_j = \frac{F(0)}{\nu^2 \cdot \Gamma(\nu)} n^\nu L_\pi(n),$$

где $\Gamma(*)$ – Гамма-функция Эйлера и $L_\pi(n) \cdot L(n) \rightarrow 1$ при $n \rightarrow \infty$.

Рассмотрим теперь случайный процесс $\{\tilde{Z}(n), n \in \mathbb{N}_0\}$, с переходными вероятностями

$$\tilde{P}_{ij}(n) = \mathbb{P}_i \{Z_n = j | \mathcal{H} > n\}.$$

Имеем следующую теорему.

Теорема 2. Пусть выполнено условие $[\mathfrak{R}_\nu]$. Тогда при $n \rightarrow \infty$

$$\nu n \cdot \tilde{P}_{ij}(n) \longrightarrow \mu_j,$$

независимо от $i \in \mathcal{S}$. Множество $\{\mu_j, j \in \mathcal{S}\}$ является инвариантной мерой для процесса $\{\tilde{Z}(n), n \in \mathbb{N}_0\}$ и, сумма

$$\sum_{j=1}^n \mu_j = \frac{1}{\nu^2 \cdot \Gamma(\nu)} n^\nu L_\mu(n),$$

где $\Gamma(*)$ – гамма-функция Эйлера и $L_\mu(n) \cdot L(n) \rightarrow 1$ при $n \rightarrow \infty$.

Авторы выражают благодарность своему научному руководителю А.А. Имомову за поддержку и внимание при выполнении данной работы.

ЛИТЕРАТУРА

1. Феллер В. Введение в теорию вероятностей и ее приложения. – М.: Мир, 1967.
2. Karamata J. Sur un mode de croissance reguliere. Theoremes fondamentaux // Bull. Soc. Math. France. – 1933. – V. 61. – P. 55–62.
3. Karamata J. Sur un mode de croissance reguliere des fonctions // Mathematica (Cluj). – 1930. – No. 4. – P. 38–53.
4. Zolotarev V.M. More exact statements of several theorems in the theory of branching processes // Theory of Prob. and Appl. – 1957. – No. 2. – P. 245–253.
5. Slack R.S. Further notes on branching processes with mean 1 // Z. Wahrscheinlichkeitsth. – 1972. – No. 25, pp.31–38.
6. Slack R.S. A branching process with mean one and possible infinite variance // Z. Wahrscheinlichkeitsth. – 1968. – No. 9. – P.139–145.
7. Seneta E. A Tauberin theorem of E. Landau and W. Feller // J. Ann. Prob. – 1973. – No. 1. – P.1057–1058.
8. Seneta E. On the invariant measures for simple branching process // J. Applied. Prob. – 1971. – No. 8. – P. 43–51.
9. Asmussen S., Hering H. Branching Processes. – Boston, 1983.
10. Bingham N.H., Goldie C.M., Teugels J.L. Regular Variation. – Cambridge, 1987.
11. Athreya K.B. and Ney P.E. Branching processes. – New York: Springer, 1972.

ВИЗУАЛИЗАЦИЯ МЕТОДА РЕКУРРЕНТНОГО ПОСТРОЕНИЯ РАСПРЕДЕЛЕНИЙ ВЕРОЯТНОСТЕЙ КОНЕЧНЫХ СЛУЧАЙНЫХ МНОЖЕСТВ НА ДУПЛЕТЕ СОБЫТИЙ

Н.А. Лукьянова^{1,2}, Д.В. Семенова^{1,2}

¹*Сибирский федеральный университет, Институт математики
и фундаментальной информатики, г. Красноярск, Россия*

²*Красноярский государственный медицинский университет
им. профессора В.Ф. Войно-Ясенецкого, г. Красноярск, Россия*

В теории вероятностей и ее приложениях в качестве случайных элементов рассматриваются различные случайные объекты, такие, как случайные величины, векторы, функции, процессы, множества, а также конечные случайные множества [1]. Под конечным случайным множеством понимается случайный элемент со значениями из совокупности всех подмножеств некоторого конечного множества U [2, 3].

Пусть (Ω, F, \mathbf{P}) – вероятностное пространство, где Ω – пространство элементарных исходов, F – σ -алгебра событий, \mathbf{P} – вероятностная мера на F . Пусть $U \subset F$ – конечное множество из N случайных событий, выбранных из алгебры этого пространства.

Определение 1. Конечное случайное множество (к.с.м.) определяется как отображение $K: \Omega \rightarrow 2^U$, измеримое относительно пары алгебр $(F, 2^{2^U})$, такое, что для всякого $X \subseteq U$ справедливо

$$K^{-1}(X) = \{\omega \in \Omega : K(\omega) = X\} \in F.$$

Отображение $K: \Omega \rightarrow 2^U$ порождает новое вероятностное пространство $(K(\Omega), K(F), \mathbf{P}_K) = (2^U, 2^{2^U}, \mathbf{P}(K^{-1}(\cdot)))$. Распределение вероятностей к.с.м. определяется вероятностной мерой \mathbf{P}_K . Для того чтобы уметь задавать распределение к.с.м., необходимо уметь задавать вероятности на прообразах $K^{-1}(\cdot)$. В работах [4–6] предлагается метод построения таких вероятностей. Распределение вероятностей к.с.м. есть набор из 2^N значений вероятностной меры на системе событий, порожденной конечным множеством U .

Распределение вероятностей I рода случайного множества K на U есть набор $\{p(X), X \subseteq U\}$ из 2^N вероятностей вида $p(X) = \mathbf{P}(\{K = X\})$. Распределение вероятностей I рода удовлетворяет следующим условиям: $0 \leq p(X) \leq 1$, $X \subseteq U$ и $\sum_{X \subseteq U} p(X) = 1$.

Распределение вероятностей II рода есть набор из 2^N вероятностей вида $\{p_X, X \subseteq U\}$, где $p_X = \mathbf{P}(\{K \supseteq X\})$.

Для полного определения распределения вероятностей к.с.м. достаточно любого одного из этих двух типов распределений, каждое из которых задано на своей системе событий, порожденных конечным множеством U . Зная одно, можно получить другое по формулам обращения Мёбиуса [4–6]

$$p_X = \sum_{Y \in 2^U: X \subseteq Y} p(Y); \quad p(X) = \sum_{Y \in 2^U: X \subseteq Y} (-1)^{|Y|-|X|} p_Y, \quad (1)$$

для всех $X \in 2^U$.

В работах [4 – 6] изложен метод рекуррентного построения распределений вероятностей на основе ассоциативных функций. В теории вероятностей и нечеткой логике широко используются разновидности ассоциативных функций, такие, как треугольные нормы и копулы [7]. Традиционно под ассоциативной функцией понимается функция 2-х аргументов на отрезке $[0, 1]$, удовлетворяющая следующим свойствам определения 2.

Определение 2. Ассоциативная функция $\text{AF}: [0, 1]^2 \rightarrow [0, 1]$, удовлетворяет следующим свойствам:

A1. Граничные условия: $\forall a \in [0, 1]$

$$\text{AF}(a, 0) = \text{AF}(0, a) = 0, \quad \text{AF}(a, 1) = \text{AF}(1, a) = a.$$

A2. Монотонность: $\forall a_1, a_2, b_1, b_2 \in [0, 1]$ таких, что $a_1 \leq a_2, b_1 \leq b_2$, справедливо $\text{AF}(a_1, b_1) \leq \text{AF}(a_2, b_2)$.

A3. Коммутативность, т.е. $\forall a, b \in [0, 1] \quad \text{AF}(a, b) = \text{AF}(b, a)$.

A4. Ассоциативность, т.е. $\forall a, b, c \in [0, 1]$

$$\text{AF}(\text{AF}(a, b), c) = \text{AF}(a, \text{AF}(b, c)).$$

A5. Условие Липшиц-непрерывности: $\forall a, b, c \in [0, 1]$

$$\text{AF}(c, b) - \text{AF}(a, b) \leq c - a, \quad a \leq c.$$

Заметим, что свойства A1 – A4 соответствуют определению t-нормы [7]. Таким образом, под ассоциативной функцией будем понимать непрерывную t-норму, удовлетворяющую условию Липшица. Методы вычисления треугольных норм применимы к распределениям вероятностей к.с.м. В качестве аргументов ассоциативной функции будем рассматривать вероятности событий (их число совпадает с мощностью базового множества). Свойства ассоциативной функции позволяют получать вероятностные распределения с заданной структурой зависимости.

В методе рекуррентного построения применительно к распределениям вероятностей II рода заданы N маргинальных вероятностей событий $\mathbf{P}(x) = p_x, x \in U, |U| = N$ и задана ассоциативная функция $\text{AF}(a, b)$. Требуется найти распределение вероятностей II рода на основе $\text{AF}(a, b)$. Суть метода – это двухэтапное решение данной задачи.

На 1 этапе осуществляется построение функции множества $\{f(X), X \subseteq U\}$, определенной на системе событий II рода как набора $|X|$ – местных функций со следующими значениями:

$$f(\emptyset) = 1; f(\{x\}) = p_x; f(\{x, y\}) = g(p_x, p_y); \dots, f(X) = g(p_x, x \in X).$$

Рекуррентное соотношение, определяющее процесс построения функции множества

$$\mathbf{P}\left(\bigcap_{x \in X} x\right) = g(p_x, x \in X) = \text{AF}(p_x, g(p_t, t \in X \setminus \{x\})). \quad (2)$$

Данное рекуррентное соотношение получено исходя из свойств ассоциативных функций.

На 2 этапе выполняется проверка достаточных условий существования распределений.

Продemonстрируем метод рекуррентного построения на примере дуплета событий. Распределение вероятностей I рода $\{p(X), X \subseteq U\}$ случайного множества K можно рассматривать как точку в $2^{|U|}$ -вершинном симплексе

$$\mathbf{S}_U = \left\{ \{p(X), X \subseteq U\} : p(X) \geq 0, \sum_{X \subseteq U} p(X) = 1 \right\},$$

каждая вершина которого занумерована подмножеством $X \subseteq U$ и ее

координате соответствует вырожденное распределение случайного множества событий, когда $p(X) = 1$, а все остальные $2^{|U|} - 1$ вероятности равны нулю.

Пусть область $\Lambda_{AF} \subset S_U$

$$\Lambda_{AF} = \left\{ \{p(X), X \subseteq U\} : p(X) = \sum_{Y \supseteq X} (-1)^{|Y|-|X|} AF(\mathbf{P}(x), x \in Y) \right\}$$

определяет множество распределений, полученных методом рекуррентного построения с ассоциативной функцией AF.

Предложенный в [4–6] рекуррентный метод позволяет выделить в S_U некоторую область, каждая точка которой определяет распределение вероятностей I рода случайного множества событий со свойствами, названными используемой ассоциативной функцией.

Лемма. Для любых $0 \leq p_x \leq 1$, $0 \leq p_y \leq 1$ и любой ассоциативной функции $AF(p_x, p_y)$ всегда существует распределение вероятностей случайного множества, заданного на дуплете событий $U = \{x, y\}$ рекуррентным методом (2).

Доказательство. Определим функцию множества $\{f(X), X \subseteq U\}$ на конечном множестве событий $U = \{x, y\}$, значения которой определяются рекуррентным методом с произвольной ассоциативной функцией AF

$$f(\emptyset) = 1, f(x) = p_x, f(y) = p_y, f(x, y) = AF(p_x, p_y).$$

Построим новую функцию множества $\{p(X), X \subseteq U\}$, преобразовав функцию множества $\{f(X), X \subseteq U\}$ по формулам обращения Мёбиуса (1):

$$\begin{aligned} & \{p(\emptyset), p(x), p(y), p(x, y)\} = \\ & = \{1 - p_x - p_y - AF(p_x, p_y), p_x, p_y, AF(p_x, p_y)\}. \end{aligned} \quad (3)$$

Из свойств ассоциативной функции непосредственно следует, что функция множества (3) является распределением I рода. Лемма доказана.

Рассмотрим дуплет событий $U = \{x, y\}$. Симплекс S_U будет иметь вид

$$S_U = \left\{ \{p(\emptyset), p(x), p(y), p(x, y)\} : p(X) \geq 0, \sum_{X \subseteq U} p(X) = 1 \right\}.$$

На рис. 1 представлен симплекс распределений вероятностей I-го рода S_U . Каждая точка в симплексе имеет координаты $(p(\emptyset), p(x), p(y), p(x, y))$ и является распределением I рода некоторого случайного множества K , заданного на дуплете событий $U = \{x, y\}$. Вершины симплекса занумерованы вероятностями событий и соответствуют вырожденным распределениям:

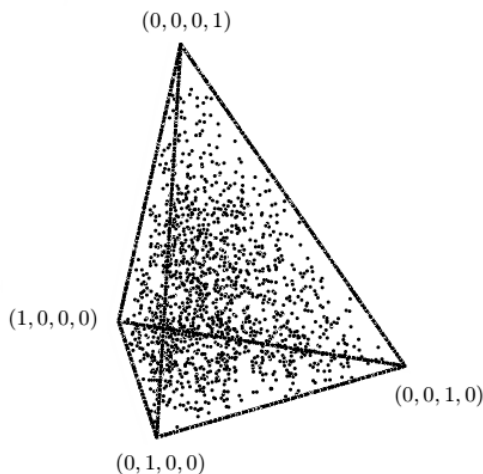


Рис. 1. Симплекс распределений вероятностей I рода S_U случайного множества K , заданного на дуплете событий $U = \{x, y\}$

– вершина $p(\emptyset)$ соответствует

$$(p(\emptyset), p(x), p(y), p(x, y)) = (1, 0, 0, 0);$$

– вершина $p(x)$ соответствует

$$(p(\emptyset), p(x), p(y), p(x, y)) = (0, 1, 0, 0);$$

– вершина $p(y)$ соответствует

$$(p(\emptyset), p(x), p(y), p(x, y)) = (0, 0, 1, 0);$$

– вершина $p(x, y)$ соответствует

$$(p(\emptyset), p(x), p(y), p(x, y)) = (0, 0, 0, 1).$$

Для случайного множества K , заданного на дуплете событий $U = \{x, y\}$, распределение вероятностей которого определяется методом рекуррентного построения с ассоциативной функцией AF, область $\Lambda_{AF} \subset S_{\{x, y\}}$ имеет вид

$$\Lambda_{AF} = \{1 - p_x - p_y - AF(p_x, p_y), p_x - AF(p_x, p_y), p_y - AF(p_x, p_y), AF(p_x, p_y)\}.$$

Далее, рассмотрим однопараметрическое семейство функций Франка:

$$\begin{aligned} Frank(p_x, p_y; \alpha) &= AF_\alpha(p_x, p_y) = \\ &= -\frac{1}{\alpha} \ln \left(1 + \frac{(e^{-\alpha p_x} - 1)(e^{-\alpha p_y} - 1)}{e^{-\alpha} - 1} \right), \quad \alpha \neq 0. \end{aligned} \quad (4)$$

В работе [5] исследование однопараметрического семейства Франка показало, что (4) дает инструмент для построения к.с.м. с заранее заданной структурой зависимостей. Для данного класса распределений исследованы предельные случаи. Проиллюстрируем предельные случаи на примере дуплета. Если конечное случайное множество K определяется распределением Π рода $\{p_\emptyset, p_x, p_y, Frank(p_x, p_y; \alpha)\}$, построенным ассоциативной функцией Франка, тогда

1) $\lim_{\alpha \rightarrow \pm 0} Frank(p_x, p_y; \alpha) = p_x \cdot p_y$ – независимо-точечное случайное множество;

2) $\lim_{\alpha \rightarrow +\infty} Frank(p_x, p_y; \alpha) = \min\{p_x, p_y\}$ – случайное множество вложенных событий;

3) $\lim_{\alpha \rightarrow -\infty} Frank(p_x, p_y; \alpha) = \max\{p_x + p_y - 1, 0\}$ – случайное множество непересекающихся событий.

На рис. 2 представлена визуализация предельных случаев на примере дуплета: структура зависимости дуплета событий от непересекающихся событий до вложенных, и симплекс распределений вероятностей I рода на основе однопараметрической ассоциативной функции Франка [5].

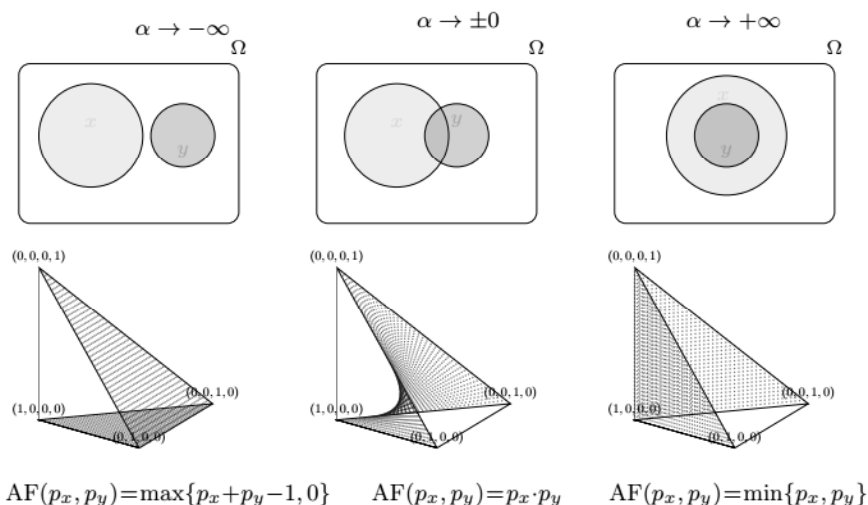


Рис. 2. Визуализация метода на примере дуплета на основе однопараметрической ассоциативной функции Франка

Использование однопараметрического семейства ассоциативных функций Франка дает набор распределений, структура зависимостей событий которых описывается параметром ассоциативной функции Франка. Различные значения параметра α приводят к различным распределениям вероятностей. На рис. 3 представлена область $\Lambda_{Frank} \subset S_{\{x, y\}}$ распределений вероятностей I-го рода случайного множества K , полученная методом рекуррентного построения на основе ассоциативной функции Франка [5] при различных значениях параметра α .

Использование однопараметрических семейств ассоциативных функций позволяет получать спектр распределений, описывающих конечные случайные множества, структура зависимостей которых регулируется параметром α .

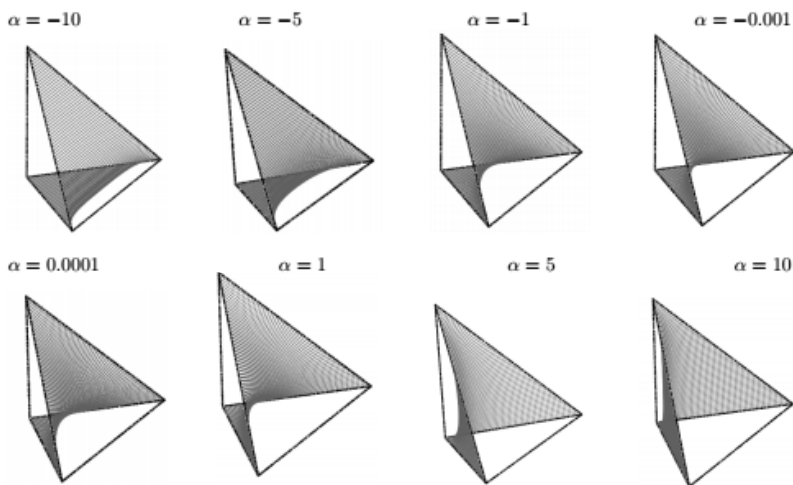


Рис. 3. Области $\Lambda_{Frank} \subseteq S_{\{x,y\}}$ распределений вероятностей I-го рода случайного множества K , полученные методом рекуррентного построения на основе ассоциативной функции Франка при различных значениях параметра α

ЛИТЕРАТУРА

1. Ширяев А.Н. Вероятность: в 2 кн. – 3-е изд., перераб. и доп. – М.: МЦНМО, 2004. – Кн. 1. – 520 с.
2. Ляшенко Н.Н. Случайное множество // Вероятность и математическая статистика. Энциклопедия / гл. ред. Ю.В. Прохоров. – М.: БРЭ, 1999. – С. 596.
3. Nguyen H.T. An Introduction to Random Sets. – Taylor & Francis Group, LLC, 2006.
4. Семенова Д.В., Лукьянова Н.А. Рекуррентное построение дискретных вероятностных распределений случайных множеств событий // Прикладная дискретная математика. – 2014. – № 4(26). – С. 47–58.
5. Лукьянова Н.А., Семенова Д.В. Ассоциативные функции Франка в построении семейств дискретных вероятностных распределений случайных множеств событий // Прикладная дискретная математика. – 2016. – № 2(32). – С. 5–19.
6. Lukyanova N.A., Semenova D.V. The study of discrete probabilistic distributions of random sets of events using associative function // J. Siberian Federal University – Mathematics and Physics. – 2014. – № 7(4). – P. 500–514.
7. Alsina S., Frank M., Schweizer B. Associative functions: Triangular Norms and Copulas. – Singapore: World Scientific Publishing Co. Pte. Ltd., 2006. – 237 p.

МЕТОД МИНИМАЛЬНОГО РАССТОЯНИЯ В ПОДГОНКЕ РАСПРЕДЕЛЕНИЙ КОНЕЧНЫХ СЛУЧАЙНЫХ МНОЖЕСТВ

Д.В. Семенова^{1,2}, Н.А. Лукьянова^{1,2}

*¹Сибирский федеральный университет, Институт математики
и фундаментальной информатики, г. Красноярск, Россия*

*²Красноярский государственный медицинский университет
им. профессора В.Ф. Войно-Ясенецкого, г. Красноярск, Россия*

В настоящее время актуальны задачи, связанные с необходимостью обработки больших массивов данных с целью поиска новых закономерностей, установления и выявления новых знаний, которые впоследствии могут быть использованы экспертами прикладной области. Подобные задачи, как правило, имеют комбинаторный характер и заключаются в поиске зависимостей между связанными событиями в виде, доступном интерпретации человеком. Извлечение полезных знаний невозможно без хорошего понимания сути данных, а успешный анализ требует качественной предобработки данных и тщательного выбора модели для интерпретации зависимостей, которые могут быть обнаружены. В работе в качестве такой математической модели сложных объектов и систем, когда число описываемых их признаков конечно и появление любого из этих признаков представляется как случайное событие рассматриваются конечные случайные множества. Вероятностные распределения конечных случайных множеств позволяют дать сжатое описание неструктурированных данных, для которых размер каждой транзакции (множества событий, произошедших одновременно) не является фиксированным.

Пусть (Ω, F, P) – вероятностное пространство. Конечное случайное множество K (к.с.м.), заданное на конечном множестве U определяется [1–3] как отображение $K: \Omega \rightarrow 2^U$, измеримое относительно пары алгебр $(F, 2^{2^U})$, такое, что для всех $X \in 2^U$ справедливо

$$K^{-1}(X) = \{\omega \in \Omega : K(\omega) = X\} \in F.$$

Измеримое событие $\{\omega \in \Omega : K(\omega) = X\}$, которое заключается в том, что случайное множество K принимает одно из своих возможных значений

$X \in 2^U$, означает, что наступившие случайные события из U образуют подмножество $X \subseteq U$, а ненаступившие – подмножество $X^c = U \setminus X$. Вместо $\{\omega \in \Omega : K(\omega) = X\}$ будем использовать обозначение $\{K = X\}$.

Распределение вероятностей к.с.м. есть набор из $2^{|U|}$ значений вероятностной меры P на системе событий, порожденной множеством U . Распределения вероятностей моделируют всевозможные способы теоретико-множественных отношений между подмножествами конечного множества событий. Для того чтобы построить распределение, необходимо задать вероятностную меру на различных системах событий, порожденных соответствующим множеством. Рассмотрены две системы событий, задающих распределения I и II рода соответственно [4–6]. На языке конечных случайных множеств эти системы событий представлены в первой строке таблицы. Вторая строка содержит их эквивалентное теоретико-множественное описание. Для полного определения распределения вероятностей конечных случайных множеств достаточно любого одного из этих типов распределений, каждое из которых задано на своей системе событий, порожденных конечным множеством событий. Зная одно, можно получить другое по формулам обращения Мёбиуса [4–6].

Распределения вероятностей к.с.м.

Системы событий	$\{K = X\}$	$\{K \supseteq X\}$
Теоретико-множественное описание	$\left(\bigcap_{x \in X} x \right) \cap \left(\bigcap_{x \in X^c} x^c \right)$	$\left(\bigcap_{x \in X} x \right)$
Вероятности	$p(X) = P(\{K = X\})$	$p_X = P(\{K \supseteq X\})$
Тип распределения вероятностей к.с.м.	I-го рода	II-го рода

Подгонкой (*fitting*) называют аналитические процедуры, позволяющие подобрать распределение, которое с достаточной степенью точности описывает наблюдаемые данные. Рассмотрим численную аппроксимацию эмпирических распределений конечных случайных множеств известными теоретическими распределениями на основе метода минимального расстояния [8]. Процесс аппроксимации выполняется поэтапно. На первом этапе выбор осуществляется среди теоретических распределений, полученных, в частности, методом рекуррентного построения распределения вероятностей конечных случайных множеств одно-

параметрическими ассоциативными функциями [4–7]. Тем самым параметр теоретического распределения определяется единственным параметром соответствующей ассоциативной функции. Предполагается, что вероятности событий p_x , $x \in U$, для теоретического и эмпирического распределений совпадают. На втором этапе для оценки параметра теоретического распределения используют традиционные для математической статистики меры близости распределений, такие, как расстояние Кульбака – Лейблера, χ^2 -расстояние Пирсона, вероятностную симметрическую χ^2 -меру, расстояние Евклида между двумя точками в $2^{|U|}$ -мерном пространстве и другие [8].

Пусть $\mathbf{p}(\alpha) = \{p(X; \alpha), X \subseteq U, \alpha \in \mathbf{Dom}_{AF}\}$ – теоретическое распределение I рода, полученное методом рекуррентного построения распределения вероятностей к.с.м. с использованием одной из однопараметрических ассоциативных функций [4–7]; \mathbf{Dom}_{AF} – область определения параметра α для используемой ассоциативной функции; $\mathbf{q} = \{q(X), X \subseteq U\}$ – эмпирическое распределение I рода. При вычислении оценок минимального расстояния по α минимизируется некоторая мера «близости» («расстояний») $D(\mathbf{p}(\alpha), \mathbf{q})$ между эмпирическим \mathbf{q} и теоретическим распределениями $\mathbf{p}(\alpha)$. В качестве оценки параметра α распределения $\mathbf{p}(\alpha)$ выбирается значение α^* , определяемое из соотношения $\alpha^* = \arg \min_{\alpha \in \mathbf{Dom}_{AF}} D(\mathbf{p}(\alpha), \mathbf{q})$.

Для анализа выбора типа расстояния при оценке параметра α ассоциативной функции, используемой для построения распределения к.с.м. рекуррентным методом, проводилось имитационное компьютерное моделирование при следующих условиях.

1. Выборки значений к.с.м. генерировались для заданного теоретического распределения случайного множества, построенного однопараметрической ассоциативной функцией. Объемы n выборок изменялись в диапазоне от n_{begin} до n_{finish} с шагом равным h . Мощность множества U , вид ассоциативной функции с параметром α , начало диапазона n_{begin} , конец диапазона n_{finish} , шаг h для каждого эксперимента задается пользователем.

2. Для каждой сформированной выборки численным методом находилось значение параметра $\alpha^* = \arg \min_{\alpha \in \mathbf{Dom}_{AF}} D(\mathbf{p}(\alpha), \mathbf{q})$.

3. Для каждого значения n из заданного диапазона проводилась серия выборок объемом M . Для каждой выборки в серии проводился расчет квадрата отклонения (ошибки) оценок α^* от истинного значения α : $S_j^2 = (\alpha_{n,j}^* - \alpha)^2$, где $\alpha_{n,j}^*$ – оценка параметра α при минимизации j -го типа расстояния.

4. В качестве критерия будем использовать оценку среднеквадратичной ошибки (СКО), вычисленную по серии M выборок: $Q_j = \frac{1}{M} \sum_{i=1}^M S_{j,i}^2$, где Q_j – оценка СКО для j -го типа расстояния.

В качестве примера рассмотрим следующий эксперимент для $|U| = 4$. Цель эксперимента – оценка и анализ СКО по параметру α для расстояния Кульбака – Лейблера и расстояния Евклида [7, 9]. Выборки генерировались для заданного теоретического распределения к.с.м., построенного однопараметрической ассоциативной функцией Франка с параметром $\alpha = 3$. Объемы выборок изменялись в диапазоне от 100 до 3900 с шагом равным $h = 200$. В каждой серии было $M = 30$ выборок. На рис. 1 по горизонтальной оси указан объём выборки n , а по вертикальной оси – величины Q_{KL} и Q_{Euc} оценок СКО для расстояния Кульбака – Лейблера и расстояния Евклида соответственно [9].

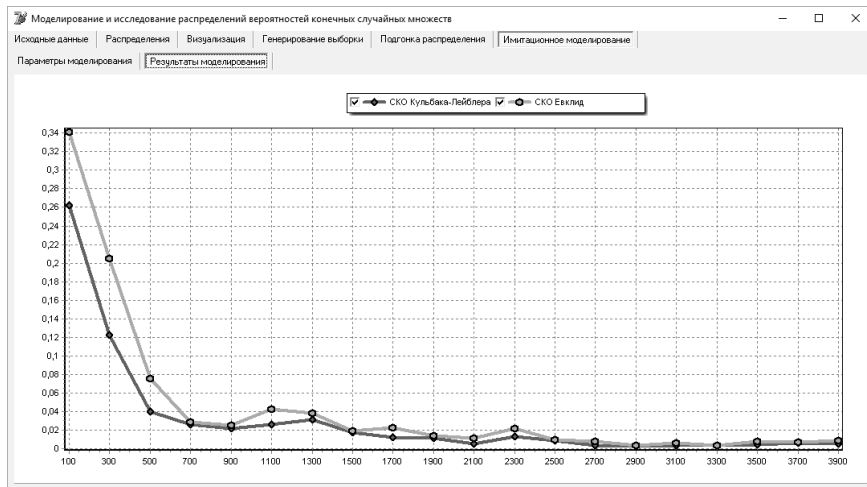


Рис. 1. Результаты имитационного моделирования: оценка СКО для расстояния Кульбака – Лейблера и расстояния Евклида

Проведенный численный анализ результатов для этих двух типов мер близости между распределениями позволил сделать выводы об области применимости расстояний: из рис. 1 видно, что при объеме выборки более 2500 СКО Q_{KL} и Q_{Euc} практически совпадают. Выбор того или иного типа расстояния в каждом конкретном случае осуществляется в зависимости от особенностей решаемой задачи случайно-множественного моделирования статистических систем.

Выполнение имитационного моделирования и численного анализа распределений к.с.м. проводились с помощью разработанного комплекса проблемно-ориентированных программ «Моделирование и исследование распределений вероятностей конечных случайных множеств» [7, 9].

ЛИТЕРАТУРА

1. *Molchanov I.* The Theory of Random Sets. – New York: Springer, 2011. – 488 p.
2. *Nguyen H.T.* An Introduction to Random Sets. – Taylor & Francis Group, LLC, 2006. – 240 p.
3. Лукьянова Н.А., Семенова Д.В. Краткий обзор по теории случайных множеств // Информационные технологии и математическое моделирование (ИТММ–2016): материалы XV Международной конференции им. А.Ф. Терпугова. Катунь, 12–16 сентября 2016 г. – Томск, 2016. – Ч. 1. – С. 172–178.
4. *Lukyanova N.A., Semenova D.V.* The study of discrete probabilistic distributions of random sets of events using associative function // J. Siberian Federal University. Mathematics & Physics. – 2014. – № 7 (4). – P. 500–514.
5. Лукьянова Н.А., Семенова Д.В. Ассоциативные функции Франка в построении семейств дискретных вероятностных распределений случайных множеств событий // Прикладная дискретная математика. – 2016. – № 2 (32). – С. 5–19.
6. Семенова Д.В., Лукьянова Н.А. Рекуррентное построение дискретных вероятностных распределений случайных множеств событий // Прикладная дискретная математика. 2014. – № 4(26). – С. 47 – 58.
7. Лукьянова Н.А., Семенова Д.В. Комплекс программ «Вероятностные распределения случайных конечных множеств событий» [Электронный ресурс]. – Объединенный фонд электронных ресурсов «Наука и образование». – Свид. о рег. № 21892, дата рег.: 01.06.2016 г.
8. Деза Е.И., Деза М.-М. Энциклопедический словарь расстояний / пер. с англ. В.И. Сычева. – М.: Наука, 2008. – 444 с.
9. Лукьянова Н.А., Семенова Д.В. Программа для ЭВМ "Fitting distribution of finite random set" // RU 2017613345, 15.03.2017, Федеральная служба по интеллектуальной собственности.

СТАТИСТИЧЕСКИЙ АНАЛИЗ ЦЕН АКЦИЙ ОПЕРАТОРОВ СВЯЗИ: МТС, МЕГАФОН, ВЫМПЕЛКОМ

С.А. Сеченова

Томский государственный университет, г. Томск, Россия

В современных условиях управленческие решения должны приниматься лишь на основе тщательного анализа имеющейся информации. Для решения подобных задач предназначен аппарат прикладной статистики, составной частью которого являются статистические методы прогнозирования. Эти методы позволяют выявить закономерности на фоне случайностей, сделать обоснованные прогнозы и выявить вероятность их выполнения.

Напомним, что под временным рядом понимается последовательность наблюдений некоторого признака (случайной величины) Y в последовательные моменты времени [1].

В рамках настоящей статьи проанализируем цены закрытия акций операторов связи: МТС, Вымпелком, Мегафон, на основе данных, взятых с сайта finam.ru за период с 1.09.2015 по 1.02.2017[2].

В пакете Statistica была проведена предварительная обработка выборок с целью определения вероятностной природы и числовых характеристик наблюдаемых величин.

Числовые характеристики цен закрытия акций операторов связи

Variable	Valid N	Mean	Median	Mode	Minimum	Maximum	Variance	Std. Dev.	Kurtosis
МТС	194	241,14	246,93	254,7	158,5	298	723,53	26,90	0,857
Вымпелком	194	98,86	99,19	99	93	109,99	2,17	1,47	19,236
Мегафон	194	950,33	959,5	1012	771	1085	5530,61	74,37	-0,979

В таблице мы видим, что компания Мегафон обладает самой высокой средней ценой на свои акции. Но цены на их акции крайне не стабильны и сильно варьируются по сравнению с ценами на акции МТС и Вымпелком. Самыми дешевыми и стабильными являются акции Вымпелком.

В программе RStudio, с помощью критерия Харка – Бера была отвергнута гипотеза о том, что функция распределения каждого временного ряда принадлежит нормальному закону распределения, так как значение p -value для каждого случая оказалось меньше уровня значимости $\alpha = 0,05$.

Посмотрим на траектории временных рядов:

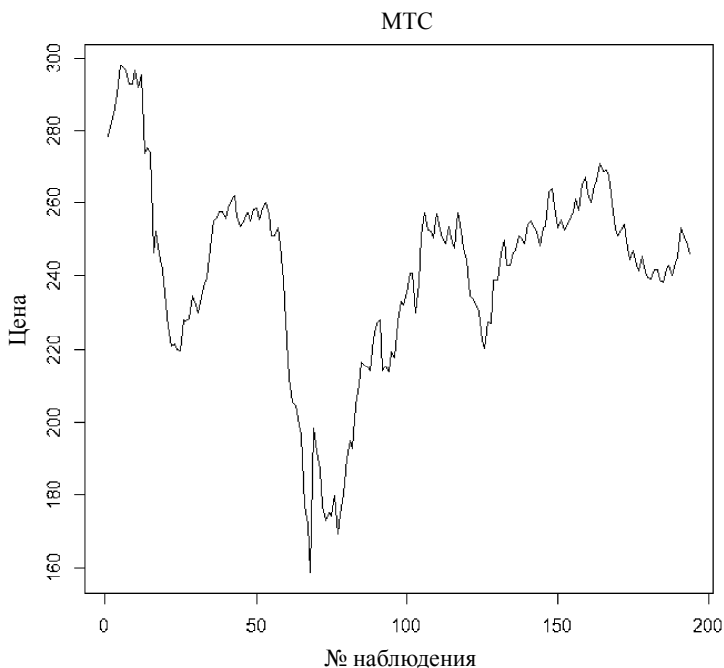


Рис. 1. Траектория цен акции компании МТС

Из рис. 1 видно, что к концу 2015 года цены на акции компании МТС резко упали, что было вызвано тем, что компания решила снизить цены на свои услуги для расширения своей доли рынка. На фоне этого компания получила преимущество перед конкурентами и с начала 2016 г. цены вновь начали расти и продолжили свой рост в 2017 г. Во многом росту выручки способствовала реализация 3-Д (Данные&Дифференциация&Дивиденды) стратегии компании в 2016 г. Кроме того, МТС выплачивает неплохие дивиденды, что делает акции компании привлекательными для вложения [3].

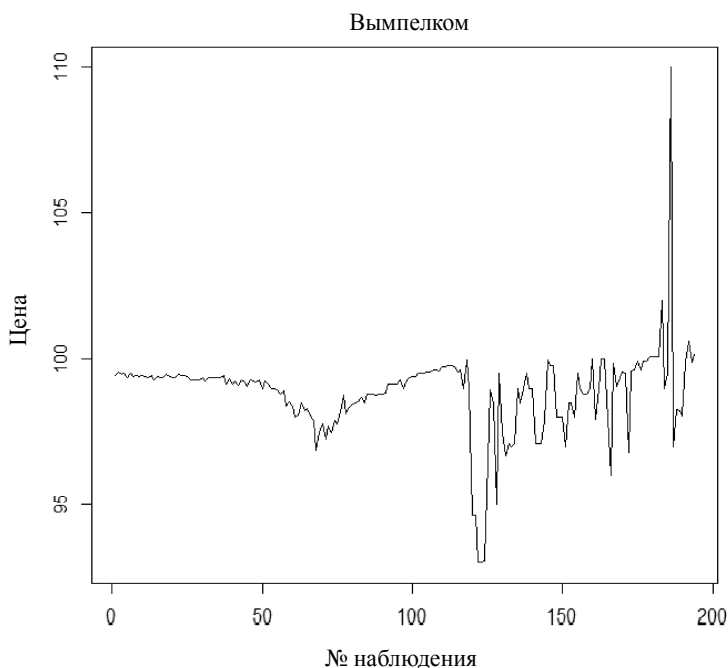


Рис. 2. Траектория цен акции компании Вымпелком

У компании Вымпелком в сентябре 2016 г. наблюдается резкий обвал цен акций (рис. 2). Это произошло на фоне отставки главы российской компании «Вымпелком» Михаила Слободина. На сегодняшний день цены акций вернулись к уровню 2015 года. Считается, что акции Вымпелкома недооценены. А ведь они представляют собой удачный вид инвестирования, как на короткий, так и на долгий срок.

Цены же акций компании Мегафон (рис. 3) в целом значительно снизились с 2015 года. Это связано с потерей крупных государственных контрактов, в выполнении которых может участвовать компания. Кроме того, компания не выдерживает конкуренции, которая присутствует сегодня на рынке операторов связи и не принимает должных управленческих решений [5, 6].

Исходя из траекторий временных рядов, можно также сделать предположение об отсутствии ярко выраженных трендов у всех выборов.

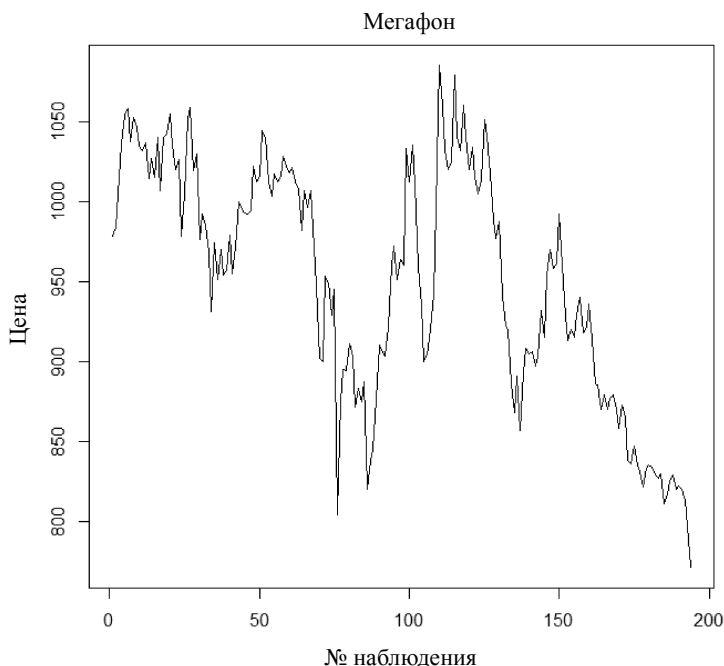


Рис. 3. Траектория цен акции компании Мегафон

Проверим, являются ли наши временные ряды однородными с помощью критерия Манна – Уитни. По умолчанию критерий проверяет гипотезу о том, что распределения двух выборок отличаются лишь постоянным и известным сдвигом, который, в свою очередь, по умолчанию равен нулю. Задать его можно при помощи параметра m . Но так как заранее нам неизвестно значение этого параметра, можно с помощью параметра `conf.int = TRUE` определить доверительный интервал для этого значения. И затем, выбрав значение данного параметра из доверительного интервала, снова применить тест. Таким образом, применив данную процедуру, мы получили, что гипотезы о том, что одно распределение отличается от другого некоторым сдвигом не были отвергнуты на уровне значимости $\alpha = 0,05$, т.е. все наши временные ряды являются однородными сами по себе и между собой. Это означает, что на цены акций всех компаний оказывают влияние одни и те же факторы.

Далее проверим стационарность наших временных рядов с помощью критерия Дикки – Фулера. В качестве нулевой гипотезы рассматривается наличие единичного корня, т.е. нестационарность ряда. Тест ADF является односторонним: в качестве альтернативной гипотезы по умолчанию считается гипотеза о стационарности ряда. В результате проверки было выявлено, что временной ряд Вымпелком является стационарным, остальные временные ряды нестационарные.

Стационарный процесс является наиболее привлекательным с точки зрения построения прогнозов, но, к сожалению, в природе таких процессов почти не существует, любой реальный процесс по мере своего развития подвержен изменениям. Но нам повезло и наш временной ряд, описывающий цены акций компании Вымпелком, оказался стационарным. Таким образом, среднее значение цены и ее дисперсия не зависят от времени.

Построим $ARMA(p,q)$ -модель. Для выбора параметров q и p построим графики автокорреляционной (ACF) и частной автокорреляционной функции (PACF):

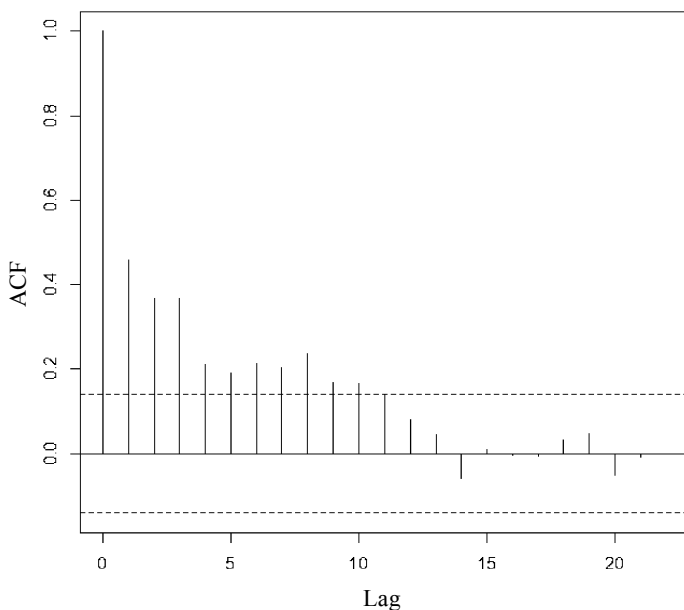


Рис. 4. Автокорреляционная функция Вымпелком

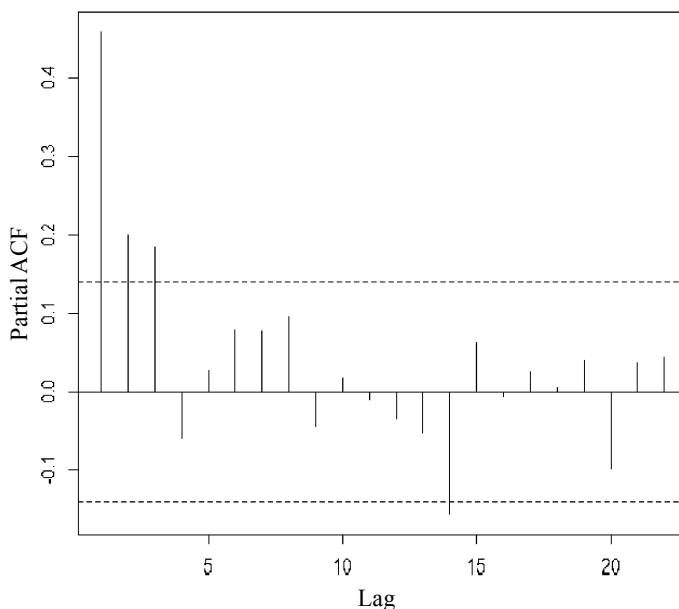


Рис. 5. Частная автокорреляционная функция Вымпелком

Выбрав пороговое значение на рис. 4 равное 0,2, получим значения параметра $q = 0, \dots, 4$. Из рис. 5 предположим значение параметра p равное $0, \dots, 3$. Таким образом, имеем семейство моделей, изменяя параметры p и q .

Для сравнения качества моделей между собой применим информационные критерии Акайке (AIC) [2]. Выберем модель с параметрами $p = 1$ и $q = 1$, для которой значение AIC минимально.

Оценим параметры модели и наша модель примет вид

$$y_t = 98,9 + 0,85y_{t-1} - 0,53\varepsilon_{t-1} + \varepsilon_t.$$

По остаткам модели осуществим тестирование качества построенной модели. Для начала проверим гипотезу о некоррелированности остатков модели, используя критерий Бокса – Льюнга. Тест Бокса – Льюнга позволяет проверить гипотезу о равенстве коэффициента автокорреляции нулю. Гипотеза была принята на уровне значимости $\alpha = 0,05$. Следовательно, ошибки некоррелированы.

Проверка на нормальность показала, что ошибки не подчиняются нормальному закону распределения.

Возьмем наши данные с 1.09.2015 по 29.01.2017. Сделаем прогноз на 3 шага, т.е. определим возможные цены акции на 30, 31 января и 1 февраля и сравним их с реальными данными. Получили, что цены предположительно будут равны 99,64, 99,53, 99,43. Реальный уровень цен: 100,62, 99,9, 100,15. Полученные значения цен акций не сильно отличаются от реальных данных.

На рис. 6 видны прогнозные значения и доверительные интервалы для прогноза с уровнями значимости 20 и 5 %.

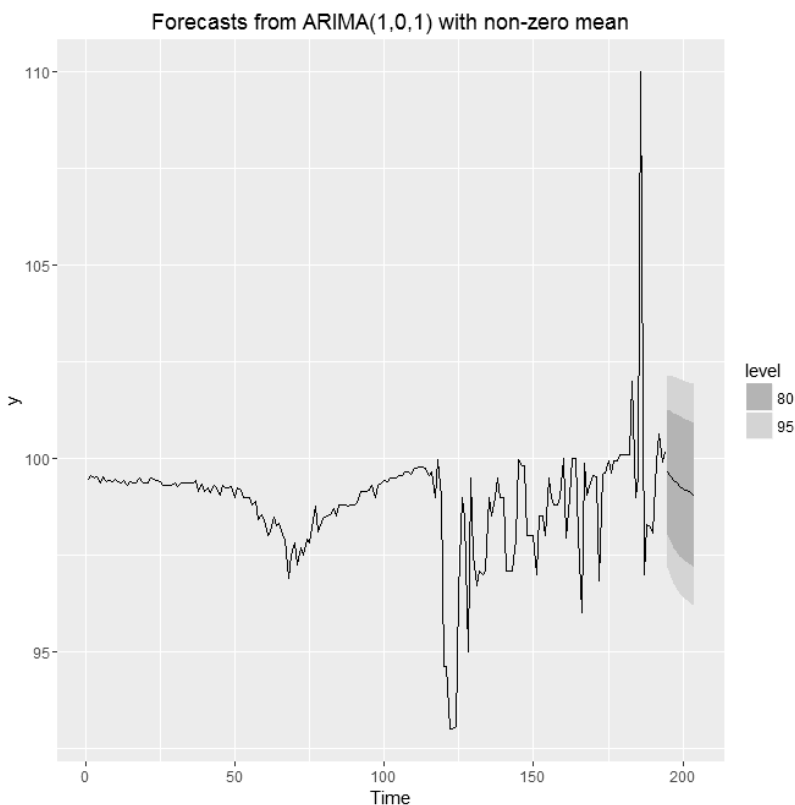


Рис. 6. Цены акций и прогноз

В процессе прогнозирования возникает задача оценки точности прогноза. Оценка точности важна и при сравнении между собой различных моделей или при определении состоятельности полученного прогноза. Известно большое количество оценок, определяющих точность прогнозирования, но для вычисления любой из них необходимо знать на каждом шаге ошибку прогнозирования.

Наиболее распространенной оценкой точности прогнозирования является среднее значение квадратов ошибок (MeanSquaredError). В нашей модели $MSE = 0,54$. Следовательно, можно сделать вывод о достаточно стабильном прогнозе.

В заключение хотелось бы отметить, что прогнозирование временных рядов в ближайшее время станет наиболее популярной областью для исследования, поскольку оно может приносить высокую материальную прибыль. Но самое трудное – это выбор хорошей модели для качественного прогноза.

ЛИТЕРАТУРА

1. *Афанасьев В.Н., Юзбашев М.М.* Анализ временных рядов и прогнозирование. – М.: Финансы и статистика, Инфра-М, 2010.
2. *Бидюк П.И., Зворыгина Т.Ф.* Структурный анализ методик построения регрессионных моделей по временным рядам наблюдений [Электронный ресурс]. – URL: <http://www.finam.ru> (дата обращения 01.06.2017).
3. *Официальный сайт АО «Инвестиционный холдинг ФИНАМ»* [Электронный ресурс]. – URL: <http://www.gmdh.net/articles/usim/Bidyuk.pdf> (дата обращения 05.09.2016).
4. «МТС представила стратегию развития 3D: данные, дифференциация, дивиденды» [Электронный ресурс]. – URL: <https://roem.ru/05-02-2014/112614/mts-predstavila-strategiyu-razvitiya-3d-dannye-differenciaciya-dividendy/> (дата обращения 02.06.2017).
5. *Дивиденды и перспективы: «Перспективы акций компании Мегафон в 2016–2019 гг.»* [Электронный ресурс]. – URL: <http://dividendo.ru/104-perspektivy-akciy-kompanii-megafon-v-2016-2019-gg.html> (дата обращения 03.06.2017).

НЕВЕРИЕ В БУДУЩЕЕ И ИНВЕСТИЦИИ В КАПИТАЛ ЗДОРОВЬЯ

О.В. Суворова, И.В. Розмаинский

*Национальный исследовательский университет
Высшая школа экономики, г. Санкт-Петербург, Россия*

Теория человеческого капитала стала одним из перспективных направлений развития экономической науки девятнадцатого века. Эффективное развитие экономики в постиндустриальном обществе подразумевает усовершенствованный подход к использованию человеческого потенциала. Увеличение инновационного производства в современном мире приводит к возрастанию роли человеческого фактора. Именно он становится самым ценным активом в конкурентной борьбе.

Резкое сокращение государственных вложений в человеческий капитал и отсутствие стимулов для частных инвестиций являются особенностью постсоветской жизни в России. Это способствует негативным воздействиям на формирование человеческого капитала. Здоровье как базовый актив человеческого капитала находится в наиболее затруднительной ситуации.

Известно, что в современной литературе термин «капитал здоровья» малоизучен, так как трактуется как составляющая человеческого капитала. Вследствие этого исследования в области накопления здоровья в России практически отсутствуют. Поэтому в работе мы будем придерживаться мнения М. Гроссмана о том, что капитал здоровья и человеческий капитал нужно рассматривать отдельно [1].

На протяжении истории люди стремились улучшить качество жизни не только в текущий момент времени, но и в будущем. Это является причиной использования здоровья как инвестиционного блага. Индекс развития человеческого потенциала (ИРЧП) является основной характеристикой для оценки качества жизни. Рассматривается средняя ожидаемая продолжительность жизни как одно из направлений построения ИРЧП.

Толчком к написанию статьи выступил тот факт, что социальная жизнь постсоветской России отличается чрезвычайно низкой продолжительностью жизни населения. Стоит заметить, что в последние годы

наблюдается тенденция к ее увеличению. В 2015 г. показатель достиг отметки в 70 лет, тем самым побив рекорд за всю историю страны, включая советское время. Но несмотря на такое увеличение, Россия всё ещё отстает не только от стран Запада, но и от некоторых стран третьего мира, таких, как Бангладеш или Иран. А по величине разницы в продолжительности жизни между женщинами и мужчинами Россия находится на первом месте. В 2015 г. эта разница составила 11,6 лет.

По нашему мнению, главной проблемой ухудшения показателей продолжительности жизни в России являются негативные инвестиции в капитал здоровья, которые порождаются таким явлением как «неверие в будущее», подразумевающее под собой настроение безысходности, пессимизм и чрезмерное беспокойство о будущем и выражающееся увеличением спроса на алкогольную продукцию и сигареты. Такие инвестиции приводят к тому, что люди, которые не верят в будущее и сосредоточены исключительно на краткосрочных задачах, «проедают» капитал здоровья [2].

Теории накопления капитала здоровья

Неоклассическая теория М. Гроссмана считается господствующей теорией в области накопления капитала здоровья. По мнению основоположника теории, запас капитала здоровья определяет время, в течение которого индивид может использовать свой человеческий капитал для получения как финансовых, так и нефинансовых выгод [1]. А объем инвестиций определяется как соотношение предельной производительности капитала и его стоимости.

Известно, что одной из предпосылок неоклассической экономической теории является гипотеза полной информации, которая означает, что экономические агенты осуществляют инвестиции, уже зная будущие выгоды и издержки от них. Но в реальной жизни агенты не знают будущего, поэтому они не могут просчитать, в течение какого времени будут поступать финансовые выгоды и какова будет их величина. Вследствие этого предельная производительность капитала не может выступать как фактор, определяющий объем вложений в него. Таким образом, главной проблемой неоклассического подхода к анализу инвестиций в капитал является пренебрежение фактором неопределенности.

На основе вышеизложенного приходим к выводу о том, что для определения инвестиций в капитал здоровья необходимо применять альтернативный подход, при котором капитал здоровья может рассматри-

ваться как актив длительного пользования. Вложения в него связаны с затратами денег и времени, и отдача от таких вложений выражается в виде большего количества времени для использования человеческого капитала.

Выбор активов длительного пользования также связан с неопределенностью, поэтому будущие выгоды и издержки не всегда могут быть рассчитаны при помощи теории вероятности. И чем выше степень неопределенности, тем больше при принятии решений индивид опирается не на знания, а на привычки, мнение большинства, а также социальные правила и нормы. При высокой степени неопределенности выбор падает на активы с большей ликвидностью и более краткосрочной окупаемостью. И так как вложения в здоровье приносят отдачу лишь в отдаленном будущем, то в результате такого выбора активов наблюдается сокращение инвестиций в капитал здоровья.

Одной из причин «проедания» капитала здоровья в постсоветской России является такое явление, как инвестиционная близорукость, означающая уменьшенный интервал планирования и прогнозирования, то есть исключение агентом из рассмотрения будущих результатов с некоторого будущего периода времени [3]. Здесь и далее инвестиционную близорукость будем использовать как синоним понятия «неверие в будущее».

Таким образом, приходим к выводу о том, что люди не заботятся о своем здоровье в условиях неверия в будущее, поскольку отдача от вложений в здоровье начинает появляться только в очень отдаленный период времени.

Построение эконометрической модели

Основная предлагаемая гипотеза заключается в том, что большие расходы на потребление алкоголя и курение, а также низкие «расходы» на занятия спортом обусловлены «неверием в будущее».

Эмпирической базой исследования служат данные, собранные путем проведения опроса в социальной сети, составленного на основе вопросов из базы RLMS.

Собственная база данных была создана для сравнения результатов относительно величины города и региона проживания, в частности относительно городов Западной России (Санкт-Петербург и Москва) и Томска. В выборке представлены индивиды в возрасте от 15 до 35 лет. Причиной этому служит тот факт, что у молодежи ещё не полностью

сформированы привычки к негативным инвестициям, поэтому легче увидеть динамику и тенденцию к увеличению или уменьшению инвестиций в капитал здоровья.

В качестве прокси для зависимой переменной в эконометрических моделях выделяют переменные, отражающие факт и частоту потребления алкоголя в течение последних 30 дней, факт потребления табака – в течение последних 7 дней, периодичность занятий спортом, а также частоту посещения врача – в течение года.

В качестве прокси для объясняющей переменной выделяются переменные, показывающие уровень жизнерадостности индивида, а также уровень удовлетворенности жизнью в целом в настоящее время. Эти переменные отражают желание индивида уйти от психологических и социальных проблем с помощью алкоголя. Для контролирования зависимости такой характеристики, как посещение врача, используется самооценка здоровья, которая имеет высокую степень корреляции с реальным состоянием здоровья индивида. Желание индивида снять напряжение и снизить ежедневный стресс с помощью курения контролируется переменной, которая отражает уровень беспокойства индивида о будущем.

В качестве контролирующих переменных выделяются переменные образования, семейного положения и материального положения семьи.

С помощью эконометрического аппарата построена эмпирическая модель для оценки зависимых переменных

$$Y_i = \beta_1 X_i + \beta_2 A_i + \varepsilon_i,$$

где Y_i – характеристика одной из зависимых переменных для индивида i , X_i – вектор социально-экономических характеристик человека (пол, возраст, уровень образования, семейное положение, род занятий и величина города проживания), A_i – вектор объясняющих переменных, $\beta_{1,2}$ – соответствующие коэффициенты, а ε_i – случайная ошибка.

В исследовании присутствуют две бинарные зависимые переменные, отражающие факт потребления алкоголя и табака. Стандартная методология для оценивания модели бинарного выбора предполагает построение logit- и/или probit-регрессии.

Для дальнейшего исследования необходимо было определиться, какая из моделей будет использована. С этой целью были вычислены информационные критерии Байеса (BIC) и Акаике (AIC).

В результате в соответствии с обоими критериями для переменной, отражающей факт потребления табака, целесообразно использовать logit-модель, а для переменной, отражающей факт потребления алкоголя, наоборот – probit-модель. Но поскольку нам важно продемонстрировать все модели, используя одну разновидность регрессии, было решено использовать probit-модель.

Остальные зависимые переменные в исследовании являются категориальными и упорядоченными. Для оценивания модели упорядоченного выбора также предполагается построение упорядоченной logit- и/или probit-регрессии. Было принято решение об использовании probit-модели.

В отличие от классической регрессионной модели, коэффициенты при параметрах модели плохо интерпретируемы. Для моделей бинарного выбора в дополнение к коэффициентам рассчитывают предельные эффекты регрессоров на ожидаемое значение объясняемой переменной. Предельный эффект не является константой, а зависит от других факторов и отвечает на вопрос о том, как изменится вероятность того, что $Y_i = 1$ с ростом регрессора на единицу. Обычно предельные эффекты рассчитываются для среднего значения каждой объясняющей переменной.

Заключение

В результате проведенного анализа были получены следующие выводы.

Снижение степени удовлетворенности жизнью оказывает влияние на частоту потребления алкогольных напитков, увеличивая ее. Также она оказывает влияние на регулярность занятий спортом, которая уменьшается с уменьшением степени удовлетворенности.

Уровень жизнерадостности индивида оказался сильно значимым при анализе зависимости потребления табачных изделий, с его уменьшением вероятность курения индивидом возрастает. Также на потребление табака влияет степень беспокойства индивидом о будущем, и чем она больше, тем выше спрос на курение.

Самооценка здоровья индивидом оказалась значимым регрессором при анализе зависимости потребления алкоголя, с ее увеличением увеличивается как вероятность, так и частота потребления алкоголя. Также она оказывает влияние на частоту посещения индивидом специалистов,

и чем выше самооценка здоровья индивидом, тем меньше вероятность посещения специалиста индивидом.

Увеличение величины города влечет за собой снижение потребления сигарет, при этом жители крупных городов потребляют алкоголь меньше, чем жители других городов.

Жители Европейской части России потребляют сигареты и алкоголь меньше, чем жители других регионов. При этом жители Томска потребляют сигареты и алкоголь меньше, чем жители Москвы и Санкт-Петербурга.

Таким образом, выдвинутая в работе гипотеза о влиянии «неверия в будущее» на инвестиции в капитал здоровья, по крайней мере, не отвергается.

В своей статье И. Розмаинский предлагает способы решения проблем в сфере здоровья и увеличения средней продолжительности жизни в нашей стране [2]. По его мнению, необходимо стимулировать у россиян нормальную, а не заниженную оценку будущего времени. Формирование институтов, помогающих укрепить уверенность в завтрашнем дне, будет способствовать борьбе с инвестиционной близорукостью. А развитие институтов гражданского общества поможет обеспечить контроль общества над властью.

Конечно, вышеперечисленные способы помогут решить часть проблем в сфере здравоохранения, но не стоит забывать и о том, что в большинстве случаев именно отдельные частные лица принимают решения о характере инвестиций в капитал здоровья. Именно поэтому необходимо просвещать население в сфере здравоохранения и пропагандировать здоровый образ жизни, бороться с высоким потреблением алкоголя и табака.

ЛИТЕРАТУРА

1. *Grossman M.* On the concept of health capital and the demand for health // *J. Political Economy*. – 1972. – V. 80. – No. 2. P. 223–255.
2. *Розмаинский И.* Почему капитал здоровья накапливается в развитых странах и «проедается» в постсоветской России? // *Вопросы экономики*. – 2011. – № 10. – С. 113–131.
3. *Розмаинский И.* «Инвестиционная близорукость» в посткейнсианской теории и в российской экономике // *Вопросы экономики*. – 2006. – № 9.

СОДЕРЖАНИЕ

МОДЕЛИ И МЕТОДЫ МАССОВОГО ОБСЛУЖИВАНИЯ

Fedorova E. Asymptotic analysis of retrial queue $M/GI/1$ with impatient calls under heavy load condition.....	3
Lisovskaya E. Total customers' capacity in the multi server queues.....	10
Mikheev P.A., Pichugina A.A., Suschenko S.P., Tkachev R.V. Optimization of pipelining and data processing.....	18
Rykov V.V. Controllable Queueing Systems from the very beginning up to nowadays	25
Багирова С.А. Об одной модели системы управления запасами с мгновенным обслуживанием заявок	27
Бородина А.В., Ефросинин Д.В., Морозов Е.В. Применение метода расщепления для имитационного моделирования управляемой деградирующей системы.....	33
Велиджанова Г.М. Алгоритм расчета одной схемы разделения каналов в сетях с широкополосными вызовами данных	37
Галилейская А.А., Лисовская Е.Ю. Исследование двухфазной бесконечнолинейной СМО с ММРР-входящим потоком требований случайного объема	43
Даммер Д.Д., Назаров А.А. Исследование математической модели страховой компании в виде системы массового обслуживания с высокоинтенсивным потоком входящих рисков и в случайной среде	50
Китаева А.В., Лившиц К.И., Ульянова Е.С. Многопродуктовая модель быстро портящихся запасов с зависящим от цены спросом	58
Клименок В.И. Ненадежная система массового обслуживания с марковским входящим потоком и повторными вызовами.....	66
Копать Д.Я., Матальцкий М.А. Исследование в нестационарном режиме сети с положительными и отрицательными заявками различных типов	74

Кочеганов В.М., Зорин А.В. Анализ потоков первичных требований в тандеме при циклическом управлении с продлением.....	81
Лившиц К.И., Ульянова Е.С. Управление темпом производства и ценой продажи скоропортящейся продукции.....	88
Медведева Е.Г. К анализу вероятности успешного установления соединения по радиоканалу случайного доступа.....	95
Моисеев А.Н., Буртовая Д.А. Имитационное моделирование циклических систем массового обслуживания	100
Моисеева С.П., Шкленник М.А., Набокова О.О. Исследование потоков в двухфазной бесконечнолинейной СМО с повторными обращениями методом предельной декомпозиции	108
Моисеева С.П., Панкратова Е.В., Фархадов М.П. Исследование неоднородной бесконечнолинейной СМО $SM M^{(n)} \infty$	115
Морозов Е.В., Потахина Л.В. Ускоренное оценивание характеристик системы с заявками случайного объема	122
Назаров А.А., Анисимова А.А. Асимптотический анализ двухфазной RQ-системы с двумя орбитами в условии большой задержки заявок на орбитах	129
Назаров А.А., Измайлова Я.Е. Исследование RQ-системы $M/GI/1$ с г-настойчивым вытеснением заявок	136
Назаров А.А., Квач А.С. Сравнение методов остаточного и истекшего времени обслуживания для исследования замкнутой RQ-системы $M/GI/1/N$ с конфликтами заявок и ненадежным прибором.....	142
Назаров А.А., Лапатин И.Л., Пауль С.В. Исследование выходящего потока в RQ-системе $M/M/1/1$ с вызываемыми заявками	150
Перепелкин Е.А. Модальный синтез системы управления длиной очереди пакетов в буфере маршрутизатора	155
Полин Е.П., Павлова Е.А. Исследование немарковской неоднородной бесконечнолинейной системы массового обслуживания в случайной среде	160
Рачинская М.А., Федоткин М.А. Квазиоптимальное управление неординарными пуассоновскими потоками	164
Рыжиков Ю.И. Метод прогонки для многоканальных систем обслуживания	172
Тананко И.Е., Фокина Н.П. Метод моделирования информационно-вычислительной сети сетью массового обслуживания с нестационарной структурой.....	180

Трегубов В.М., Кирпичников А.П., Якимов И.М., Шакирзянов Р.М. Дискретная стохастическая оптимизация систем массового обслуживания на имитационных моделях.....	188
Федоткин М.А., Кудрявцев Е.В. Аналитические и численные методы исследования системы адаптивного управления конфликтными потоками.....	197
Якимов И.М., Кирпичников А.П., Мокшин В.В., Яхина З.Т. Сравнение систем структурного и имитационного моделирования вероятностных объектов.....	204

МОДЕЛИРОВАНИЕ ТЕЛЕКОММУНИКАЦИОННЫХ СЕТЕЙ И ТЕЛЕТРАФИКА

Blinov A.I., Vasilyev S.A., Sevastianov L.A. Using Dobrushin mean-field approach for large-scale transport networks analysis	216
Vasilyev S.A., Tsareva G.O. Dobrushin mean-field approach and large-scale networks with a small parameter	221
Vasilyeva D.G., Vasilyev S.A., Sevastianov L.A. Stability analysis of the inverted pendulum under stochastic driving forces.....	226
Huseynova G.N., Mustafayeva E.A. Hidden information transmission through WhatsApp.....	231
Yarkina N., Popovskaya N., Khalina V., Gaidamaka A., Samouylov K. Modeling end-to-end business processes of a telecom company with a BCMP queueing network.....	234
Zaripova E.R., Smirnova N.M., Alekseeva A.S. Modeling of seven client-oriented business processes using queueing network theory	242
Агеев К.А., Сопин Э.С. Анализ показателей качества трафика реального времени в беспроводной сети с помощью ресурсных СМО	248
Дараселия А.В., Сопин Э.С. О задаче оптимизации энергопотребления облачной инфраструктуры	255
Коннов И.В., Кашина О.А. Модели распределения соединений в беспроводных телекоммуникационных сетях	262
Коннов И.В., Кашуба А.Ю. Методы распределения ресурсов в сетях телекоммуникации	267
Печеный Е.А., Нуриев Н.К. Построение концептуальной модели трафика в беспроводных сетях	275

ПРИКЛАДНОЙ ВЕРОЯТНОСТНЫЙ АНАЛИЗ

Мейлиев А.Х., Тухтаев Э.Э. О применении тауберовой теоремы для производящих функций в теории ветвящихся случайных процессов	281
Лукьянова Н.А., Семенова Д.В. Визуализация метода рекуррентного построения распределений вероятностей конечных случайных множеств на дуплете событий	286
Семенова Д.В., Лукьянова Н.А. Метод минимального расстояния в подгонке распределений конечных случайных множеств	294
Сеченова С.А. Статистический анализ цен акций операторов связи: Мтс, Мегафон, Вымпелком	299
Суворова О.В., Розмаинский И.В. Неверие в будущее и инвестиции в капитал здоровья	307

ДЛЯ ЗАМЕТОК

Научное издание

**ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ
И МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ
(ИТММ-2017)**

**Материалы
XVI Международной конференции
имени А. Ф. Терпугова
29 сентября – 3 октября 2017 г.**

Часть 1

Редактор *Т.С. Портнова*

Дизайн, верстка *Д.В. Фортеса*

ООО «Издательство научно-технической литературы»
634050, Томск, пл. Новособорная, 1, тел. (3822) 533-335

Изд. лиц. ИД № 04000 от 12.02.2001. Подписано к печати 09.09.2017.
Формат 60 × 84 ¹/₁₆. Бумага офсетная. Печать офсетная. Гарнитура «Таймс».
Усл. п. л. 18,48. Уч.-изд. л. 20,70. Тираж 100 экз. Заказ № 30.
