

НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
КАРШИНСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
СИБИРСКИЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ
РОССИЙСКИЙ УНИВЕРСИТЕТ ДРУЖБЫ НАРОДОВ
ИНСТИТУТ ПРОБЛЕМ УПРАВЛЕНИЯ
им. В.А. ТРАПЕЗНИКОВА РАН
НИЖЕГОРОДСКИЙ ГОСУДАРСТВЕННЫЙ
УНИВЕРСИТЕТ им. Н.И. ЛОБАЧЕВСКОГО
БАКИНСКИЙ ИНЖЕНЕРНЫЙ УНИВЕРСИТЕТ

**ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ
И МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ
(ИТММ-2024)**

**МАТЕРИАЛЫ
XXIII Международной конференции
имени А. Ф. Терпугова
20–26 октября 2024 г.**

ТОМСК
Издательство Томского
государственного университета
2024

УДК 519
ББК 22.17
И74

Информационные технологии и математическое моделирование (ИТММ-2024): Материалы XXIII Международной конференции имени А. Ф. Терпугова (20–26 октября 2024 г.). — Томск: Издательство Томского государственного университета, 2024. — 613 с.

ISBN 978–5–907890–08–4

Сборник содержит избранные материалы XXIII Международной конференции имени А. Ф. Терпугова по следующим направлениям: теория массового обслуживания и ее приложения, прикладной вероятностный анализ, интеллектуальные системы, информационные технологии и программная инженерия, математическое и компьютерное моделирование. Также в сборник вошли материалы международного симпозиума "Современные стохастические модели и проблемы актуарной математики" (МАМОИТ-2024).

Для специалистов в области информационных технологий и математического моделирования.

УДК 519
ББК 22.17

Р е д к о л л е г и я:

С.П. Моисеева, доктор физико-математических наук, профессор
Д.В. Семенова, кандидат физико-математических наук, доцент
А.Н. Моисеев, доктор физико-математических наук, доцент
Е.А. Фёдорова, кандидат физико-математических наук
А.С. Шкуркин, кандидат технических наук, доцент
О.Д. Лизюра

ISBN 978–5–907890–08–4

© Авторы. Текст, 2024
© Томский государственный университет. Оформление.
Дизайн, 2024

NATIONAL RESEARCH TOMSK STATE UNIVERSITY

KARSHI STATE UNIVERSITY

SIBERIAN FEDERAL UNIVERSITY

PEOPLES' FRIENDSHIP UNIVERSITY OF RUSSIA

V.A. TRAPEZNIKOV INSTITUTE OF CONTROL
SCIENCES OF RUSSIAN ACADEMY OF SCIENCES

LOBACHEVSKY STATE UNIVERSITY
OF NIZHNI NOVGOROD

BAKU ENGINEERING UNIVERSITY

**INFORMATIONAL TECHNOLOGIES
AND MATHEMATICAL MODELLING
(ITMM-2024)**

PROCEEDINGS
of the 23rd International Conference
named after A. F. Terpugov
2024 October, 20–26

TOMSK
Tomsk State
University Publishing
2024

UDC 519
LBC 22.17
I60

Informational technologies and mathematical modelling (ITMM-2024):
Proceedings of the 23rd International Conference named after A. F.
Terpugov (2024 October, 20–26). — Tomsk: Tomsk State University
Publishing, 2024. — 613 p.

ISBN 978–5–907890–08–4

This volume presents selected papers from the 23rd International Conference named after A.F. Terpugov. The papers are devoted to new results in the following areas: queuing theory and its applications, applied probabilistic analysis, intelligent systems, information technology and software engineering, mathematical and computer modeling. The collection also presents the proceedings of symposium "Modern Stochastic Models and Problem of Actuarial Mathematics"(MAMMOTH-2024).

UDC 519
LBC 22.17

E d i t o r s:

S.P. Moiseeva, Doctor of Physical and Mathematical Sciences,
Professor.

D.V. Semenova, Candidate of Physical and Mathematical Sciences,
Associate Professor.

A.N. Moiseev, Doctor of Physical and Mathematical Sciences,
Associate Professor.

E.A. Fedorova, Candidate of Physical and Mathematical Sciences.

A.S. Shkurkin, Candidate of Technical Sciences, Associate Professor.

O.D. Lizyura.

ISBN 978–5–907890–08–4

© Authors. Text, 2024
© Tomsk State University
Publishing. Design, 2024

Математическая теория
телетрафика и теория
массового обслуживания

INVESTIGATION OF M/G/1//N SYSTEM WITH IMPATIENT CUSTOMERS, UNRELIABLE PRIMARY AND A BACKUP SERVER

Á. Tóth, J. Sztrik

University of Debrecen, Debrecen, Hungary

This paper investigates a finite-source retrial queueing system characterized by request collisions, primary server unreliability, and the inclusion of a backup server. In cases of collisions, when a new job arrives while the service facility is occupied, both jobs are sent to a virtual waiting area called the orbit. Customers in the orbit make further attempts to access the server after a random interval. During server breakdowns, the customer at the server is transferred to the orbit. The system consists of a backup facility when the primary server is unreachable to process requests while the main service unit is under repair. The novelty of this study lies in the implementation of the impatience of the customers and conducting a sensitivity analysis using various service time distributions for the primary customers. We examined a scenario where key performance measures are visually represented, highlighting the observed disparities. **Keywords:** *simulation, queueing system, finite-source model, sensitivity analysis, backup server, collisions, unreliable operation, impatience.*

Introduction

In the current era of increasing traffic volumes and growing user bases, analyzing communication systems and designing optimal configurations present significant challenges. Information exchange plays a crucial role in all aspects of life, making it essential to develop or adapt mathematical and simulation models for telecommunication systems to meet these evolving demands. Retrial queues are particularly effective and suitable for modeling real-world scenarios encountered in telecommunication systems, networks, mobile networks, call centers, and similar fields. Numerous scholarly works, such as those referenced in [2] and [3], have extensively investigated various aspects of retrial queueing systems characterized by retrial calls.

In certain contexts, researchers assume the perpetual availability of service units; however, operational interruptions or unexpected events may occur, leading to the rejection of incoming customers. Devices across various industries are prone to malfunctions, making the presumption of their

infallible operation overly optimistic and impractical. Similarly, in wireless communication environments, diverse factors can affect transmission rates, causing interruptions during packet delivery. The inherent unreliability of retrial queuing systems significantly impacts system functionality and performance metrics. Additionally, halting production entirely is unfeasible, as it may cause delays in order fulfillment. Therefore, during such occurrences, machines or operators with lower processing capacities may continue operating to maintain smoother functionality. Furthermore, the authors explore the feasibility of incorporating a backup server capable of providing services at a reduced rate when the primary server is unavailable. Numerous recent studies have extensively examined retrial queuing systems with unreliable servers, as demonstrated by references such as [5].

Waiting is a ubiquitous phenomenon in various aspects of life, often leading to dissatisfaction due to the time spent in queues. This dissatisfaction can result in early departures of requests from the system without being served, a behavior termed impatience. Such behavior is observed in diverse domains including healthcare applications, call centers, and telecommunication networks. The impatience mechanism is a crucial aspect of the model, as it influences the overall system performance by potentially reducing the number of customers waiting in the system and affecting the service dynamics. Studies examining these behaviors include [6].

In technological contexts such as Ethernet networks or constrained communication sessions, job collisions are likely to occur. Multiple entities within the source may initiate asynchronous attempts, leading to signal interference and necessitating retransmissions. Consequently, it is crucial to incorporate this phenomenon into research aimed at developing effective strategies to mitigate conflicts and the resultant message delays. Publications that address findings related to collisions include [7].

The objective of this study is to perform a sensitivity analysis using various service time distributions for the primary server, in order to evaluate the main performance metrics in scenarios that incorporate the feature of impatience of the customers. When the primary server fails, customer service is transferred to the backup facility. During this period, new customers are directed to the backup unit or to the orbit if the backup unit is busy. Our investigation focuses on the impact of the impatient feature, with results obtained through simulation using Simpack [4]. The simulation program is developed based on fundamental code elements that facilitate the computation of desired metrics across a range of input parameters. Graphical representations are provided to illustrate the effects of different parameters and distributions on key performance indicators.

1. System model

We investigate a finite-source retrial queueing system denoted as type $M/G/1//N$ (as depicted in Figure 6), which integrates an unreliable primary service unit, collision occurrences, and a backup service unit. This model features a finite-source, where each of the N individuals generates requests to the system following an exponential distribution with parameter λ . Arrival times adhere to an exponential distribution with a mean of $\lambda * N$. In the absence of queues, arriving jobs are serviced immediately based on a gamma, hypo-exponential, hyper-exponential, Pareto, or log-normal distribution, each characterized by distinct parameters but sharing equivalent mean and variance values (η).

In cases of server busyness, an arriving customer causes a collision with the customer currently being serviced, resulting in both customers being transferred to the orbit. Jobs in the orbit subsequently initiate further attempts to access the server after an exponentially distributed random time with parameter σ . Additionally, random breakdowns occur, with failure times modeled by exponential random variables. The failure time parameter is γ_0 when the server is occupied and γ_1 when idle.

Upon the primary service unit's failure, repair begins immediately, with the repair duration following an exponential distribution with parameter γ_2 . If the server fails while busy, the customer is promptly moved to the orbit. During the primary server's unavailability, all customers in the source continue to generate requests, which are then directed to the backup server. The backup server operates at a reduced rate, described by an exponentially distributed random variable with parameter μ , and is assumed to be reliable, functioning only when the primary server is unavailable. Incoming requests are directed to the orbit if the backup server is busy, and collisions do not occur at the backup service unit.

Each primary customer in the system is characterized by an impatience property, which reflects their potential decision to leave the system if not served within a certain time frame. This decision to abandon the system is made after a random time period, which follows an exponential distribution with rate parameter τ .

The model presumes complete independence among all random variables in its formulation.

2. Simulation results

We utilized a statistical module class equipped with an advanced statistical analysis tool to quantitatively estimate the mean and variance of observed variables via the batch mean method. This technique involves ag-

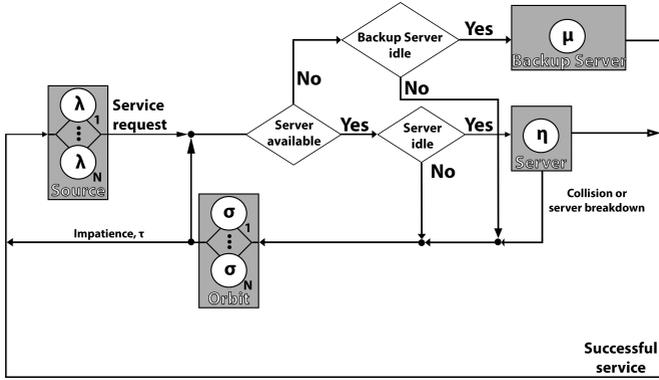


Figure 1. System model

gregating n successive observations from a steady-state simulation to generate a sequence of approximately independent samples. The batch mean method is widely recognized for its effectiveness in establishing confidence intervals for the steady-state mean of a process. To ensure the sample averages are approximately independent, large batch sizes are essential. Detailed information on the batch mean method is provided in [1]. Our simulations were conducted with a confidence level of 99.9%, and the simulation run was terminated when the relative half-width of the confidence interval reached 0.00001.

Table 1

Numerical values of model parameters

N	γ_0	γ_1	γ_2	σ	μ	τ
100	0.1	0.1	1	0.05	0.1	0.01

In this section, our goal was to establish service time parameters for each distribution such that they exhibit equivalent mean values and variances. Four distinct distributions were analyzed to evaluate their impact on performance metrics. The hyper-exponential distribution was specifically chosen to ensure a squared coefficient of variation exceeding one. The input parameters for the various distributions are detailed in Table 2, and Table 1 enumerates the values of other pertinent parameters.

Figure 2 illustrates the relationship between the mean response time of successfully served customers and the arrival intensity. Under successfully served customers we mean those customers which do not depart from the

Table 2

Parameters of service time of primary customers

Distribution	Gamma	Hyper-exponential	Pareto	Lognormal
Parameters	$\alpha = 0.011$ $\beta = 0.011$	$p = 0.494$ $\lambda_1 = 0.989$ $\lambda_2 = 1.011$	$\alpha = 2.005$ $k = 0.501$	$m = -2.257$ $\sigma = 2.125$
Mean	1			
Variance	90.25			
Squared coefficient of variation	90.25			

system earlier because of impatience. The Pareto distribution demonstrates the highest mean response time, whereas the differences among the other distributions become more pronounced. Notably, the gamma distribution results in the lowest mean response time.

An interesting observation is that as the arrival intensity increases, the mean response time initially rises but subsequently decreases after reaching a specific threshold. This phenomenon is characteristic of retrial queuing systems with a finite-source and tends to manifest under appropriate parameter configurations. In the extended version of our paper, we plan to present results utilizing an alternative parameter configuration where the squared coefficient of variation is less than one and focusing more on how impatience has a role in modifying the performance measures.

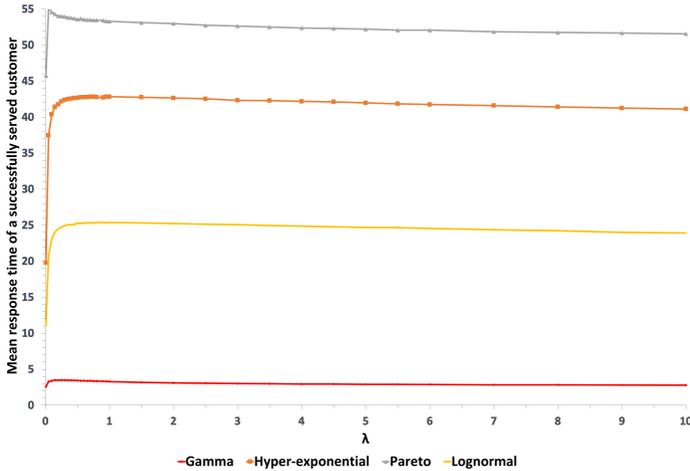


Figure 2. Mean response time vs. arrival intensity

3. Conclusion

We conducted simulations of a retrial queuing system based on the $M/G/1//N$ model, which incorporates an unreliable primary server and a backup service unit. Our program facilitated a sensitivity analysis on various performance metrics, including the mean response time of a successfully served customer. Notably, when the squared coefficient of variation exceeds one, significant deviations are observed among distributions across multiple aspects of the investigated metrics. Future studies will aim to further explore the impact of server blocking, other types of customer impatience in alternative models, and conduct sensitivity analyses for other variables, such as failure rates.

REFERENCES

1. *Chen J.E., Kelton D.W.* A Procedure for Generating Batch-Means Confidence Intervals for Simulation: Checking Independence and Normality // SIMULATION. 2007. Vol. 83. № 10. P. 683–694.
2. *Dragieva V. I.* Number of retrials in a finite source retrial queue with unreliable server // Asia-Pac. J. Oper. Res. 2014. Vol. 31. № 2. P. 1–23.
3. *Fiems D., Phung-Duc T.* Light-traffic analysis of random access systems without collisions // Annals of Operations Research. 2019. Vol. 277. P. 311–327.
4. *Fishwick P.A.* SimPack: Getting Started With Simulation Programming In C And C++ // In 1992 Winter Simulation Conference, New York: Association for Computing Machinery, 1992. P. 154–162.
5. *Gharbi N., Nemmouchi B., Mokdad L., Ben-Othman J.* The impact of breakdowns disciplines and repeated attempts on performances of small cell networks // Journal of Computational Science. 2014. Vol. 5. № 4. P. 633–644.
6. *Kumar R., Jain N., Mokdad L., Som B.* Optimization of an $M/M/1/N$ Feedback Queue with Retention of Reneged Customers // Operations Research and Decisions. 2014. Vol. 24. P. 45–58.
7. *Nazarov A., Kvach A., Yampolsky V.* Asymptotic Analysis of Closed Markov Retrial Queuing System with Collision // Information Technologies and Mathematical Modelling: 13th International Scientific Conference, ITMM 2014, Anzhero-Sudzhensk: Springer International Publishing, 2014. P. 334–341.

Ádám Tóth — PhD, assistant professor, Department of Informatics Systems and Networks. E-mail: toth.adam@inf.unideb.hu

János Sztrik — Doctor of the Hungarian Academy of Sciences, full professor, Department of Informatics Systems and Networks. E-mail: sztrik.janos@inf.unideb.hu

TANDEM OF A SELF-SERVICE QUEUEING SYSTEM AND A FINITE CAPACITY QUEUEING SYSTEM PROCESSING IN WHICH IS DESCRIBED BY THE MARKOV CHAIN

A. Dudin, O. Dudina

Belarusian State University, Minsk, Belarus

A dual tandem queueing system with an infinite-server first stage is considered. Arrivals are defined by the Markov arrival process, the service time is exponentially distributed. After receiving service at the first stage, a customer either departs from the system permanently or transits to the second stage. Processing of a customer at the second stage is described by a system having a finite capacity. Service of customers at the second stage is defined by an arbitrary Markov process. If the second stage is full at the service completion moment of a customer at the first stage, this customer is lost or repeats service at the first stage. The dynamics of the system is described by the four-dimensional Markov chain. This chain belongs to the class of asymptotically quasi-Toeplitz Markov chains. This allows to implement the stationary analysis of tandem states and compute the main performance measures of the tandem.

Keywords: *tandem queue, infinite-server system, system of a finite capacity.*

Introduction

Tandem queueing systems are the simplest case of queueing networks and suit for modelling various real-world systems in telecommunication, transportation, manufacturing, etc. industries. Therefore, they are the subject of intensive research. Here, we consider a specific tandem system. The first stage of the system is described by the infinite server system (similar to $MAP/M/\infty$ type system). Infinite-server queues are one of the simplest kinds of queueing systems, for which sometimes it is possible to obtain even the transient distribution of the system states (see, e.g., [1]). Motivation for consideration of tandems with the infinite-server first stage lies in the wide distribution of such kinds of tandems in real life. Service at the first stage may be interpreted as a preliminary self-service of a customer before receiving service at some service facility. The term "self-service" was used in this context, e.g., in the paper [2]. Self-service can mean the preliminary preparation of a customer to service via an Internet-based facility, such as the

preparation of some files, wish lists, or pre-checks. In medical applications, self-service can be interpreted, e.g., as a diet compliance or a provision of emergency medical care outside of a hospital. Service at the second stage is the essential service of a customer. It is natural that the capacity of any real service system is finite. Thus, here we make an assumption that the second stage of the tandem is described by a quite arbitrary queueing system of a finite capacity. Only a restricted number of customers can present in the system (obtain service or wait in the finite buffer or the orbit). We omit details of the service process at the second stage. It can be a single- or multi-server system with a finite buffer or no buffer, a system with some kind of limited processor-sharing service discipline, a polling system with finite buffers, or a hybrid system. We assume only that a customer processing at the second stage is described by a two-dimensional continuous-time Markov chain with the known matrices of the chain transition intensities that do not lead or lead to departure of an arbitrary customer from the second stage and its transition probabilities at the moments of a new customer transfer from the first stage.

1. Model description

The Markovian arrival process (*MAP*) arriving at the first stage of the tandem is defined by the irreducible underlying Markov chain ν_t having a finite state space $\{1, 2, \dots, W\}$ and matrices D_0 and D_1 defining the transition rates of this chain without and with the generation of new customers. The average arrival rate is denoted by λ . More details about the *MAP* can be found, e.g., in the paper [3] and books [4], [5], and [6].

The first stage contains an unlimited number of independent, identical servers. All arriving customers are admitted to the system and immediately begin service. The service time of an arbitrary customer at an arbitrary server has the exponential distribution with the rate μ .

After the end of service at the first stage, with the probability p , the customer transits for service at the second stage or, with the complementary probability, the customer finishes service and departs from the system. For example, after the search of information about some entities from the database in some available indices, the user may discover that he/she already found the required information and does not need to be processed using the tables of the relational database.

The second stage of a tandem has a finite capacity. The integer limit N is fixed. A customer, which finishes service at the first stage when the number of customers at the second stage is less than N , transits to the second stage. If this number is equal to N , the customer cannot enter the second stage. With the probability q , this customer is lost. With the

complementary probability, the customer returns for the repeated service at the first stage. Duration of the repeated services also has the exponential distribution with the rate μ .

For generality of the model, we do not suppose any concrete detailed mechanism of operation of the second stage. We assume that all customers are processed at this stage independently of each other. The operation of the second stage is completely described by the two-dimensional Markov chain. One of the components of this chain is the number, n_t , of customers at the second stage. The second component, η_t , defines the auxiliary processes, such the process $\{n_t, \eta_t\}$ is the Markovian one. Actually, the component η_t itself can consist of a finite number of sub-components. For example, if service times of customers at the second stage have a phase-type distribution (for definition, see e.g., [4] and [5]), these sub-components may define the current phase of service of each customer or the number of servers providing service at each phase.

Customer's processing at the second stage is characterized by the three sets of matrices. Under the fixed number n of customers processing at the second stage (receiving service or waiting in a buffer), $n \in \{0, 1, \dots, N\}$, the process η_t admits values from the set of cardinality K_n , $n \in \{0, 1, \dots, N\}$. The non-diagonal entries of the square matrix H_n , $n \in \{0, 1, \dots, N\}$, of size K_n describe the transition rates of the process η_t within its state space that do not lead to departure of a customer from the second stage. The diagonal entries of the matrix H_n are negative. The module of such an entry defines the rate of the process η_t exit from the corresponding state. The matrix H_n^- , $n \in \{1, 2, \dots, N\}$, of size $K_n \times K_{n-1}$ describes transition rates of the process η_t when one of n customers processed at the second stage departs from this stage. The stochastic matrix H_n^+ , $n \in \{0, 1, \dots, N-1\}$, of size $K_n \times K_{n+1}$ describes transition probabilities of the process η_t when a new customer arrives to the second stage.

Here, $H_n \mathbf{e} + H_n^- \mathbf{e} = \mathbf{0}^T$, $n \in \{1, 2, \dots, N\}$, $H_n^+ \mathbf{e} = \mathbf{e}$, $n \in \{0, 1, \dots, N-1\}$, where \mathbf{e} and $\mathbf{0}$ are column and row vectors of 1's and 0's of a suitable size, respectively.

2. Markov chain describing behavior of the tandem

Let us consider the four-dimensional process $\xi_t = \{i_t, n_t, \nu_t, \eta_t\}$, $i_t \geq 0, n_t \in \{0, 1, \dots, N\}$, $\nu_t \in \{1, 2, \dots, W\}$, $\eta_t \in \{1, 2, \dots, K_{n_t}\}$, $t \geq 0$, where, at the moment t , i_t is the number of customers at the first stage, n_t is the number of customers at the second stage, ν_t is the state of the underlying process of arrivals, and η_t is the state of the underlying process of customers processing at the second stage.

It is obvious that the process ξ_t is the Markov chain. Let us enumerate the states of this chain in some order. The most reasonable is the direct lexicographic order. However, sometimes the states of the process η_t are preferred to be enumerated in the reverse lexicographic order. For example, this takes place when the second stage is a multi-server queueing system without a buffer and with a phase-type distribution of service times in the servers. This allows to apply for writing down the explicit form of the matrices H_n , H_n^- , and H_n^+ the corresponding recursive formulas presented in [7] and [8].

Let us call the set of the states of the Markov chain ξ_t having the value (i, n) of the first two components as the sub-level (i, n) of the chain, and the set $\{(i, 0), (i, 1), \dots, (i, N)\}$ is the level i , $i \geq 0$.

Let $\mathbf{Q}_{i,j}^{(n,n')}$ be the matrix consisting of the transition rates from the sub-level (i, n) to the sub-level (j, n') , $i, j \geq 0$, $|i - j| \leq 1$, $n, n' \in \{0, 1, \dots, N\}$, $|n - n'| \leq 1$. By default, the diagonal entries of the matrix $\mathbf{Q}_{i,i}^{(n,n)}$ are negative, and their moduli define the rates of the exit from the corresponding state of the Markov chain ξ_t . The matrix $\mathbf{Q}_{i,j}$ consisting of the blocks $\mathbf{Q}_{i,j}^{(n,n')}$, $n, n' \in \{0, 1, \dots, N\}$, $|n - n'| \leq 1$, defines transition rates between the levels i and j . The infinite-size matrix \mathbf{Q} consisting of the blocks $\mathbf{Q}_{i,j}$ is the generator of the Markov chain ξ_t .

Theorem 1. The generator \mathbf{Q} of the Markov chain ξ_t is the block tri-diagonal matrix with the blocks $\mathbf{Q}_{i,j}$ that have the following form:

- The matrix $\mathbf{Q}_{i,i}$ has the diagonal blocks of the form

$$\mathbf{Q}_{i,i}^{(n,n)} = -i\mu I_{W I_{K_n}} + D_0 \oplus H_n, \quad n \in \{0, 1, \dots, N - 1\},$$

$$\mathbf{Q}_{i,i}^{(N,N)} = -i\mu(1 - p(1 - q))I_{W K_N} + D_0 \oplus H_N,$$

and subdiagonal blocks of the form

$$\mathbf{Q}_{i,i}^{(n,n-1)} = I_W \otimes H_n^-, \quad n \in \{1, 2, \dots, N\}.$$

- The matrix $\mathbf{Q}_{i,i+1}$ is the diagonal matrix with the blocks of the form

$$\mathbf{Q}_{i,i+1}^{(n,n)} = D_1 \otimes I_{K_n}, \quad n \in \{0, 1, \dots, N\}.$$

- The matrix $\mathbf{Q}_{i,i-1}$ has the diagonal blocks of the form

$$\mathbf{Q}_{i,i-1}^{(n,n)} = i\mu(1 - p)I_{W I_{K_n}}, \quad n \in \{0, 1, \dots, N - 1\},$$

$$\mathbf{Q}_{i,i-1}^{(N,N)} = i\mu(pq + 1 - p)I_W \otimes I_{K_N},$$

and updiagonal blocks of the form

$$\mathbf{Q}_{i,i-1}^{(n,n+1)} = i\mu p I_W \otimes H_n^+, \quad n \in \{0, 1, \dots, N-1\},$$

where \otimes and \oplus are symbols of the Kronecker product and sum of matrices.

Proof of the theorem is implemented by analysing all possible transitions of the considered Markov chain during an interval of infinitesimal length.

Theorem 2. The Markov chain ξ_t belongs to the class of asymptotically quasi-Toeplitz Markov chains (AQTMCs), for definition and relevant properties see [5] and [9].

Proof. Let R_i be the diagonal matrix with the diagonal entries defined by the moduli of the diagonal entries of the matrix $\mathbf{Q}_{i,i}$. It is easy to verify that the following limits exist:

$$Y_0 = \lim_{i \rightarrow \infty} R_i^{-1} \mathbf{Q}_{i,i-1}, \quad Y_1 = \lim_{i \rightarrow \infty} R_i^{-1} \mathbf{Q}_{i,i} + I, \quad Y_2 = \lim_{i \rightarrow \infty} R_i^{-1} \mathbf{Q}_{i,i+1}$$

where Y_1 and Y_2 are zero matrices and the matrix Y_0 is the stochastic matrix with the diagonal blocks $(Y_0)^{(n,n)} = (1-p)I_{WK_n}$, $n \in \{0, 1, \dots, N-1\}$, $(Y_0)^{(N,N)} = I_{WK_N}$ and updiagonal blocks $pI_W \otimes H_n^+$, $n \in \{0, 1, \dots, N-1\}$.

According to the definition given in [9], the existence of limits Y_k , $k = 0, 1, 2$, implies that the Markov chain ξ_t belongs to the class of AQTMCs. Theorem 2 is proven.

Corollary 1. The Markov chain ξ_t is ergodic for any set of system parameters.

Proof. The sufficient condition for ergodicity of AQTMC, see [5] and [9], is the fulfillment of the inequality $\mathbf{y}Y_0\mathbf{e} > \mathbf{y}Y_2\mathbf{e}$ where the vector \mathbf{y} is the solution of equations $\mathbf{y} = \mathbf{y}(Y_0 + Y_1 + Y_2)$, $\mathbf{y}\mathbf{e} = 1$. Because Y_0 is the stochastic matrix while $Y_2 = O$, this inequality is always true. The corollary is proven.

3. Calculation of the stationary distribution of the Markov chain and performance measures of the tandem

Let us denote by $\boldsymbol{\pi}(i, n)$ the row vector of the stationary probabilities of the states of the chain ξ_t that belongs to the sub-level (i, n) and $\boldsymbol{\pi}_i = (\boldsymbol{\pi}(i, 0), \boldsymbol{\pi}(i, 1), \dots, \boldsymbol{\pi}(i, N))$, $i \geq 0$. Numerically stable algorithms for computation of the vectors $\boldsymbol{\pi}_i$, $i \geq 0$, can be found in [9, 10, 11].

Having the vectors $\boldsymbol{\pi}_i$, $i \geq 0$, been computed, we can calculate the values of various performance measures of the system. We present expressions for computing some of them.

The mean numbers of customers receiving service at the first and second stages at an arbitrary moment are

$$L_1 = \sum_{i=1}^{\infty} i\pi_i \mathbf{e}, \quad L_2 = \sum_{i=0}^{\infty} \sum_{n=1}^N n\pi(i, n) \mathbf{e}.$$

The rates of customers that received service at the first stage (and do not need further service) and the second stages are

$$r_1 = \mu(1-p)L_1, \quad r_2 = \sum_{i=0}^{\infty} \sum_{n=1}^N \pi(i, n) H_n^- \mathbf{e}.$$

The loss probability of an arbitrary customer is given by the formulas

$$P_{loss} = \frac{pq}{\lambda} \sum_{i=1}^{\infty} i\mu\pi(i, N) \mathbf{e} = 1 - \frac{r_1 + r_2}{\lambda}.$$

Having fixed, e.g., the costs of maintenance of an unit of the capacity of the second stage and the penalty for a customer loss, various optimization problems related to the choice of the parameter N can be solved.

4. Conclusion

We considered a dual tandem of an infinite-server system and a finite state queueing system processing admitted customers at which is described by the Markovian process. A multidimensional Markov chain describing the behavior of the tandem system is analysed. The imposed assumption that the arrival flow is described by the *MAP* essentially increases the adequacy of the considered tandem to the real-world systems. The considered model creates the framework for analysis of plenty of tandem systems with self-service at the first stage. Customers processing at the second stage can be described by the single- and multi-server system, or a system with limited processor sharing, a system with servers breakdowns, customers impatience, etc. When the discipline of service at the second stage is fixed, the matrices H_n , H_n^- , and H_n^+ , $n \in \{0, 1, \dots, N\}$, have to be computed and, then, the presented above results can be directly applied for performance evaluation of a corresponding tandem system.

REFERENCES

1. *Nazarov A., Dudin A., Moiseev A.* Pseudo steady-state period in non-stationary infinite-server queue with state dependent arrival intensity // *Mathematics*. 2022. Vol. 10. № 15. P. 2661.

2. *Chakravarthy S., Shajin D., Krishnamoorthy A.* Infinite server queueing-inventory models // Journal of the Indian Society for Probability and Statistics. 2020. Vol. 21. № 1. P. 43-68.
3. *Lucantoni D.* New results on the single server queue with a batch Markovian arrival process // Commun.-Stat.-Stoch. Model. 1991. Vol. 7. № 51 P. 1-46.
4. *Chakravarthy S.* Introduction to Matrix-Analytic Methods in Queues 2: Queues and Simulation. ISTE Ltd, London and John Wiley and Sons, New York, 2022. 442 p.
5. *Dudin A.N., Klimenok V.I., Vishnevsky V.M.* The theory of queuing systems with correlated flows. Springer Nature. Cham, 2020. 420 p.
6. *Naumov V., Gaidamaka Y., Yarkina N., Samouylov K.* Matrix and Analytical Methods for Performance Analysis of Telecommunication Systems. Springer Nature. Berlin/Heidelberg, 2021.
7. *Ramaswami V.* Independent Markov processes in parallel // Commun.-Stat.-Stoch. Model. 1985. Vol. 1. № 1 P. 419-432.
8. *Kim C. et al* Mathematical model of operation of a cell of a mobile communication network with adaptive modulation schemes and handover of mobile users // IEEE Access. 2021. Vol. 9. P. 106933-106946.
9. *Klimenok V., Dudin A.* Multi-dimensional asymptotically quasi-Toeplitz Markov Chains and their application in queueing theory // Queueing Systems. 2006. Vol. 54. P. 245-259.
10. *Dudin S. et al* Effective algorithm for computation of the stationary distribution of multi-dimensional level-dependent Markov chains with upper block-Hessenberg structure of the generator // Journal of Computational and Applied Mathematics. 2020. Vol. 366. 112425.
11. *Dudin S., Dudina O.* Retrial multi-server queueing system with PHF service time distribution as a model of a channel with unreliable transmission of information // Applied Mathematical Modelling. 2019. Vol. 65. P. 676-695.

Dudin Alexander — Doctor of Sciences, Professor, Head, laboratory of applied probabilistic analysis. E-mail: dudin@bsu.by

Dudina Olga — Candidate of Sciences, Leading Researcher, laboratory of applied probabilistic analysis. E-mail: dudina@bsu.by

A COMPUTATIONAL ALGORITHM DEVELOPMENT FOR STATIONARY CHARACTERISTICS OF RESOURCE LOSS SYSTEM WITH RESOURCE-DEPENDENT SERVICE TIME

A. Nazarin¹, E. Sopin^{1,2}

¹*Peoples' Friendship University of Russia (RUDN University), Moscow, Russian Federation*

²*Institute of Informatics Problems, Federal Research Center Computer Science and Control of Russian Academy of Sciences, Moscow, Russian Federation*

We analyze a multi-server resource loss system model with service times that depends on the resource requirements. We develop a recurrence algorithm to calculate the stationary probability distribution and its characteristics, such as the blocking probability and the average amount of occupied resources. In numerical example, we calculate these characteristics under the assumption of the Erlang distribution for the service time and the geometric distribution for the resource requirements. **Keywords:** *resource loss system, random requirements, blocking probability, normalization constant.*

Introduction

Resource loss systems are widely used in wireless communication network systems performance analysis [1]. An important element of resource loss systems in application to the analysis of wireless communication network systems is the limited amount of available resources [2]. However, systems with infinite amount of available resources also have applications [3].

In [2], a multi-server resource loss system with limited amount of resources is introduced with the total amount of resources occupied by all customers taken for tracking. In [3], a model of the system with service time that depends on the amount of resources required is considered. This model can be used to analyze data transmission characteristics under conditions of elastic traffic.

Note that the formulas for calculating stationary probabilities and characteristics presented in [2] are too complex for analytical calculations due to multiple convolutions of resource requirements' cumulative distribution

The research was funded by the Russian Science Foundation, project no.22-79-10128, <https://rscf.ru/en/project/22-79-10128/>

function. In [4], a recurrence algorithm is presented for calculating stationary probabilities and system characteristics based on [2], with a service time independent of the resource requirements.

This paper presents the development of a recurrence algorithm for the case with resource-dependent service times. We introduce formulas for calculating the blocking probability and the average amount of occupied resources and provide a numerical analysis example.

1. Model Description

Consider a multi-server resource loss system with the maximum number of servers N and R resource units, $R < \infty$. Customers arrive according to the Poisson flow with arrival rate λ . The service times are mutually independent. Let $H_j(x)$ be the conditional cumulative distribution function of the service time provided that j discrete resources are required. The distribution of resource requirements is determined by $\{p_j\}$, $j = 1, 2, \dots, R$.

The state of the system at time t is described by the stochastic process $X(t) = (\xi(t), \gamma(t))$. Here $\xi(t)$ is the number of customers in the system at time t . Each customer occupies an integer number of $\gamma(t)$ resource units, $\gamma(t) < R$. The newly arrived i -th customer requires a random number of resources $r_i \geq 0$. If at the arrival instant t_i , $\xi(t_i) = N$ or $(R - \gamma(t_i) < r_i)$, then the customer is lost. At the end of the service, the customer leaves the system and releases the amount of resources it occupied. If $\xi(t) < N$ or $R - \gamma(t) \geq r_i$ then the customer is accepted for service.

Let $b_i = \int_0^\infty x d(h_i(x))$ be the conditional average service time, provided that the customer requires i resource units.

Then the average service time b can be found as:

$$b = \sum_{i=0}^R b_i p_i. \quad (1)$$

Denote

$$g_i = \frac{1}{b} b_i p_i. \quad (2)$$

Then, based on [3], the stationary distribution of the process $X(t)$ can be found as

$$q_0 = \lim_{t \rightarrow \infty} P\{\xi(t) = 0\} = \left(1 + \sum_{k=1}^N \frac{\rho^k}{k!} \sum_{r=0}^R g_r^{(k)}\right)^{-1}, \quad (3)$$

$$q_k(r) = \lim_{t \rightarrow \infty} P\{\xi(t) = k, \gamma(t) = r\} = q_0 \cdot g_r^{(k)} \cdot \frac{\rho^k}{k!}, \quad (4)$$

$k = 0, 1, 2, \dots, N$, $r = 0, 1, 2, \dots, R$, $g_r^{(k)}$ — k -fold convolution of the distribution g_r , $\rho = \lambda b$.

We introduce recurrence algorithm for a model with a service time that depends on the resource requirements.

$$G(n, r) = \sum_{k=0}^n \frac{\rho^k}{k!} \sum_{j=0}^r g_j^{(k)}. \quad (5)$$

Note that by definition $G(N, R) = q_0^{-1}$ is a normalization constant.

Proposition 1. Functions $G(n, r)$ with $G(0, r) = 1$ for $r = 0, 1, 2, \dots, R$ and $G(1, r) = 1 + \rho \sum_{j=0}^r g_j$ for $j = 0, 1, 2, \dots, r$ satisfy the following recurrence relation:

$$G(n, r) = G(n-1, r) + \frac{\rho}{n} \sum_{j=0}^r g_j (G(n-1, r-j) - G(n-2, r-j)), \quad (6)$$

$2 \leq n \leq N$.

Proof.

$$\begin{aligned} G(n, r) - G(n-1, r) &= \sum_{k=0}^n \frac{\rho^k}{k!} \sum_{j=0}^r g_j^{(k)} - \sum_{k=0}^{n-1} \frac{\rho^k}{k!} \sum_{j=0}^r g_j^{(k)} = \\ &= \frac{\rho^n}{n!} \sum_{j=0}^r \sum_{i=0}^j g_j g_{j-i}^{(n-1)} = \frac{\rho^n}{n!} \sum_{i=0}^r g_i \sum_{i=j}^r g_{j-i}^{(n-1)} \\ &= \frac{\rho}{n} \sum_{i=0}^r g_i \left(\frac{\rho^{n-1}}{(n-1)!} \sum_{i=0}^{r-i} g_{j-i}^{(n-1)} \right) = \\ &= \frac{\rho}{n} \sum_{i=0}^r g_i (G(n-1, r-i) - G(n-2, r-i)). \quad \square \end{aligned}$$

Proposition 2. The blocking probability π can be expressed as:

$$\pi = 1 - G^{-1}(N, R) \sum_{i=0}^R g_i G(N-1, R-i). \quad (7)$$

Proof.

$$1 - q_0 \sum_{k=0}^{N-1} \frac{\rho^k}{k!} \sum_{j=0}^R g_j^{(k+1)} = 1 - G^{-1}(N, R) \sum_{k=0}^{N-1} \frac{\rho^k}{k!} \sum_{j=0}^R g_j^{(k+1)} =$$

$$\begin{aligned}
&= 1 - G^{-1}(N, R) \sum_{k=0}^{N-1} \frac{\rho^k}{k!} \sum_{j=0}^R \sum_{i=0}^j g_i g_{j-i}^{(k)} = \\
&= 1 - G^{-1}(N, R) \sum_{k=0}^{N-1} \frac{\rho^k}{k!} \sum_{i=0}^j g_i \sum_{j=0}^R g_{j-i}^{(k)} = \\
&= 1 - G^{-1}(N, R) \sum_{k=0}^{N-1} \frac{\rho^k}{k!} \sum_{i=0}^j g_i \sum_{r=0}^{R-i} g_j^{(k)} = \\
&= 1 - G^{-1}(N, R) \sum_{i=0}^j g_i G(N-1, R-i).
\end{aligned}$$

Proposition 3. The average number of occupied resources B can be expressed as:

$$B = R - G^{-1}(N, R) \sum_{i=1}^R G(N, R-i). \quad (8)$$

Proof. Let \bar{B} be the average number of unoccupied resources. Then, by definition:

$$\begin{aligned}
\bar{B} &= R - \bar{B} \\
\bar{B} &= \sum_{j=0}^R (R-j) \sum_{k=0}^N q_0 \frac{\rho^k}{k!} g_j^{(k)} = G^{-1}(N, R) \sum_{k=0}^N \frac{\rho^k}{k!} \sum_{j=0}^R (R-r) g_j^{(k)} = \\
G^{-1}(N, R) \sum_{k=0}^N \frac{\rho^k}{k!} \sum_{r=0}^R \sum_{i=1}^{R-j} g_j^{(k)} &= G^{-1}(N, R) \sum_{k=0}^N \frac{\rho^k}{k!} \sum_{i=0}^R \sum_{j=0}^{R-i} g_j^{(k)} = \\
G^{-1}(N, R) \sum_{i=1}^R G(N, R-i). \quad \square
\end{aligned}$$

2. Numerical example

Consider a case where the amount of resources allocated to the customer follows the truncated geometric distribution. Let $N = 100, R = 100, k = 2, \lambda = [0.3, 0.65], p = 0.4$ and 0.6 .

We suppose that the service time follows the Erlang distribution $\Gamma(k, \theta_j)$ with density $h_j(x) = \frac{x^{k-1} \theta_j^k}{(k-1)!} e^{-\theta_j x}$, where $\theta_j = \frac{j}{R}, j = 1, 2, \dots, R$. Take $k = 2$, therefore $b_j = \frac{2R}{j}$.

The results are depicted in Fig. 1. Initially, the probability of losses at $p = 0.4$ is lower than at $p = 0.6$. But with the increase of the arrival intensity, the system takes more "heavy" customers, which, according to the condition, are serviced faster. Thus, at $p = 0.4$, the system receives more customers that do not require many resource units than at $p = 0.6$. Customers that are not resource-demanding but time-demanding fill the system, and when a new one arrives, a loss occurs.

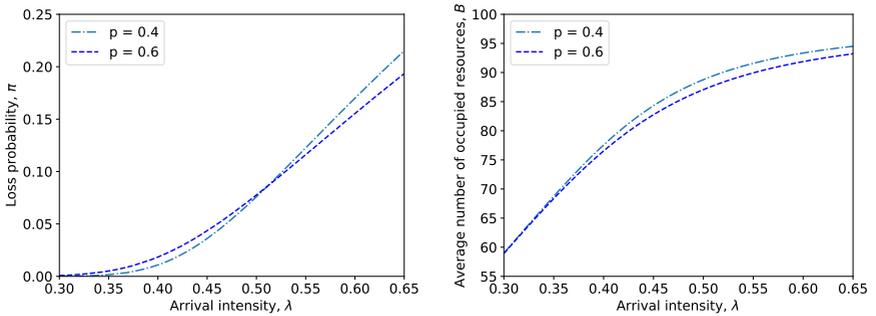


Figure 1. Numerical example results

3. Conclusion

In this paper, we considered a resource loss system with resource-dependent service times. We developed a recurrence algorithm to evaluate the system's characteristics. The recurrence algorithm can be applied to models of wireless networks.

REFERENCES

1. *Moltchanov D., Sopin E., Begishev V., Samuylov A., Koucheryavy Y., Samouylov K.* A Tutorial on Mathematical Modeling of 5G/6G Millimeter Wave and Terahertz Cellular Systems // *IEEE Communications Surveys and Tutorials*, 2022. V. 24, P. 1072–1116.
2. *Tikhonenko O., Ziolkowski M.* Queueing Systems with Random Volume Customers and their Performance Characteristics. *Journal of information and organizational sciences* // *Journal of information and organizational sciences*, 2021. V. 45. № 1 P. 21–38.
3. *Lisovskaya E., Pankratova E., Moiseeva S., Pagano M.* Analysis of a Resource-Based Queue with the Parallel Service and Renewal Arrivals // *Distributed Computer and Communication Networks*, 2020. P. 335—349.

4. *Naumov V., Samuilov K., Samuilov A.* On the total amount of resources occupied by serviced customers // *Avtomat. i Telemekh.*, 2016. V. 8. P. 125–135.
5. *Naumov V.A., Samuilov K.E.* O svyazi resursnyh sistem massovogo ob-sluzhivaniya s setyami Erlanga // *Inform. i eyo primen.*, 2016. V. 10. № 3. P. 9–14.
6. *Sopin E., Ageev K., Markova E., Vikhrova O., Gaidamaka Yu.* Performance Analysis of M2M Traffic in LTE Network Using Queuing Systems with Random Resource Requirements // *Automatic Control and Computer Sciences*, 2018. V. 52. № 5 P. 345–353.

Nazarin Artem — Master’s Student, Probability Theory and Cybersecurity Department, RUDN University. E-mail: *nazaryin-ai@pfur.ru*

Sopin Eduard — Candidate of Physical and Mathematical Sciences, associate professor, Probability Theory and Cybersecurity Department, RUDN University. E-mail: *sopin-es@pfur.ru*

ANALYSIS OF A K-OUT-OF-N SYSTEM WITH A SINGLE SERVER HAVING MULTIPLE VACATIONS AND EXTERNAL NON-PREEMPTIVE SERVICE

Binumon Joseph¹, K. P. Jose²

¹*Govt. Engineering College Idukki, Painavu-685603, Kerala, India*

²*Dept.of Mathematics, St. Peter's College, Kolenchery-682311, Kerala, India*

This paper investigates the reliability of a k-out-of-n system with a single server taking multiple vacations. The server provides service to failed system components and extends the service to external clients. The arrival of externally failed customers and the failure rate of internal components follow Poisson distribution. The service times of both internal and external customers, are exponentially distributed. The external and internal services are governed by N policy with non-preemptive external services. If the server is on vacation, the vacation stops and internal service starts when the number of failed system components reaches N. The model is examined using the Matrix Geometric Method. Several system performance measures and an appropriate cost function are determined.

Keywords: *k-out-of-n system, vacation, N-Policy, non-preemptive service, matrix-geometric method.*

Introduction

A k-out-of-n system with n components is functional if and only if at least k of those components are functional. Chakravarthy et al. [2], analysed a k-out-of-n system with an unreliable server that takes frequent vacations. A system with N machines having an exponential failure rate and being served by an unreliable server was examined by Chakravarthy et al. [1]. Dudin et al. [3] utilised the idle time by offering service to external clients.

Krishnamoorthy et al. [5] used the Matrix Geometric Method to assess the reliability of a k-out-of-n system servicing external clients and to derive relevant performance measurements. N-policy with non-preemption regulates the switching of servers between internal and external customers. A repairable system with non-identical components was examined by Wang et al. [7] under phase-type distributed multiple vacations of a single server. On the reliability study of a multistate system under phase-type distributed

The authors acknowledge the financial support provided by FIST Program, Dept.of.Science and Technology, Govt.of India, to the PG & Research Dept.of Mathematics through SR/FST/College-2018- XA 276(C).

repeated vacations of a single server, Liu et al. [6] presented a mixed redundancy technique. Joseph and Jose [4] analysed a k-out-of-n system with server vacation and extended service to external customers. In this study, we examine a k-out-of-n system with server vacation, where one server provides both internal and external services under non-preemptive service. This problem can be modeled for service cooperation between different companies that prioritize their own customers.

1. Mathematical Modelling and Analysis of the Problem

At the beginning, all components of a k-out-of-n system work well. The failure time of system components is distributed exponentially with a value of λ_s/i , where i is the number of functional components in the system. The average rate of system component failures per unit of time is thus λ_s . Failed component's service time is distributed exponentially, with μ_s as the parameter. The arrival of failed external customers to the system is characterised by an exponential distribution with parameter λ_e . An exponential distribution with parameter μ_e characterises the service duration of externally failed customers. By using the idle time for external services, the server's performance will increase as a consequence of greater experience from various external service scenarios. N-policy regulates the servicing of external failed customers and internal failed components. When the number of internal failed components reaches level N, then the server begins the internal service. When the number of internal failed system components reaches level N and the server is busy providing services for external customers, it switches from providing external to internal services after the ongoing external services are finished.

If the system has less than N internal failed components and no external failed customers, the server takes a vacation. With parameter ν , the vacation time is exponentially distributed. Following the servicing of internal components that failed, the server goes on vacation. The server takes another vacation after the first one ends if there are no external failed customers and fewer failed internal components than N. Also, when the system failure count reaches N, the server's vacation terminates, and it immediately fixes each of the N internal failures one by one. When the server is in use by internal components, external failed customers do not join the system for service. An external customer does not join the system while the server is in external service and there are N or more failed system components. If not, an infinite-length queue is formed to which the external failed customers are added.

Let $E(t)$ represent the number of external failed components in the system, $I(t)$ be the number of internal failed components, and $S(t)$ be the

status of the server at time t . Let $S(t) = 0$, represents the server vacation, $S(t) = 1$ if the server services the internal components, and $S(t) = 2$ if the server services the external customers. Then $\{X(t), t \geq 0\}$, where $X(t) = (E(t), S(t), I(t))$ is a continuous time Markov chain with the state space $\Omega = \{(i_1, 0, i_2)/i_1 \geq 0, 0 \leq i_2 \leq N-1\} \cup \{(i_1, 1, i_2)/i_1 \geq 0, 1 \leq i_2 \leq n-k+1\} \cup \{(i_1, 2, i_2)/i_1 \geq 1, 0 \leq i_2 \leq n-k+1\}$. The following notations are used in the sequel. I_n denotes the n^{th} order identity matrix. E_k is the k^{th} order square matrix such that $E_k(i, i) = -1$, if $1 \leq i \leq k$, 1 if $j = i + 1$, and all other elements are zeros. E_k^t is the transpose of E_k , $r_k(i)$ is a $1 \times k$ order row matrix with i^{th} element is 1 and all other elements are zeros. $c_k(i)$ is the transpose of $r_k(i)$, \mathbf{e} is a column matrix of appropriate order with all elements being 1, and \otimes is the Kronecker product of matrices.

The block tridiagonal infinitesimal generator matrix of $\{X(t), t \geq 0\}$ is

$$Q = \begin{pmatrix} B_1 & B_0 & & & \\ B_2 & A_1 & A_0 & & \\ & A_2 & A_1 & A_0 & \\ & & \ddots & \ddots & \ddots \end{pmatrix}, \text{ where } B_1 = \begin{pmatrix} B_{11} & B_{12} \\ B_{13} & B_{14} \end{pmatrix}, B_{11} = \lambda_s E_N - \lambda_e I_N, B_{12} = \lambda_s (C_N(N) \otimes r_{n-k+1}(N)), B_{13} = \mu_s (C_{n-k+1}(1) \otimes r_N(1)), B_{14} = \lambda_s E_{n-k+1} + \mu_s E_{n-k+1}^t + \lambda_s (C_{n-k+1}(n-k+1) \otimes r_{n-k+1}(n-k+1)), B_0 = \begin{pmatrix} \lambda_e I_N & 0_{N \times (n-k+1)} & 0_{N \times (n-k+2)} \\ 0_{(n-k+1) \times N} & 0_{(n-k+1) \times (n-k+1)} & 0_{(n-k+1) \times (n-k+2)} \end{pmatrix}, B_2 = \begin{pmatrix} 0_{N \times N} & 0_{N \times (n-k+1)} \\ 0_{(n-k+1) \times N} & 0_{(n-k+1) \times (n-k+1)} \end{pmatrix}, \text{ where } B_{21} = \begin{pmatrix} \mu_e I_N \\ 0_{(n-k+2-N) \times N} \end{pmatrix}, B_{22} = \begin{pmatrix} 0_{N \times (N-1)} & 0_{N \times (n-k+2-N)} \\ 0_{(n-k+2-N) \times (N-1)} & \mu_e I(n-k+2-N) \end{pmatrix}, A_1 = \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{14} & A_{15} & A_{16} \\ A_{17} & A_{18} & A_{19} \end{pmatrix}, \text{ where } A_{11} = B_{11} - \nu I_N, A_{13} = (\nu I_N \quad 0), A_{12} = \lambda_s (C_N(N) \otimes r_{n-k+1}(N)), A_{14} = \mu_s (C_{n-k+1}(1) \otimes r_N(1)), A_{15} = \lambda_s E_{n-k+1} + \mu_s E_{n-k+1}^t + \lambda_s (r_{n-k+1}(n-k+1) \otimes C_{n-k+1}(n-k+1)), A_{16} = 0_{(n-k+1) \times (n-k+2)}, A_{17} = 0_{(n-k+2) \times N}, A_{18} = 0_{(n-k+2) \times (n-k+1)}, A_{19} = \lambda_s E_{n-k+2-N} + \mu_s E_{n-k+2-N}^t + \lambda_s (r_{n-k+2-N}(n-k+2-N) \otimes C_{n-k+2-N}(n-k+2-N)),$$

$$\begin{aligned}
A_{19} &= \begin{pmatrix} \lambda_s E_N - (\lambda_e + \mu_e) I_N & \lambda_1 (C_N(N) \otimes r_{(n-k+2-N)}) \\ 0_{(n-k+2-N) \times N} & A_{191} \end{pmatrix}. \\
A_0 &= \begin{pmatrix} \lambda_e I_N & 0_{N \times (n-k+1)} & 0_{N \times (n-k+2)} \\ 0_{(n-k+1) \times N} & 0_{(n-k+1) \times (n-k+1)} & 0_{(n-k+1) \times (n-k+2)} \\ 0_{(n-k+2) \times N} & 0_{(n-k+2) \times (n-k+1)} & A_{01} \end{pmatrix}, \\
\text{where } A_{01} &= \begin{pmatrix} \lambda_e I_N & 0_{N \times (n-k+2-N)} \\ 0_{(n-k+2-N) \times N} & 0_{(n-k+2-N) \times (n-k+2-N)} \end{pmatrix}. \\
A_2 &= \begin{pmatrix} 0_{N \times N} & 0_{N \times (n-k+1)} & 0_{N \times (n-k+2)} \\ 0_{(n-k+1) \times N} & 0_{(n-k+1) \times (n-k+1)} & 0_{(n-k+1) \times (n-k+2)} \\ 0_{(n-k+2) \times N} & A_{21} & A_{22} \end{pmatrix}, \text{ where} \\
A_{21} &= B_{22}, A_{22} = \begin{pmatrix} \mu_e I_N & 0_{N \times (n-k+2-N)} \\ 0_{(n-k+2-N) \times N} & 0_{(n-k+2-N) \times (n-k+2-N)} \end{pmatrix}.
\end{aligned}$$

2. Stability Condition

Let $\Pi = (\Pi_1, \Pi_2, \Pi_3)$ be the steady state probability vector of the generator matrix $A = A_2 + A_1 + A_0$, where $\Pi_1 = ((\pi_{(1,1)}, \pi_{(1,2)}, \pi_{(1,3)} \dots, \pi_{(1,N)})$, $\Pi_2 = (\pi_{(2,1)}, \pi_{(2,2)}, \pi_{(2,3)} \dots, \pi_{(2,n-k+1)})$, $\Pi_3 = (\pi_{(3,1)}, \pi_{(3,2)}, \pi_{(3,3)} \dots, \pi_{(3,n-k+2)})$. From $\Pi A = 0$ and $\Pi e = 1$ we obtain,

$$\Pi_1 A_{11}^* + \Pi_2 A_{21}^* = 0, \quad \Pi_1 A_{12}^* + \Pi_2 A_{22}^* + \Pi_3 A_{32}^* = 0, \quad \Pi_1 A_{13}^* + \Pi_3 A_{33}^* = 0.$$

where $A_{11}^* = \lambda_s E_N - \nu I_N$, $A_{12}^* = A_{12}$, $A_{13}^* = A_{13}$, $A_{21}^* = A_{14}$, $A_{22}^* = A_{15}$, $A_{32}^* = A_{21}$, and $A_{33}^* = \begin{pmatrix} \lambda_s E_N & \lambda_1 (C_N(N) \otimes r_{(n-k+2-N)}) \\ 0_{(n-k+2-N) \times N} & A_{191} \end{pmatrix}$.

The above Markov chain is stable if and only if $\Pi A_0 e < \Pi A_2 e$. We partitioned Π_3 as Π_{31} and Π_{32} , where $\Pi_{31} = (\pi_{(3,1)}, \pi_{(3,2)}, \dots, \pi_{(3,N)})$ and $\Pi_{32} = (\pi_{(3,N+1)}, \pi_{(3,N+2)}, \dots, \pi_{(3,n-k+2)})$. Using the structure of matrices, we obtain

$$\Pi A_0 e = \lambda_e [\Pi_1 e + \Pi_{31} e] = \pi_{(2,1)} \frac{N \lambda_e \mu_s}{\lambda_s}. \quad (1)$$

$$\begin{aligned}
\Pi A_2 e &= \mu_e [\Pi_{31} e + \Pi_{32} e], \\
&= \pi_{(2,1)} \mu_e \left[\frac{N \mu_s}{\lambda_s} + \left(\frac{\mu_s}{\mu_e} - \frac{\mu_s}{\nu} \right) \left(1 - \left(\frac{\lambda_s}{\lambda_s + \nu} \right)^N \right) \right]. \quad (2)
\end{aligned}$$

Theorem 1. The system is stable if and only if

$$N \left(\frac{\lambda_e - \mu_e}{\lambda_s} \right) < \left(\frac{\nu - \mu_e}{\nu} \right) \left[1 - \left(\frac{\lambda_s}{\lambda_s + \nu} \right)^N \right]. \quad (3)$$

Proof. Using equation(10) and equation(7) in the stability condition $\Pi A_0 e < \Pi A_2 e$, we obtain equation(3) ■

3. Steady State Probability Vector

The Markov process $\{X(t), t \geq 0\}$ is a level-independent QBD process. The stationary distribution, when it exists, has a matrix geometric solution. Let $\mathbf{x} = (x_0, x_1, x_2, \dots)$ be the probability steady state vector of Q , the generator matrix of the process. Then \mathbf{x} satisfies the equations $\mathbf{x}Q = 0$ and the normalizing condition $\mathbf{x}\mathbf{e} = 1$. Here \mathbf{e} represents the column matrix of 1's with infinite order. Then $x_{i+1} = x_i R, \forall i \geq 1$, where R is the minimal nonnegative solution of the matrix equation $A_0 + RA_1 + R^2 A_2 = 0$. The boundary probability vectors (x_0, x_1) are obtained from the equations $x_0 B_0 + x_1 B_2 = 0, x_0 B_1 + x_1(RA_2 + A_1) = 0$. Using normalization, $x_0 \mathbf{e} + x_1 (I - R)^{-1} \mathbf{e} = 1$, one can solve the equations for x_0 and x_1 .

4. System Performance Measures

The important performance measures of the system are given below.

- 1) Portion of time the system is down,

$$P_F = \sum_{i=0}^{\infty} x(i, 1, n - k + 1) + \sum_{i=1}^{\infty} x(i, 2, n - k + 1).$$

- 2) Reliability of the system, $P_R = 1 - P_F$.

- 3) The average number of outside failed units in the queue,

$$N_Q = \sum_{i=0}^{\infty} i \sum_{j=0}^{N-1} x(i, 0, j) + \sum_{i=0}^{\infty} i \sum_{j=1}^{n-k+1} x(i, 1, j) + \sum_{i=1}^{\infty} (i-1) \sum_{j=0}^{n-k+1} x(i, 2, j).$$

- 4) The average number of failed main components,

$$N_{IF} = \sum_{j=0}^{N-1} j \sum_{i=0}^{\infty} x(i, 0, j) + \sum_{j=1}^{n-k+1} j \sum_{i=0}^{\infty} x(i, 1, j) + \sum_{j=0}^{n-k+1} j \sum_{i=1}^{\infty} x(i, 2, j).$$

- 5) Probability that the server was found on vacation, $P_v = \sum_{i=0}^{\infty} \sum_{j=0}^{N-1} x(i, 0, j)$.

- 6) Expected rate of external customer loss,

$$E_{EL} = \lambda_2 \sum_{i=0}^{\infty} \sum_{j=1}^{n-k+1} x(i, 1, j) + \lambda_2 \sum_{i=1}^{\infty} \sum_{j=N}^{n-k+1} x(i, 2, j).$$

5. Cost Function

The cost per unit of time incurred if the system fails is shown by C_1 . Holding cost of each external customer within the queue for one unit of time is denoted by C_2 ; the cost due to the lose of one external customer is represented by C_3 ; Holding cost of each failed system component for one unit of time is represented by C_4 ; and the cost/unit of time if the server

is on vacation is represented by C_5 . Then the expected cost/unit time $= C_1P_F + C_2N_Q + C_3E_{EL} + C_4N_{IF} + C_5P_v$.

6. Conclusion

This work investigates a k-out-of-n system where a reliable server additionally offers service to outside clients. The model is analysed using the Matrix Geometric Method. The system's stability state and key performance indicators are obtained. A cost function is developed to optimise the value of N, the N policy level. The study in this paper can be extended by considering working vacations for the server.

REFERENCES

1. *Chakravarthy S.R., Agarwal A.* Analysis of a machine repair problem with an unreliable server and phase type repairs and services // Naval Research Logistics (NRL). 2003 Aug. Vol. 50. № 5. P. 462–480.
2. *Chakravarthy S.R., Krishnamoorthy A., Ushakumari P.V.* A k-out-of-n reliability system with an unreliable server and phase type repairs and services: the (N, T) policy // Journal of Applied Mathematics and Stochastic Analysis. 2001. Vol. 14. № 4. P. 361–380.
3. *Dudin A.N., Krishnamoorthy A., Narayanan V.C.* Idle time utilization through service to customers in a retrial queue maintaining high system reliability // Journal of Mathematical Sciences. 2013 Jun. Vol. 191. № P. 506–517.
4. *Joseph B., Jose K.P.* A k-out-of-n reliability system with internal and external service, N policy and server vacation // Informational Technologies and Mathematical Modelling 2023. Vol. 117. № . P. 12.
5. *Krishnamoorthy A., Sathian M.K., Cviswanath N.* Reliability of A k-out-of-n system with a single server extending non-preemptive service to external customers-part I // Reliability: Theory & Applications. 2016. Vol. 11. № 3(42). P. 62–75.
6. *Liu B., Wen Y., Shi H., Chen J.* Reliability analysis for multi-state systems under K-mixed redundancy strategy considering switching failure // Reliability Engineering & System Safety. 2022 Dec 1. Vol. 288:108814.
7. *Wang G., Hu L., Zhang T., Wang Y.* Reliability modeling for a repairable (k1, k2)-out-of-n: G system with phase-type vacation time // Applied Mathematical Modelling. 2021 Mar 1. Vol. 91. № . P. 311–321.

Binumon Joseph — Assistant Professor , Govt. Engineering College Idukki, Painavu, Kerala, India. E-mail: jbinumonjoseph@gmail.com

K.P. Jose — Associate Professor & Head, Dept.of Mathematics, St. Peter's College, Kolenchery, Kerala, India. E-mail: kpjspc@gmail.com

A PERISHABLE INVENTORY SYSTEM WITH UNRELIABLE SERVER

Bobina J. Mattam, K. P. Jose

*Post Graduate & Research Dept. of Mathematics, St. Peter's College,
Kolenchery-682311, Kerala, India.*

In this work, a single unreliable server and a perishable inventory system with clients arriving at exponentially distributed time intervals and demanding a single item are considered. The inventory is acquired from an external vendor, and the policy for control is (s, Q) . The server may seize up while in service and the failure time follows an exponential distribution. When a consumer arrives and discovers that the server is busy or malfunctioning, the customer either leaves the system or proceeds to an infinitely large waiting area called orbit. Then further retrial of customers from the orbit is also considered. An algorithmic solution to the problem is obtained using Matrix Analytic Method. The mean number of customer loss before and after entering the system, the average rate of breakdown and repair, and some other relevant performance measures of the system are derived, and a suitable profit function is obtained.

Keywords: *buffer, matrix analytic method, perishable inventory, retrials, unreliable server.*

1. Introduction

The majority of perishable inventory models that are currently in use in the literature do not take the server's annoyances into account. Businesses and sectors would suffer as a result. Generally, in order to supply the products on time, a company needs to be ready to handle any unforeseen limitations brought on by the server. Through appropriate server repair and maintenance, a business can effectively prevent missed sales and backlogged cases. In this study, we present an unreliable server inventory model for decaying objects. A thorough overview of perishable inventory modeling can be found in the review paper [4]. At first, there wasn't been much focus in the literature on inventory systems that have favorable service times and customer retrials. Artalejo et al. [2] introduced an alternative to classical approaches based either on backlogged or lost sale cases. The authors considered a continuous review (s, S) inventory system in which primary customer arrives in stock out period, leave the server and retries after some random time.

The authors acknowledge the financial support provided by FIST Program, Dept. of Science and Technology, Govt. of India, to the P. G. & Research Dept. of Mathematics, through SR/FST/College-2018-XA 276(C).

Reshmi and Jose [6] studied a queueing inventory system with perishable items where the items in the inventory perish at a linear rate and all the underlying processes are assumed to be exponential. Bobina and Jose [3] studied an M/M/1 perishable inventory system with production where the service may experience some interruption while serving.

2. Mathematical Model

Consider an (s, Q) inventory system with a single server retrial facility and perishable items. The lifetime of an item in the inventory is exponentially distributed with parameter $j\omega$, when there are j items in the inventory. Customers arrive at the system following the Poisson distribution with parameter λ , and each customer demands a single item. If the server is idle on arrival epoch of a customer, then that customer is immediately taken for service, provided the inventory is non-empty. The service time duration follows a negative exponential distribution with parameter μ .

When on-hand inventory level drops to s , an order of quantity Q is placed and is replenished after a random lead-time, and it is exponentially distributed with parameter β . The server may be subject to breakdown while at service, and it follows Poisson distribution with parameter δ_1 and its repair times are exponentially distributed with parameter δ_2 . Any arriving customer, on finding inventory level zero or the server busy or breakdown, either proceeds join a waiting space of infinite capacity called orbit with probability γ or exits the system with probability $1 - \gamma$. All customers who enter the orbit; generate requests for service at exponentially distributed time intervals with mean $\frac{1}{\theta}$ independently. The retrial customers who find the inventory out of stock or the server busy or breakdown, return to the orbit with probability δ or exit the system with probability $1 - \delta$.

Let $N(t)$ and $I(t)$ denote the number of customers in the orbit at time t and the inventory level at time t , respectively. Further, let

$$C(t) : \begin{cases} 0, & \text{if the server idle at time } t \\ 1, & \text{if the server busy at time } t \\ 2, & \text{if the server is breakdown at time } t. \end{cases}$$

Now, $\mathbf{X} = \{(N(t), C(t), I(t)) | t \geq 0\}$ constitutes a continuous time Markov chain with state space $F_0 \cup F_1 \cup F_2$, where $F_0 = \{(i, 0, j) | i \geq 0; 0 \leq j \leq S\}$, $F_1 = \{(i, 1, j) | i \geq 0; 1 \leq j \leq S\}$ and $F_2 = \{(i, 2, j) | i \geq 0; 1 \leq j \leq S\}$.

The generator matrix of the process is

$$Q_1 = \begin{bmatrix} A_{10} & A_0 & & & \\ A_{21} & A_{11} & A_0 & & \\ & A_{22} & A_{12} & A_0 & \\ & & \ddots & \ddots & \ddots \end{bmatrix},$$

We introduce some notations to describe the terms in the above matrix.

- I_m denotes an identity matrix of order m .
- e is a column matrix of appropriate order where all entries are 1.
- $r_m(i)$ denotes a $1 \times m$ row matrix whose i th entry is 1 and all other entries are zeros and $c_m(i)$ denotes the transpose of $(r_m(i))$.
- \otimes denotes the Kronecker product of matrices.

Thus the sub-matrices given in the matrix A_N are as follows.

$$\begin{aligned}
 H_{11} &= \beta(c_{S+1}(1) \otimes r_{S+1}(1)) + \sum_{i=2}^{s+1} (-\lambda - \beta - (i-1)\omega - N\theta)c_{S+1}(i) \otimes r_{S+1}(i) \\
 &+ \sum_{i=s+2}^{S+1} (-\lambda - (i-1)\omega - N\theta)c_{S+1}(i) \otimes r_{S+1}(i) + \sum_{i=2}^{S+1} i\omega(c_{S+1}(i) \otimes r_{S+1}(i-1)) \\
 &+ \sum_{i=1}^{s+1} \beta(c_{S+1}(i) \otimes r_{S+1}(Q+i)), H_{12} = (\lambda + N\theta) \begin{bmatrix} 0_{1 \times S} \\ I_S \end{bmatrix}, H_{21} = \mu [I_S \quad 0_{S \times 1}], \\
 H_{22} &= (-\beta - \mu - \delta_1)(c_S(1) \otimes r_S(1)) + \sum_{i=2}^s (-\beta - \mu - (i)\omega - \delta_1)c_S(i) \otimes r_S(i) \\
 &+ \sum_{i=s+1}^S (-\mu - (i)\omega - \delta_1)c_S(i) \otimes r_S(i) + \sum_{i=2}^S ((i)\omega)c_S(i) \otimes r_S(i-1) \\
 &+ \sum_{i=1}^s \beta c_S(i) \otimes r_S(Q+i), H_{23} = \delta_1 I_S, H_{32} = \delta_2 I_S, \\
 H_{33} &= (-\beta - \delta_2)(c_S(1) \otimes r_S(1)) + \sum_{i=2}^s (-\beta - (i)\omega - \delta_2)c_S(i) \otimes r_S(i) \\
 &+ \sum_{i=s+1}^S (-(i)\omega - \delta_2)c_S(i) \otimes r_S(i) + \sum_{i=2}^S ((i)\omega)c_S(i) \otimes r_S(i-1) \\
 &+ \sum_{i=1}^s \beta c_S(i) \otimes r_S(Q+i).
 \end{aligned}$$

The stationary probability vector π_N of A_N be partitioned as $\pi_N = (\pi_0^N, \pi_1^N, \pi_2^N)$;

$$\pi_0^N = (\pi_{N,0,0}, \dots, \pi_{N,0,S}), \pi_1^N = (\pi_{N,1,1}, \dots, \pi_{N,1,S}), \pi_2^N = (\pi_{N,2,1}, \dots, \pi_{N,2,S}).$$

Then the relation $\pi_N A_N = 0$ along with the normalizing condition $\pi_N e = 1$ gives rise to the following equations:

$$\begin{aligned}
 \pi_0^N H_{11} + \pi_1^N H_{21} &= 0, \rightarrow & \pi_0^N &= -(\pi_1^N H_{21})H_{11}^{-1}, \\
 \pi_0^N H_{12} + \pi_1^N H_{22} + \pi_2^N H_{32} &= 0, \rightarrow & \pi_1^N &= -(\pi_0^N H_{12} + \pi_2^N H_{32})H_{22}^{-1}, \\
 \pi_1^N H_{23} + \pi_2^N H_{33} &= 0. \rightarrow & \pi_2^N &= -(\pi_1^N H_{23})H_{33}^{-1}.
 \end{aligned}$$

The matrices $H_{ii}; i = 1, 2, 3$ in the above set of equations are all diagonally dominant and hence invertible. Thus, all the terms in the previous set of equations exists and hence by Block Gauss-Seidel iteration, we can find the vector π_N .

Thus, the stability condition can be asserted as, the truncated system is stable if and only if $\pi_N A_0 e < \pi_N A_2 N e$. By rearranging and using the limiting technique used by Krishnamoorthy et.al. [5], as $N \rightarrow \infty$, we get $\lim_{N \rightarrow \infty} (\frac{\pi_N A_0 e}{\pi_N A_2 N e}) < 1$, where

$$\begin{aligned}
 \pi_N A_0 e &= \pi_0^N (\lambda \gamma c_{S+1}(1) \otimes r_{S+1}(1))e + (\lambda \gamma (\pi_1^N + \pi_2^N))e, \\
 \pi_N A_2 N e &= (\pi_0^N (N\theta(1-\delta))(c_{S+1}(1) \otimes r_{S+1}(1))e \\
 &+ ((\pi_0^N)N\theta \sum_{i=1}^s c_{S+1}(i+1) \otimes r_{S+1}(i))e + ((\pi_1^N + \pi_2^N)N\theta(1-\delta))e.
 \end{aligned}$$

4. Steady State Analysis

Since \mathbf{X} is a level dependent quasi-birth-death process, to calculate the steady state probability vector, we use the method described by Neuts-Rao [5]. Consider the steady state probability vector $\mathbf{x} = (x_0, x_1, x_2, \dots)$ of Q , where,

$$x_i = (z_{i,0,0}, z_{i,0,1}, \dots, z_{i,0,S}, z_{i,1,1}, z_{i,1,2}, \dots, z_{i,1,S}, z_{i,2,1}, z_{i,2,2}, \dots, z_{i,2,S}), (i \geq 0)$$

satisfies the relation, $x_{N+k-1} = x_{N-1} R^k$, $k \geq 1$ where the matrix R is the unique non-negative solution of the matrix quadratic equation $R^2 A_2 + R A_1 + A_0 = \mathbf{0}$ with $A_1 = A_{1N}$, $A_2 = A_{2N}$ and $R = \lim_{n \rightarrow \infty} R_n$, where $\{R_n\}$ is defined such that $R_{n+1} = -A_0 A_1^{-1} - R_n A_2 A_1^{-1}$; $n \geq 0$ and $R_0 = \mathbf{0}$. The vectors x_0, x_1, \dots, x_{N-1} corresponding to boundary portion of Q are obtained using Gauss-Siedel method subject to normalizing condition $\sum_{i=0}^{\infty} x_i \mathbf{e} = 1$.

5. Performance Measures

Using the above probability vector, some important performance measures are given below,

- 1) Expected inventory level in the system,

$$E_{inv} = \sum_{i=0}^{\infty} \sum_{k=0}^2 \sum_{j=1}^S j z_{i,k,j} + \sum_{i=0}^{\infty} \sum_{k=0}^2 \sum_{j=1}^S j z_{i,k,j}.$$

- 2) Mean number of customers in the orbit,

$$E_{orbit} = \sum_i \sum_k \sum_j j z_{i,k,j}$$

- 3) Expected reorder rate

$$E_{ro} = \mu \sum_{i=0}^{\infty} z_{i,1,s+1} + (s+1) \omega \sum_{i=0}^{\infty} \sum_{k=0}^2 z_{i,k,s+1}.$$

- 4) Expected perishability rate,

$$E_p = \omega \left(\sum_{i=0}^{\infty} z_{i,0,1} + \sum_{i=0}^{\infty} \sum_{k=0}^2 \sum_{j=2}^S j z_{i,k,j} \right).$$

- 5) Average number of departures after service completion,

$$E_{ds} = \mu \sum_{i=0}^{\infty} \sum_{j=1}^S z_{i,1,j}.$$

- 6) Average number of customers lost prior to entering the orbit,

$$E_{la} = \lambda(1 - \gamma) \sum_{i=0}^{\infty} \left(z_{i,0,0} + \sum_{j=1}^S z_{i,1,j} + \sum_{j=1}^S z_{i,2,j} \right).$$

- 7) Average number of customers lost during retrials,

$$E_{lr} = \theta(1 - \delta) \sum_{i=0}^{\infty} i \left(z_{i,0,0} + \sum_{j=1}^S z_{i,1,j} + \sum_{j=1}^S z_{i,2,j} \right).$$

- 8) Overall rate of retrials,

$$\theta_1^* = \theta \left(\sum_{i=1}^{\infty} i x_i \right) \mathbf{e}.$$

- 9) Average rate of breakdown,

$$A_{br} = \delta_1 \sum_{i=0}^{\infty} \left(\sum_{j=1}^S z_{i,1,j} \right).$$

- 10) Average rate of repair,

$$A_{rr} = \delta_2 \sum_{i=0}^{\infty} \left(\sum_{j=1}^S z_{i,2,j} \right).$$

6. Profit Analysis

The long-run profit function for this model, defined as the total profit per unit time is given by

$$E_{profit} = c_0 E_{ds} - c_1 E_{inv} - c_2 E_{orbit} - c_3 E_{ro} - c_4 E_p - c_5 (E_{la} + E_{lr}) - c_6 A_{br} - c_7 A_{rr},$$

where c_0 =revenue per unit purchase, c_1 = holding cost of inventory/unit/unit time, c_2 = holding cost of customers/unit/unit time, c_3 =reorder cost, c_4 =cost due to decay of items/unit/unit time, c_5 =cost due to loss of customers/unit/unit time, c_6 =penalty due to breakdown of server/unit time, c_7 =repair cost/unit time.

7. Conclusion

In this paper, we studied a perishable inventory system with possible server interruption and an infinite orbit for accommodating retrial customers. A suitable profit function is constructed by using system's performance measures. This work can be extended by considering the addition of a buffer or incorporating server vacation periods to further enhance system efficiency.

REFERENCES

1. *Artalejo J.R., Krishnamoorthy A. , Lopez-Herrero M. J.* Numerical analysis of (s, S) inventory systems with repeated attempts // *Annals of Operations Research*, 2006. Vol. 141. № 1. P. 67–83.
2. *Bobina J.M., Jose K.P.* A Perishable Inventory System with Production and Unreliable Server // *Informational Technologies and Mathematical Modelling (ITMM-2023)*, Tomsk: Tomsk State University Publishing, 2023. P. 18–24.
3. *Fred Raafat* Survey of literature on continuously deteriorating inventory models // *Journal of the Operational Research society (Springer)*. 1991. Vol. 42. № 1. P. 27–37.
4. *Krishnamoorthy A., Sathian M. K., C Viswanath N.* Reliability Of a k-out-of-n System With A Single Server Extending Non-Preemptive Service-Part II. // *Reliability: Theory And Applications*. 2016. Vol. 11. № 3. P. 76–88.
5. *Neuts M.F., Rao B.* Numerical investigation of a multiserver retrial model // *Queueing systems (Springer)*. 1990. Vol. 7. № 2. P. 69–190.
6. *Reshmi P.S., Jose K.P.* A queueing inventory system with perishable items and retrial of customers // *Malaya Journal of Matematik*. 2019. Vol. 7. № 2. P. 165–170.

Bobina J. Mattam — Research Scholar, P.G.& Research Dept. of Mathematics, St. Peter's College, Kolenchery-682311, India. E-mail: *itsboby7@gmail.com*

K. P. Jose — Ph.D, Assoc. Prof. & Head, Post Graduate Dept. of Mathematics, St. Peter's College, Kolenchery-682311, India. E-mail: *kpjpsc@gmail.com*

SYNTHESIS OF A NEW MARKOV PROCESS BASED ON THE CHARACTERISTICS OF OTHER TWO PROCESSES

G. Tsitsiashvili

*Institute for Applied Mathematics, Far Eastern Branch of Russian Academy
Sciences, Vladivostok, Radio street 7, Russia*

The paper considers two models of queuing with a variable structure based on the introduction of additional transient intensities into known models or their combinations, which create stationary distributions convenient for calculation. In the first model, it is a probabilistic mixture of known stationary distributions with given weights. In the second model, this uniform distribution is repeatedly used in physical statistics. Both models are based on the selection of states, between which additional transient intensities are introduced. The algorithms used in the work for introducing new transient intensities are closely related to the concept of flow in a deterministic transport network. **Keywords:** *queuing systems, variable structure, transient intensities, mixture of marginal distributions.*

1. Introduction

Recently, due to the development of an information-theoretic approach to modeling self-organization processes in open physical systems [1], [2] with stochasticity, it has become possible to combine the dynamic and information (control) parts of their description into a single system. In this paper, this idea is the basis for building a queuing model with a variable structure. Two models based on the construction of discrete ergodic Markov processes with continuous time are used to construct a mathematical model of a queuing system with a variable structure. This idea is the basis for building a queuing model with a variable structure, considered in this paper. Therefore, it becomes necessary to try to extend the above synergistic approach to queuing systems. To construct a mathematical model of queuing systems with a variable structure, the construction of discrete ergodic Markov processes with continuous time is used, and the model itself is a combination of several queuing systems. In this case, the set of states of the combined

The research was carried out within the state assignment for IAM FEB RAS (N 075-00459-24-00).

system is the union of the sets of states of the combined systems. This is the difference from the classical queuing systems, in which the set of states of the system is a direct product of the sets of states of the combined systems.

The first model is a combination of several models of service systems. The set of states of a combined system is the union of sets of states of combined systems. The transient intensities between the states of different systems can be considered as an input control, which is selected so that the limiting distribution of the combined system is a mixture of the limiting distributions of the combined systems with different weights determined by the introduced transient intensities. In particular, when supercritical values of control parameters ("informants in the terminology of G. Haken [1]) are reached, forced transitions between stationary states occur in self-organizing open systems), in which ordered space-time macroscopic formations arise (dissipative structures) that persist only in the presence of pumping energy, matter, etc. from the environment.

The second model is based on the introduction of control in a single-channel queuing system with losses in order to design a queuing system with uniform distribution over a variety of states. To do this, we solve the problem of determining transient intensities in a single-channel queuing system with losses, with a finite number of waiting places and the possibility of cleaning and restoring the system from applications if all waiting places are filled. This queuing system is a special case of stochastic cleaning systems that accept and accumulate input data of random variables over random time intervals until pre-defined criteria are met. Then there is an instant cleanup of some or all of these inputs (see, for example, [3] - [5]). At the same time, there is a whole series of anatomical and psychological studies on the transformation of short-term memory into long-term memory with the participation of the hippocampus [6], [7].

2. Main results

We call a homogeneous Markov discrete process $X(t)$ with a counting set of states \mathcal{X} and transient intensities $\lambda(x, y)$, $x, y \in \mathcal{X}$ irreducible if there exists such a state $x^* \in \mathcal{X}$, which is reachable from any other state. I.e. for any $x \neq x^*$ or $\lambda(x, x^*) > 0$, or there will be such a chain of states x_1, \dots, x_n , that $\lambda(x, x_1) > 0$, $\lambda(x_1, x_2) > 0, \dots, \lambda(x_{n-1}, x_n) > 0$, $\lambda(x_n, x^*) > 0$.

Theorem 1. An irreducible homogeneous Markov process is ergodic if and only if the system of equations

$$\sum_{y \neq x} u(y) \lambda(y, x) = \sum_{y \neq x} u(x) \lambda(x, y), \quad x \in \mathcal{X} \quad (1)$$

has at least one solution $u(x)$, $x \in \mathcal{X}$, for which the sum $0 < \sum_{x \in \mathcal{X}} |u(x)| < \infty$.

The formulation of Theorem 1 was given in the classical monograph [8].

Using this theorem, the following theorem is proved, which is the main result of this paper. Let us first consider the model based on two discrete, homogeneous and irreducible (and therefore ergodic) Markov processes $X_k(t)$, $t \geq 0$, $k = 1, 2$, with counting sets of states \mathcal{X}_k , and with transient intensities $\lambda_k(x_k, y_k) \geq 0$, and stationary probabilities $P_k(x_k)$, $x_k, y_k \in \mathcal{X}_k$. From the stationary Kolmogorov-Chapman equations, the equalities follow

$$\sum_{y_k \in \mathcal{X}_k} P_k(x_k) \lambda_k(x_k, y_k) = \sum_{y_k \in \mathcal{X}_k} P_k(y_k) \lambda_k(y_k, x_k), \quad x_k \in \mathcal{X}_k. \quad (2)$$

Theorem 2. Suppose that transitional intensities are introduced between some states $x_1^* \in \mathcal{X}_1$, $x_2^* \in \mathcal{X}_2$ of these two processes $\Lambda(x_1^*, x_2^*) > 0$, $\Lambda(x_2^*, x_1^*) > 0$, satisfying for some c_1, c_2 , $0 < c_1, c_2 < 1$, $c_1 + c_2 = 1$, the conditions of local equilibrium $c_1 P_1(x_1^*) \Lambda(x_1^*, x_2^*) = c_2 P_2(x_2^*) \Lambda(x_2^*, x_1^*)$. Then the process $X(t)$ with a set of states $\mathcal{X}_1 \cup \mathcal{X}_2$ and transient intensities $\lambda_k(x_k, y_k)$, $x_k, y_k \in \mathcal{X}_k$, $k = 1, 2$, $\Lambda(x_1^*, x_2^*)$, $\Lambda(x_2^*, x_1^*)$, is also ergodic and its stationary probabilities have the form $\pi(x_k) = c_k P_k(x_k)$, $x_k \in \mathcal{X}_k$, $k = 1, 2$, and satisfy the stationary Kolmogorov-Chapman equations for $k, j = 1, 2$, $k \neq j$

$$\begin{aligned} \pi(x_k) \sum_{y_k \in \mathcal{X}_k} \lambda_k(x_k, y_k) &= \sum_{y_k \in \mathcal{X}_k} \pi(y_k) \lambda_k(y_k, x_k), \quad x_k \in \mathcal{X}_k, \quad x_k \neq x_k^*, \\ \pi(x_k^*) \left(\sum_{y_k \in \mathcal{X}_k} \lambda_k(x_k^*, y_k) + \Lambda(x_k^*, x_j^*) \right) &= \\ = \sum_{y_k \in \mathcal{X}_k} \pi(y_k) \lambda_k(y_k, x_k^*) + \pi(x_j^*) \Lambda(x_j^*, x_k^*). \end{aligned} \quad (3)$$

EXAMPLE 1. Suppose that the discrete Markov processes $X_1(t)$, $X_2(t)$, considered in the theorem 2 describe the number of applications in single-channel queuing systems $M|M|1|\infty$ with Poisson input streams with intensities λ_1, λ_2 and with service intensities μ_1, μ_2 , $\rho_1 = \lambda_1/\mu_1 < 1$, $\rho_2 = \lambda_2/\mu_2 < 1$. Then discrete Markov processes $X_1(t)$, $X_2(t)$, describing the number of customers in the first and second queuing systems have marginal distributions $P_1(x_1 = k) = (1 - \rho_1) \rho_1^k$, $P_2(x_2 = l) = (1 - \rho_2) \rho_2^l$, $k, l = 0, 1, \dots$. Select the states $x_1^* = 0$, $x_2^* = 0$, select c_1, c_2 , $0 < c_1, c_2$, $c_1 + c_2 = 1$ and determine the intensity of transitions $\Lambda_1 = \Lambda(x_1^* = 0, x_2^* = 0)$, $\Lambda_2 = \Lambda(x_2^* = 0, x_1^* = 0)$ from the condition

$c_1(1-\rho_1)\Lambda(x_1^* = 0, x_2^* = 0) = c_2(1-\rho_2)\Lambda(x_2^* = 0, x_1^* = 0)$. For this example, the positive intensities of transitions between states are determined by Fig. 1.

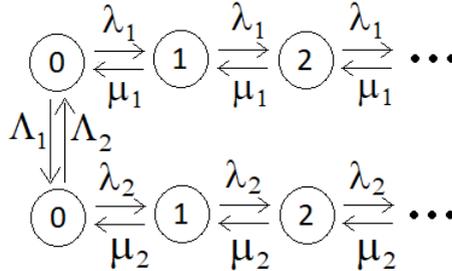


Fig. 1. Transient intensities in the queuing system from Example 1.

Let's now move on to the second model, based on a single-server queuing system with losses $M|M|1|m$ when the number of customers exceeds the upper limit of m . We describe this queuing system by a discrete Markov process $x(t)$ with states $0, 1, \dots, m$, characterizing the number customers in the system. Suppose that the Poisson input flow to this queuing system has an intensity λ , and the service intensity is μ . Then the transition intensities between the states of the system have the form $\gamma(i, i+1) = \lambda$, $\gamma(i+1, i) = \mu$, $i = 0, 1, \dots, m-1$.

Denote $\rho = \lambda/\mu$ coefficient of system loads, then stationary probability $p(i)$ oncoming systems in state i satisfies equalities

$$p(i) = \frac{(1-\rho)\rho^i}{1-\rho^{m+1}}, \rho \neq 1; \quad p(i) = \frac{1}{m+1}, \rho = 1; \quad i = 0, 1, \dots, m.$$

Let us now consider the problem of converting a system with losses $M|M|1|m$ into a system with a uniform distribution of stationary probabilities. We will introduce additional transition intensities into the $M|M|1|m$ system, characterizing the cleaning of the system from customers and the replenishment of the system with deleted customers. $\gamma(0, m) = \lambda$, $\gamma(m, 0) = \mu$.

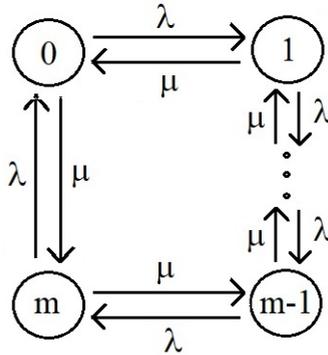


Fig. 2. Transient intensities in a system with clearance of customers and uniform stationary distribution.

An elementary calculation shows that the states $0, 1, \dots, m, 0$, form a cycle in which the sum of the intensities entering the state i , is equal to the sum of the intensities leaving the state (and is equal to $\lambda + \mu$). It follows that the stationary probabilities of being in the states of the extended system satisfy the equality

$$p(i) = \frac{1}{m+1}, \quad i = 0, 1, \dots, m, \quad (4)$$

Remark 1. It should be noted that the transition intensities $\gamma(0, m) = \mu$, $\gamma(m, 0) = \lambda$ introduced in Fig. 2 may be different. Indeed, at the vertices $1, \dots, m-1$ of a graph depicting the states of the $M|M|1|m$ system and its transient intensities, the difference between the sum of the output intensities and the sum of the input intensities is zero. At the vertex 0, this difference is equal to $\lambda - \mu$, and at the vertex m it is equal to $\mu - \lambda$. Therefore, introducing between the vertices 0, m transient intensities $\gamma(0, m) = \mu'$, $\gamma(m, 0) = \lambda'$, $\lambda + \mu' = \lambda' + \mu$, it is possible to obtain equality to zero of the corresponding differences. Moreover, with $\lambda > \mu$, you can select $\lambda' = \lambda - \mu + \varepsilon$, $\mu' = \varepsilon$, $\varepsilon > 0$. Similarly, if $\mu > \lambda$, then the transient intensities introducing between the vertices 0, m may be selected as $\gamma(0, m) = \varepsilon$, $\gamma(m, 0) = \mu - \lambda + \varepsilon$. The corresponding differences of transit intensities connecting with the states 0, m are equal to zero. It should be noted that in order to weaken the connection $0 \rightarrow m$, leading to frequent switches between these states 0, m , it is enough to require the condition $\varepsilon \ll 1$.

3. Discussion

The number of examples in the first model can be significantly expanded using well-known formulas for marginal distributions in various queuing systems, the number of applications in which the process of death and birth is described by [8]. Similarly, for this purpose, random processes can also be used that describe the number of requests in nodes of an open [9] or closed [10] queuing network, the marginal distributions in which obey multiplicative theorems.

The design of the managed queuing network given in the second model allows for numerous generalizations based on the model of a deterministic transport network with single input and output nodes [11]. By determining the total flow to the output node in a deterministic network and building an edge from the output node to the input node with this flow, it is possible to obtain a system with a uniform stationary distribution.

4. Conclusion

In this paper, two models of managed queuing systems with a variable structure are presented. The first model is based on a combination of several well-known models with switching between them. The second model is based on the deterministic transport network model and uses the introduction of feedback in a single-channel queuing system with losses. Stationary distributions are calculated for both models. For the first model, this distribution is a mixture of two stationary distributions with given weights. For the second model, this is a uniform distribution.

The research was carried out within the state assignment for IAM FEB RAS (N 075-00459-24-00). The author thanks Yu.N. Kharchenko for a help in physical interpretation of obtained results.

REFERENCES

1. *Haken H.* Information and Self-Organization: A Macroscopic Approach to Complex Systems. Springer Series in Synergetics, 2006. 196 p.
2. *Kolesnichenko A.V.* Continuum models of natural and space environments: Problems of thermodynamic design. Moscow: LENAND, 2017. 400 p.(In Russian).
3. *Ghosh S., Hassin R.* Inefficiency in stochastic queueing systems with strategic customers // European J. Oper. Res.. 2021. Vol. 295. № 1. P. 1–11.
4. *He Q.M., Bookbinder J.H., Cai Q.* Optimal policies for stochastic clearing systems with time-dependent delay penalties // Naval Res. Logist. 2020. Vol. 67. № 7. P. 487–502.

5. *Missbauer S.J., Stolletz R., Schneckenreither M.* Order release optimisation for time-dependent and stochastic manufacturing systems // *Int. J. Prod. Res.* 2024. Vol. 62. № 7. P. 1–20.
6. *Lupien N., Evans A. et al.* Hippocampal volume is as variable in young as in older adults: Implications for the notion of hippocampal atrophy in humans // *Neuroimage.* 2007. Vol. 34. № 2. P. 479–485.
7. *Squire L.R.* The legacy of patient H.M. for neuroscience // *Neuron.* 2009. Vol. 61. № 1. P. 6–9.
8. *Ivchenko G.I., Kashtanov V.A., Kovalenko I.N.* Theory of queuing: A textbook for universities. Moscow: Higher School, 1982. 255 p. (In Russian).
9. *Jackson J.R.* Networks of Waiting Lines // *Oper. Res.* 1957. Vol. 5. № 4. P. 518–521.
10. *Gordon K.D., Newell G.F.* Closed Queuing Systems with Exponential Servers // *Oper. Research.* 1967. Vol. 15. № 2. P. 254–265.
11. *Ford L.R., Fulkerson D.R.* Maximal flow through a network // *Canadian Journal of Mathematics.* 1956. Vol. 8. P. 399–404.
12. *Klimontovich Y.L.* Statistical theory of open systems. M.: Janus LLP, 1995. 624 p.

Tsitsiashvili Gurami — doctor of physical and mathematical sciences, professor, main researcher of the Institute for Applied Mathematics, Far Eastern Branch of Russian Academy Sciences. E-mail: guram@iam.dvo.ru

ANALYSIS OF A DISCRETE-TIME PERISHABLE INVENTORY MODEL WITH POSITIVE SERVICE AND LEAD TIMES UNDER MODIFIED N-POLICY FRAMEWORK

Jijo Joy¹, K. P. Jose²

¹*Post Graduate Department of Mathematics, St. Aloysius College Edathua, Alappuzha, Kerala, India*

²*Post Graduate Department of Mathematics, St. Peter's College Kolenchery, Ernakulam, Kerala, India*

In this article, we examine a discrete-time perishable inventory model with positive service and lead times under a modified N-policy framework. We define a suitable cost function for the model by assuming that the demand process follows a Bernoulli process and that both service and lead times are geometrically distributed. This model is formulated as a level-independent quasi-birth-death process. We investigate the model using the Matrix Analytic Method. **Keywords:** *discrete-time inventory, perishable products, n-policy, matrix analytic method, quasi-birth-death process.*

Introduction

Effective inventory management is vital for manufacturing and service companies to ensure smooth operations. Managing perishable inventory is particularly challenging due to the deterioration and limited shelf life of products such as food items, pharmaceuticals, and chemicals. This research examines a discrete-time perishable inventory model with positive service and lead times under a modified N-policy framework minimize total expected cost. For example, a grocery store handling perishable items like fruits, vegetables, and dairy products must manage these items efficiently to minimize waste, ensure availability, and maintain profitability. The model in the study determines optimal order quantities and timing, considering the positive service and lead times common in real-world scenarios where processing and delivering orders take time.

Key contributions to this field include Ghare and Schrader's (1963)[3] model for exponentially decaying inventory and Lian and Liu's (1999)[8]

The authors acknowledge the financial support provided by FIST program, Dept. of Science and Technology, Govt. of India, to the PG and Research Dept. of Mathematics through SR/FST/College-2018-XA 276(C).

work on approximating continuous-review models using discrete-time models. Viswanath and Krishnamoorthy (2011)[13] analyzed production inventory systems with service interruptions, highlighting disruption impacts. Balagopal et al. (2021)[2] studied discrete-time inventory systems with positive service and lead times, while Selvakumar and Thangaraj (2020)[9] used Markov Decision Processes to optimize systems with Bernoulli arrivals. Tan and Weng (2012)[12] explored perishable inventory control with constant deterioration and demand rates.

Jose and Anilkumar (2020)[1] and Shajin et al. (2018)[11] investigated queuing inventory models and service time dynamics, respectively. Sigman and Simchi-Levi (1992)[10] analyzed stochastic lead times, emphasizing inherent variability. The N -Policy, introduced by Yadin and Naor (1963)[14] and further developed by Kella (1989)[6], focuses on controlling service initiation and cessation. Krishnamoorthy and Anbazhagan (2008)[7] applied N -Policy to perishable inventory systems, showing its effectiveness in integrating perishability into decision-making. Jose and Anilkumar (2020)[5] demonstrated improved efficiency and cost-effectiveness by incorporating N -Policy in a discrete-time Geo/Geo/1 inventory system. Recent works by Jijo Joy and K. P. Jose (2020)[4] highlight the growing importance of perishable inventory models in modern inventory management.

1. Mathematical Modelling and Analysis

In this study, we model a discrete-time perishable inventory system with positive service and lead times under a modified N -policy framework. Customer inter-arrival times follow a geometric distribution with parameter a . The server goes on vacation when inventory is empty, and all remaining inventory is removed. Upon N customer arrivals, batch service of size N begins, with service times geometrically distributed with parameter b_2 . Additional arriving customers are served individually, with service times geometrically distributed with parameter b_1 . No customers enter when inventory is zero to prevent stockouts.

Inventory replenishment lead times follow a geometric distribution with parameter c , occurring at slot boundaries. Inventory perishability also follows a geometric distribution with parameter d . During batch service, inventory is restocked to maximum levels, ensuring demand is met.

$$A_4 = \begin{cases} \bar{a}b_2, & \text{for } i = 1, j = S - N + 2 \\ 0, & \text{otherwise} \end{cases},$$

$$A_1 = \begin{cases} \bar{a}\bar{b}_2, & \text{for } i = 1, j = 1 \\ [D_1], & \text{otherwise} \end{cases}, \quad A_0 = \begin{cases} \bar{a}\bar{b}_2, & \text{for } i = 1, j = 1 \\ [D_0], & \text{otherwise} \end{cases},$$

$$C_0 = (a, 0, \dots, 0), \quad C_1 = (\bar{a}),$$

$$D_1 = \begin{bmatrix} \bar{c} & & & & & & & c \\ \bar{c}ab_1 & \bar{c}\bar{a}\bar{b}_1 & & & & & & \beta \\ dab_1\bar{c} & \alpha & \bar{d}\bar{a}\bar{b}_1\bar{c} & & & & & \beta \\ & \ddots & \ddots & \ddots & & & & \vdots \\ & & dab_1\bar{c} & \alpha & \bar{d}\bar{a}\bar{b}_1\bar{c} & & & \beta \\ & & & dab_1 & \alpha_1 & \bar{d}\bar{a}\bar{b}_1 & & \\ & & & & & \ddots & \ddots & \\ & & & & & dab_1 & \alpha_1 & \bar{d}\bar{a}\bar{b}_1 \end{bmatrix},$$

$$A_2 = \begin{bmatrix} \bar{a}b_1\bar{c} & & & & & & & \bar{a}b_1c \\ \bar{d}\bar{a}\bar{b}_1\bar{c} & \bar{d}\bar{a}\bar{b}_1\bar{c} & & & & & & \bar{a}b_1c \\ & \ddots & \ddots & & & & & \vdots \\ & & \bar{d}\bar{a}\bar{b}_1\bar{c} & \bar{d}\bar{a}\bar{b}_1\bar{c} & & & & \bar{a}b_1c \\ & & & \bar{d}\bar{a}\bar{b}_1 & \bar{d}\bar{a}\bar{b}_1 & & & \\ & & & & & \ddots & \ddots & \\ & & & & & \bar{d}\bar{a}\bar{b}_1 & \bar{d}\bar{a}\bar{b}_1 & \end{bmatrix}.$$

2. Stability Condition and Steady State Analysis

Since customers are not allowed when the inventory is zero, the system considered is stable if $a \leq b_1$.

Matrix Analytic Method is used to analyze the model.

3. Performance Measures and Cost Function

Let

$$\pi_i = \begin{cases} (\pi_{i,0,0}, \pi_{i,1,0}, \pi_{i,1,1}, \dots, \pi_{i,1,s}), & \text{for } 1 \leq i \leq N, \\ (\pi_{i,2,S}, \pi_{i,1,0}, \pi_{i,1,1}, \dots, \pi_{i,1,S}), & \text{for } i \geq N + 1. \end{cases}$$

Some important performance measures of the system under steady state are

- 1) Expected Re order Rate, ERR , is given by

$$ERR = b_1 [\sum_{i=1}^{\infty} \pi_{i,1,s+1} + d \sum_{i=1}^{\infty} \pi_{i,1,s+2}],$$
- 2) Expected Inventory Level, EIL , is given by

$$EIL = \sum_{i=1}^{\infty} \sum_{k=1}^S k \pi_{i,1,k} + S \sum_{i=N+1}^{\infty} \pi_{i,2,S},$$
- 3) Expected Loss Rate of Customers, ELC , is given by

$$ELC = \sum_{i=1}^{\infty} a \pi_{i,1,0} + \sum_{i=0}^N a \pi_{i,0,0},$$
- 4) Expected Number of Customers waiting in the system for single service when the inventory level is zero, $EW0$, is given by

$$EW0 = \sum_{i=1}^{\infty} i \pi_{i,1,0},$$
- 5) Expected Number of Customers Waiting in the System for Single Service, ECQ , is given by

$$ECQ = \sum_{i=1}^{\infty} \sum_{k=0}^S (i-1) \pi_{i,1,k} + \sum_{i=N}^{\infty} (i-N) \pi_{i,2,S},$$
- 6) Expected Rate of Departure after Completing Service, EDS is given by

$$EDS = b_1 \sum_{i=1}^{\infty} \sum_{k=1}^S \pi_{i,1,k} + N b_2 \sum_{i=N}^{\infty} \pi_{i,2,S},$$
- 7) Expected Perishable Quantity, EPQ , is given by

$$EPQ = d \sum_{i=1}^{\infty} \sum_{k=1}^S k \pi_{i,1,k}.$$

Define the expected total cost of the system per unit time as

$$ETC = c_0(ERR) + \sum_{k=1}^s c(S-k)x_k c_1 + c_2(EIL) + c_3(ELC) + c_4(EW0) + c_5(ECQ) + c_6(EPQ) + c_7(EDS),$$

where x_k is the probability that the inventory level is k during single service, c_0 - the setup cost/order, c_1 - procurement cost/unit/unit time, c_2 - inventory holding cost/unit/unit time, c_3 - cost due to loss of customers/unit/unit time, c_4 - holding cost of a customer/unit time when the inventory level is zero, c_5 - holding cost of a customer/unit time when the server is busy, c_6 - expected perishability cost/unit item, c_7 - service cost/unit/unit time.

Conclusion

This study has analyzed a discrete-time perishable inventory model with positive service and lead times under a modified N-policy framework. The underlying model incorporates geometric distributions for inter-arrival times, service times, lead times, and perishability, providing a realistic representation of inventory dynamics. The modified N-policy effectively manages inventory and service processes, obtaining some relevant performance

measures and minimizing total expected cost. Future research will include extensive numerical experiments and the evaluation of the (s, S) policy pair to further refine and validate the model's applicability in various real-world scenarios.

REFERENCES

1. *Anilkumar M. P., Jose K. P.* A Stochastic optimization of local purchase quantities in a Geo/Geo/1 production inventory system // Communications in Computer and Information Science. 2020. Vol.1337. P. 206–220.
2. *Balagopal N., Selvakumar B.,Tangaraj P.* Analysis of discrete-time inventory system with positive service time and lead time under (s, S) and (s, Q) policies // Operations Research Letters. 2021. Vol. 49. № 4. P. 569–575.
3. *Ghare P. M., Schrader G. F.* A model for exponentially decaying inventory // Journal of Industrial Engineering. 1963. Vol. 14. № 3. P. 238–243.
4. *Jijo Joy, Jose K. P.* A Discrete time perishable inventory model with positive service time // Malaya Journal of Matematik. 2020. Vol. S. № 1. P. 39–44.
5. *Jose K. P., Anilkumar M. P.* A discrete time Geo/Geo/1 inventory system with modified N -Policy // Malaya Journal of Matematik. 2020. Vol. S No 1. P. 868–876.
6. *Kella O.* The threshold policy in the $M/G/1$ queue with server vacations // Nav. Res. Logist. 1989. Vol. 36. P. 111–123.
7. *Krishnamoorthy A., Anbazhagan N.* Perishable inventory system at service facility with N -Policy // Stoch. Anal. Appl. 2008. Vol. 26. № 1. P. 120–135.
8. *Lian Z., Liu L.* Approximation of Continuous-review inventory models using discrete-time models // Operations Research. 1999. Vol. 47. № 5. P. 709–721.
9. *Selvakumar B., Tangaran P.* Discrete time service facility system with Bernoulli arrivals under (s, S) policy: A Markov Decision Process Approach // Journal of Industrial and Management Optimization. 2020. Vol. 16. № 3. P. 1295–1314.
10. *Sigman K., Simchi-Levi D.* Analysis of Stochastic lead times in inventory models // Operations Research. 1992. Vol. 40. № 4. P. 760–773.
11. *Shajin D., Krishnamoorthy A.* Discrete queueing inventory models with positive service time and priority customers // International Journal of Production Economics. 201. Vol. 204. P. 246–259.
12. *Tan B., Weng Z.* Discrete-time inventory control for perishable products with constant deterioration and demand rates // European Journal of Operational Research. 2012. Vol. 219. № 2. P. 337–346.
13. *Viswanath A., Krishnamoorthy A.* Analysis of (s, S) production inventory system with positive service time and service interruption // Computers and Industrial Engineering. 2011. Vol. 61. № 3. P. 580–588.

14. *Yadin M., Naor P.* Queueing system with a removable service station // *Oper. Res.* 1963. Vol. 14. P. 393–405.

Jijo Joy — M. Sc., Assistant Professor, Department of Mathematics, St. Aloysius College, Edathua, Alappuzha, Kerala - 689573, India. E-mail: *jjoyjac@gmail.com*

K. P. Jose — M. Sc., Ph. D., Associate Professor and Head of the Department, Department of Mathematics, St. Peter's College, Kolenchery, Ernakulam, Kerala - 682311, India. E-mail: *kpjpsc@gmail.com*

TWO-WAY COMMUNICATION MAP/GI/1 RETRIAL QUEUE WITH MULTIPLE TYPES OF OUTGOING CALLS

O. Lizyura

Tomsk State University, Tomsk, Russian Federation

In this paper, we consider a retrial queue with two types of calls: incoming and outgoing calls. The input flow is Markovian arrival process and the service times are arbitrary distributed random variables. In its idle time the server can initiate the service itself and make outgoing calls. We consider the system with multiple types of outgoing calls. We use asymptotic analysis method under the condition of low rate of retrials to study such system and propose the approximation of the stationary probability distribution of the number of calls in the orbit. **Keywords:** *retrial queue, Markovian arrival process, two-way communication, incoming calls, outgoing calls, diffusion analysis.*

Introduction

Retrial queues are the models of telecommunication and service systems [2, 6]. They differ from other queueing models by the following behavior: customers who cannot occupy a server upon arrival join the orbit and retry to access the server after some random delay. Retrial queues with two-way communication are used for modeling blended call centers, where the operator can provide both inbound and outbound calls. In paper [1], the authors describe the impact of using retrial queues as models of call centers. Studies [4, 5, 7] are devoted to the QoS requirements in telephone services. In [3], the authors provide the investigation of input processes in real call centers.

In literature on retrial queues with two-way communication the models with both Markovian arrival process and arbitrary distribution of service times are not considered due to their complex structure. The main complication is that the analysis requires considering four-dimensional Markov chain. Therefore, we use the asymptotic-diffusion analysis method and build a modification of it for MAP/GI/1 type of retrial queues with two-way communication.

Similar asymptotic results without using diffusion analysis are presented by Sakurai and Phung-Duc in [9]. They study asymptotic behavior of

M/G/1 retrial queue with outgoing calls. Nazarov et al. [8] study the slow retrial limit condition for Markovian retrial queues.

1. Model Description

We consider single server retrial queue with MAP input. The MAP is driven by Markov chain $m(t)$ with finite set of states $1, 2, \dots, M$ defined by the generator $\mathbf{Q} = [q_{vm}]$. Let the diagonal matrix $\mathbf{\Lambda}$ denotes of the intensities λ_m , where λ_m is the rate of arrivals in m -th state of the MAP. The matrix $\mathbf{D} = [d_{vm}]$ of the probabilities of the arrival at the moment of the underlying Markov chain transition from state v to state m . Service times of incoming calls follow an arbitrary distribution with cumulative distribution function $B_1(x)$. If the server is idle at the instant of arrival, then an incoming call occupies it for service.

If the server is busy upon arrival, the incoming call joins the orbit. Incoming calls in orbit make an exponentially distributed delay and reattempt to access the server. The rate of retrials is σ for each incoming call in the orbit. After delay, the incoming call behaves the same way as if it first arrived at the system. Thus, it can receive service or make a delay again.

When the server is idle, it makes outgoing calls. We consider the system with multiple types of outgoing calls. The rate of making outgoing calls of type n is equal to α_n . Service times of outgoing calls of type n are defined by distribution function $B_n(x)$. We assume $n = 2, 3, \dots, N$ to simplify derivations.

We denote $i(t)$ as the number of incoming calls in the orbit at instant t . Process $k(t)$ reflects the state of the server at instant t as follows:

$$k(t) = \begin{cases} 0, & \text{if the server is idle,} \\ 1, & \text{if an incoming call is in service,} \\ n, & \text{if an outgoing call of type } n \text{ is in service, } n = \overline{2, N}. \end{cases}$$

We also consider the residual time of service $z(t)$, which occurs when the server is busy. Thus, the random process $\{k(t), i(t), m(t), z(t)\}$ with a variable number of components is under investigation.

We denote

$$P\{k(t) = 0, i(t) = i, m(t) = m\} = P_0(i, m, t),$$

$$P\{k(t) = k, i(t) = i, m(t) = m, z(t) < z\} = P_k(i, m, z, t)$$

and introduce partial characteristic functions

$$H_0(u, m, t) = \sum_{i=0}^{\infty} e^{ju_i} P_0(i, m, t),$$

$$H_k(u, m, z, t) = \sum_{i=0}^{\infty} e^{ju_i} P_k(i, m, z, t), \quad k = 1, \dots, N,$$

where j is the imaginary unit. Deriving the system of Kolmogorov differential equations, we then transform it to the system for the partial characteristic functions, which yields

$$\begin{aligned} \frac{\partial \mathbf{H}_0(u, t)}{\partial t} &= \mathbf{H}_0(u, t) \left(\mathbf{Q} - \mathbf{A} - \sum_{n=2}^N \alpha_n \mathbf{I} \right) + j\sigma \frac{\partial \mathbf{H}_0(u, t)}{\partial u} + \\ &\quad + \sum_{k=1}^N \frac{\partial \mathbf{H}_k(u, 0, t)}{\partial z}, \\ \frac{\partial \mathbf{H}_1(u, z, t)}{\partial t} - \frac{\partial \mathbf{H}_1(u, z, t)}{\partial z} &= \mathbf{H}_1(u, z, t) (\mathbf{Q} + (e^{ju} - 1) \mathbf{A}) - \\ &\quad - \frac{\partial \mathbf{H}_1(u, 0, t)}{\partial z} + B_1(z) \mathbf{H}_0(u, t) \mathbf{A} - j\sigma e^{-ju} B_1(z) \frac{\partial \mathbf{H}_0(u, t)}{\partial u}, \\ \frac{\partial \mathbf{H}_n(u, z, t)}{\partial t} - \frac{\partial \mathbf{H}_n(u, z, t)}{\partial z} &= \mathbf{H}_n(u, z, t) (\mathbf{Q} + (e^{ju} - 1) \mathbf{A}) - \\ &\quad - \frac{\partial \mathbf{H}_n(u, 0, t)}{\partial z} + \alpha_n B_n(z) \mathbf{H}_0(u, t), \quad n = 2, \dots, N, \end{aligned} \quad (1)$$

where

$$\begin{aligned} \mathbf{H}_0(u, t) &= \{H_0(u, 1, t), H_0(u, 2, t), \dots, H_0(u, M, t)\}, \\ \mathbf{H}_k(u, z, t) &= \{H_k(u, 1, z, t), H_k(u, 2, z, t), \dots, H_k(u, M, z, t)\}, \\ \mathbf{A} &= \mathbf{\Lambda} + \mathbf{D} \circ \mathbf{Q}, \end{aligned}$$

and \mathbf{I} is the identity matrix. Also, we derive the additional equation obtained by summing up the equations of the system and taking the limit by $z \rightarrow \infty$.

$$\frac{\partial \mathbf{H}(u, t)}{\partial t} \mathbf{e} = (e^{ju} - 1) \left\{ j\sigma e^{-ju} \frac{\partial \mathbf{H}_0(u, t)}{\partial u} \mathbf{e} + (\mathbf{H}(u, t) - \mathbf{H}_0(u, t)) \mathbf{A} \mathbf{e} \right\}, \quad (2)$$

where $\mathbf{H}(u, t) = \mathbf{H}_0(u, t) + \sum_{k=1}^N \mathbf{H}_k(u, \infty, t)$, \mathbf{e} is the vector of ones.

Using the diffusion analysis method under low rate of retrials condition and the system above, the diffusion approximation of the number of calls in the orbit is derived.

2. Diffusion Approximation for the Number of Calls in the Orbit

From system (1) and equation (2), the following equation was derived:

$$\frac{\partial \Phi(w, \tau)}{\partial \tau} = w \frac{\partial \Phi(w, \tau)}{\partial w} a'(x) + \frac{(jw)^2}{2} \Phi(w, \tau) b(x). \quad (3)$$

Here $\Phi(w, \tau)$ is the asymptotic characteristic function of some diffusion process approximating the number of calls in the orbit of considered system, $a(x)$ is the drift coefficient of the diffusion process and $b(x)$ is the diffusion coefficient. $a(x)$ is given by

$$a(x) = -x\mathbf{R}_0(x)\mathbf{e} + (\mathbf{R} - \mathbf{R}_0(x))\mathbf{A}\mathbf{e}, \quad (4)$$

vectors $\mathbf{R}_k(x)$ are the solution of the following system:

$$\begin{aligned} \mathbf{R}_1(x)\mathbf{Q}\mathbf{v}_k + (1 - B_1^*(s_k))\mathbf{R}_0(x)(\mathbf{A} + x\mathbf{I})\mathbf{v}_k &= 0, \\ \mathbf{R}_1(x)\mathbf{e} &= b_1\mathbf{R}_0(x)(\mathbf{A} + x\mathbf{I})\mathbf{e}, \\ \mathbf{R}_n(x)\mathbf{Q}\mathbf{v}_k + \alpha_n(1 - B_n^*(s_k))\mathbf{R}_0(x)\mathbf{v}_k &= 0, \quad n = \overline{2, N}, \\ \mathbf{R}_n(x)\mathbf{e} &= \alpha_n b_n \mathbf{R}_0(x)\mathbf{e}, \\ \sum_{k=0}^N \mathbf{R}_k(x) &= \mathbf{R}. \end{aligned}$$

Here vectors $\mathbf{R}_k(x)$, $k = \overline{0, N}$ are the stationary distribution of two-dimensional process $\{k(t), m(t)\}$, s_m and \mathbf{v}_m are eigenvalues and eigenvectors of matrix $-\mathbf{Q}$, b_n is the mean service time for calls of type n . Vector \mathbf{R} is the stationary distribution of underlying process $m(t)$, which is the solution of system $\mathbf{R}\mathbf{Q} = \mathbf{0}$, $\mathbf{R}\mathbf{e} = 1$. Functions $B_k^*(s)$ are the Laplace-Stieltjes transforms of $B_k(x)$. $b(x)$ is given by

$$b(x) = a(x) + 2[-x\mathbf{g}_0\mathbf{e} + (\mathbf{g} - \mathbf{g}_0)\mathbf{A}\mathbf{e} + x\mathbf{R}_0\mathbf{e}], \quad (5)$$

vectors $\mathbf{g}_k(x)$ are the solution of system

$$\begin{aligned} \mathbf{g}_1(x)\mathbf{Q}\mathbf{v}_m + \mathbf{g}_0(x)(\mathbf{A} + x\mathbf{I})(1 - B_1^*(s_m))\mathbf{v}_m &= x(1 - B_1^*(s_m))\mathbf{R}_0(x)\mathbf{v}_m - \\ &- (\mathbf{R}_1(x) - \mathbf{R}_1^*(s_m))(\mathbf{A} - a(x)\mathbf{I})\mathbf{v}_m, \\ \mathbf{g}_1(x)\mathbf{e} - b_1\mathbf{g}_0(x)(\mathbf{A} + x\mathbf{I})\mathbf{e} &= -b_1x\mathbf{R}_0\mathbf{e} - \frac{d\mathbf{R}_1^*(0)}{ds}(\mathbf{A} - a(x)\mathbf{I})\mathbf{e}, \\ \mathbf{g}_n(x)\mathbf{Q}\mathbf{v}_m + \alpha_n(1 - B_n^*(s_m))\mathbf{g}_0(x)\mathbf{v}_m &= \end{aligned}$$

$$\begin{aligned}
&= -(\mathbf{R}_n(x) - \mathbf{R}_n^*(s_m))(\mathbf{A} - a(x)\mathbf{I})\mathbf{v}_m, \quad n = \overline{2, N}, \\
\mathbf{g}_n(x)\mathbf{e} - \alpha_n b_n \mathbf{g}_0(x)\mathbf{e} &= -\frac{d\mathbf{R}_n^*(0)}{ds}(\mathbf{A} - a(x)\mathbf{I})\mathbf{e}, \quad n = \overline{2, N}, \\
\sum_{k=0}^N \mathbf{g}_k(x)\mathbf{e} &= 0,
\end{aligned}$$

where s_m and \mathbf{v}_m are eigenvalues and eigenvectors of matrix $-\mathbf{Q}$. Vectors $\mathbf{R}_n^*(s_m)$, $m = \overline{2, M}$, $n = \overline{1, N}$ satisfy the system

$$\begin{aligned}
\mathbf{R}_1^*(s_m, x)(\mathbf{Q} + s_m\mathbf{I}) &= \mathbf{R}_1(x)\mathbf{Q} + (1 - B_1^*(s_m))\mathbf{R}_0(x)(\mathbf{A} + x\mathbf{I}), \\
\mathbf{R}_1^*(s_m, x)\mathbf{v}_m &= -B_1^{*'}(s_m)\mathbf{R}_0(x)(\mathbf{A} + x\mathbf{I})\mathbf{v}_m, \\
\mathbf{R}_n^*(s_m, x)(\mathbf{Q} + s_m\mathbf{I}) &= \mathbf{R}_n(x)\mathbf{Q} + \alpha_n(1 - B_n^*(s_m))\mathbf{R}_0(x), \\
\mathbf{R}_n^*(s_m, x)\mathbf{v}_m &= -\alpha_n B_n^{*'}(s_m)\mathbf{R}_0(x)\mathbf{v}_m.
\end{aligned}$$

Vectors $-\frac{d\mathbf{R}_n^*(0)}{ds}$, $n = \overline{1, N}$ are the partial means of three-dimensional process $\{k(t), m(t), z(t)\}$ and satisfy the system

$$\begin{aligned}
-\frac{d\mathbf{R}_1^*(0)}{ds}\mathbf{Q} &= \mathbf{R}_1 - b_1\mathbf{R}_0(\mathbf{A} + x\mathbf{I}), \\
-\frac{d\mathbf{R}_1^*(0)}{ds}\mathbf{e} &= \frac{1}{2}b_1^{(2)}\mathbf{R}_0(\mathbf{A} + x\mathbf{I})\mathbf{e}, \\
-\frac{d\mathbf{R}_n^*(0)}{ds}\mathbf{Q} &= \mathbf{R}_n - \alpha_n b_n \mathbf{R}_0, \quad n = \overline{2, N}, \\
-\frac{d\mathbf{R}_n^*(0)}{ds}\mathbf{e} &= \frac{1}{2}\alpha_n b_n^{(2)}\mathbf{R}_0\mathbf{e}, \quad n = \overline{2, N}.
\end{aligned}$$

Applying to equation (3) the inverse Fourier transform, we obtain the Fokker–Planck equation for density function $P(y, \tau)$ of some diffusion process $y(\tau)$

$$\frac{\partial P(y, \tau)}{\partial \tau} = -a'(x)\frac{\partial(yP(y, \tau))}{\partial y} + \frac{b(x)}{2}\frac{\partial^2 P(y, \tau)}{\partial y^2}. \quad (6)$$

Thus, process $y(\tau)$ is the solution of stochastic differential equation

$$dy(\tau) = a'(x)y(\tau)d\tau + \sqrt{b(x)}d\omega(\tau), \quad (7)$$

where $\omega(\tau)$ is Wiener process.

We introduce diffusion process $z(\tau) = x(\tau) + \varepsilon y(\tau)$, where $\varepsilon = \sqrt{\sigma}$ and function $x(\tau)$ is the solution of differential equation $dx(\tau) = a(x)d\tau$. Then process $z(\tau)$ is the solution of the following stochastic differential equation:

$$dz(\tau) = [a(x) + \varepsilon a'(x)y(\tau)]d\tau + \varepsilon\sqrt{b(x)}d\omega(\tau).$$

We transform the coefficients of the obtained equations as follows:

$$a(x) + \varepsilon a'(x)y = a(x + \varepsilon y) + O(\varepsilon^2) = a(z) + O(\varepsilon^2),$$

$$\varepsilon\sqrt{b(x)} = \varepsilon\sqrt{b(x + \varepsilon y - \varepsilon y)} = \varepsilon\sqrt{b(z - \varepsilon y)} = \varepsilon\sqrt{b(z)} + O(\varepsilon^2),$$

which yields the stochastic differential equation

$$dz(\tau) = a(z)d\tau + \varepsilon\sqrt{b(z)}d\omega(\tau). \quad (8)$$

Hence, we can write the Fokker–Planck equation for density function $\Pi(z, \tau)$ of process $z(\tau)$

$$\frac{\partial \Pi(z, \tau)}{\partial \tau} = -\frac{\partial(a(z)\Pi(z, \tau))}{\partial z} + \frac{\varepsilon^2}{2} \frac{\partial^2(b(z)\Pi(z, \tau))}{\partial z^2}. \quad (9)$$

Making reverse substitution $\sigma = \varepsilon^2$ and transform the equation for the stationary distribution of process $z(\tau)$

$$-(a(z)\Pi(z))' + \frac{\sigma}{2}(b(z)\Pi(z))'' = 0.$$

It is easy to see that the solution of the last equation is given by

$$\Pi(z) = \frac{C}{b(z)} \exp \left\{ \frac{2}{\sigma} \int_0^z \frac{a(x)}{b(x)} dx \right\},$$

where C is an integration constant.

Based on the obtained distribution, we build the approximation of probability distribution of the number of incoming calls in the orbit

$$PD(i) = \frac{\Pi(i\sigma)}{\sum_{n=0}^{\infty} \Pi(n\sigma)}. \quad (10)$$

3. Conclusion

We have considered MAP/GI/1 retrial queue with two-way communication and multiple types of outgoing calls. The state of the system is four-dimensional Markov chain, which we investigated using asymptotic-diffusion approach. We have derived drift and diffusion coefficients of diffusion process approximating the number of incoming calls in the orbit. Based on the density function of diffusion process, we have build the approximation (10) of the steady state probability distribution of the number of calls in the orbit.

REFERENCES

1. *Aguir S., Karaesmen F., Aksin O. Z., Chauvet F.* The impact of retrials on call center performance // OR Spectrum. 2004. Vol. 26. № 3. P. 353–376.
2. *Artalejo J. R., Karaesmen F., Aksin O., Chauvet F.* A classified bibliography of research on retrial queues: progress in 1990–1999 // Top. 1999. Vol. 7. № 2. P. 187–211.
3. *Avramidis A., Deslauriers A., L'Ecuyer P.* Modeling daily arrivals to a telephone call center // Management Science. 2004. Vol. 50. № 7. P. 896–908.
4. *Bernett H. G., Fischer M. J., Masi D. M. B.* Blended call center performance analysis // IT professional. 2002. Vol. 4. № 2. P. 33–38.
5. *Bhulai S., Koole G.* A queueing model for call blending in call centers // IEEE Transactions on Automatic Control. 2003. Vol. 48. № 8. P. 1434–1438.
6. *Falin G., Templeton J. G. C.* Retrial queues. London: CRC Press, 1997. 320 p.
7. *Gilmore A., Moreland L.* Call centres: how can service quality be managed? // Irish Marketing Review. 2000. Vol. 13. № 1. P. 3–11.
8. *Nazarov A., Phung-Duc T., Paul S.* Slow retrial asymptotics for a single server queue with two-way communication and markov modulated poisson input // Journal of Systems Science and Systems Engineering. 2019. Vol. 28. № 2. P. 181–193.
9. *Sakurai H., Phung-Duc T.* Scaling limits for single server retrial queues with two-way communication // Annals of Operations Research. 2016. Vol. 247. № 1. P. 229–256.

MULTI-SERVER JOB MODEL: LUMPING AND NUMERICAL ANALYSIS

S. Astafiev

Institute of Applied Mathematical Research, KarRC, Petrozavodsk, Russia

Present paper studies a model of a multi-server system that works at several processing speeds with randomized speed scaling policy upon customer arrivals and departures. This system is also known as Cluster Model or Multi-server Job Model. The possibility of reducing the phase space size of the model is substantiated. An implementation in the R language is presented for the numerical compilation of the generator matrix of the Markov chain and its subsequent application to calculation of the stationary distribution.

Keywords: *supercomputer model, matrix-analytic method, quasi-birth-death process, numerical solution.*

Introduction

Most modern computing systems have many processors and their incoming jobs can use several cores at once. Supercomputers are one of these classes which is widely used for tasks requiring large computing resources and have massive energy consumption [1]. There are several energy saving methods, one of which is frequency scaling [2]. Using frequency scaling implies that the computing cluster (i.e. system) switches between different speed modes that determine performance and power consumption. Present paper discusses simultaneous frequency switching between predetermined values on arrival and departure events with certain probabilities.

The present paper is concentrated on a method (somewhat similar to state space lumping [3]) for decreasing the (large in general) size of phase space of the Quasi-Birth-Death (QBD) continuous-time Markov chain that describes behavior of the supercomputer model [9]. This allows one to compute the stationary distribution and some related characteristics such as distribution of the number of customers or mean power consumption. Our method allows to construct infinitesimal generator matrix algorithmically with reduced phase space at once.

Author thanks Alexander Rumyantsev for constructive criticism.

1. Model

We briefly describe the Multi-server Job model (MJM) [9] below. The system consists of c identical servers and a FCFS queue. Arrivals follow Poisson process with rate λ . The k -th customer has two parameters:

- A random number of required servers N_k (class), drawn from a discrete distribution $\mathcal{P} = \{(j, p_j) : 1 \leq j \leq c, p_j = \mathbb{P}(N_k = j)\}$. Customer k is served by exactly N_k customers (seized/released at once).
- A random exponentially distributed job amount S_k (with mean $\mu_{N_k}^{-1}$).

The system works in one of n regimes (modes) which affects performance and power consumption. Regime i , where $1 \leq i \leq n$, means that all servers run at speed (frequency) f_i . Switching from regime i to j occurs when customers arrive or depart, with probabilities $p_{i,j}^{(a)}$ and $p_{i,j}^{(d)}$, respectively. Such systems can be studied as a QBD, which is shown in the next section.

2. Quasi-birth-death process

Models similar to the one presented in Section 1 can be studied as the well-known QBD processes, see e.g. [4]. Consider the random process $X(t) = (x(t), \varphi(t))_{t \geq 0}$ in continuous time $t \in \mathbb{R}_0$ where:

- $x(t)$ is a *level* (enumerable), e.g. number of customers in the system ($|x(t_2) - x(t_1)| \leq 1$ where $t_1 < t_2$ are the epochs of consecutive events);
- $\varphi(t)$ is a *phase* (finite), i.e. internal state of system. We call the phase space the set of all possible states of the $\varphi(t)$.

For more details on QBD see e.g. [5, pp. 117-132].

Evolution of such process describes by equation $\frac{d\pi(t)}{dt} = \pi(t)Q$ where Q is a three-diagonal infinitesimal generator matrix. The stationary distribution π is expressed from an infinite system of linear equations $\pi Q = \mathbf{0}$ and can be calculated by using well-known Matrix-Analytic Method (MAM) [7]. The analytical solution in MAM method is usually possible when number of phases is small (see [8], [9]). Since one of the main problems in analyzing QBD processes is the curse of dimensionality, we propose compact phase description in the next section.

3. Description of the phase

There are several ways to describe the phase $\varphi(t)$ of process $X(t)$. Further, the notation in the form of tuples is used to describe the phase. Since the set of possible phase states depends on the level $x(t)$ and its cardinality increases with increasing level (from 0 to c), in this section we consider the set of phase states at level $x(t) = c$.

Consider the pair (\mathbf{m}, \mathbf{u}) first, where \mathbf{m} is a current regime, \mathbf{u} is a class for the first c customers in the system. This method is used for explicit expression of the stationary distribution in [9] for small systems ($c \leq 2$). Size of the phase space is $|\{(\mathbf{m}, \mathbf{u}) : \mathbf{m} \in 1..n, \mathbf{u} \in \{1..c\}^{\times c}\}| = nc^c$ for levels greater than or equal to c .

Present paper is focused on the pair (\mathbf{m}, \mathbf{s}) where \mathbf{m} is a current regime, \mathbf{s} is number of customers at service of each class. It was first used for a single-speed version of MJM in [6] for obtaining stability condition. It's easy to understand that $|\{(\mathbf{m}, \mathbf{s}) : \mathbf{m} \in 1..n, \mathbf{s} \in \{0..c\}^{\times c} : 0 < \sum i\mathbf{s}_i \leq c\}| \leq nc^c$ but an explicit expression for size of this set is unknown (see [10, A026905] and [10, A000041]). At the time of writing, to the best of our knowledge, this is a description of the phase with the minimum set of states, for which it was possible to obtain transition matrices. Further in this section theoretical foundations are described that allow one to move from (\mathbf{m}, \mathbf{u}) description to (\mathbf{m}, \mathbf{s}) .

3.1. Reducing phase description

Consider the description of the first type, (\mathbf{m}, \mathbf{u}) . The article [4] demonstrates that this description can be reduced: the distribution of customers classes is the same for any customer in the queue, except for the first one. We denote such state description as $(\mathbf{m}, \hat{\mathbf{u}}, \mathbf{q})$ where \mathbf{m} is a current regime, $\hat{\mathbf{u}}$ is a list of currently served customers and \mathbf{q} is a first customer in queue. Note that due to the memoryless property [11, pp. 191] of the exponential distribution, the order in the vector $\hat{\mathbf{u}}$ is not important.

Let $\xi_1 \sim \text{Exp}(\lambda_1)$ and $\xi_2 \sim \text{Exp}(\lambda_2)$, then $\text{min}(\xi_1, \xi_2) \sim \text{Exp}(\lambda_1 + \lambda_2)$ (see e.g. [11, p. 193]). This property allows to collect individual customers of the same class together and get the form $(\mathbf{m}, \mathbf{s}, \mathbf{q})$, where \mathbf{m} , \mathbf{s} and \mathbf{q} were described earlier. Note that \mathbf{q} has distribution \mathcal{P} for some subsets from $\{(\mathbf{m}, \mathbf{s}, \mathbf{q})\}$ (all servers busy or queue is empty) with fixed \mathbf{m} and \mathbf{s} . We use notation $\mathbf{q} = 0$ for such cases. Further reduction is possible due to the removal of the \mathbf{q} component, which is based on the following theorem:

Theorem 1. Let the initial state $X(0)$ of the Markov chain $X(t) = (x(t), \varphi(t))$ where $\varphi(t) \in \{(\mathbf{m}, \mathbf{s}, \mathbf{q})\}$ be such that $\varphi(0) = (m(0), s(0), 0)$ (i.e. $q(0) = 0$). Consider the set of all possible phase states $\{(\mathbf{m}', \mathbf{s}', \mathbf{q}')\}$ at time t where \mathbf{m}', \mathbf{s}' are fixed. Let $v = c - \sum_{i=1}^c i\mathbf{s}'_i$ i.e. is a number of idle servers. Then $\mathbb{P}(\mathbf{q}' = i | \mathbf{s}') = \mathbb{P}(\mathbf{q}' = i | v) = \frac{p_i}{w}$ for all $v < i \leq c$ where $w = \sum_{i=v+1}^c p_i$.

Proof. To prove Theorem 1 we consider changes of component \mathbf{q} and number of free servers step by step, when customer leaves or tries to start service. For convenience, the proof is divided into two parts:

I) Let the class of customer that tries to enter service have the initial distribution \mathcal{P} , i.e. $q^{(0)} = 0$ where $q^{(0)}$ is a state of component \mathbf{q} before trial.

Let $v^{(0)}$ is a number of idle servers before trial, i.e. $v^{(0)} = c - \sum_{i=1}^c i s_i^{(0)}$ where $s^{(0)} \in \mathfrak{s}$ are currently served customers. Then

$$q^{(1)} = \begin{cases} 0, & \text{w.p. } \sum_{i=1}^{v^{(0)}} p_i, \\ k, & \text{w.p. } p_k, k > v^{(0)}, \end{cases}, v^{(1)} = \begin{cases} v^{(0)} - k, & \text{w.p. } p_k, k \leq v^{(0)}, \\ v^{(0)}, & \text{w.p. } \sum_{i=v^{(0)}+1}^c p_i, \end{cases},$$

where $q^{(1)}, v^{(1)}$ is the class of first customer in queue and number of idle servers respectively after trial. Then, by the Bayes theorem [11, pp. 79-86]:

$$\mathbb{P}(q^{(1)} = k | v^{(1)} = v^{(0)}) = \frac{\mathbb{P}(v^{(1)} = v^{(0)} | q^{(1)} = k) \mathbb{P}(q^{(1)} = k)}{\mathbb{P}(v^{(1)} = v^{(0)})} = \frac{p_k}{\sum_{i=v^{(0)}+1}^c p_i}, \quad (1)$$

where $k > v^{(1)}$. $\mathbb{P}(v^{(1)} = v^{(0)} | q^{(1)} = k) = 1$ because customers come to service in sequence, and $q^{(1)} \neq 0$ means that a next customer joined/remained in the queue. Denote $\mathcal{D}(v) = \{(k, d_k) : v < k \leq c, d_k = p_k / \sum_{j=v+1}^c p_j\}$. Note that $\mathcal{D}(0) = \mathcal{P}$ which corresponds to all busy servers.

II) Let $\bar{v}^{(-1)}$ and $\bar{v}^{(0)}$ servers were idle before and after customer leaving the system, respectively. Let before leaving queue was not empty and $\bar{q}^{(-1)} \sim \mathcal{D}(\bar{v}^{(-1)})$. Then after customer leaving becomes: $\bar{q}^{(0)} \sim D(\bar{v}^{(-1)})$ and

$$\bar{q}^{(1)} = \begin{cases} 0, & \text{w.p. } \sum_{i=\bar{v}^{(0)}+1}^{\bar{v}^{(0)}} d_i, \\ k, & \text{w.p. } d_k, k > \bar{v}^{(0)}, \end{cases}, \bar{v}^{(1)} = \begin{cases} \bar{v}^{(0)}, & \text{w.p. } \sum_{i=\bar{v}^{(0)}+1}^c d_i, \\ \bar{v}^{(0)} - k, & \text{w.p. } d_k, k \leq \bar{v}^{(0)}, \end{cases},$$

where $\bar{q}^{(1)}, \bar{v}^{(1)}$ are the class of first customer in queue and the number of idle servers respectively after trial. Then, by the Bayes theorem:

$$\mathbb{P}(\bar{q}^{(1)} = k | \bar{v}^{(1)} = \bar{v}^{(0)}) = \frac{\mathbb{P}(\bar{v}^{(1)} = \bar{v}^{(0)} | \bar{q}^{(1)} = k) \mathbb{P}(\bar{q}^{(1)} = k)}{\mathbb{P}(\bar{v}^{(1)} = \bar{v}^{(0)})} = \frac{d_k}{\sum_{i=\bar{v}^{(0)}+1}^c d_i} = \frac{p_k}{\sum_{j=\bar{v}^{(-1)}+1}^c p_j} \div \sum_{i=\bar{v}^{(0)}+1}^c \frac{p_i}{\sum_{j=\bar{v}^{(-1)}+1}^c p_j} = \frac{p_k}{\sum_{i=\bar{v}^{(1)}+1}^c p_i}.$$

That is, if not enough servers are idle after a customer leaves, then $\mathcal{D}(\bar{v}^{(1)})$ is the distribution of the first customer in the queue, and otherwise its distribution is reduced to the first case considered. The probability $\mathbb{P}(\bar{v}^{(1)} = \bar{v}^{(0)} | \bar{q}^{(1)} = k) = 1$ because $k > 0$ means that the next client remains in the queue and the number of idle servers has not changed. This completes the proof of the Theorem 1. ■

4. Implementation

To test the functionality of the concept, algorithms were created for the numerical construction of the matrix Q with a phase description in the form (\mathbf{m}, \mathbf{s}) . The R language¹ implementation is available as `MJMrrss` package². The matrix R that needs for MAM method was calculated by logarithmic reduction algorithm [12, pp. 188-189].

To validate the calculation results, a discrete-event simulation model of the system was used. Confidence intervals for the probabilities of being at each level were calculated by using the regenerative estimation method [13]. In total, 10^3 independent trajectories were launched with a length of 10^4 events of the customers arrivals.

The distribution of the number of customers in the system with 16th servers, 2 regimes and 4 classes is showed at Figure 1 (only for levels 0..20). Normalized root-mean-square deviation (NRMSD) is approximately 0.433420%. Figure 2 shows the relative deviations of the 99% confidence intervals for mean values obtained from the simulation.

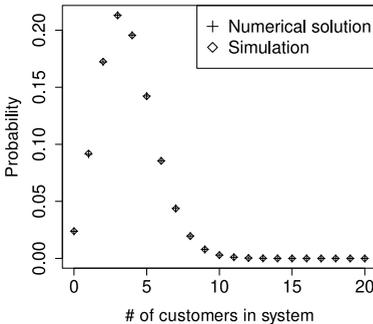


Figure 1. Distribution of the number of customers in the system.

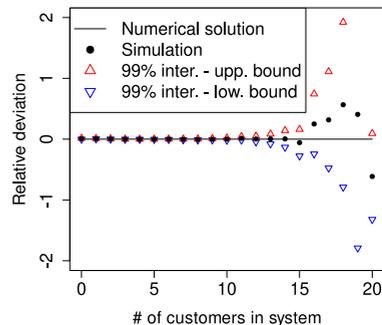


Figure 2. Relative deviation for 99% confidence intervals of distribution.

The Figures and *NRMSD* show that the values obtained numerically are close to the values obtained by simulation. The expansion of confidence intervals on the right side of Figure 2 is associated with the proximity of numerical values to zero.

5. Conclusion

The method presented in the article allows one to numerically obtain the stationary distribution of the MJM with up to several dozens of servers

¹<https://www.r-project.org/>

²<https://github.com/ProgGrey/MJMrrss>

(depending on memory and computation time limits), which is an order of magnitude higher than in earlier works. Thus, the MJM_{rss} R package presented in present paper can be used as another tool for validating models of larger systems.

REFERENCES

1. *Basmadjian R.* Flexibility-Based Energy and Demand Management in Data Centers: A Case Study for Cloud Computing // *Energies*. 2019. Vo 12. № 17.
2. *Harrison P. G., Patel N. M., Knottenbelt W. J.* Energy–Performance Trade-Offs via the EP Queue // *ACM Trans. Model. Perform. Eval. Comput. Syst.* 2016. Vo 1. № 2. P. 1–31
3. *Tian J. P., Kannan D.* Lumpability and Commutativity of Markov Processes // *Stochastic Analysis and Applications*. 2006. Vo 24. № 3. P. 685–702
4. *Kim S.* M/M/s Queueing System Where Customers Demand Multiple Server Use. PhD Thesis, Southern Methodist University, 1979. 103 p.
5. *Gautam N.* Analysis of Queues : Methods and Applications. Boca Raton: CRC Press, 2012. 802 p.
6. *Rumyantsev A., Morozov E.* Stability criterion of a multiserver model with simultaneous service // *Annals of Operations Research*. 2017. № 1. P. 29–39.
7. *Latouche G., Ramaswami V.* Introduction to Matrix Analytic Methods in Stochastic Modeling. Philadelphia: Society for Industrial and Applied Mathematics, 1999. 331 p.
8. *Garimella R. M., Rumyantsev A.* On an exact solution of the rate matrix of G/M/1-type Markov process with small number of phases // *Journal of Parallel and Distributed Computing*. 2018. Vol. 119. P. 172–178.
9. *Rumyantsev A., Basmadjian R., Astafiev S., Golovin A.* Three-level modeling of a speed-scaling supercomputer // *Annals of Operations Research*. 2022. Vol. 331. P. 649–677.
10. <https://oeis.org> — The On-Line Encyclopedia of Integer Sequences. 2024.
11. *Sheldon M. R.* Introduction to Probability and Statistics for Engineers and Scientists – Sixth Edition. Los Angeles: Academic Press, 2020. 730 p.
12. *Bini Dario, Latouche G., Meini B.* Numerical methods for structured Markov chains. Oxford: Oxford University Press, 2005. 336 p.
13. *Henderson S. G., Glynn P. W.* Regenerative steady-state simulation of discrete-event systems // *ACM Transactions on Modeling and Computer Simulation*. 2001. № 4. P. 313–345

AGE OF INFORMATION ANALYSIS FOR WIRELESS SYSTEMS WITH TIME-CORRELATED TRAFFIC ARRIVAL PROCESS

V. Manaeva¹, E. Zhabankova¹, E. Markova^{1,2}

¹*People's Friendship University of Russia named after Patrice Lumumba, Moscow, Russian Federation*

²*Federal Research Center "Computer Science and Control" of the Russian Academy of Sciences, Moscow, Russian Federation*

Term "Internet of Things" (IoT) describes a type of technical system in which various devices and sensors are connected with each other and cloud computing resources via wireless networks. IoT found its applications in various industrial spheres such as energy grids, smart factories, network load monitoring and many others. For the purposes of monitoring and manipulating in industrial IoT applications it is important for information to be as relevant as possible, which is where Age of Information steps in. Age of Information (AoI) and its upper bound Peak Age of Information (PAoI) are metrics of information freshness in communication systems. In this study we propose a queuing system (QS) for describing wireless network of industrial IoT application which will be analysed through imitation modeling. The QS in question will have several update types arriving in batches at the system according to switched Poisson process and servicing in discrete time. For the purposes of the study, various service disciplines will be analysed utilizing mean PAoI metric.

Keywords: *AoI, Age of Information, 5G, Internet of Things.*

Introduction

Nowadays communication systems are actively employed in different spheres of life. Internet of Things (IoT) is a type of communication system which includes so-called "things" and wireless networks connecting them. The term "things" in IoT can describe any physical object that can be modified with sensors and/or communication devices to provide remote monitoring, status, manipulation and control. To provide an example, in consumer applications IoT can be utilized to create smart homes, smart household appliances and connected health.

The research was funded by the Russian Science Foundation, project №.22-79-10053, (<https://rscf.ru/en/project/22-79-10053/>). This publication has been supported by the RUDN University Scientific Projects Grant System, project №.021937-2-000. Authors thank prof. Molchanov Dmitri for his extensive assistance with system model analysis.

IoT found a lot of applications in the industry. As a matter of fact, industrial Internet of Things (IIoT) can be utilized in energy grids, network load control, smart factories, crop monitoring and various other industries. IIoT devices acquire and analyze data from connected equipment, different locations, sensors and devices in order to maintain, control or manipulate the work of the devices remotely and/or automatically.

To adequately assess the information in the remote control center (CC), the information in system should be as relevant as possible. This is where the Age of Information (AoI) metric is utilized. AoI quantitatively measures the relevance of information in CC. AoI's upper bound is measured by Peak Age of Information (PAoI). The mean PAoI across a large enough amount of time can be utilized as an optimization metric. Mean PAoI for any system can be divided into two components: interarrival time and sojourn time [1].

Previous literature review [2,3] showed that the analysis of metrics related to latency, which is a component of sojourn time, is often complicated by orthogonal frequency-division multiple access (OFDMA). This leads to the need to use packet service.

Consequently, to take into account the features of wireless systems with OFDMA technology support, this study suggests model serving an arbitrary number of devices. Each device be in either of two states: "regular" and "emergency", where the "emergency" state corresponds to a higher intensity of update generation as compared to "regular". Model is described as a queuing system in discrete time-space with single server, limited buffer, several batch income streams and batch service.

1. System model

The considered system model is based on IIoT application. System model is presented in Fig. 1. It describes the up-link channel on base station (BS). Information from N devices is sent in batches of updates and is collected in a buffer space on BS. Maximum number of updates that can be in buffer space at BS is denoted as R .

Each of N devices can be in either of two states, i , $i = 1, 2$. For each device let the difference in states correspond to different update generation intensity denoted as λ_{ni} , $n = \overline{1, N}$, $i = 1, 2$, where n refers to the number of device updates are generated at and i refers to the state that the system is in. Consider that in state $i = 2$ the device will have update generation intensity higher than in state $i = 1$.

Updates arriving in system form unanimous income stream and are ordered based on the number of device they are generated at. Upon arriving updates may be blocked if buffer space is nit enough for them. Blocked

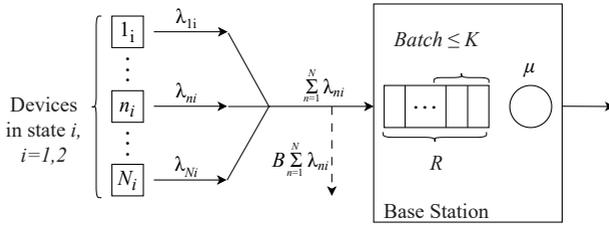


Figure 1. System model

are only the updates that are exceeding the buffer space available, $B \sum_{n=1}^N$ denoting the blocked updates.

The service time μ for a batch of updates is fixed and equals to 1 ms. K in system denotes the service batch size. After a previous service time is finished, the buffer is checked for updates:

- if it has no updates, the service in this time slot will not be done;
- if buffer has less than K updates, they form a batch and the batch is served in this slot;
- if buffer has K or more updates, the batch of K updates is formed from the first K updates in the buffer and the batch is served in a slot.

2. Queuing system

2.1. Income stream

The updates arrive in batches from N devices in system, which corresponds to N group income streams in the system. Each of the income streams has two states, i , $i = 1, 2$, in which the update generating intensity is varied. As per technical system requirements, time in system is slotted, slot duration is fixed. At the instant new time slot starts each device can either switch its state or conserve it.

Let α_n and β_n , $n = \overline{1, N}$, denote the probability of n -th device switching from state 1 to state 2 and from state 2 to state 1 respectively.

Since each of N devices can be in either of two states, let $S_n(t)$, $n = \overline{1, N}$ denote the state of n -th device at instance right after slot t starts. Thus, to account for all the devices, let $\mathbf{S}(t)^\top = (S_1(t), \dots, S_N(t))$ denote the vector of states of N devices at the start of the slot t .

Devices generate updates at the end of each time slot. In each of the two states the updates are generated in batches of size varying in accordance to Poisson distribution law with the mean given as λ_{ni} , $n = \overline{1, N}$; $i = 1, 2$. In

system λ_{ni} , $n = \overline{1, N}, i = 1, 2$, denotes the update generation intensity of n -th device in state i , which corresponds to the mean size of update batch sent to BS.

Consider updates generated at arbitrary slot. At each device a batch of updates is generated. Let $W_n(t)$, $n = \overline{1, N}$ denote the size of batch generated at n -th device at the end of time slot t . Generated updates from N devices are ordered based on the number of device they were sent from and are combined into unified batch of size $W(t) = \sum_{i=1}^N W_i(t)$ for the QS to receive.

2.2. Buffer space and update processing

Batches of generated updates are arriving at the system and, due to delayed access, attempt to access the buffer space. Let $Q(t)$ denote the occupied buffer space at instance new slot starts. Then $R - Q(t)$ denotes the buffer space accessible for updates. As the unoccupied buffer space is limited, update losses are inevitable. Thus, if number $W(t)$ of updates generated at time slot t is greater than $R - Q(t)$, $W(t) - R + Q(t)$ last updates are not accepted in buffer and are considered lost.

The accepted updates are placed in buffer space and, based on their position, either form a batch and are sent for processing, or wait for processing in buffer.

3. Simulation model

Simulation model for the considered system was written using Python programming language. Imitation model flow diagram is presented in fig. 2.

For the purposes of imitating the considered queuing system, several input parameters are required. Input parameters include those mentioned in section 1 (N, R, K), section 2.1 ($\alpha = (\alpha_1, \dots, \alpha_N)^\top, \beta = (\beta_1, \dots, \beta_N)^\top, \lambda_1 = (\lambda_{1,1}, \dots, \lambda_{N,1})$ and $\lambda_2 = (\lambda_{1,2}, \dots, \lambda_{N,2})$), as well as simulation time constraint, which is denoted as t_{max} .

For each of the updates generated in imitation model several parameters are stored. Those parameters later form final statistics for the imitation run. Parameters are as follows: global sequential number of the update, n ; number of source device, *type*; time slot the update was generated at, G_n ; variable storing if update was accepted to system, E_n ; time slot update was processed by the system, F_n ; time slot previous accepted update with the same *type* was generated at, G_n^{prev} ; interarrival time for the n -th update, Y_n ; sojourn time for the n -th update, T_n ; PAoI for the n -th update, A_n .

Out of those parameters, n , *type* and G_n are fixed from the moment update was generated at the device. Then update is sent to the buffer and,

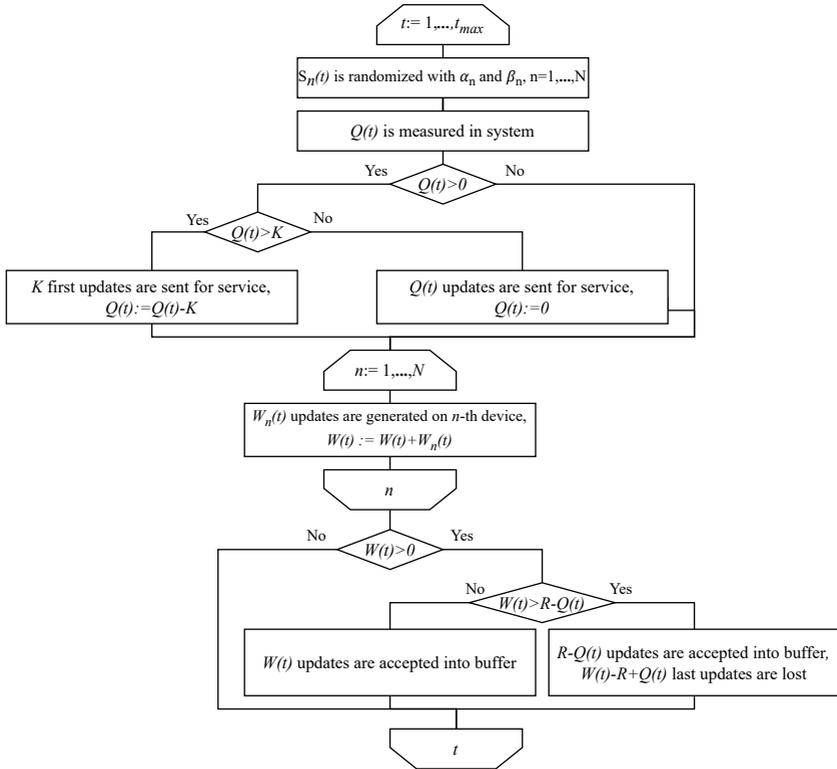


Figure 2. Flow diagram for the simulation model (main cycle)

if accepted, E_n value is set to *True*. If accepted, F_n is modified at the time slot update was included in service batch and sent to processing. After processing, previous statistics is checked and G_n^{prev} value is set based on the G_n of the previous processed update with the same *type*. If none of the processed updates have the same *type* value, G_n^{prev} is set to 0. For an update, Peak Age of Information, as calculated in [1], includes interarrival time and sojourn time.

From imitation model data and calculated parameters, final statistics for a run is collected and can be exported in any format. In addition to information about each of the updates generated across imitation model run, several values are calculated. Those values are as follows: number of updates generated across the imitation model run, U ; number of updates accepted in system, E ; number of lost updates, L ; update loss probability,

B ; mean interarrival time across the whole system and across a fixed *type*, $M[Y]$ and $M[Y]_n$ respectively; mean sojourn time across the whole system and across a fixed *type*, $M[T]$ and $M[T]_n$ respectively; mean PAoI across the whole system and across a fixed *type*, $M[A]$ and $M[A]_n$ respectively.

Number of generated, accepted and lost updates are the direct results of imitation model run. Update loss probability B is calculated as a ratio of number L of updates lost to number U of updates generated. The means in statistics are calculated as averages across either all the generated packets, or the packets with a fixed *type* value.

4. Conclusion

The considered system more closely represents the work of 5G NR wireless communication systems. The queuing system is analysed, imitation model is constructed and thoroughly described in this study. Peak Age of Information for the considered system is to be evaluated utilizing the provided imitation model.

REFERENCES

1. *Kosta A., Pappas N., Angelakis V.* Age of Information: A New Concept, Metric, and Tool // Foundations and Trends in Networking. 2017 Vol. 12. № 3. P. 162–259.
2. *Zhbankova E., Khakimov A., Markova E., Gaidamaka Y.* The Age of Information in Wireless Cellular Systems: Gaps, Open Problems, and Research Challenges // Sensors. 2023. Vol. 23. № 19. article № 8238.
3. *Abbas Q., Hassan S.A., Qureshi H.K., Dev K., Jung H.* A comprehensive survey on age of information in massive IoT networks // Computer Communications. 2023. Vol. 197. P. 199–213.

Manaeva Varvara — B.Sc., student, probability theory and cybersecurity department. E-mail: *manaeva_ve@pfur.ru*

Zhbankova Elena — M.Sc., PhD student, probability theory and cybersecurity department. E-mail: *zhbankova_ea@pfur.ru*

Markova Ekaterina — PhD degree in applied mathematics and computer sciences, associate professor, probability theory and cybersecurity department. E-mail: *markova_ev@pfur.ru*

ANALYSIS OF A BATCH ARRIVAL QUEUE WITH POWER SAVING MODE

Y. Sakai, T. Phung-Duc

University of Tsukuba, Tsukuba, Japan

Queue models incorporating working vacations have been extensively studied. This research further develops the M/M/1 model with two types of working vacations by introducing a scenario where customers arrive in batches. Using the probability generating function method, performance indicators such as steady-state probabilities are derived. Additionally, numerical experiments are conducted to discuss the average number of customers in the system when the batch size follows an arbitrary distribution. **Keywords:** *working vacation, batch arrival, probability generating function, arbitrary distribution.*

Introduction

In recent years, with the proliferation of IoT and AI, there has been a daily transmission of massive amounts of data, leading to concerns about the increased power consumption of IT devices that need to process this data. As a countermeasure, it has become common for recent IT devices to be equipped with power-saving modes. The analysis of queueing models incorporating working vacations has been conducted as a method for evaluating the performance of systems equipped with power-saving modes. Recent studies have focused on models incorporating multiple working vacations to evaluate the performance of systems equipped with multi-stage power-saving modes.

Servi and Finn [5] studied an M/M/1 queue with working vacations. Xu et al. [6] analyzed an $M^x/M/1$ queue with working vacations. Baba [1] analyzed a model introducing multiple working vacations into an $M^x/M/1$ model. Ibe and Isijola [2] examined an M/M/1 queueing system with multiple differentiated vacations. Phung-Duc [4] analyzed a model introducing two types of working vacations into an M/M/1 model. Ye [7] analyzed an $M^x/M/1$ queue with a two-stage vacation policy. Manickam and Kalidass [3] studied the optimality of bulk entry queues with differentiated hiatuses. In this paper, we extend the model discussed by Phung-Duc [4] and introduce two types of working vacations into an $M^x/M/1$ model with batch arrivals. Using the probability generating function method, we derive exact

solutions for performance evaluation indices such as steady-state probabilities and the average number of customers in the system. Additionally, as a numerical experiment, we compare and discuss the simulation and analytical results for the average number of customers in the system when the batch size follows an arbitrary distribution.

1. Mathematical Model

The model in this study introduces two types of working vacations with different service rates into an $M^x/M/1$ model where customers arrive in batches. Batches of customers arrive at the server according to a Poisson process with rate λ . The probability that the batch size $B = i$ is b_i , where i represents the number of customers in a batch. Service times follow an exponential distribution, with an average service time of $1/\mu$. When the server finishes serving a customer and there are no customers in the system, the server starts Vacation 1. The service time during Vacation 1 follows an exponential distribution, with an average service time of $1/\mu_1$. The vacation time of Vacation 1 follows an exponential distribution with an average time of $1/\gamma_1$. If there are no customers in the system when Vacation 1 ends, the server starts Vacation 2. The service time during Vacation 2 follows an exponential distribution, with an average service time of $1/\mu_2$. The vacation time of Vacation 2 follows an exponential distribution with an average time of $1/\gamma_2$. Vacation 2 will continue to repeat as long as there are no customers in the system. When each vacation ends, the server returns to the normal state and resumes service at a rate of μ . The state transition diagram of the model in this study is shown in Figure 1.

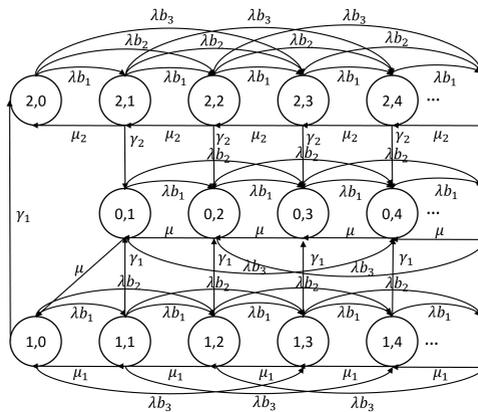


Figure 1. State transition diagram ($b_1 + b_2 + b_3 = 1$).

2. Analysis

Let $N(t)$ be the number of customers in the server at time t and $S(t)$ be the state of the server

$$S(t) = \begin{cases} 0 & \text{if the server in normal,} \\ 1 & \text{if the server in vacation 1,} \\ 2 & \text{if the server in vacation 2.} \end{cases}$$

Then, $(S(t), N(t))$ forms a two dimensional Markov chain on the state space $\{0, 1, 2\} \times \mathbb{Z}^+$, where $\mathbb{Z}^+ = \{0, 1, 2, \dots\}$.

The steady-state probability of the system being in state (i, j) is denoted as $\pi_{i,j}$. From the state transition diagram in Figure 1, the following balance equations can be derived:

$$(\lambda + \mu)\pi_{0,1} = \gamma_1\pi_{1,1} + \gamma_2\pi_{2,1} + \mu\pi_{0,2}, \quad (1)$$

$$(\lambda + \mu)\pi_{0,n} = \gamma_1\pi_{1,n} + \gamma_2\pi_{2,n} + \mu\pi_{0,n+1} + \lambda \sum_{k=1}^{n-1} b_k\pi_{0,n-k}, \quad (n \geq 2), \quad (2)$$

$$(\lambda + \gamma_1)\pi_{1,0} = \mu_1\pi_{1,1} + \mu\pi_{0,1}, \quad (3)$$

$$(\lambda + \gamma_1 + \mu_1)\pi_{1,1} = \mu_1\pi_{1,2} + \lambda b_1\pi_{1,0}, \quad (4)$$

$$(\lambda + \gamma_1 + \mu_1)\pi_{1,n} = \mu_1\pi_{1,n+1} + \lambda b_n\pi_{1,0} + \lambda \sum_{k=1}^{n-1} b_k\pi_{1,n-k}, \quad (n \geq 2), \quad (5)$$

$$\lambda\pi_{2,0} = \mu_2\pi_{2,1} + \gamma_1\pi_{1,0}, \quad (6)$$

$$(\lambda + \gamma_2 + \mu_2)\pi_{2,1} = \mu_2\pi_{2,2} + \lambda b_1\pi_{2,0}, \quad (7)$$

$$(\lambda + \gamma_2 + \mu_2)\pi_{2,n} = \mu_2\pi_{2,n+1} + \lambda b_n\pi_{2,0} + \lambda \sum_{k=1}^{n-1} b_k\pi_{2,n-k}, \quad (n \geq 2). \quad (8)$$

The probability generating function of the system in state (i, j) is defined as $\Pi_i(z) = \sum_{n=1}^{\infty} \pi_{i,n}z^n$ for $i = 0, 1, 2$. Additionally, the probability generating function of the batch size B is defined as $B(z) = \sum_{n=1}^{\infty} b_nz^n$.

Multiplying both sides of equations (1) and (2) by z^n , summing over n , and transforming into generating functions, we obtain:

$$\Pi_0(z) = \frac{z[\gamma_1\Pi_1(z) + \gamma_2\Pi_2(z) - \mu\pi_{0,1}]}{(\lambda + \mu)z - \mu - \lambda zB(z)}.$$

Similarly, from equations (4) and (5), we obtain:

$$\Pi_1(z) = \frac{z[\lambda\pi_{1,0}B(z) - \mu_1\pi_{1,1}]}{(\lambda + \gamma_1 + \mu_1)z - \mu_1 - \lambda zB(z)}.$$

Since there exists a unique $0 < z_1 < 1$ such that the denominator becomes zero, we find z_1 . When substituting $z = z_1$, the numerator also becomes zero, yielding the following equation:

$$\pi_{1,1} = \frac{\lambda}{\mu_1}\pi_{1,0}B(z_1).$$

Similarly, from equations (6) and (7), we obtain:

$$\Pi_2(z) = \frac{z[\lambda\pi_{2,0}B(z) - \mu_2\pi_{2,1}]}{(\lambda + \gamma_2 + \mu_2)z - \mu_2 - \lambda zB(z)}.$$

Similarly, we obtain the following equation:

$$\pi_{2,1} = \frac{\lambda}{\mu_2}\pi_{2,0}B(z_2),$$

where $0 < z_2 < 1$ is a unique value that the denominator of $\Pi_2(z)$ is zero.

The steady-state probability $\lim_{z \rightarrow 1} \Pi_0(z)$ is derived as follows:

$$\begin{aligned} \Pi_0(1) &= \lim_{z \rightarrow 1} \frac{z[\gamma_1\Pi_1'(z) + \gamma_2\Pi_2'(z)] + [\gamma_1\Pi_1(z) + \gamma_2\Pi_2(z) - \mu\pi_{0,1}]}{(\lambda + \mu) - \lambda zB'(z) - \lambda B(z)} = \\ &= \frac{\gamma_1\Pi_1'(1) + \gamma_2\Pi_2'(1) + [\gamma_1\Pi_1(1) + \gamma_2\Pi_2(1) - \mu\pi_{0,1}]}{\mu - \lambda B'(1)}. \end{aligned}$$

From the total probability formula, we obtain:

$$\pi_{1,0} + \Pi_1(1) + \pi_{2,0} + \Pi_2(1) + \Pi_0(1) = 1.$$

From the state transition diagram, the following equations are obtained:

$$\gamma_1\pi_{1,0} = \gamma_2\Pi_2(1), \quad \mu\pi_{0,1} = \gamma_1\pi_{1,0} + \gamma_1\Pi_1(1).$$

The unknowns $\pi_{1,0}$, $\pi_{2,0}$, and $\pi_{0,1}$ can be solved from the above three equations as follows:

$$A = \mu\gamma_1(\gamma_1 + \gamma_2) + \lambda\gamma_2(1 - B(z_1))(\mu - \mu_1) - \gamma_1^2\mu_2,$$

$$\pi_{1,0} = \frac{\lambda\gamma_1\gamma_2(\mu - \lambda B'(1))(1 - B(z_2))}{\lambda A(1 - B(z_2)) + \mu\gamma_1^2\gamma_2},$$

$$\pi_{2,0} = \frac{\gamma_1^2\gamma_2(\mu - \lambda B'(1))}{\lambda A(1 - B(z_2)) + \mu\gamma_1^2\gamma_2},$$

$$\pi_{0,1} = \frac{\lambda\gamma_1\gamma_2(\mu - \lambda B'(1))(1 - B(z_2))(\gamma_1 + \lambda(1 - B(z_1)))}{\lambda\mu A(1 - B(z_2)) + \mu^2\gamma_1^2\gamma_2}.$$

Let $\Pi(z)$ be the generating function representing the number of customers in the system. $\Pi(z)$ is derived from the following equation:

$$\Pi(z) = \Pi_0(z) + \Pi_1(z) + \Pi_2(z) + \pi_{1,0} + \pi_{2,0}.$$

The average number of customers in the system $E[L]$ is obtained from $\Pi'(1)$:

$$E[L] = \Pi'(1) = \Pi'_0(1) + \Pi'_1(1) + \Pi'_2(1).$$

3. Numerical examples

In this experiment, we investigated the behavior of the average number of customers in the system $E[L]$ with respect to the traffic density $\rho = \frac{\lambda B'(1)}{\mu}$ and compared the analytical results with the simulation results. The experiment was conducted with the following parameter settings: $\mu = 1.0$, $\mu_1 = 0.05$, $\mu_2 = 0.1$, $\gamma_1 = 1.0$, $\gamma_2 = 1.0, 0.5, 0.25, 0.1$, and $B'(1) = 2.0$. Figures 2(a), 2(b) and 2(c) show the graphs when the batch size follows a geometric distribution $B(z) = \frac{pz}{1 - (1-p)z}$, a Poisson distribution $B(z) = e^{\lambda(z-1)}$ and a uniform distribution $B(z) = \frac{1}{N} \sum_{k=1}^N z^k$, respectively. In all graphs, we confirmed that the analytical results matched the simulation results.

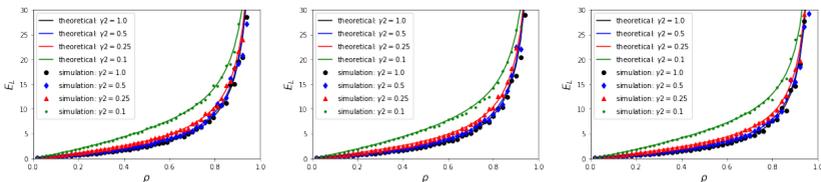


Figure 2. Comparison of $E[L]$ against ρ .

4. Conclusion

In this study, we considered a model that introduces two types of working vacations with different service rates into an $M^x/M/1$ model with batch arrivals. We derived the steady-state probabilities and the average number of customers in the system. Additionally, through numerical experiments, we confirmed that the analytical results for the average number of customers in the system match the simulation results when the batch size follows an arbitrary distribution. For future research, an extension to an $M^x/M/C$ model with two types of working vacations can be considered.

REFERENCES

1. *Baba Y.* The $M^x/M/1$ queue with multiple working vacations // American Journal of Operations Research. 2012. Vol. 2. № 2. P. 217–224.
2. *Ibe O.C., Isijola O.A.* M/M/1 multiple vacation queueing systems with differentiated vacations // Modelling and Simulation in Engineering. 2014. Article ID 158247, 6 p.
3. *Manickam V., Kalidass K.* A discussion on the optimality of bulk entry queue with differentiated hiatuses // Operations Research and Decisions. 2022. Vol. 32. № 2. P. 137–150.
4. *Phung-Duc T.* Single-Server Systems with Power-Saving Modes and Batch Arrivals // Performance Evaluation Review. 2015. Vol. 43. № 3. P. 15–27.
5. *Servi L.D., Finn S.G.* M/M/1 queues with working vacations (M/M/1/WV) // Performance Evaluation. 2002. Vol. 50. № 1. P. 41–52.
6. *Xu X.-L., Zhang Z.-J., China P.-R., Tian N.-S.* Analysis for the $M^x/M/1$ working vacation queue // International Journal of Information and Management Science. 2009. Vol. 20. № 3. P. 379–394.
7. *Ye Q.* The analysis of $M^x/M/1$ queue with two-stage vacations policy // Communications and Statistics - Theory and Methods. 2019. Vol. 48. № 18. P. 4492–4510.

Yuta Sakai — Master student, Department of Policy and Planning Sciences, University of Tsukuba. E-mail: s2320432@u.tsukuba.ac.jp

Tuan Phung-Duc — Associate Professor, Department of Policy and Planning Sciences, University of Tsukuba. E-mail: tuan@sk.tsukuba.ac.jp

АНАЛИЗ МОДЕЛИ С ОТНОСИТЕЛЬНЫМ ПРИОРИТЕТОМ ПЕРЕРАСПРЕДЕЛЕНИЯ РЕСУРСА МЕЖДУ ДВУМЯ СЕГМЕНТАМИ СЕТИ С ЭЛАСТИЧНЫМ ТРАФИКОМ

А. С. Власкина¹, С. А. Бурцева¹, И. А. Кочеткова^{1,2}

¹ *Российский университет дружбы народов, г. Москва, Россия*

² *Федеральный исследовательский центр «Информатика и управление»
Российской академии наук, г. Москва, Россия*

Технология нарезки сети в беспроводных сетях 5G оптимизирует работу сети и облегчает управление ее ресурсами. Чтобы повысить пропускную способность сети, следует принимать во внимание относительный приоритет услуг. В рамках исследования рассматривается проблема распределения ресурсов между двумя сегментами беспроводной сети с учетом приоритета одного из них. **Ключевые слова:** 5G, нарезка сети, эластичный трафик, относительный приоритет, система массового обслуживания.

1. Введение

Технология нарезки сети (network slicing) в беспроводных сетях 5G помогает оптимизировать работу сети и упростить управление ее ресурсами [1, 2]. Для увеличения пропускной способности сети необходимо учитывать относительный приоритет услуг. В данной работе исследуется задача распределения ресурса между двумя сегментами беспроводной сети [3] с учетом приоритета одного из них.

2. Системная модель

Рассмотрим ситуацию, когда основной оператор выделяет свои ресурсы V двум виртуальным операторам (рис. 1). Для этого между ними заключается соглашение об уровне обслуживания (service level agreement, SLA), в котором определяется рекомендуемый объем ресурса для каждого сегмента: \tilde{V}_1, \tilde{V}_2 . Управление перераспределением ресурса между сегментами осуществляет контроллер. Он периодически отправляет сигналы, чтобы проверить необходимость такого перераспределения. Поток сигналов является пуассоновским с интенсивностью δ .

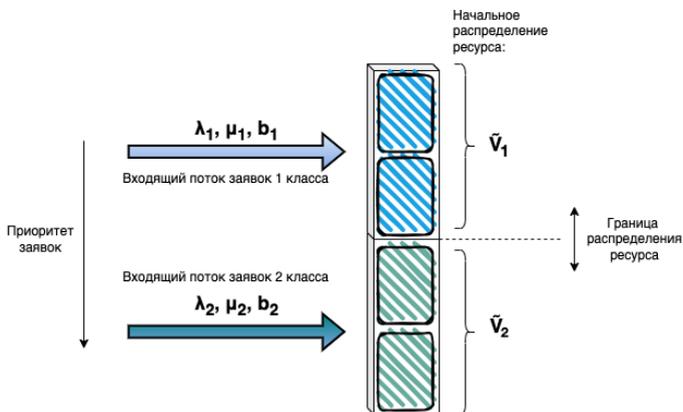


Рис. 1. Схема системы массового обслуживания

Каждый виртуальный оператор предоставляет своим пользователям услугу передачи данных. У первого оператора приоритетный доступ к ресурсу, и объем ресурса сегмента не может быть меньше \tilde{V}_1 . Кроме того, пользователи должны обслуживаться с минимальной скоростью b_1 и b_2 . Входящие потоки заявок на передачу трафика являются пуассоновскими с интенсивностями λ_1 и λ_2 . Объем трафика распределен по экспоненциальному закону с параметрами μ_1 и μ_2 .

3. Система массового обслуживания

Рассмотрим систему как систему массового обслуживания с относительным приоритетом и сигналами. Опишем функционирование системы случайным процессом $\mathbf{X}(t)$ с состояниями $\mathbf{x} = (V_1, n_1, n_2)$, где V_1 объем ресурса 1-сегмента, n_1 число обслуживаемых заявок 1-класса, n_2 число обслуживаемых заявок 2-класса. Пространство состояний системы выглядит следующим образом

$$\mathcal{X} = \{(V_1, n_1, n_2) : 0 \leq n_1 b_1 \leq V_1, 0 \leq n_2 b_2 \leq V - V_1, \tilde{V}_1 \leq V_1 \leq V\}.$$

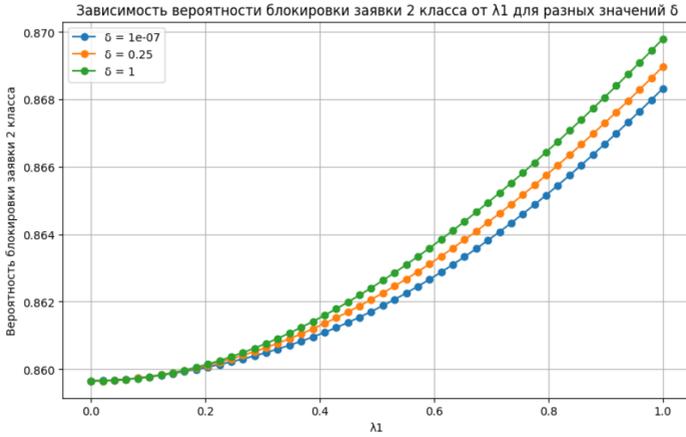


Рис. 2. Вероятность блокировки заявки 2 класса

Элементы матрицы интенсивностей переходов имеют вид

$$a(\mathbf{x}, \mathbf{x}') = \begin{cases} \lambda_1, & \mathbf{x}' = (V_1, n_1 + 1, n_2), \quad n_1 b_1 + b_1 \leq V_1, \\ \lambda_2, & \mathbf{x}' = (V_1, n_1, n_2 + 1), \quad n_2 b_2 + b_2 \leq V - V_1, \\ V_1 \mu_1, & \mathbf{x}' = (V_1, n_1 - 1, n_2), \quad n_1 > 0, \\ (V - V_1) \mu_2, & \mathbf{x}' = (V_1, n_1, n_2 - 1), \quad n_2 > 0, \\ \delta, & \mathbf{x}' = (V_1 + 1, n_1, n_2), \quad n_2 b_2 < V - V_1, \\ \delta, & \mathbf{x}' = (V_1 - 1, n_1, n_2), \quad n_2 b_2 = V - V_1, \\ & n_1 b_1 < \tilde{V}_1, \quad V_1 > \tilde{V}_1. \end{cases}$$

Проиллюстрируем, как вероятность блокировки заявок зависит от интенсивности поступления заявок двух классов λ_1 и λ_2 . Исходные данные: $\lambda_1 = 0,8$, $\lambda_2 = 0,6$, $\mu_1 = 0,4$, $\mu_2 = 0,1$. На рис. 2 и 3 показаны зависимости вероятности блокировки заявок второго класса от интенсивности поступления заявок первого класса и второго классов.

Закключение

В данной работе проанализирована модель, в которой ресурс между двумя сегментами сети распределяется с учетом относительного приоритета. Рассмотрено влияние интервала нарезки на вероятность блокировки заявок первого и второго классов.

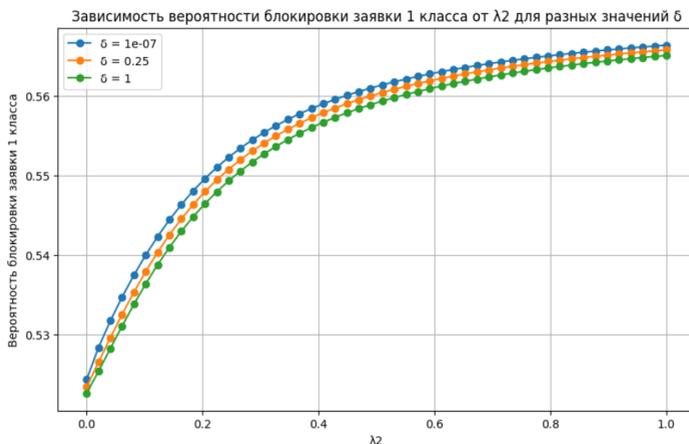


Рис. 3. Вероятность блокировки заявки 1 класса

СПИСОК ЛИТЕРАТУРЫ

1. Sun G., Ayepah-Mensah D., Boateng G. O., Kuadey N. A., Omer M. B., Liu G. Holistic Roadmap of Trends in Radio Access Network Slicing: A Survey // IEEE Communications Magazine. 2023. Vol. 61. No. 12. P. 118–124.
2. Yarkina N., Correia L. M., Moltchanov D., Gaidamaka Y., Samouylov K. Multi-Tenant Resource Sharing with Equitable-Priority-Based Performance Isolation of Slices for 5G Cellular Systems // Computer Communications. 2022. Vol. 188. P. 39–51.
3. Kochetkova I., Leonteva K., Ghebrial I., Vlaskina A., Burtseva S., Kushchazli A., Samouylov K. Controllable Queuing System with Elastic Traffic and Signals for Resource Capacity Planning in 5G Network Slicing // Future Internet. 2024. Vol. 16. No. 1. art. no. 18.

Власкина Анастасия Сергеевна — к.ф.-м.н., ассистент кафедры теории вероятностей и кибербезопасности РУДН. E-mail: vlaskina-as@rudn.ru

Бурцева София Артемовна — аспирант кафедры теории вероятностей и кибербезопасности РУДН. E-mail: burtseva-sa@rudn.ru

Кочеткова Ирина Андреевна — к.ф.-м.н., доцент, доцент кафедры теории вероятностей и кибербезопасности РУДН, старший научный сотрудник ФИЦ ИУ РАН. E-mail: kochetkova-ia@rudn.ru

АСИМПТОТИЧЕСКИЙ АНАЛИЗ RQ-СИСТЕМЫ M|M₁ С ДВУМЯ ОРБИТАМИ И ВЫСОКОЙ ИНТЕНСИВНОСТЬЮ ОБРАЩЕНИЯ ЗАЯВОК СО ВТОРОЙ ОРБИТЫ

А. А. Назаров, Я. Е. Измайлова

*Национальный исследовательский Томский государственный университет,
г. Томск, Россия*

В работе рассматривается система массового обслуживания с повторными вызовами. На вход поступает простейший поток заявок. Если прибор свободен, то пришедшая заявка обслуживается, если занят, то переходит на первую орбиту. После экспоненциально распределенного времени задержки заявки на орбите, она вновь обращается к прибору с повторной попыткой захвата. Если прибор занят, то заявка переходит на вторую орбиту, в которой интенсивность обращения заявок высокая. Выполнив случайную задержку, заявка обращается к прибору с повторной попыткой его захвата. Если прибор свободен, то заявка начинает обслуживание, если занят, то мгновенно возвращается на вторую орбиту для осуществления следующей случайной задержки. После окончания обслуживания заявки покидают систему. Исследование проводится методом асимптотического анализа. Получены распределение вероятностей числа заявок на первой орбите и распределение вероятностей числа заявок на второй орбите и приборе. **Ключевые слова:** RQ-система, орбита, высокая интенсивность обращения заявок, асимптотический анализ.

Введение

Образование очередей перед стойками бронирования билетов, в больницах, банках, торговых центрах – это сценарии с которыми современный человек сталкивается практически ежедневно. Иногда из-за загруженности клиенты лишаются обслуживания и пытаются повторить попытку обслужиться чуть позже. Такой сценарий поведения и привел к возникновению систем с повторными вызовами (RQ-систем). Большая работа была проделана в области исследования систем с повторными вызовами, и соответствующую литературу можно найти в обзорных статьях J.R. Artalejo [1], Г. Фалина [2].

Ситуации, когда клиенты могут быть заинтересованы в ожидании обслуживания в более удобном месте (то есть на премиальной орбите), в то время как обычные клиенты будут ожидать обслуживание на обычной орбите, описывают математические модели RQ-систем с двумя орбитами. Двухорбитальные системы массового обслуживания с повторными вызовами были рассмотрены в работах [3], [4], [5], [6].

В данной работе же рассматривается RQ-система, в которой клиенты, обнаружив прибор занятым, переходят на первую орбиту и после обращения с первой орбиты к прибору, обнаружив его занятым, переходят на вторую орбиту. Интенсивность обращения со второй орбиты высокая, что уменьшает время ожидания обслуживания клиента (заявки).

1. Математическая модель и постановка задачи

На вход системы поступает простейший поток заявок с интенсивностью λ (Рис. 1). Если прибор свободен, то заявка начинает обслуживаться в течение времени, распределенному по экспоненциальному закону с параметром μ . Если прибор занят, то поступившая заявка уходит на первую орбиту, где осуществляет случайную экспоненциально распределенную задержку с параметром σ_1 . После случайной задержки, заявка обращается к прибору с повторной попыткой его захвата. Если он свободен, то заявка обслуживается, после чего покидает систему. Если прибор занят, то она переходит на вторую орбиту, с которой интенсивность обращения заявок к прибору высокая. На второй орбите заявка осуществляет экспоненциальную задержку с параметром σ_2 . После задержки заявка обращается к прибору. Если свободен, то обслуживается. Если он занят, то мгновенно возвращается обратно на вторую орбиту.

Обозначим $n(t)$ – число заявок на второй орбите и приборе; $i(t)$ – число заявок на первой орбите; $k(t)$ – состояние прибора. Если свободен, то принимает значение 0, если занят, то значение 1; $P(k, n, i)$ – распределение вероятностей числа i заявок на первой орбите, числа n заявок на второй орбите и приборе, состояния прибора; $R(k, n)$ – распределение вероятностей числа n заявок на второй орбите и приборе и состояния прибора.

Ставится задача нахождения распределения вероятностей числа заявок на первой орбите, распределения вероятностей числа заявок на второй орбите и приборе. Исследование проводим методом асимптотического анализа в условии большой задержки заявок на первой орбите ($\sigma_1 \rightarrow 0$).

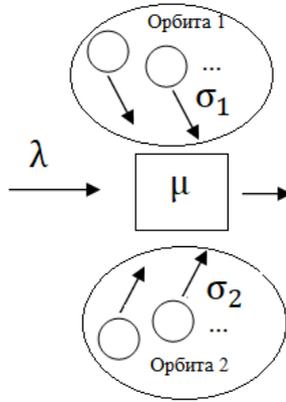


Рис. 1. Математическая модель RQ-системы $M|M|1$ с двумя орбитами.

2. Система уравнений Колмогорова

Для распределения вероятностей $P(k, n, i)$, запишем систему уравнений Колмогорова в стационарном режиме:

$$\begin{aligned} & -(\lambda + i\sigma_1 + n\sigma_2)P(0, n, i) + \mu P(1, n + 1, i) = 0, \\ & -(\lambda + i\sigma_1 + \mu)P(1, n, i) + \lambda P(0, n - 1, i) + \\ & \quad + n\sigma_2 P(0, n, i) + \lambda P(1, n, i - 1) + \\ & + (i + 1)\sigma_1 P(0, n - 1, i + 1) + (i + 1)\sigma_1 P(1, n - 1, i + 1) = 0. \end{aligned}$$

Введем частичные характеристические функции

$$H(k, n, u) = \sum_{i=0}^{\infty} e^{ju} P(k, n, i), \quad H(1, 0, u) = 0. \quad (1)$$

Запишем систему уравнений для характеристических функций:

$$\begin{aligned} & -(\lambda + n\sigma_2)H(0, n, u) + \mu H(1, n + 1, u) + j\sigma_1 \frac{\partial H(0, n, u)}{\partial u} = 0, \\ & -(\lambda + \mu)H(1, n, u) + \lambda e^{ju} H(1, n, u) + \lambda H(0, n - 1, u) + \\ & \quad + n\sigma_2 H(0, n, u) + j\sigma_1 \frac{\partial H(1, n, u)}{\partial u} - \\ & - j e^{-ju} \sigma_1 \frac{\partial H(0, n - 1, u)}{\partial u} - j e^{-ju} \sigma_1 \frac{\partial H(1, n - 1, u)}{\partial u} = 0. \end{aligned}$$

Данную систему будем решать методом асимптотического анализа. Основные результаты, полученные при решении, сформулированы ниже.

3. Распределение вероятностей числа заявок на второй орбите и приборе

Введем частичную производящую функцию распределения вероятностей числа заявок на второй орбите и приборе:

$$G(k, z) = \sum_{n=k}^{\infty} z^n R(k, n), \quad G(z) = G(0, z) + G(1, z).$$

Лемма 1. Распределение вероятностей состояний прибора имеет вид

$$R(0) = 1 - \frac{\lambda}{\mu}, \quad R(1) = \frac{\lambda}{\mu}.$$

Сформулируем следующую теорему.

Теорема 1. Производящая функция $G(z)$ распределения вероятностей числа заявок на второй орбите и приборе имеет вид

$$G(z) = R(0) \left(\frac{1 - \rho}{1 - \rho z} \right)^{\alpha} + R(1) z \left(\frac{1 - \rho}{1 - \rho z} \right)^{\alpha+1}, \quad (2)$$

где $\alpha = \frac{\lambda(\mu + \lambda)}{\sigma_2 \mu}$, а $\rho = \left(\frac{\lambda}{\mu} \right)^2$.

Таким образом, получаем, что производящая функция числа заявок на второй орбите и приборе имеет вид производящей функции взвешенной суммы двух отрицательно-биномиальных распределений.

4. Распределение вероятностей числа заявок на первой орбите

Для нахождения распределения вероятностей числа заявок на первой орбите воспользуемся методом асимптотического анализа в условии $\sigma_1 \rightarrow 0$.

Теорема 2. В предельном условии $\sigma_1 \rightarrow 0$ характеристическую функцию $H(u) = \sum_{n=0}^{\infty} \sum_{k=0}^1 H(k, n, u)$ распределения вероятностей числа заявок на первой орбите можно аппроксимировать

$$h(u) = \exp \left(ju \frac{\varkappa_1}{\sigma_1} + \frac{(ju)^2}{2} \frac{\varkappa_2}{\sigma_1} \right), \quad (3)$$

где $\varkappa_1 = \frac{\lambda^2}{\mu}$, $\varkappa_2 = \frac{\lambda^2(\mu + 2\lambda)}{\mu(\mu + \lambda)}$.

Таким образом, распределение вероятностей числа заявок на первой орбите можно аппроксимировать гауссовским распределением вероятностей с параметрами $\frac{\varkappa_1}{\sigma_1}$ и $\sqrt{\frac{\varkappa_2}{\sigma_1}}$.

Заключение

В работе рассмотрена система с повторными вызовами с простейшим входящим потоком и двумя орбитами. После первого обращения заявок с первой орбиты к прибору, они переходят на вторую, с которой высокая интенсивность обращения заявок к прибору. Предложенная математическая модель расширяет класс двухорбитальных RQ-систем и методов исследования их. Результаты решения задачи могут быть использованы в практических приложениях, где необходимо ускорение в ожидании на обслуживание клиента.

СПИСОК ЛИТЕРАТУРЫ

1. *Artalejo J. R.* Accessible bibliography on retrial queues // *Mathematical and Computer Modelling*. 1999. V. 30. № 3–4. P. 1–6.
2. *Falin G.* A survey of retrial queues // *Queueing Syst.* 1990. V. 7. P. 127–167.
3. *Avrachenkov K., Dudin A., Klimenok V.* Retrial Queueing Model MMAP/M 2/1 with Two Orbits // *Multiple Access Communications. MACOM 2010. Lecture Notes in Computer Science, Berlin, Heidelberg: Springer*, 2010. P. 107–118.
4. *Jain M., Bhagat A., Shekhar C.* Double orbit finite retrial queues with priority customers and service interruptions // *Applied Mathematics and Computation*. 2015. V. 253. P. 324–344.
5. *Singh Sanga S., Jain M.* FM/FM/1 double orbit retrial queue with customers' joining strategy: A parametric nonlinear programming approach // *Applied Mathematics and Computation*. 2019. V. 362.
6. *Dhibar S., Jain M.* Strategic behaviour for M/M/1 double orbit retrial queue with imperfect service and vacation // *International Journal of Mathematics in Operational Research*. 2023. V. 25. № 3. P. 369–385.

Назаров Анатолий Андреевич — доктор технических наук, профессор, профессор кафедры теории вероятностей и математической статистики НИ ТГУ. E-mail: nazarov.tsu@gmail.com

Измайлова Яна Евгеньевна — кандидат физико-математических наук, доцент кафедры теории вероятностей и математической статистики НИ ТГУ. E-mail: evgenevna.92@mail.ru

О НЕКОТОРЫХ СООТНОШЕНИЯХ ДЛЯ ВРЕМЕНИ РАЗГРУЗКИ ОЧЕРЕДИ ПРИ ЦИКЛИЧЕСКОМ УПРАВЛЕНИИ ФОРМИРУЕМЫМИ В СЛУЧАЙНОЙ СРЕДЕ ПОТОКАМИ

А. В. Зорин

Национальный исследовательский Нижегородский государственный университет им. Н.И. Лобачевского, г. Нижний Новгород, Россия

В работе рассматривается система обслуживания конечного числа входных потоков по циклическому алгоритму с фиксированным ритмом переключения. Входные потоки формируются во внешней случайной среде с конечным числом состояний, синхронизированной с обслуживающим устройством. При каждом из состояний среды каждый входной поток является пуассоновским неординарным потоком с параметрами, зависящими от случайной среды. Построена математическая модель в виде многомерной счетной цепи Маркова, получены необходимые и достаточные условия существования стационарного распределения. Построены мартингалы последовательности, с помощью которых можно изучать среднее время до опустошения выбранной очереди. **Ключевые слова:** конфликтные потоки, случайная среда, циклический алгоритм обслуживания, время опустошения очереди, цепь Маркова, условие стационарности, мартингалы.

Введение

Представляющей практический интерес задачей является исследование времени достижения очередью нулевого уровня в системе обслуживания конфликтных потоков. В настоящей работе будет найдено выражение для среднего времени достижения нулевого уровня очереди в системе массового обслуживания в классе циклических алгоритмов для входных потоков, формируемых в случайной среде. Имея в виду приложения к задачам управления транспортными потоками, случайную среду можно интерпретировать как погодные условия, влияющие как на интенсивности поступления автотранспорта, так и на структуру транспортных пачек. Поскольку погодные условия меняются гораздо медленнее, чем происходят переключения светофора, выберем в качестве математической модели погоды цепь Маркова, возможные смены

состояний которой могут происходить только вместе со сменой состояний сигналов светофора.

Управление входными потоками с аналогичной переменной вероятностной структурой рассматривалось в работах [1, 2] где были установлены условия существования стационарного режима при дополнительном предположении о сохранении интенсивности потоков при разных состояниях внешней среды. В целом, исследование систем обслуживания с формируемыми в случайной среде потоками имеет важное теоретическое и прикладное значение [6]. Задача о времени опустошения очереди при циклическом управлении рассматривалась в [3]. Используемые мартингалные методы восходят к работам [4, 5].

1. Постановка задачи и математическая модель

В систему поступают $m < \infty$ конфликтных потоков. Вероятностная структура всех потоков зависит от состояния внешней случайной среды с конечным числом возможных состояний $\{e^{(1)}, e^{(2)}, \dots, e^{(d)}\}$, $d < \infty$. При состоянии среды $e^{(k)}$, требования потока Π_j поступают группами с известными параметрами $\lambda_j^{(k)} > 0$ — интенсивность поступлений групп и $g_j^{(k)}(b)$ — вероятность того, что группа содержит $b = 1, 2, \dots$ требований. Будем предполагать, что при каждой фиксированной паре (j, k) наименьший шаг решетки, на которой сосредоточено распределение $\{g_j^{(k)}(b); b = 1, 2, \dots\}$, равен единице. Смена состояний случайной среды может происходить только в моменты смены состояний обслуживающего устройства, и пусть $a_{k,l}$ есть вероятность смены состояния $e^{(k)}$ на состояние $e^{(l)}$, а состояния связаны в эргодическую цепь Маркова. Требования потока Π_j помещаются в накопитель O_j неограниченной вместимости. Обслуживающее устройство имеет $2m$ состояний $\Gamma^{(1)}, \Gamma^{(2)}, \dots, \Gamma^{(2m)}$. Длительность пребывания в состоянии $\Gamma^{(r)}$ неслучайна и равна T_r , $r = 1, 2, \dots, 2m$. В состоянии вида $\Gamma^{(2j)}$, $j = 1, 2, \dots, m$, требования не обслуживаются. В состоянии $\Gamma^{(2j-1)}$ обслуживаются требования только из очереди O_j . Максимальное число обслуженных требований при этом не может превышать заданного неслучайного порога $\ell_j \in \{1, 2, \dots\}$. Обслуженные требования покидают систему.

Будем считать, что все определяемые ниже случайные величины и случайные элементы определены на общем вероятностном пространстве $(\Omega, \mathfrak{F}, \mathbb{P})$. Обозначим через τ_i , $i = 0, 1, \dots$ моменты смены состояний обслуживающего устройства, причем $\tau_0 = 0$. Пусть $\Gamma_i \in \{\Gamma^{(1)}, \Gamma^{(2)}, \dots, \Gamma^{(2m)}\}$ есть состояние обслуживающего устройства на промежутке $(\tau_{i-1}, \tau_i]$, обозначим через $X_{j,i}$ число требований в очереди O_j в момент τ_i , через $n_{j,i}$ — число требований входного потока Π_j , по-

ступивших на промежутке $(\tau_i, \tau_{i+1}]$, $\xi_{j,i}$ — число требований j -го потока насыщения на промежутке $(\tau_i, \tau_{i+1}]$, $\bar{\xi}_{j,i}$ — число обслуженных требований из очереди O_j на промежутке $(\tau_i, \tau_{i+1}]$. Пусть $r \oplus 1 = r + 1$ для $r = 1, 2, \dots, 2m - 1$, $(2m) \oplus 1 = 1$, $r \oplus x = (r \oplus (x - 1)) \oplus 1$ для целых неотрицательных x . Наконец, обозначим через χ_i состояние случайной среды на промежутке $(\tau_i, \tau_{i+1}]$. Из постановки задачи устанавливаются следующие функциональные соотношения:

$$X_{j,i+1} = X_{j,i} + \eta_{j,i} - \bar{\xi}_{j,i} = \max\{0, X_{j,i} + \eta_{j,i} - \xi_{j,i}\}, \quad (1)$$

$$\bar{\xi}_{j,i} = \min\{\xi_{j,i}, X_{j,i} + \eta_{j,i}\}, \quad (2)$$

$$\Gamma_{i+1} = u(\Gamma_i), \quad \text{где } u(\Gamma^{(r)}) = \Gamma^{(r \oplus 1)}, \quad (3)$$

$$\tau_{i+1} = \tau_i + v(\Gamma_i), \quad \text{где } v(\Gamma^{(r)}) = T_{r \oplus 1}. \quad (4)$$

Также из постановки задачи следует, что при фиксированном значении $(\Gamma^{(r)}, e^{(k)})$ вектора (Γ_i, χ_i) , набор $(\chi_{i_1}, \eta_{1,i}, \eta_{2,i}, \dots, \eta_{m,i})$ условно независим с $(\xi_{1,i}, \xi_{2,i}, \dots, \xi_{m,i})$ и оба набора не зависят от событий до момента τ_i . При этом, для произвольных целых неотрицательных $b_1, b_2, \dots, b_m, \bar{b}_1, \bar{b}_2, \dots, \bar{b}_m$ условная вероятность события

$$\begin{aligned} \{\omega: \chi_{i+1} = e^{(l)}, \eta_{1,i} = b_1, \eta_{2,i} = b_2, \dots, \eta_{m,i} = b_m\} \cap \\ \cup \{\omega: \xi_{1,i} = \bar{b}_1, \xi_{2,i} = \bar{b}_2, \dots, \xi_{m,i} = \bar{b}_m\} \end{aligned}$$

равна

$$a_{k,l} \varphi_1(b_1; T_{r \oplus 1}, k) \varphi_2(b_2; T_{r \oplus 1}, k) \times \varphi_m(b_m; T_{r \oplus 1}, k)$$

если либо нечетно r и $\bar{b}_1 = \bar{b}_2 = \dots = \bar{b}_m = 0$, либо четно $r = 2j$ и $\bar{b}_s = \ell_{j \oplus 1} \delta_{s,j \oplus 1}$, $s = 1, 2, \dots, m$; в остальных случаях эта вероятность равна 0. Здесь $\delta_{s,j}$ — дельта-функция Кронекера, принимающая значение 1 для $s \neq j$ и значение 1 при $s = j$, а функции $\varphi_j(\cdot; \cdot, \cdot)$, $j = 1, 2, \dots, m$, определяются из разложений производящих функций для простого неординарного потока:

$$\sum_{b=0}^{\infty} z^b \varphi_j(b; t, k) = \exp\left\{ \lambda_j^{(k)} t \left(\sum_{b=1}^{\infty} g_j^{(k)}(b) z^b - 1 \right) \right\}, \quad t > 0, |z| < 1.$$

Теорема 1. Многомерная случайная последовательность

$$\{(\Gamma_i, \chi_i, X_{1,i}, X_{2,i}, \dots, X_{m,i}); i = 0, 1, \dots\} \quad (5)$$

при заданном распределении вектора $(\Gamma_0, \chi_0, X_{1,0}, X_{2,0}, \dots, X_{m,0})$ является однородной цепью Маркова с единственным классом сообщающихся периодических состояний с периодом $2m$.

Пусть α_k — стационарная вероятность состояния среды $e^{(k)}$, $k = 1, 2, \dots, d$. Обозначим через $\mu_j^{(k)} = \sum_{b=1}^{\infty} b g_j^{(k)}(b) < \infty$ средний размер группы требований по потоку Π_j при состоянии среды $e^{(k)}$, и пусть $\bar{\lambda}_j^{(k)} = \lambda_j^{(k)} \mu_j^{(k)}$.

Теорема 2. Для существования единственного стационарного распределения цепи Маркова (5) необходимо и достаточно выполнения неравенства:

$$\max_{1 \leq j \leq m} \left\{ (T_1 + \dots + T_{2m}) \sum_{k=1}^d \alpha_k \bar{\lambda}_j^{(k)} - \ell_j \right\} < 0. \quad (6)$$

Доказательство теоремы 2 основано лемме 1, содержащей рекуррентные соотношения для производящих функций

$$\Psi_{j,i}(z; r, k) = \mathbb{E}(z^{X_{j,i}} I(\Gamma_i = \Gamma^{(r)}, \chi_i = e^{(k)})), \quad |z| \leq 1.$$

Пусть

$$f_j^{(k)}(z) = \sum_{b=1}^{\infty} z^b g_j^{(k)}(b), \quad q_j(z; t, e^{(k)}) = \exp\left\{ \lambda_j^{(k)} t (f_j^{(k)}(z) - 1) \right\}, \quad |z| \leq 1.$$

Лемма 1. Имеют место рекуррентные по $i = 0, 1, \dots$ соотношения:

$$\begin{aligned} \Psi_{j,i+1}(z; r \oplus 1, l) &= \sum_{k=1}^d a_{k,l} q_j(z; T_{r \oplus 1}, e^{(k)}) \Psi_{j,i}(z; r, k), \quad r \oplus 1 \neq 2j - 1; \\ \Psi_{j,i+1}(z; r \oplus 1, l) &= \sum_{k=1}^d a_{k,l} \left(z^{-\ell_j} q_j(z; T_{r \oplus 1}, e^{(k)}) \Psi_{j,i}(z; r, k) + \right. \\ &+ \left. \sum_{x=0}^{\ell_j} \mathbb{P}(\Gamma_i = \Gamma^{(r)}, X_{j,i} = x, \chi_i = e^{(k)}) \sum_{b=0}^{\ell_j - 1 - x} \varphi_j(b; T_{r \oplus 1}, k) (1 - z^{x+b-\ell_j}) \right). \end{aligned}$$

В остальной части работы будем считать условия существования стационарного распределения выполненными.

2. Изучение среднего времени разгрузки очереди

Всюду далее, не уменьшая общности, положим $j = 1$. Введем случайный момент

$$\nu = \begin{cases} \infty, & \text{если } X_{1,i} > 0 \text{ для всех } i = 1, 2, \dots; \\ \min\{i \geq 1: X_{1,i} = 0\} & \text{в противном случае.} \end{cases}$$

Под периодом разгрузки очереди O_1 будем понимать случайный промежуток времени $[0, \tau_\nu)$.

Введем матрицы $A(z; t) = (a_{k,l} a_1^{(k)}(z; t))_{k,l=\overline{1,d}}$, $t > 0$. При $0 < z < 1$ элементы неразложимой (в силу эргодичности внешней случайной среды) матрицы $A(z, T_r)$ неотрицательны, следовательно, по теореме Фробениуса она имеет простое положительное собственное число $\zeta_r(z)$, мажорирующее модули всех остальных ее собственных чисел и такое что $\zeta_r(1) = 1$. Распространим эту функцию на комплексный круг $|z| \leq 1$.

Лемма 2. Пусть $\zeta_r(z)$ — регулярная в области $|z| < 1$ ветвь функции, удовлетворяющей характеристическому уравнению

$$\det(A(z, T_r) - \zeta I_d) = 0, \quad (I_d = (\delta_{k,l})_{k,l=\overline{1,d}}), \quad (7)$$

такая что $\zeta_r(1) = 1$. Тогда $|\zeta_r(z)| < 1$ при $|z| \leq 1$, $z \neq 1$ и $\zeta_r'(1) = T_r \sum_{k=1}^d \alpha_k \bar{\lambda}_1^{(k)}$.

Заметим, что одно из неравенств (6), соответствующее первой очереди ($j = 1$) примет вид: $\sum_{r=1}^{2m} \zeta_r'(1) < \ell_1$.

Лемма 3. Уравнение $z^{\ell_1} = \zeta_1(z)\zeta_2(z) \times \dots \times \zeta_{2m}(z)$ имеет $\ell_1 - 1$ ноль $\beta_1, \beta_2, \dots, \beta_{\ell_1-1}$ в круге $|z| < 1$ и ноль $\beta_{\ell_1} = 1$ на границе.

Пусть $\psi^{(r)}(z)$ — правый собственный вектор матрицы $A(z, T_r)$, соответствующий собственному числу $\zeta_r(z)$ и удовлетворяющий условию $\psi^{(r)}(1) = \bar{1}$.

Теорема 3. При $|z| \leq 1$ случайная последовательность

$$M_i(z) = z^{X_{1,i} + \sum_{s=0}^{i-1} \min\{\xi_{1,s}, X_{1,s} + \eta_{1,s}\}} \frac{\prod_{s=0}^i \sum_{r=1}^{2m} I(\Gamma_s = \Gamma^{(r \oplus 1)}) \psi_{\chi_s}^{(r \oplus 1)}(z)}{\prod_{s=0}^{i-1} \sum_{r=1}^{2m} I(\Gamma_s = \Gamma^{(r)}) \zeta_{r \oplus 1}(z) \psi_{\chi_s}^{(r \oplus 1)}(z)}, \quad (8)$$

$i = 0, 1, \dots$, является мартингалом относительно потока

$$\tilde{\mathfrak{F}}_i = \sigma(X_{1,0}, \Gamma_0, \eta_{1,0}, \eta_{1,1}, \dots, \eta_{1,i-1}, \chi_0, \chi_1, \dots, \chi_i), \quad i = 0, 1, \dots$$

Следствие 1. Пусть $x = 1, 2, \dots$. Имеет место следующее соотношение для момента ν опустошения очереди:

$$\mathbb{E} \left(z^{X_{1,\nu-1} + \eta_{1,\nu-1}} \left(\frac{z^{\ell_1}}{\zeta_1(z) \times \dots \times \zeta_{2m}(z)} \right)^{\lfloor \frac{\nu-1}{2m} \rfloor} \times \right.$$

$$\begin{aligned} & \times \frac{\prod_{s=0}^{\nu} \psi_{\chi_s}^{(r \oplus s)}(z)}{\prod_{s=0}^{\nu-1} \psi_{\chi_s}^{(r \oplus (s+1))}(z)} \left| \left\{ \Gamma_0 = \Gamma^{(r)}, \chi_0 = e^{(k)}, X_{1,0} = x \right\} \right. = \\ & = z^x \psi_k^{(r)}(z) \zeta_1(z) \begin{cases} 1, & r = 2m, \\ \zeta_{r \oplus 1}(z) \times \dots \times \zeta_{2m}(z), & r = 1, 2, \dots, 2m - 1. \end{cases} \end{aligned}$$

Заключение

В настоящей работе построен мартигал, связанный с колебаниями длины выбранной очереди, а также с состояниями внешней случайной среды и обслуживающего устройства для системы обслуживания формируемых в случайной среде конфликтных потоков по циклическому алгоритму. В дальнейшем изучение этого мартинагала должно позволить изучить время, необходимое для опустошения выбранной очереди.

СПИСОК ЛИТЕРАТУРЫ

1. Куделин, А.Н. Управление конфликтными потоками в случайной среде по информации о наличии очереди / А.Н. Куделин, М.А. Федоткин. — Нижегородский государственный университет им. Н.И. Лобачевского. Деп. в ВИНТИ, № 1717–В96. — Н. Новгород, 1996, — 22 с.
2. Куделин, А.Н. Предельные теоремы для систем управления потоками в случайной среде в классе алгоритмов с упреждением / А.Н. Куделин, М.А. Федоткин. — Нижегородский государственный университет им. Н.И. Лобачевского. Деп. в ВИНТИ, №2593–В96. — Н. Новгород, 1996. — 40 с.
3. Зорин А. В. К определению периода занятости при нелокальном описании потоков // Информатика и ее применения. 2024. № 3 (принято в печать).
4. Baccelli F., Makowski A. M. Dynamic, transient and stationary behavior of the M/GI/1 queue via martingales // The Annals of Probability. 1989. V. 17. №. 4. P. 1691–1699.
5. Baccelli F., Makowski A. M. Martingale relations for the M/GI/1 queue with Markov modulated Poisson input // Stochastic processes and their applications. 1991. V. 38. P. 99–133.
6. Dudin A. N., Klimenok V. I., Vishnevsky V. M. The theory of queuing systems with correlated flows Springer Cham, 2020. 430 p.

Зорин Андрей Владимирович — доктор физико-математических наук, доцент, заведующий кафедрой теории вероятностей и анализа данных Нижегородского государственного университета им. Н.И.Лобачевского. E-mail: andrei.zorine@itmm.unn.ru

МОДЕЛЬ СЕТИ МАССОВОГО ОБСЛУЖИВАНИЯ С ТРЕМЯ УЗЛАМИ ДЛЯ РАСЧЕТА ТРЕБУЕМОЙ ВЫЧИСЛИТЕЛЬНОЙ МОЩНОСТИ СЕТИ ИНТЕРНЕТА ВЕЩЕЙ

Б. Э. Бабаева, Е. Ю. Лисовская

*Национальный исследовательский
Томский государственный университет, г. Томск, Россия*

В данной работе проводится исследование сети массового обслуживания, состоящей из трех узлов, каждый из которых представляет суммарную вычислительную мощность разных уровней сети Интернета вещей (IoT): граничный (edge), туманный (fog) и облачный (cloud). В результате исследования предлагается инструмент, позволяющий рассчитать требуемую вычислительную мощность на узлах сети. В работе приводится демонстрационный пример, в котором вычисляется необходимое количество процессоров для обработки запросов на каждом из уровней вычислений для семи типов устройств, чтобы минимизировать вероятность отказа и простоя обслуживания.

Ключевые слова: *сеть массового обслуживания, интернет вещей, СЛАУ, оптимизация.*

Введение

Рассмотрим некоторую территорию города, на которой необходимо развернуть вычислительную сеть трех уровней: граничный, туманный и облачный [1].

Предполагается, что подобные сети уже были развернуты в других районах города, поэтому становится возможным статистическая оценка параметров математической модели. На рассматриваемой территории будут обслуживаться устройства следующих категорий: 1 – датчики, 2 – активаторы, 3 – умные устройства для дома, 4 – здравоохранение и носимая электроника, 5 – промышленные устройства, 6 – транспортные устройства, 7 – сельское хозяйство.

Количество вычислительных устройств на каждом из уровней вычислений ограничено. Ставится задача поиска оптимального количества вычислительных устройств на узлах сети для рассматриваемых типов запросов с целью уменьшения вероятности отказа и простоя вычислительных устройств.

Данная работа является частью исследований (такие как [2, 3] и множество других), посвященных технологии интернета вещей, и предлагает еще один подход к моделированию сети интернета вещей и вычислению требуемой вычислительной мощности на вычислительных узлах.

1. Математическая модель

В качестве математической модели сети устройств интернета вещей рассматривается открытая СеМО [6] с тремя узлами, каждый из которых представляет различный уровень вычислений (граничный, туманный и облачный), с марковской маршрутизацией и ограниченным числом приборов в узлах (N_1, N_2, N_3 , соответственно). В сеть поступает простейший поток заявок с интенсивностью λ [4]. Каждая из заявок с вероятностью p_k поступает в k -ый узел, где находится в течение случайного времени, имеющего экспоненциальную функцию распределения с параметром μ_k , $k \in \{1, 2, 3\}$. После чего заявка либо переходит в узел l с вероятностью r_{kl} , которая является элементом матрицы маршрутизации, либо покидает сеть с вероятностью

$$r_{k0} = 1 - \sum_{i=1}^3 r_{ki}.$$

Иллюстрация математической модели представлена на рисунке 1.

Обозначим $n_k(t)$ – число заявок, которые находятся на обслуживании в k -ом узле сети в момент времени t . Тогда обозначим стационарные (финальные) вероятности состояний из множества $\mathbb{N} = \{(n_1, n_2, n_3)\} : 0 \leq n_k \leq N_k\}$ случайного процесса $\{n_1(t), n_2(t), n_3(t)\}$:

$$\lim_{t \rightarrow \infty} P\{n_1(t) = n_1, n_2(t) = n_2, n_3(t) = n_3\} = \pi(n_1, n_2, n_3).$$

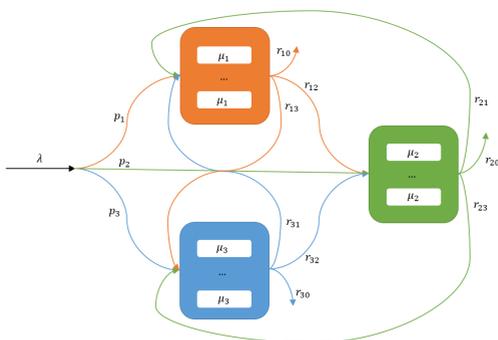


Рис. 1. Схема СеМО с тремя узлами

Составим систему уравнений равновесия для стационарных вероятностей состояний исследуемого трехмерного процесса:

$$\begin{aligned} & \pi(n_1, n_2, n_3)[(\lambda p_1 + n_2 \mu_2 r_{21} + n_3 \mu_3 r_{31})I(n_1 + 1 \leq N_1) + \\ & + n_1 \mu_1 r_{10} + n_2 \mu_2 r_{20} + n_3 \mu_3 r_{30} + (\lambda p_2 + n_1 \mu_1 r_{12} + n_3 \mu_3 r_{32})I(n_2 + 1 \leq N_2) \\ & + (\lambda p_3 + n_1 \mu_1 r_{13} + n_2 \mu_2 r_{23})I(n_3 + 1 \leq N_3)] = \\ & = \lambda p_1 \pi(n_1 - 1, n_2, n_3) + \lambda p_2 \pi(n_1, n_2 - 1, n_3) + \lambda p_3 \pi(n_1, n_2, n_3 - 1) + \\ & + (n_1 + 1) \mu_1 r_{10} \pi(n_1 + 1, n_2, n_3) I(n_1 + 1 \leq N_1) + \\ & + (n_2 + 1) \mu_2 r_{20} \pi(n_1, n_2 + 1, n_3) I(n_2 + 1 \leq N_2) + \\ & + (n_3 + 1) \mu_3 r_{30} \pi(n_1, n_2, n_3 + 1) I(n_3 + 1 \leq N_3) + \\ & + (n_1 + 1) \mu_1 (r_{12} \pi(n_1 + 1, n_2 - 1, n_3) + r_{13} \pi(n_1 + 1, n_2, n_3 - 1)) I(n_1 + 1 \leq N_1) + \\ & + (n_2 + 1) \mu_2 (r_{21} \pi(n_1 - 1, n_2 + 1, n_3) + r_{23} \pi(n_1, n_2 + 1, n_3 - 1)) I(n_2 + 1 \leq N_2) + \\ & + (n_3 + 1) \mu_3 (r_{31} \pi(n_1 - 1, n_2, n_3 + 1) + r_{32} \pi(n_1, n_2 - 1, n_3 + 1)) I(n_3 + 1 \leq N_3), \end{aligned}$$

где

$$I(A) = \begin{cases} 1, & \text{если } A \text{ истина,} \\ 0, & \text{если } A \text{ ложь.} \end{cases}$$

Решая матричное уравнение $\mathbf{A}\boldsymbol{\pi} = \mathbf{b}$ совместно с условием нормировки $\boldsymbol{\pi}\mathbf{e} = 1$, где \mathbf{A} – матрица коэффициентов системы, \mathbf{b} – вектор-столбец свободных членов, \mathbf{e} – единичный вектор-столбец, найдем вектор-строку вероятностей состояний $\boldsymbol{\pi} = [\pi(n_1, n_2, n_3)]_{(n_1, n_2, n_3) \in \mathbb{N}}$.

2. Решение системы уравнений

Систему уравнений для стационарного распределения вероятностей состояний сети будем решать численным методом с помощью библиотеки `numpy` языка программирования Python согласно алгоритмам 1 и 2.

Алгоритм 1. Вычисление стационарного распределения вероятностей состояний сети массового обслуживания

Вход: $N_1, N_2, N_3, \lambda, p_1, p_2, p_3, \mu_1, \mu_2, \mu_3,$
 $r_{10}, r_{12}, r_{13}, r_{20}, r_{21}, r_{23}, r_{30}, r_{31}, r_{32}$

Выход: $\pi = [\pi(n_1, n_2, n_3)]_{(n_1, n_2, n_3) \in \mathbb{N}}$

- 1: Вычислить общее количество состояний $|\mathbb{N}| = (N_1 + 1)(N_2 + 1)(N_3 + 1)$.
 - 2: Установить взаимно-однозначное соответствие между индексом состояния $i \in [0, |\mathbb{N}|]$ и состоянием (n_1, n_2, n_3) .
 - 3: Заполнить матрицу \mathbf{A} коэффициентов системы уравнений по алгоритму 2.
 - 4: Заполнить последнюю строку матрицы \mathbf{A} единицами.
 - 5: Заполнить вектор свободных членов \mathbf{b} , у которого все элементы равны нулю, кроме последнего – он равен единице.
 - 6: Решить систему линейных алгебраических уравнений $\mathbf{A}\pi = \mathbf{b}$ с помощью метода `numpy.linalg.solve`.
 - 7: **возвратить** π
-

Алгоритм 2. Заполнение матрицы коэффициентов системы

Вход: $|\mathbb{N}|, N_1, N_2, N_3, \lambda, p_1, p_2, p_3, \mu_1, \mu_2, \mu_3,$
 $r_{10}, r_{12}, r_{13}, r_{20}, r_{21}, r_{23}, r_{30}, r_{31}, r_{32}$

Выход: **A**

```

1: цикл
2: Для каждого состояния  $(n_1, n_2, n_3)$  найти соответствующее  $i$ 
   и выполнить следующие шаги:
3: для всех  $(n_1, n_2, n_3)$ 
4:    $A[i, i]^- = n_1\mu_1r_{10}, \quad A[i, i]^- = n_2\mu_2r_{20}, \quad A[i, i]^- = n_3\mu_3r_{30}$ 
5: если  $n_1 + 1 \leq N_1$  то
6:    $A[i, i]^- = \lambda p_1 + n_2\mu_2r_{21} + n_3\mu_3r_{31}$ 
7:    $A[i, (n_1 + 1, n_2, n_3)]^+ = (n_1 + 1)\mu_1r_{10}$ 
8: если  $n_2 + 1 \leq N_2$  то
9:    $A[i, i]^- = \lambda p_2 + n_1\mu_1r_{12} + n_2\mu_2r_{32}$ 
10:   $A[i, (n_1, n_2 + 1, n_3)]^+ = (n_2 + 1)\mu_2r_{20}$ 
11: если  $n_3 + 1 \leq N_3$  то
12:   $A[i, i]^- = \lambda p_3 + n_1\mu_1r_{13} + n_2\mu_2r_{23}$ 
13:   $A[i, (n_1, n_2, n_3 + 1)]^+ = (n_3 + 1)\mu_3r_{30}$ 
14: если  $n_1 > 0$  то
15:   $A[i, (n_1 - 1, n_2, n_3)]^+ = \lambda p_1$ 
16: если  $n_2 > 0$  то
17:   $A[i, (n_1, n_2 - 1, n_3)]^+ = \lambda p_2$ 
18: если  $n_3 > 0$  то
19:   $A[i, (n_1, n_2, n_3 - 1)]^+ = \lambda p_3$ 
20: если  $n_1 + 1 \leq N_1$  and  $n_2 > 0$  то
21:   $A[i, (n_1 + 1, n_2 - 1, n_3)]^+ = (n_1 + 1)\mu_1r_{12}$ 
22: если  $n_1 + 1 \leq N_1$  and  $n_3 > 0$  то
23:   $A[i, (n_1 + 1, n_2, n_3 - 1)]^+ = (n_1 + 1)\mu_1r_{13}$ 
24: если  $n_2 + 1 \leq N_2$  and  $n_1 > 0$  то
25:   $A[i, (n_1 - 1, n_2 + 1, n_3)]^+ = (n_2 + 1)\mu_2r_{21}$ 
26: если  $n_2 + 1 \leq N_2$  and  $n_3 > 0$  то
27:   $A[i, (n_1, n_2 + 1, n_3 - 1)]^+ = (n_2 + 1)\mu_2r_{23}$ 
28: если  $n_3 + 1 \leq N_3$  and  $n_1 > 0$  то
29:   $A[i, (n_1 - 1, n_2, n_3 + 1)]^+ = (n_3 + 1)\mu_3r_{31}$ 
30: если  $n_3 + 1 \leq N_3$  and  $n_2 > 0$  то
31:   $A[i, (n_1, n_2 - 1, n_3 + 1)]^+ = (n_3 + 1)\mu_3r_{32}$ 
32: возвратить A

```

3. Численный эксперимент

Вернемся к рассмотрению территории города. В таблице 1 представлены параметры сети для каждой из категорий устройств интернета ве-

щей. Параметры подобраны таким образом, чтобы отразить специфику поведения запросов от соответствующих устройств. Рассмотренная математическая модель является моделью процесса обработки трафика для одной из категорий. Для каждого из типов устройств находим такое количество приборов N_1, N_2, N_3 , чтобы вероятность отказа на каждом блоке была не более 0,1, а коэффициент загрузки R всей сети был максимальным. Результаты эксперимента также приведены в таблице 1.

Таблица 1

Данные для эксперимента и результаты

№	λ	μ_1	p_1	$r_{12}/r_{13}/r_{10}$	N_1	R
		μ_2	p_2	$r_{21}/r_{23}/r_{20}$	N_2	
		μ_3	p_3	$r_{31}/r_{32}/r_{30}$	N_3	
1	50	100	0,6	0,7/0,2/0,1	2	0,817
		10	0,3	0,2/0,5/0,3	10	
		1	0,1	0,1/0,4/0,5	50	
2	50	100	0,3	0,7/0,2/0,1	2	0,829
		10	0,5	0,2/0,5/0,3	10	
		1	0,2	0,1/0,4/0,5	52	
3	50	50	0,4	0,6/0,3/0,1	3	0,886
		5	0,3	0,3/0,4/0,3	14	
		0,5	0,3	0,1/0,3/0,6	96	
4	20	20	0,4	0,6/0,3/0,1	3	0,886
		2	0,3	0,3/0,4/0,3	14	
		0,2	0,3	0,1/0,3/0,6	96	
5	80	200	0,6	0,7/0,2/0,1	2	0,867
		10	0,3	0,2/0,5/0,3	14	
		1	0,1	0,1/0,4/0,5	78	
6	80	50	0,4	0,6/0,3/0,1	4	0,679
		20	0,3	0,3/0,4/0,3	8	
		5	0,3	0,1/0,4/0,5	20	
7	20	20	0,6	0,7/0,2/0,1	6	0,730
		10	0,3	0,3/0,4/0,3	10	
		2	0,1	0,2/0,3/0,5	25	

Заключение

В данной работе построена модель сети массового обслуживания с тремя узлами и ограниченным числом приборов на узлах, составлена система уравнений, решение которой, дает распределение вероятностей числа заявок на узлах сети. Приведен алгоритм для вычисления распределения вероятностей состояний сети для произвольного числа приборов на узлах сети, которое позволяет найти оптимальное количество устройств на каждом из уровней вычислений.

СПИСОК ЛИТЕРАТУРЫ

1. *Bouaouad A. E., Cherradi A., Assoul S., Souissi N.* Architectures and emerging trends in Internet of Things and Cloud computing: a literature review // 2020 4th International Conference on Advanced Systems and Emergent Technologies (IC ASET), Hammamet, Tunisia, 2020. P. 147–151.
2. *Maiti P., Apat H. K., Kumar A., Sahoo B., Turuk A. K.* Deployment of Multi-tier Fog Computing System for IoT Services in Smart City // 2019 IEEE International Conference on Advanced Networks and Telecommunications Systems (ANTS), Goa, India, 2019. P. 1–6.
3. *Islam J., Kumar T., Kovacevic I., Harjula E.* Resource-Aware Dynamic Service Deployment for Local IoT Edge Computing: Healthcare Use Case // IEEE Access. 2021. Vol. 9. P. 115868–115884.
4. *Мусеев А. Н., Назаров А. А.* Бесконечнолинейные системы и сети массового обслуживания // Томск: НТЛ, 2015. 240 с.
5. *Назаров А. А., Терпугов А. Ф.* Теория массового обслуживания // Томск: НТЛ, 2010. 228 с.

Бабаева Басти Эльдаровна — студент ИПМКН НИ ТГУ. E-mail: *basti.babaeva@yandex.ru*

Лисовская Екатерина Юрьевна — канд. физ.-мат. наук, доцент каф. ТВиМС ИПМКН НИ ТГУ. E-mail: *ekaterina_lisovs@mail.ru*

АСИМПТОТИЧЕСКИЙ АНАЛИЗ СИСТЕМЫ МАССОВОГО ОБСЛУЖИВАНИЯ С ОЖИДАЮЩИМИ ОТРИЦАТЕЛЬНЫМИ ЗАЯВКАМИ

Д. А. Королев, С. П. Моисеева, А. Н. Моисеев

*Национальный исследовательский Томский государственный университет,
г. Томск, Россия*

В работе рассматривается математическая модель системы с ожидающими отрицательными заявками, а также асимптотический анализ этой системы. Получен вид характеристической функции числа положительных и отрицательных заявок в данной системе. **Ключевые слова:** *система массового обслуживания, отрицательные заявки, асимптотический анализ.*

Введение

В современном мире сферы обслуживания, такие как клиентские сервисы, транспортные системы, медицинские учреждения и многие другие, сталкиваются с растущими потоками заявок, где как положительные, так и отрицательные обращения требуют эффективного управления. Подобные ситуации нередко включают в себя не только обслуживание клиентов, но и управление нежелательными событиями или запросами. Информационные технологии играют все более важную роль в повседневной деятельности как предприятий, так и частных лиц. От электронной коммерции до социальных сетей, наша цифровая инфраструктура подвергается постоянному потоку запросов, которые можно рассматривать как заявки в системе массового обслуживания (СМО). В реальной жизни, сфера информационных технологий сталкивается с различными вызовами, связанными с обработкой разнообразных заявок и выявлением оптимальных стратегий управления. Например, атаки злоумышленников или сбои в работе могут негативно отразиться на производительности и безопасности информационных систем. Отрицательные заявки, как и в реальной жизни, в системах обслуживания могут не только потреблять ресурсы, но и влиять на обработку положительных заявок. Например, они могут вызывать задержки или

даже блокировки процессов обслуживания, что отрицательно сказывается на общей эффективности системы.

Рассмотрим процесс обработки сетевых запросов в центре обработки данных (Data Center). В этой системе положительные заявки представляют собой запросы на доступ к определенным данным или ресурсам, которые обрабатываются центром обработки данных. Такие запросы поступают непрерывно в систему в течение случайного времени. С другой стороны, отрицательные заявки могут быть представлены в виде злонамеренных атак или DDoS-атак (атак распределенного отказа в обслуживании). Когда такая отрицательная заявка поступает, она не взаимодействует с предыдущей обрабатываемой положительной заявкой, а начинает свою деструктивную активность по перегрузке сетевой инфраструктуры. Когда поступают новые запросы на доступ к данным, помехи, вызванные DDoS-атаками, могут заблокировать или уничтожить эти запросы.

В [6] представлена СМО с отрицательными заявками и бункером для вытеснения положительных заявок. В работе [14] была рассмотрена бесконечнолинейная СМО с отрицательными заявками с ожиданием, в которой приходящая отрицательная заявка мгновенно уничтожает обслуживающую положительную заявку. В отличие от упомянутой работы в представленной статье рассматривается случай, когда входящая отрицательная заявка не взаимодействует с имеющимися положительными.

1. Математическая модель

Рассмотрим систему массового обслуживания с неограниченным числом серверов. Пусть на вход системы поступают два простейших потока заявок, по одному из них с интенсивностью λ поступают обычные (положительные) заявки, по другому с интенсивностью α – отрицательные. Положительная заявка, приходящая в систему, выбирает любой из свободных приборов и сразу начинает обслуживание в течение случайного времени, экспоненциально распределенного с параметром μ . В задаче с ожиданием отрицательная заявка, приходящая в систему, не взаимодействует с уже обслуживаемыми положительными заявками, а ожидает прихода новой положительной заявки. Как только такая заявка появляется, отрицательная заявка уничтожает её и они обе покидают систему. В случае если система пуста, отрицательная заявка ожидает поступления положительной для выполнения указанной процедуры. Если в системе присутствует несколько отрицательных заявок, то только одна (любая) реагирует описанным образом на поступление новой положительной.

В работе исследуется случайный процесс $\{i(t), l(t)\}$, где $i(t)$ – число положительных, а $l(t)$ – число отрицательных заявок в системе в момент времени t .

Для распределения вероятностей $P(i, l, t) = P\{i(t) = i, l(t) = l\}$ построим систему дифференциальных уравнений Колмогорова в стационарном режиме:

$$\begin{aligned} -(\lambda + \alpha)P(0, 0) + \mu P(1, 0) + \lambda P(0, 1) &= 0, \\ -(\lambda + \alpha + i\mu)P(i, 0) + \lambda P(i - 1, 0) + (i + 1)\mu P(i + 1, 0) + \lambda P(i, 1) &= 0, \\ -(\lambda + \alpha)P(0, l) + \mu P(1, l) + \alpha P(0, l - 1) + \lambda P(0, l + 1) &= 0, \\ -(\lambda + \alpha + i\mu)P(i, l) + \alpha P(i, l - 1) + (i + 1)\mu P(i + 1, l) + \lambda P(i, l + 1) &= 0. \end{aligned}$$

Используя метод частичных характеристических функций [15], можно получить следующую систему:

$$\begin{aligned} j\mu(1 - e^{-ju}) \frac{dH(u, 0)}{du} + (\lambda e^{ju} - (\lambda + \alpha))H(u, 0) + \lambda H(u, 1) &= 0, \\ j\mu(1 - e^{-ju}) \frac{dH(u, l)}{du} - (\lambda + \alpha)H(u, l) + \alpha H(u, l - 1) + \lambda H(u, l + 1) &= 0, \end{aligned}$$

где $H(u, l) = \sum_{i=0}^{\infty} e^{ju i} P(i, l)$, $l = \overline{0, \infty}$.

2. Метод асимптотического анализа

Поскольку прямое решение полученной системы дифференциальных уравнений не представляется возможным, применим для её решения метод асимптотического анализа [7] в предельном условии эквивалентно высокой интенсивности входящих потоков. Это условие можно определить как

$$\lambda = \lambda N, \quad \alpha = \alpha N,$$

где $N \rightarrow \infty$ – параметр высокой интенсивности потоков.

В результате асимптотического анализа была получена характеристическая функция совместного распределения числа положительных и отрицательных заявок:

$$H(u, l) = \exp \left\{ juN \varkappa_1 + \frac{(ju)^2 N \varkappa_2}{2} \right\} \left(1 - \frac{\alpha}{\lambda} \right) \left(\frac{\alpha}{\lambda} \right)^l,$$

где \varkappa_1 и \varkappa_2 имеют вид:

$$\varkappa_1 = \frac{\lambda - \alpha}{\mu},$$

$$\varkappa_2 = \varkappa_1 \left(\frac{1}{2} + \frac{\lambda}{\alpha} \right) + \frac{\lambda}{\mu} \left(\frac{3}{2} \left(1 - \frac{\alpha}{\lambda} \right) \right).$$

Если просуммировать по всем l , то можно получить характеристическую функцию одномерного процесса числа положительных заявок:

$$H(u) = \exp \left\{ juN\varkappa_1 + \frac{(ju)^2 N\varkappa_2}{2} \right\}.$$

Таким образом, асимптотическое стационарное распределение вероятностей числа положительных заявок в рассматриваемой системе является гауссовским с математическим ожиданием $N\varkappa_1$ и дисперсией $N\varkappa_2$.

При $u = 0$, можно получить распределение вероятностей числа отрицательных заявок в системе:

$$P(l) = \left(1 - \frac{\alpha}{\lambda} \right) \left(\frac{\alpha}{\lambda} \right)^l.$$

То есть одномерное распределение числа отрицательных заявок в системе в стационарном режиме в указанном асимптотическом условии является геометрическим.

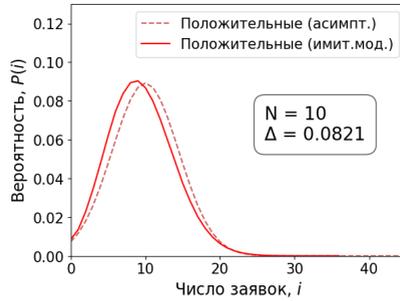
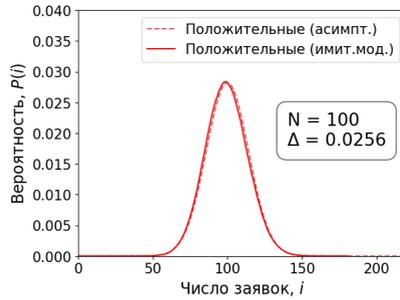
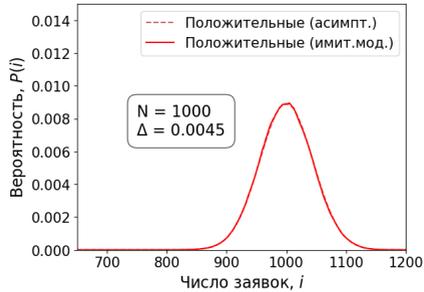
3. Анализ области применимости

Чтобы оценить точность полученной аппроксимации и установить границы ее применимости, была проведена серия численных экспериментов, в которых асимптотическое распределение сравнивалось с эмпирическим, полученным по результатам имитационного моделирования. На языке программирования Python была разработана программа для имитационного моделирования СМО с ожидающими отрицательными заявками.

Для экспериментов были выбраны следующие значения параметров: $\lambda = 2$, $\alpha = 1$, $\mu = 1$. В этом случае $\varkappa_1 = 1$ и асимптотическое среднее число положительных заявок в системе равно N , что удобно для понимания области применимости результатов. На рисунках 1–3 изображены графики асимптотических распределений вероятностей числа положительных заявок и распределений вероятностей, полученных с помощью имитационной модели при различных значениях параметра высокой интенсивности потоков N .

Для оценки погрешности полученного асимптотического распределения вероятностей числа положительных заявок будем использовать расстояние Колмогорова, которое вычисляется по формуле

$$\Delta = \max_{0 \leq i < \infty} \left| \sum_{n=0}^i P_{imit}(n) - P_{asympt}(n) \right|,$$

Рис. 1. Распределения вероятностей числа положительных заявок, $N = 10$ Рис. 2. Распределения вероятностей числа положительных заявок, $N = 100$ Рис. 3. Распределения вероятностей числа положительных заявок, $N = 1000$

где P_{imit} – распределение вероятностей, полученное в результате имитационного моделирования, P_{asympt} – асимптотическое распределение вероятностей. В таблице 1 представлены значения расстояний Колмогорова между асимптотическим распределением и распределением, по-

лученным с помощью имитационной модели, для разных значений параметра μ . Как видно из таблицы и графиков, при увеличении параметра высокой интенсивности потоков N и при уменьшении параметра обслуживания μ расстояния Колмогорова уменьшаются. Если в качестве приемлемой погрешности выбрать $\Delta \leq 0,05$, тогда для $\mu \leq 1$ получаем область применимости полученной аппроксимации $N \geq 20$.

Таблица 1

Расстояние Колмогорова при разных значениях N

N	$\mu = 1$	$\mu = 0.5$
10	0,082	0,051
20	0,050	0,034
50	0,029	0,021
100	0,025	0,013
200	0,017	0,010
500	0,009	0,008
1000	0,004	0,002

Заключение

В статье рассмотрена математическая модель системы массового обслуживания с ожидающими отрицательными заявками. Был выполнен асимптотический анализ этой системы, получен вид характеристической функции совместного распределения числа положительных и отрицательных заявок, выведены одномерные асимптотические распределения вероятностей числа положительных и отрицательных заявок в условии эквивалентного роста интенсивностей входящих потоков. Проведена серия численных экспериментов и сравнение с результатами имитационного моделирования, в результате чего установлены границы применимости полученных аппроксимаций.

СПИСОК ЛИТЕРАТУРЫ

1. Печинкин А. В., Разумчик Р. В. Система массового обслуживания с отрицательными заявками и бункером для вытесненных заявок в дискретном времени // Автоматика и телемеханика. 2009. Т. 38. № 12. С. 109–120.
2. Назаров А. А., Феропонтова Н. М. Исследование бесконечнолинейной системы обслуживания с отрицательными заявками и их ожиданием в пустой системе // Известия высших учебных заведений. Физика, 2015. 7 с

3. *Назаров А. А., Терпугов А. Ф.* Теория вероятностей и случайных процессов: учебное пособие. Томск: Изд-во НТЛ, 2006. 200 с.
4. *Назаров А. А., Моисеева А. А.* Метод асимптотического анализа в теории массового обслуживания. Томск: Изд-во НТЛ, 2006. 112 с.

Королев Данил Анатольевич — аспирант, НИ ТГУ. E-mail:
danilkorolev9999@gmail.com

Моисеева Светлана Петровна — д.ф.-м.н., профессор, НИ ТГУ. E-mail:
smoiseeva@mail.ru

Моисеев Александр Николаевич — д.ф.-м.н., доцент, НИ ТГУ. E-mail:
moiseev.tsu@gmail.com

АСИМПТОТИЧЕСКИЙ АНАЛИЗ RQ-СИСТЕМЫ С S-НАСТОЙЧИВЫМИ ЗАЯВКАМИ И ОБРАТНОЙ СВЯЗЬЮ В УСЛОВИИ БОЛЬШОЙ ЗАДЕРЖКИ

Д. Ю. Николаева, Е. А. Фёдорова

*Национальный исследовательский Томский государственный университет,
г. Томск, Россия*

В работе исследуется СМО с повторными вызовами (RQ-система) M/M/1. Если прибор занят в момент обращения заявки с орбиты, она покидает систему с заданной вероятностью (s-настойчивые заявки). По окончании обслуживания заявка с заданной вероятностью возвращается на орбиту для осуществления повторного обслуживания в будущем таким образом реализуя отложенную обратную связь. Для нахождения стационарного распределения вероятностей числа заявок на орбите предлагается метод асимптотического анализа в условии большой задержки заявок на орбите. Приведено доказательство, что асимптотическая характеристическая функция числа заявок на орбите в данном условии имеет вид нормального распределения.

Ключевые слова: RQ-система, s-настойчивые заявки, обратная связь, большая задержка, асимптотический анализ.

Введение

RQ-системы представляют собой математические модели в теории массового обслуживания (ТМО), используемые для моделирования call-центров, компьютерных и телекоммуникационных сетей и т.д., где заявки повторно пытаются получить обслуживание в случае недоступности обслуживающего устройства. Например, если клиенты звонят в call-центр и их звонок не может быть обработан (из-за нехватки свободных операторов или других причин), они вынуждены повторно обращаться через некоторое, обычно случайное время. Существует вероятность, что после нескольких неудачных попыток клиент прекратит звонить, однако возможно он попробует связаться снова спустя некоторое время для получения дополнительной информации. Такие особенности поведения систем в ТМО описываются терминами "s-настойчивость заявок" и "обратная связь".

Наиболее широкое исследование RQ-систем приведено в монографиях Г.И. Фалина [1] и Дж. Арталехо [2]. RQ-системы с ненастойчивыми заявками исследовались, например, в [4], а модели с обратной связью предлагались в работах [4]. Метод асимптотического анализа, применяемый в работе, описан в [4].

1. Описание модели

В работе исследуется RQ-система M|M|1. На вход поступает простейший поток заявок с интенсивностью λ . Если прибор свободен, то заявка занимает его на случайное время, распределенное по экспоненциальному закону с параметром μ . В случае занятости прибора поступившая заявка попадает на орбиту, где осуществляет случайную задержку, продолжительность которой имеет экспоненциальное распределение с параметром σ . После случайной задержки заявка снова пытается получить доступ к прибору. Если прибор не занят, то заявка с орбиты занимает его на случайное время обслуживания, если же он занят, то заявка с вероятностью $1 - s$ покидает систему или с вероятностью s возвращается на орбиту. После завершения обслуживания заявка покидает систему с вероятностью $1 - r$ или переходит на орбиту с вероятностью r для осуществления повторного обслуживания в будущем. Структура этой модели представлена на Рис. 1.

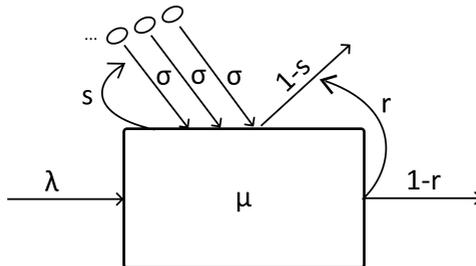


Рис. 1. RQ-система с s -настойчивыми заявками и обратной связью

Пусть $i(t)$ – число заявок на орбите, а процесс $k(t)$ описывает состояние прибора:

$$k(t) = \begin{cases} 0, & \text{прибор свободен,} \\ 1, & \text{прибор занят.} \end{cases}$$

Обозначим $P\{k(t) = k, i(t) = i\} = P(k, i, t)$ – вероятность того, что в момент времени t прибор находится в состоянии k , а на орбите i заявок.

Для распределения вероятностей $P(k, i, t)$ состояний рассматриваемой RQ-системы составим систему дифференциальных уравнений Колмогорова:

$$\begin{cases} \frac{\partial P(0, i, t)}{\partial t} = -(\lambda + i\sigma)P(0, i, t) + \mu P(1, i, t)(1 - r) + r\mu P(1, i - 1, t), \\ \frac{\partial P(1, i, t)}{\partial t} = -(\lambda + \mu + (1 - s)i\sigma)P(1, i, t) + \sigma(i + 1)P(0, i + 1, t) + \\ + \lambda P(0, i, t) + \lambda P(1, i - 1, t) + \sigma(1 - s)(i + 1)P(1, i + 1, t). \end{cases} \quad (1)$$

Для стационарного режима система уравнений имеет следующий вид:

$$\begin{cases} -(\lambda + i\sigma)P(0, i) + \mu P(1, i)(1 - r) + r\mu P(1, i - 1) = 0, \\ -(\lambda + \mu + (1 - s)i\sigma)P(1, i) + \sigma(i + 1)P(0, i + 1) + \lambda P(0, i) + \\ + \lambda P(1, i - 1) + \sigma(1 - s)(i + 1)P(1, i + 1) = 0. \end{cases} \quad (2)$$

Составим систему уравнений, задающих частичные характеристические функции $H(k, u) = \sum_i e^{ju} P(k, i)$:

$$\begin{cases} -\lambda H(0, u) + j\sigma \frac{\partial H(0, u)}{\partial u} + \mu(1 - r)H(1, u) + \mu r e^{ju} H(1, u) = 0, \\ -(\lambda + \mu)H(1, u) + j\sigma(1 - s) \frac{\partial H(1, u)}{\partial u} - j\sigma \frac{\partial H(0, u)}{\partial u} e^{-ju} + \\ + \lambda H(0, u) + \lambda e^{ju} H(1, u) - j\sigma(1 - s) \frac{\partial H(1, u)}{\partial u} e^{-ju} = 0. \end{cases} \quad (3)$$

2. Метод асимптотического анализа

Для решения системы (3), предлагается применить метод асимптотического анализа в условии большой задержки заявок на орбите, то есть при $\sigma \rightarrow 0$.

2.1. Первая асимптотика

Введем следующие обозначения:

$$\sigma = \varepsilon, u = \varepsilon w, H(0, u) = F(0, w, \varepsilon), H(1, u) = F(1, w, \varepsilon). \quad (4)$$

Тогда из системы (3) получаем следующие асимптотические уравнения:

$$\begin{cases} -\lambda F(0, w, \varepsilon) + j \frac{\partial F(0, w, \varepsilon)}{\partial w} + \mu(1 - r)F(1, w, \varepsilon) + \\ + \mu r e^{j\varepsilon w} F(1, w, \varepsilon) = 0, \\ -(\lambda + \mu)F(1, w, \varepsilon) + j(1 - s) \frac{\partial F(1, w, \varepsilon)}{\partial w} - j \frac{\partial F(0, w, \varepsilon)}{\partial w} e^{-j\varepsilon w} + \\ + \lambda F(0, w, \varepsilon) + \lambda e^{j\varepsilon w} F(1, w, \varepsilon) - j(1 - s) \frac{\partial F(1, w, \varepsilon)}{\partial w} e^{-j\varepsilon w} = 0. \end{cases} \quad (5)$$

Суммируя уравнения, после некоторых преобразований имеем:

$$e^{-j\varepsilon w} \left(j \frac{\partial F(0, w, \varepsilon)}{\partial w} + j(1-s) \frac{\partial F(1, w, \varepsilon)}{\partial w} \right) + (\mu r + \lambda) F(1, w, \varepsilon) = 0. \quad (6)$$

Запишем (5) и (6) при $\varepsilon \rightarrow 0$:

$$\begin{cases} -\lambda F(0, w) + j \frac{\partial F(0, w)}{\partial w} + \mu F(1, w) = 0, \\ -\mu F(1, w) + \lambda F(0, w) - j \frac{\partial F(0, w)}{\partial w} = 0, \\ j \frac{\partial F(0, w)}{\partial w} + j(1-s) \frac{\partial F(1, w)}{\partial w} + (\mu r + \lambda) F(1, w) = 0. \end{cases} \quad (7)$$

Пусть

$$F(k, w) = R_k \exp\{jw\kappa_1\}.$$

Подставляя вид функции $F(k, w)$ в систему (7), получаем следующие уравнения:

$$\begin{cases} -\lambda R_0 - \kappa_1 R_0 + \mu R_1 = 0, \\ -\kappa_1 R_0 - (1-s)\kappa_1 R_1 + (\mu r + \lambda) R_1 = 0. \end{cases} \quad (8)$$

Добавляя условие нормировки $R(0) + R(1) = 1$, получим:

$$R_1 = \frac{\lambda + \kappa_1}{\mu + \lambda + \kappa_1}; \quad R_0 = \frac{\mu}{\mu + \lambda + \kappa_1}. \quad (9)$$

Подставляя (9) в последнее уравнение системы (8), получим, что κ_1 имеет вид:

$$\kappa_1 = \frac{-(\mu(1-r) - s\lambda) + \sqrt{(\mu(1-r) - s\lambda)^2 + 4(\mu r + \lambda)\lambda(1-s)}}{2(1-s)}. \quad (10)$$

Возвращаясь к заменам, асимптотическая характеристическая функция первого порядка $H_1(u) = e^{ju \frac{\kappa_1}{\sigma}}$.

2.2. Вторая асимптотика

В системе (3) выполним замену:

$$H(0, u) = e^{ju \frac{\kappa_1}{\sigma}} H_2(0, u), \quad H(1, u) = e^{ju \frac{\kappa_1}{\sigma}} H_2(1, u).$$

Получаем следующие уравнения:

$$\begin{cases} -(\lambda + \varkappa_1)H_2(0, u) + j\sigma \frac{\partial H_2(0, u)}{\partial u} + \\ + (\mu + \mu r(e^{ju} - 1))H_2(1, u) = 0, \\ -((\lambda(1 - e^{ju}) + \mu) + (1 - s)(1 - e^{-ju})\varkappa_1)H_2(1, u) + \\ + j\sigma(1 - s)(1 - e^{-ju}) \frac{\partial H_2(1, u)}{\partial u} + \\ + (\lambda + \varkappa_1 e^{-ju})H_2(0, u) - j\sigma \frac{\partial H_2(0, u)}{\partial u} e^{-ju} = 0. \end{cases} \quad (11)$$

Сложим уравнения системы (11):

$$\begin{aligned} -\varkappa_1 H_2(0, u) + \sigma j \frac{\partial H_2(0, u)}{\partial u} + (e^{ju}(\mu r + \lambda) - (1 - s)\varkappa_1)H_2(1, u) + \\ + j\sigma(1 - s) \frac{\partial H_2(1, u)}{\partial u} = 0. \end{aligned} \quad (12)$$

Выполним асимптотические замены:

$$\sigma = \varepsilon^2, \quad u = \varepsilon w, \quad H(0, u) = F(0, w, \varepsilon), \quad H(1, u) = F(1, w, \varepsilon). \quad (13)$$

Поставив (13) в (11), получим:

$$\begin{cases} -(\lambda + \varkappa_1)F(0, w, \varepsilon) + \varepsilon j \frac{\partial F(0, w, \varepsilon)}{\partial w} + \\ + (\mu(1 - r) + \mu r e^{j\varepsilon w})F(1, w, \varepsilon) = 0, \\ -(\lambda + \mu)F(1, w, \varepsilon) - (1 - s)(1 - e^{-j\varepsilon w})\varkappa_1 F(1, w, \varepsilon) + \\ + j\varepsilon(1 - s)(1 - e^{-j\varepsilon w}) \frac{\partial F(1, w, \varepsilon)}{\partial w} + (\lambda + \varkappa_1 e^{-j\varepsilon w})F(0, w, \varepsilon) - \\ - j\varepsilon \frac{\partial F(0, w, \varepsilon)}{\partial w} e^{-j\varepsilon w} + \lambda e^{j\varepsilon w} F(1, w, \varepsilon) = 0. \end{cases} \quad (14)$$

Из системы (12) имеем

$$\begin{aligned} -\varkappa_1 F(0, w, \varepsilon) + \varepsilon j \frac{\partial F(0, w, \varepsilon)}{\partial w} + \\ + ((\mu r + \lambda)e^{j\varepsilon w} - (1 - s)\varkappa_1)F(1, w, \varepsilon) + j\varepsilon(1 - s) \frac{\partial F(1, w, \varepsilon)}{\partial w} = 0. \end{aligned} \quad (15)$$

Пусть

$$F(k, w, \varepsilon) = \Phi(w) (R_k + j\varepsilon w f_k) + O(\varepsilon^2).$$

Подставляем в уравнения (14)-(2), получаем при $\varepsilon \rightarrow 0$:

$$\begin{cases} jw\Phi(w)[-(\lambda + \varkappa_1)f_0 + \mu r R_1 + \mu f_1] + j\Phi'(w)R_0 = 0, \\ jw\Phi(w)[- \mu f_1 + (\lambda - (1 - s)\varkappa_1)R_1 + \\ + (\lambda + \varkappa_1)f_0 - \varkappa_1 R_0] - j\Phi'(w)R_0 = 0. \end{cases} \quad (16)$$

Нетрудно заметить, что функция $\Phi(w)$ имеет вид: $\Phi(w) = \exp\left\{\frac{(jw)^2}{2\kappa_2}\right\}$, где f_k и κ_2 определяются системой:

$$\begin{cases} [-(\lambda + \kappa_1)f_0 + \mu r R_1 + \mu f_1] - R_0 \kappa_2 = 0, \\ -\mu f_1 + (\lambda - (1-s)\kappa_1)R_1 + (\lambda + \kappa_1)f_0 - \kappa_1 R_0 + \kappa_2 R_0 = 0, \\ \mu r R_1 - (\lambda + (1-s)\kappa_1)R_1 - \kappa_1 R_0 = 0. \end{cases} \quad (17)$$

Откуда

$$\kappa_2 = \frac{-\kappa_1 f_0 + (\mu r + \lambda)f_1 + (\mu r + \lambda)R_1 - (1-s)\kappa_1 f_1}{R_1(1-s) + R_0}. \quad (18)$$

Возвращаясь к замене, $H_2(u) = \exp\left\{\frac{(ju)^2 \kappa_2}{2\sigma}\right\}$.

Таким образом, было доказано, что асимптотическая характеристическая функция числа заявок на орбите имеет форму характеристической функции гауссовского распределения:

$$H(u) = H_1(u)H_2(u) = \exp\left\{ju \frac{\kappa_1}{\sigma} + \frac{(ju)^2 \kappa_2}{2\sigma}\right\}. \quad (19)$$

Заключение

В статье рассмотрена RQ-система M/M/1 с s -настойчивыми заявками и отложенной обратной связью. Проведен асимптотический анализ системы в условии большой задержки заявок на орбите ($\sigma \rightarrow 0$). В результате доказано, что асимптотическая характеристическая функция числа заявок на орбите имеет вид гауссовского распределения с найденными формулами для асимптотического математического ожидания и дисперсии ((10) и (18)).

СПИСОК ЛИТЕРАТУРЫ

1. *Falin G. I., Tempeton J. G. C. Retrial Queues* Retrial Queues. London: Chapman and Hall, 1997. 328 p.
2. *Artalejo J. R., Gomez-Corral A. Retrial queueing systems: A computational approach*. — Berlin: Springer, 2008. 267 p.
3. *Полховская А.В., Моисеева С.П.* Асимптотический анализ RQ-системы M|M|1 с коллизиями и H1, H2 настойчивыми заявками // Информационные технологии и математическое моделирование (ИТММ-2022) : материалы XXI Международной конференции имени А. Ф. Терпугова, 25-29 октября 2022 г. Томск: Изд-во Том. ун-та, 2023. С. 136–142.
4. *Nazarov A. A., Rozhkova S. V., Titarenko E. Yu.* Asymptotic Analysis of RQ-System with Feedback and Batch Poisson Arrival Under the Condition of Increasing Average Waiting Time in Orbit // Communications in Computer and Information Science, 2020. Vol. 1337 P. 327–339.

5. Назаров А. А., Моисеева С. П. Метод асимптотического анализа в теории массового обслуживания: учебное пособие. Томск: Изд-во НТЛ, 2006. 112 с.

Николаева Дарья Юрьевна — студентка четвертого курса Института прикладной математики и компьютерных наук Томского государственного университета. E-mail: *nik-fa01@yandex.ru*

Фёдорова Екатерина Александровна — кандидат физ.-мат. наук, доцент кафедры Теории вероятности и математической статистики Института прикладной математики и компьютерных наук Томского государственного университета. E-mail: *ekat_fedorova@mail.ru*

АНАЛИЗ ВРЕМЕННЫХ ХАРАКТЕРИСТИК ПОЛЛИНГОВОЙ СИСТЕМЫ С ДВУМЯ ОЧЕРЕДЯМИ И ПУАССОНОВСКИМИ ВХОДЯЩИМИ ПОТОКАМИ 2-ГО РОДА

Д. И. Николаев¹, Ю. В. Гайдамака^{1,2}

¹ *Кафедра теории вероятностей и кибербезопасности,
Российский университет дружбы народов им. П. Лумумбы,
Россия, 117198, Москва, ул. Миклухо-Маклая, д.6*

² *Федеральный исследовательский центр «Информатика и управление»
Российской академии наук (ФИЦ ИУ РАН)
Россия, 119333, Москва, ул. Вавилова, д. 44-2*

Системы поллинга являются важной темой для изучения в области теории массового обслуживания, поскольку они широко используются для моделирования работы современных сетей и систем беспроводной широкополосной связи. В данной работе рассматривается модель поллингового обслуживания с двумя очередями и пуассоновскими входящими потоками 2-го типа, для которой были получены аналитические выражения для временных характеристик системы. К ним относятся преобразования Лапласа-Стилтьеса, средние и дисперсии времён пребывания заявок в системе для произвольных распределений времён обслуживания. Кроме того, представлены функции распределения, а также начальные и центральные моменты любого порядка для времён пребывания заявок в системе с экспоненциальными временами обслуживания. **Ключевые слова:** *поллинг, система массового обслуживания, временные характеристики.*

Введение

Системы поллинга широко исследуются как в русскоязычной [1, 2, 3], так и в англоязычной [4, 5] литературе. Далее рассмотрим отличительные признаки поллинговых моделей, непосредственно касающиеся предлагаемой модели.

Во-первых, поллинговые системы могут работать как в дискретном времени, когда время делится на равные промежутки времени (называемые тактами дискретизации), так и в непрерывном времени.

Во-вторых, если случайные процессы, характеризующие работу системы (поступление и обслуживание заявок и длительности переключения прибора между очередями), стохастически эквивалентны для всех очередей, то такая система называется симметричной [6]. В других случаях система называется несимметричной.

Также, если времена переключения прибора между очередями равны нулю, то мы говорим о мгновенном переключении сервера между очередями, в противном случае — не мгновенном переключении.

Циклический порядок обслуживания очередей, когда каждая очередь в рамках цикла обслуживается один раз, является одним из наиболее часто встречающихся порядков обслуживания в системах поллинга.

Если прибор во всех очередях обслуживает только поступившие на момент начала цикла заявки, то такая дисциплина называется глобально-шлюзовой. В случае если очередь обслуживается до тех пор, пока она полностью не опустеет, то дисциплина называется исчерпывающей. В работах [7, 8] были определены средние времена ожидания обслуживания для шлюзовой и исчерпывающей дисциплин.

Схема предлагаемой в данной работе модели представлена на Рис. 1, а её особенности описаны в таблице 1.

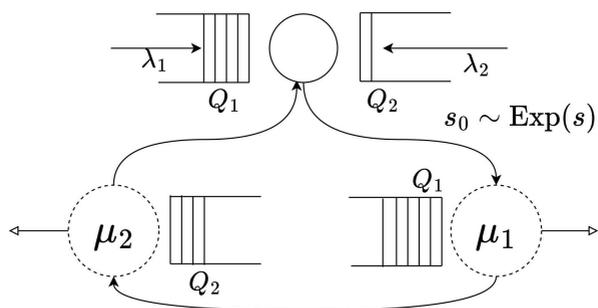


Рис. 1. Поллинговая модель $M_2|M_2|1$ с двумя очередями и входящими потоками 2-го рода

Так как в предлагаемой модели очереди заполняются непостоянно, то имеем входящий пуассоновский поток 2-го рода [9]. Тогда, если Ω — множество состояний поступления заявок, а ζ отображает текущее состояние системы, то параметр λ пуассоновского потока зависит от состояния системы и выражается как

$$\lambda(\zeta) = \begin{cases} \lambda, & \zeta \in \Omega, \\ 0, & \zeta \notin \Omega. \end{cases} \quad (1)$$

Учитывая данную особенность имеем сочетание глобально-шлюзовой и исчерпывающей дисциплин обслуживания, которое назовём глобально-исчерпывающей дисциплиной.

Таблица 1

Особенности модели $M_2|GI_2|1$ с двумя очередями и входящими потоками 2-го рода

Порядок обслуживания	Циклический
Стохастическая эквивалентность характеризующих систему СП	Несимметричная
Дисциплина обслуживания очереди(ей)	Глобально-исчерпывающая
Функционирование во времени	Непрерывное
Время переключения между очередями	Нулевое время переключения внутри цикла и ненулевое время переключения между циклами ($s_1 = s_0, s_2 = 0$)
Входящие пуассоновские потоки	2-го рода
Дисциплина обслуживания заявок	First Come First Served (FCFS)

1. Математическая модель

Поллингвая модель, описываемая таблицей 1 в предыдущем разделе, была построена в работе [10] применительно к функционированию граничного узла сети интегрированного доступа и транзита (IAB), для которой были найдены вероятностные характеристики.

В данной системе случайные величины (СВ) b_{ij} — времена обслуживания j -й заявки (так как накопители бесконечные, то $j = 0, 1, 2, \dots$) в i -й очереди (то есть из i -го входящего потока, $i = 1, 2$), где b_{ij} независимы и одинаково распределены (b_i) с функциями распределения (ФР) $B_i(t)$, $t \geq 0$. Также $b_i^{(n)} = \mathbb{E}(b_i^n)$ — начальный момент порядка n ($n \geq 1$) времени обслуживания заявки в очереди Q_i , а $b_i^{(1)} = \bar{b}_i = \mathbb{E}(b_i)$ — среднее время обслуживания заявки в очереди Q_i . Аналогично преобразование Лапласа-Стилтьеса (ПЛС) СВ b_i обозначим как $\tilde{B}_i(w)$.

Для нахождения временных характеристик системы и ФР времён пребывания заявки в системе (от момента начала обслуживания первой очереди до ухода заявки из системы) введём ряд новых обозначений:

- θ_i — СВ, равная времени пребывания заявки в очереди Q_i без учёта времени, затраченного с начала цикла обслуживания на обслуживание очередей до Q_i , имеющая ФР $\Theta_i(t)$ с начальным моментом n -го порядка $\theta_i^{(n)}$, средним $\bar{\theta}_i = \theta_i^{(1)}$ и ПЛС $\tilde{\Theta}_i(w)$
- v_i — СВ, равная времени пребывания заявки в системе, учитывая время пребывания до начала обслуживания очереди Q_i , имеющая ФР $V_i(t)$ с начальным моментом n -го порядка $v_i^{(n)}$, средним $\bar{v}_i = v_i^{(1)}$ и ПЛС $\tilde{V}_i(w)$. Заметим, что $v_i = \sum_{j=1}^i \theta_j = v_{i-1} + \theta_i$, в частности,

$v_1 = \theta_1, v_2 = v_1 + \theta_2 = \theta_1 + \theta_2, v_3 = v_2 + \theta_3 = \theta_1 + \theta_2 + \theta_3, \dots$ Данные обозначения проиллюстрированы на Рис. 2.

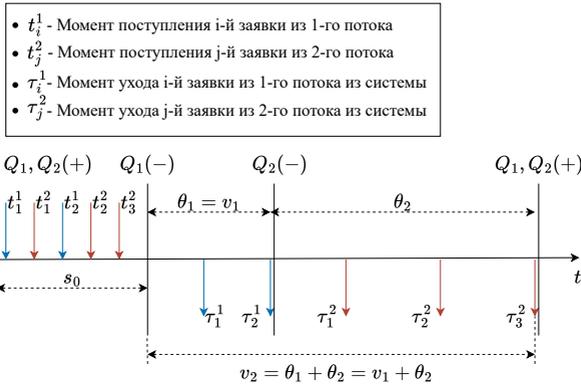


Рис. 2. Пример траектории СП поступления и окончания обслуживания заявок для системы $M_2|GI_2|1$ с входящими потоками 2-го рода

Сформулируем далее следующую теорему.

Теорема 1. Для системы поллингового обслуживания $M_2|GI_2|1$ с входящими потоками 2-го рода и временем переключения s_0 , распределённым согласно экспоненциальному закону $S(t) = 1 - e^{-st}, t \geq 0$, время пребывания v_i заявки в системе, пришедшей в i -м потоке, $i = 1, 2$, имеет ПЛС, среднее и дисперсию, представленные в следующей форме:

$$\widetilde{V}_1(w) = \frac{s}{s + \lambda_1(1 - \widetilde{B}_1(w))}, \quad \widetilde{V}_2(w) = \frac{s^2}{(s + \lambda_1(1 - \widetilde{B}_1(w)))(s + \lambda_2(1 - \widetilde{B}_2(w)))}, \quad (2)$$

$$\overline{v}_1 = v_1^{(1)} = \frac{\lambda_1 \overline{b}_1}{s}, \quad \overline{v}_2 = v_2^{(1)} = \frac{\lambda_1 \overline{b}_1 + \lambda_2 \overline{b}_2}{s}, \quad (3)$$

$$\text{Var}(v_i) = \frac{1}{s^2} \sum_{j=1}^i \lambda_j (s b_j^{(2)} + \lambda_j \overline{b}_j^2), \quad i = 1, 2. \quad (4)$$

В случае если времена обслуживания имеют экспоненциальное распределение с параметрами $\mu_i, i = 1, 2$, то есть $b_i \sim \text{Exp}(\mu_i)$ и $B_i(t) = 1 - e^{-\mu_i t}, t \geq 0$, то мы получим дополнительные к теореме 1 результаты.

Теорема 2. Для системы поллингового обслуживания $M_2|M_2|1$ с входящими потоками 2-го рода и временем переключения s_0 , распределённым согласно экспоненциальному закону $S(t) = 1 - e^{-st}, t \geq 0$, время пребывания v_i заявки в системе, пришедшей в i -м потоке, $i = 1, 2$, имеет ПЛС $\widetilde{V}_i(w)$, начальные $v_i^{(n)}$ и центральные $v_i^{(n)}$ моменты порядка n

($n \in \mathbb{N}$) и ФР $V_i(t)$, представленные в следующей форме:

$$\widetilde{V}_1(w) = \frac{s(w + \mu_1)}{w(\lambda_1 + s) + \mu_1 s}, \widetilde{V}_2(w) = s^2 \prod_{j=1}^2 \frac{w + \mu_j}{w(\lambda_j + s) + \mu_j s}, \quad (5)$$

$$v_1^{(n)} = \frac{(\lambda_1 + u(1-n)s)n!}{\lambda_1 + s} \left(\frac{\lambda_1 + s}{\mu_1 s} \right)^n, n \in \mathbb{N}_0, u(x) = \begin{cases} 1, & x > 0 \\ 0, & x \leq 0, \end{cases} \quad (6)$$

$$v_2^{(n)} = \frac{n!}{s^n} \sum_{m_1=0}^n \prod_{j=1}^i \frac{\lambda_j + u(1-m_j)s}{\lambda_j + s} \left(\frac{\lambda_j + s}{\mu_j} \right)^{m_j}, \quad m_2 = n - m_1, \quad (7)$$

$$v_i^{\circ(n)} = \frac{n!}{s^n} \sum_{m=0}^n (-1)^m \cdot \left(\sum_{r_1=0}^m \cdots \sum_{r_{i-1}=0}^{m-r_1-\cdots-r_{i-2}} \prod_{j=1}^i \frac{\rho_j^{r_j}}{(r_j)!} \right) \cdot \left(\sum_{m_1=0}^{n-m} \prod_{j=1}^i \frac{\lambda_j + u(1-m_j)s}{\lambda_j + s} \left(\frac{\lambda_j + s}{\mu_j} \right)^{m_j} \right), \rho_j = \frac{\lambda_j}{\mu_j},$$

$$V_1(t) = 1 - \Lambda_1 e^{-c_1 t}, \quad V_2(t) = 1 - \gamma_1 e^{-c_1 t} - \gamma_2 e^{-c_2 t}, \quad t \geq 0, \quad (8)$$

где

$$\Lambda_i = \frac{\lambda_i}{\lambda_i + s}, \quad c_i = \frac{\mu_i s}{\lambda_i + s}, \quad (9)$$

$$\gamma_1 = \Lambda_1 \frac{\mu_2(\lambda_1 + s) - \mu_1 s}{\mu_2(\lambda_1 + s) - \mu_1(\lambda_2 + s)}, \gamma_2 = \Lambda_2 \frac{\mu_1(\lambda_2 + s) - \mu_2 s}{\mu_1(\lambda_2 + s) - \mu_2(\lambda_1 + s)}. \quad (10)$$

Заключение

В результате работы описаны особенности модели поллингового обслуживания с двумя очередями и пуассоновскими входящими потоками 2-го рода, функционирующей в непрерывном времени. Выведены формулы временных характеристик системы, включая аналитические выражения для ПЛС, средних и дисперсий времён пребывания заявок в системе при произвольных распределениях времён обслуживания, а также ФР, начальные и центральные моменты любого порядка при экспоненциальных распределениях времён обслуживания.

Разработка и изучение модели обусловлены растущим спросом на эффективные системы обслуживания в телекоммуникациях, общественном транспорте и бизнес-процессах. Результаты могут быть использованы для оценки и анализа эффективности телекоммуникационных сетей, что позволит оптимизировать их и повысить качество обслуживания

клиентов. В будущем планируется расширение модели на произвольное число очередей и учёт дополнительных факторов, таких как случайный порядок обслуживания.

СПИСОК ЛИТЕРАТУРЫ

1. *Вшивневский В. М., Семёнова О. В.* Системы поллинга: Теория и применение в широкополосных беспроводных сетях. Москва: Техносфера, 2007. 312 с.
2. *Семёнова О. В.* Методы и алгоритмы анализа моделей массового обслуживания с поллингом, ненадежными обслуживающими приборами и коррелированными потоками. Автореферат дисс. На соиск. уч.ст. д.ф.-м.н. Москва: Изд-во РУДН, 2022, 38 с.
3. *Рыков В. В.* К анализу поллинг-систем // Автоматика и телемеханика, 2009. № 6. С. 90–114.
4. *Takagi H.* Analysis of polling systems. MIT Press, 1986, 175 p.
5. *Takagi H., Kleinrock L.* A tutorial on the analysis of polling systems. Computer Science Department, University of California, Los Angeles. Report No. CSD-850005, 1985, 172 p.
6. *Takagi H.* Mean message waiting times in symmetric multiqueue systems with cyclic service // Название конференции или сборника, Performance Evaluation, 1985, P. 271–277.
7. *Гайдамака Ю. В., Зарипова Э. Р.* Модель SIP-сервера с дисциплинами шлюзового и исчерпывающего обслуживания очередей // Вестник РУДН. 2013. Серия: математика, информатика, физика. № 1. С. 52–57.
8. *Зарипова Э. Р.* Методы анализа показателей эффективности телекоммуникационной сети серверов протокола установления сессий. Автореферат дисс. На соиск. уч.ст. к.ф.-м.н. Москва: Изд-во РУДН, 2015, 16 с.
9. *Вашарин Г. П.* Лекции по математической теории телетрафика: Учеб. пособие. — Изд. 3-е, испр. и доп. Москва: РУДН, 2009. 346 с.
10. *Николаев Д. И., Гайдамака Ю. В.* Построение модели ретрансляционного узла сети интегрированного доступа и транзита в виде системы поллинга с двумя очередями // Информационно-телекоммуникационные технологии и математическое моделирование высокотехнологичных систем (ИТТММ 2024). Москва: РУДН, 2024, С. 21–28.

Николаев Дмитрий Иванович — студент, кафедра теории вероятностей и кибербезопасности, Российский университет дружбы народов им. П. Лумумбы. E-mail: nikolaev-di@rudn.ru

Гайдамака Юлия Васильевна — д.ф.-м.н., профессор, кафедра теории вероятностей и кибербезопасности, Российский университет дружбы народов им. П. Лумумбы; ст.н.с., Федеральный исследовательский центр «Информатика и управление» Российской академии наук (ФИЦ ИУ РАН). E-mail: gaydatamaka-yuv@rudn.ru

АНАЛИЗ G-СЕТИ С КОНТРОЛЬНЫМИ И КАРАНТИННЫМИ ОЧЕРЕДЯМИ И НЕТЕРПЕЛИВЫМИ ПОЛОЖИТЕЛЬНЫМИ И ОТРИЦАТЕЛЬНЫМИ ЗАЯВКАМИ

Д. Я. Копать

*Гродненский государственный университет имени Янки Купалы, г. Гродно,
Республика Беларусь*

Объектом исследования в статье является G-сеть, состоящая из карантинных, контрольных, обслуживающих очередей в системах с нетерпеливыми положительными и отрицательными заявками. Положительные заявки имеют случайное время ожидания в двух очередях (кроме карантина), по истечении которого переходят в контрольную очередь другой системы (СМО). Отрицательные заявки по истечении случайного времени и уничтожения одной положительной заявки могут как покидать сеть, так и перемещаться между системами сети. Данная G-сеть является математической моделью компьютерной сети (КС), в каждом устройстве которой установлено антивирусное программное обеспечение (АПО), с возможностями распространения вирусов между СМО сети и управлением нагрузкой в сети. Предполагается, что с определённой вероятностью вирус может обмануть АПО и причинить КС вред. Такое возможно до тех пор, пока вирус не будет обнаружен АПО определённой СМО. С помощью метода последовательных приближений, совмещенного с методом рядов (МПП) найдены переходные вероятности состояний данной сети. **Ключевые слова:** *G-сеть, нетерпеливые положительные и отрицательные заявки, системы с карантинными и контрольными очередями, МПП, переходные вероятности состояний.*

Введение

В [1] была исследована G-сеть с нетерпеливыми отрицательными заявками в переходном режиме, а в стационарном – в [2]. Механизм нетерпеливых положительных заявок исследовался в [3]. В [4] данная сеть исследовалась в стационарном режиме. Первая математическая модель АПО с помощью СеМО с отрицательными заявками в переходном режиме была рассмотрена в [5]. В данной статье в каждой СМО сети имеется контрольная очередь и очередь на обслуживание, но карантинная очередь одна на всю сеть. В [6] предполагается наличие в каждой

системе массового обслуживания (СМО) контрольной (КонО) и карантинной очередей (КарО), первая из которых проверяет заявку на стандартность и в случае успешного прохождения данной проверки заявка попадает в очередь на обслуживание. Иначе она попадает в КарО и проходит лечение. В случае его успешности она возвращается в ту же СМО на обслуживание, иначе покидает СеМО. В случае если отрицательной заявке удаётся обмануть АПО, она уничтожает одну положительную заявку и может перемещаться между системами сети, пока не будет обнаружена в КонО. Данная статья обобщает модели [1, 3, 6] в одну сеть. Для решения системы РДУ будет применен МПП, впервые введённый в рассмотрение в [7].

1. Описание сети

Рассмотрим G-сеть [1], состоящую из n СМО. В СМО S_i поступают простейшие потоки положительных и отрицательных заявок с интенсивностями $\lambda_{0i}^+, \lambda_{0i}^-, i = \overline{1, n}$, соответственно. Первоначально поступившая в i -ю СМО заявка становится в КонО, где проверяется на стандартность, причём время проверки имеет показательное распределение (п.р.) с параметром $\mu_i^{(v)}, i = \overline{1, n}$. Время ожидания положительной заявкой начала обслуживания в очереди на обслуживание ограничено случайной величиной (СВ) с п.р. с параметром $\theta_i, i = \overline{1, n}$, по истечении которого в КонО с вероятностью $q_{ij}^{(v)}$ она перемещается в КонО j -той СМО $j = \overline{1, n}$ или с вероятностью $q_{i0}^{(v)} = 1 - \sum_{j=1}^n q_{ij}^{(v)}$ покидает сеть. Время активации отрицательной заявки в i -й СМО ограничено СВ с п.р. с параметром $\mu_i^-, i = \overline{1, n}$, по истечении которого в КонО с вероятностью $n_{ij}^{(v)}$ она перемещается в КонО j -той СМО $j = \overline{1, n}$ или с вероятностью $n_{i0}^{(v)} = 1 - \sum_{j=1}^n n_{ij}^{(v)}$ покидает сеть. После проверки на стандартность в i -й СМО путь заявок следующий: для положительной заявки с вероятностью p_i^+ она признаётся таковой и поступит в очередь на обслуживание в этой СМО, иначе будет признана отрицательной заявкой и отправится в КарО на лечение; для отрицательной заявки с вероятностью p_i^- она признаётся таковой и переходит в КарО на лечение, иначе она признаётся положительной заявкой и поступит в очередь на обработку, где она уничтожает одну положительную заявку в непустой системе после чего с вероятностью n_{i0} покидает сеть или с вероятностью n_{ij} переходит в КонО j -й СМО, $\sum_{j=0}^n n_{ij} = 1, \sum_{j=0}^n n_{ij} = 1, i = \overline{1, n}$. В пустой очереди на обслуживание отрицательная заявка не оказывает никакого влияния на СМО. Пусть длительности обслуживания положительных заявок в СМО имеют п.р. с параметром $\mu_i, i = \overline{1, n}$ по завершении ко-

того с вероятностью p_{ij}^+ переходит в КонО СМО как положительная заявка, с вероятностью p_{ij}^- – как отрицательная заявка и с вероятностью $p_{i0} = 1 - \sum_{j=1}^n (p_{ij}^+ + p_{ij}^-)$ покидает сеть, $i, j = \overline{1, n}$. После завершения ожидания положительная заявка с вероятностью q_{ij}^+ она переходит в КонО j -той СМО как положительная заявка, с вероятностью q_{ij}^- – как отрицательная заявка или с вероятностью $q_{i0} = 1 - \sum_{j=1}^n (q_{ij}^+ + q_{ij}^-)$ покидает сеть. Пусть $\theta_i^{(sv)}$ – интенсивность досрочного выхода заявки из очереди на обслуживание, если она прибыла из КонО i -й СМО. Тогда она удовлетворяет следующей системе уравнений:

$$\theta_i^{(sv)-1} = \theta_i^{-1} - (\lambda_{0i}^+ + \lambda_{0i}^- - \mu_i^{(v)}) + \sum_{j=1}^n \mu_j (p_{ji}^+ + p_{ji}^-) - \sum_{j=1}^n \theta_j^{(sv)} (q_{ji}^+ + q_{ji}^-) (\mu_i^{(v)})^{-1}.$$

В КарО заявки, признанные нестандартными, становятся в очередь на лечение. Предположим, что длительность лечения заявки в i -м узле имеет п.р. с параметром $\mu_i^{(c)}$, $i = \overline{1, n}$. Лечение успешное с вероятностью $p_i^{(s)}$, и заявка переходит в очередь на обслуживание в i -й СМО, иначе зараженная заявка оказывается вирусом и покидает сеть. При описании КарО мы предполагаем, что вирус не может обмануть его при его лечении, и заявка не может покинуть КарО. В случае перехода заявки в очередь на обслуживание его время ожидания в очереди уменьшается на величину $(\mu_i^{(v)} t(p_i^- + 1 - p_i^+ - \mu_i^{(c)})) (\mu_i^{(c)})^{-1}$. Пусть $1/\theta_i^{(v)}$ – среднее время ожидания начала обслуживания положительной заявкой в очереди на обслуживание. Его можно выразить следующим образом: $\theta_i^{-1} - \theta_i^{(sv)-1} p_i^+ + ((1 - p_i^+) \mu_i^{(v)} t(p_i^- + 1 - p_i^+ - \mu_i^{(c)})) (\mu_i^{(c)})^{-1}$. Состояние сети описывается вектором:

$$(\vec{k}, \vec{l}, t) = (\vec{k}_1, \vec{k}_2, \dots, \vec{k}_n, \vec{l}_1, \vec{l}_2, \dots, \vec{l}_n, t), \quad (1)$$

где $(\vec{k}_i, \vec{l}_i, t) = (k_i^{(p)}, k_i^{(s)}, l_i^{(n)}, l_i^{(s)}, l_i^{(c)}, t)$; $k_i^{(p)}$ и $l_i^{(n)}$ – число положительных и отрицательных заявок, соответственно, находящихся в КонО i -й СМО; $l_i^{(s)}$ – количество отрицательных заявок в очереди на обслуживание i -й СМО; $k_i^{(s)}$ – количество положительных заявок на обслуживание в i -й СМО; $l_i^{(c)}$ – количество заявок на КарО в i -й СМО. Пусть заявки выбираются на проверку на стандартность из очереди случайным образом. Как в [3] положим, что вероятность проверки в КонО положительной заявки равна $q_i^+ = (\lambda_{0i}^+ + \sum_{j=1}^n \mu_j p_{ji}^+) (\lambda_{0i}^+ + \lambda_{0i}^{(1)} + \sum_{j=1}^n \mu_j (p_{ji}^+ + p_{ji}^-))^{-1}$.

2. РДУ Колмогорова для переходных вероятностей состояний

Пусть \tilde{I}_i – нулевой вектор размерности $2n$, кроме компоненты с номером i равной 1, I_i – нулевой вектор размерности $3n$, кроме компоненты с номером i , равной 1. $u(x)$ – функция Хевисайда. Возможны следующие переходы нашей цепи Маркова в состояние (\vec{k}, \vec{l}, t) за время Δt :

1) из состояния $(\vec{k} - \tilde{I}_{2i-1}, \vec{l}, t + \Delta t)$ с вероятностью $\lambda_{0i}^+ u(k_i^{(p)}) \Delta t + o(\Delta)$ в КонО i -й СМО извне поступит положительная заявка;

2) из состояния $(\vec{k}, \vec{l} - I_{3i-2}, t + \Delta t)$ с вероятностью $\lambda_{0i}^- u(l_i^{(n)}) \Delta t + o(\Delta)$ в КонО i -й СМО извне за время поступит отрицательная заявка;

3) из состояния $(\vec{k} + \tilde{I}_{2i-1} - \tilde{I}_{2i}, \vec{l}, t)$ с вероятностью $\mu_i^{(v)} q_i^+ p_i^+ u(k_i^{(s)}) \Delta t + o(\Delta t)$ положительная заявка после проверки на стандартность в i -й СМО будет признана таковой и перейдет в очередь для обслуживания;

4) из состояния $(\vec{k} + \tilde{I}_{2i-1} - \tilde{I}_{2j-1}, \vec{l}, t)$ с вероятностью $\theta_i^{(v)} q_{ij}^{(v)} \Delta t + o(\Delta t)$ время ожидания в КонО положительной заявки завершено и перейдет в КонО j -той СМО;

5) из состояния $(\vec{k} + \tilde{I}_{2i-1}, \vec{l}, t)$ с вероятностью $\theta_i^{(v)} q_{i0}^{(v)} \Delta t + o(\Delta t)$ по завершении ожидания в КонО положительной заявки, она покинет сеть;

6) из состояния $(\vec{k} + \tilde{I}_{2i-1}, \vec{l} - I_{3i}, t)$ с вероятностью $\mu_i^{(v)} q_i^+ (1 - p_i^+) u(k_i^{(c)}) \Delta t + o(\Delta t)$ положительная заявка после некорректной проверки на стандартность в i -й СМО перейдет в карантин для лечения;

7) из состояния $(\vec{k}, \vec{l} + I_{3i-2} - \tilde{I}_{3i}, t)$ с вероятностью $\mu_i^{(v)} (1 - q_i^+) p_i^- \Delta t + o(\Delta t)$ отрицательная заявка после проверки на стандартность в i -й СМО будет признана таковой и перейдет в карантин для лечения;

8) из состояния $(\vec{k}, \vec{l} + I_{3i-2} - I_{3i-1}, t)$ с вероятностью $\mu_i^{(v)} (1 - q_i^+) (1 - p_i^-) u(k_i^{(s)}) \Delta t + o(\Delta t)$ отрицательная заявка после некорректной проверки на стандартность в i -той СМО перейдет в очередь на обслуживание;

9) из состояния $(\vec{k}, \vec{l} + I_{3i-2} - I_{3j-2}, t)$ с вероятностью $\mu_i^- n_{ij}^{(v)} \Delta t + o(\Delta t)$ после ожидания в КонО отрицательной заявки в i -й СМО она перейдет в КонО j -той СМО;

10) из состояния $(\vec{k}, \vec{l} + I_{3i-2}, t)$ с вероятностью $\mu_i^- n_{i0}^{(v)} \Delta t + o(\Delta t)$ ожидание в КонО отрицательной заявки завершено и она покинет сеть;

11) из состояния $(\vec{k} + \tilde{I}_{2i}, \vec{l} + I_{3i-1}, t)$ с вероятностью $(\mu_i^- - \theta_i^{(v)}) n_{i0} \Delta t + o(\Delta t)$ отрицательная заявка активируется в i -той СМО удалит одну положительную заявку, покидая сеть;

12) из состояния $(\vec{k} + \tilde{I}_{2i}, \vec{l} + I_{3i-1} - I_{3j-2}, t)$ с вероятностью $(\mu_i^- - \theta_i^{(v)}) n_{ij} \Delta t + o(\Delta t)$ отрицательная заявка активируется в i -той СМО и удалит положительную заявку, перейдя в КонО j -той СМО;

13) из состояния $(\vec{k}, \vec{l} + I_{3i-1}, t)$ с вероятностью $(\mu_i^- - \theta_i^{(v)}) (1 - u(k_i)) n_{i0} \Delta t + o(\Delta t)$ отрицательная заявка активируется в пустой i -той СМО и покинет сеть;

14) из состояния $(\vec{k} + \tilde{I}_{2i}, \vec{l} + I_{3i-1} - I_{3j-2}, t)$ с вероятностью $(\mu_i^- - \theta_i^{(v)})n_{ij}\Delta t + o(\Delta t)$ отрицательная заявка активируется в пустой i -той СМО и перейдет в КонО j -той СМО;

15) из состояния $(\vec{k} + \tilde{I}_{2i}, \vec{l}, t)$ с вероятностью $(\mu_i p_{i0} + \theta_i^{(v)} q_{i0})\Delta t + o(\Delta t)$ после обслуживания или ожидания заявки в i -й СМО она покидает сеть;

16) из состояния $(\vec{k} + \tilde{I}_{2i} - \tilde{I}_{2j-1}, \vec{l}, t)$ с вероятностью $(\mu_i p_{ij}^+ + \theta_i^{(v)} q_{ij}^+)u(k_j^{(p)})\Delta t + o(\Delta t)$ время обслуживания или ожидания заявки в i -й СМО закончилось, и она направится в КонО j -й СМО снова как положительная заявка;

17) из состояния $(\vec{k} + \tilde{I}_{2i}, \vec{l} - I_{3j-2}, t)$ с вероятностью $(\mu_i p_{ij}^- + \theta_i^{(v)} q_{ij}^-)u(l_j^{(n)})\Delta t + o(\Delta t)$ время обслуживания или ожидания заявки в i -й СМО закончилось, и она направляется в КонО j -й СМО как отрицательная заявка;

18) из состояния $(\vec{k} - \tilde{I}_{2i}, \vec{l} + I_{3i}, t)$ с вероятностью $\mu_i^{(c)} p_i^{(s)} u(k_i^{(s)})\Delta t + o(\Delta t)$ КарО i -й СМО удастся вылечить зараженную отрицательную заявку, и она отправляется в очередь на обслуживание в i -ю СМО;

19) из состояния $(\vec{k}, \vec{l} + I_{3i}, t)$ с вероятностью $\mu_i^{(c)}(1 - p_i^{(s)})\Delta t + o(\Delta t)$ КарО не удастся вылечить отрицательную заявку и она покидает сеть;

20) из состояния (\vec{k}, \vec{l}, t) с вероятностью $1 - \sum_{i=1}^n [\lambda_{0i}^+ u(k_i^{(p)}) + \lambda_{0i}^- u(l_i^{(n)})] + (\mu_i^{(v)} + \mu_i^{(c)} + \theta_i + \mu_i + \mu_i^-)\Delta t + o(\Delta t)$ состояние сети не изменится.

Можно показать, что нестационарные вероятности состояний удовлетворяют следующим РДУ:

$$\begin{aligned} \frac{dP(\vec{k}, \vec{l}, t)}{dt} = & - \sum_{i=1}^n \left[\lambda_{0i}^+ u(k_i^{(p)}) + \lambda_{0i}^- u(l_i^{(n)}) + \right. \\ & \left. + (\mu_i^{(v)} + \mu_i^{(c)} + \theta_i + \mu_i + \mu_i^-) \right] P(\vec{k}, \vec{l}, t) + \\ & + \sum_{i=1}^n \left\{ \lambda_{0i}^+ u(k_i^{(p)}) P(\vec{k} - \tilde{I}_{2i-1}, \vec{l}, t) + \lambda_{0i}^- u(l_i^{(n)}) P(\vec{k}, \vec{l} - I_{3i-2}, t) + \right. \\ & + \mu_i^{(v)} \left[q_i^+ (p_i^+ u(k_i^{(s)})) P(\vec{k} + \tilde{I}_{2i-1} - \tilde{I}_{2i}, \vec{l}, t) + \right. \\ & + (1 - p_i^+) u(l_i^{(c)}) P(\vec{k} + \tilde{I}_{2i-1}, \vec{l} - I_{3i}, t) \left. \right] + \\ & + (1 - q_i^+) p_i^- u(l_i^{(c)}) P(\vec{k}, \vec{l} + I_{3i-2} - I_{3i}, t) \left. \right] + \\ & + (\mu_i^- - \theta_i^{(v)})(1 - u(k_i^{(s)}))n_{i0} P(\vec{k}, \vec{l} + I_{3i-1}, t) \left. \right\} + \\ & + \mu_i^{(v)}(1 - q_i^+)(1 - p_i^-)u(k_i^{(s)})P(\vec{k}, \vec{l} + I_{3i-2} - I_{3i-1}, t) + \\ & + \mu_i^{(c)} (p_i^{(s)} u(k_i^{(s)})) P(\vec{k} - \tilde{I}_{2i}, \vec{l} + I_{3i}, t) + (1 - p_i^{(s)}) P(\vec{k}, \vec{l} + I_{3i}, t) \left. \right) + \end{aligned}$$

$$\begin{aligned}
 & +\theta_i^{(v)} q_{i0}^{(v)} P(\vec{k} + \tilde{I}_{2i-1}, \vec{l}, t) + \mu_i^- n_{i0}^{(v)} P(\vec{k}, \vec{l} + I_{3i-2}, t) + \\
 & + (\mu_i^- - \theta_i^{(v)}) n_{i0} u(k_i^{(s)}) P(\vec{k} + \tilde{I}_{2i-1}, \vec{l} + I_{3i-1}, t) + \\
 & + (\mu_i p_{i0} + \theta_i^{(v)} q_{i0}) P(\vec{k} + \tilde{I}_{2i}, \vec{l}, t) + \\
 & + \sum_{i,j=1}^n \left\{ \theta_i^{(v)} q_{ij}^{(v)} P(\vec{k} + \tilde{I}_{2i-1} - \tilde{I}_{2j-1}, \vec{l}, t) + \right. \\
 & \quad \left. + \mu_i^- n_{ij}^{(v)} P(\vec{k}, \vec{l} + I_{3i-2} - I_{3j-2}, t) + \right. \\
 & + (\mu_i^- - \theta_i^{(v)}) u(k_i^{(s)}) n_{ij} P(\vec{k} + \tilde{I}_{2i}, \vec{l} + I_{3i-1} - I_{3j-2}, t) + \\
 & + (\mu_i p_{ij}^+ + \theta_i^{(v)} q_{ij}^+) u(k_j^{(p)}) P(\vec{k} + \tilde{I}_{2i} - \tilde{I}_{2j-1}, \vec{l}, t) + \\
 & + (\mu_i^- - \theta_i^{(v)}) n_{ij} (1 - u(k_i^{(s)})) P(\vec{k}, \vec{l} + I_{3i-1} - I_{3j-2}, t) + \\
 & \left. + (\mu_i p_{ij}^- + \theta_i^{(v)} q_{ij}^-) u(l_j^{(n)}) P(\vec{k} + \tilde{I}_{2i}, \vec{l} - I_{3j-2}, t) \right\}. \tag{2}
 \end{aligned}$$

Система (2) является частным случаем системы РДУ:

$$\begin{aligned}
 \frac{dP(\vec{k}^*, \vec{l}, t)}{dt} = & -\Lambda(\vec{k}^*, \vec{l}) P(\vec{k}^*, \vec{l}, t) + \sum_{\alpha, \beta, \gamma, \theta, \eta=0}^{3n} \Phi_{i^* j^* \alpha \beta \gamma \theta \eta}(\vec{k}^*, \vec{l}) \times \\
 & \times P(\vec{k}^* + \tilde{I}_\alpha + \tilde{I}_\beta - \tilde{I}_\gamma, \vec{l} + \tilde{I}_\eta - \tilde{I}_\theta, t), \tag{3}
 \end{aligned}$$

если $\vec{k}^* = (\mathbf{k}, 0, \dots, 0_n)$, n нулей в конце.

$$\begin{aligned}
 \Lambda(\vec{k}^*, \vec{l}) = & - \sum_{i=1}^n \left[\lambda_{0i}^+ u(k_i^{(p)}) + \lambda_{0i}^- u(l_i^{(n)}) + \left(\mu_i^{(v)} + \mu_i^{(c)} + \theta_i + \mu_i + \mu_i^- \right) \right], \\
 \Phi_{i^* j^* \alpha \beta \gamma \theta \eta}(\vec{k}^*, \vec{l}) = & \lambda_{0i}^+ u(k_i^{(p)}) \delta_{\gamma(2i-1)} \delta_\theta \delta_\alpha \delta_\beta \delta_\theta + \\
 & + \lambda_{0i}^- u(l_i^{(n)}) \delta_{\theta(3i-2)} \delta_\eta \delta_\alpha \delta_\gamma \delta_\beta \delta_\theta + \mu_i^{(v)} \left[q_i^+ p_i^+ u(k_i^{(s)}) \delta_\alpha \delta_\beta \delta_{(2i-1)} \delta_\gamma \delta_{(2i)} \delta_\theta \eta + \right. \\
 & + q_i^+ (1 - p_i^+) u(l_i^{(c)}) \delta_{\alpha(2i-1)} \delta_\beta \delta_\gamma \delta_\eta \delta_{\theta(3i)} + \\
 & \left. + (1 - q_i^+) p_i^- u(l_i^{(c)}) \delta_\alpha \delta_\beta \delta_\gamma \delta_\eta \delta_{(3i-2)} \delta_\theta \delta_{(3i)} \right] + \\
 & + \mu_i^{(c)} \delta_\alpha \delta_\eta \delta_\theta \delta_{\theta(3i)} [p_i^{(s)} u(k_i^{(s)}) \delta_\gamma \delta_{(2i)} \delta_\beta \delta_\theta + \\
 & + (1 - p_i^{(s)}) \delta_\beta \delta_\gamma] + \mu_i^{(v)} (1 - q_i^+) (1 - p_i^-) u(k_i^{(s)}) \delta_\alpha \delta_\beta \delta_\theta \delta_{(3i-1)} \delta_\eta \delta_{(3i-2)} + \\
 & + \theta_i^{(v)} q_{i0}^{(v)} \delta_{\alpha(2i-1)} \delta_\beta \delta_\gamma \delta_\theta \eta + \mu_i^- n_{i0}^{(v)} \delta_\alpha \delta_\beta \delta_\gamma \delta_\eta \delta_{(3i-2)} \delta_\theta \delta_\theta + \\
 & + (\mu_i^- - \theta_i^{(v)}) n_{i0} u(k_i^{(s)}) \delta_{\alpha(2i-1)} \delta_\beta \delta_\gamma \delta_\eta \delta_{(3i-1)} \delta_\theta \delta_\theta + (\mu_i p_{i0} + \theta_i^{(v)} q_{i0}) \delta_\alpha \delta_{(2i)} \delta_\beta \delta_\gamma \delta_\eta \delta_\theta + \\
 & + (\mu_i^- - \theta_i^{(v)}) (1 - u(k_i^{(s)})) n_{i0} \delta_\alpha \delta_\beta \delta_\gamma \delta_\eta \delta_{(3i-1)} \delta_\theta \delta_\theta + \\
 & + \theta_i^{(v)} q_{ij}^{(v)} \delta_{\alpha(2i-1)} \delta_\gamma \delta_{(2j-1)} \delta_\beta \delta_\theta \delta_\eta \delta_\theta + \mu_i^- n_{ij}^{(v)} \delta_\alpha \delta_\gamma \delta_\beta \delta_\eta \delta_{(3i-2)} \delta_\theta \delta_{(3j-2)} +
 \end{aligned}$$

$$\begin{aligned}
& + (\mu_i^- - \theta_i^{(v)}) u(k_i^{(s)}) n_{ij} \delta_{\alpha(2i)} \delta_{\gamma\beta} \delta_{\eta(3i-1)} \delta_{\theta(3j-2)} + (\mu_i p_{ij}^+ + \theta_i^{(v)} q_{ij}^+) u(k_j^{(p)}) \times \\
& \times \delta_{\alpha(2i)} \delta_{\gamma(2j-1)} \delta_{\beta\theta} \delta_{\eta\theta} + (\mu_i^- - \theta_i^{(v)}) n_{ij} (1 - u(k_i^{(s)})) \delta_{\alpha\theta} \delta_{\gamma\beta} \delta_{\eta(3i-1)} \delta_{\theta(3j-2)} + \\
& + (\mu_i p_{ij}^- + \theta_i^{(v)} q_{ij}^-) u(l_j^{(n)}) \delta_{\alpha(2i)} \delta_{\gamma\beta} \delta_{\eta\theta} \delta_{\theta(3j-2)},
\end{aligned}$$

где δ_{ij} – символ Кронекера.

Пусть $P_q(\vec{k}, \vec{l}, t)$ – приближение $P(\vec{k}, \vec{l}, t)$ на q -ой итерации, и $P_{q+1}(\vec{k}, \vec{l}, t)$ – решение системы (2), полученное МПП. Каждое приближение представимо сходящимся рядом с бесконечным радиусом сходимости [8]: $P_q(\vec{k}, \vec{l}, t) = \sum_{l=0}^{\infty} d_{ql}^{+-}(\vec{k}, \vec{l}) t^l$, с коэффициентами удовлетворяющими соотношениям [8]:

$$\begin{aligned}
d_{q+1z}^{+-}(\vec{k}, \vec{l}) &= \frac{[-\Lambda(\vec{k}, \vec{l})]^z}{z!} \left(P(\vec{k}, \vec{l}, 0) + \sum_{u=0}^{z-1} (-1)^{u+1} u! \frac{D_{qu}^{+-}(\vec{k}, \vec{l})}{[\Lambda(\vec{k}, \vec{l})]^{u+1}} \right), \\
d_{q0}^{+-}(\vec{k}, \vec{l}) &= P(\vec{k}, \vec{l}, 0), d_{0z}^{+-}(\vec{k}, \vec{l}) = P(\vec{k}, \vec{l}, 0) \delta_{z0}, z \geq 0, \\
D_{qz}^{+-}(\vec{k}, \vec{l}) &= \sum_{\alpha, \beta, \gamma, \theta, \eta=0}^{3n} \Phi_{\alpha\beta\gamma\theta\eta}(\vec{k}, \vec{l}) d_{qz}^{+-}(\vec{k} + \tilde{I}_\alpha + \tilde{I}_\beta - \tilde{I}_\gamma, \vec{l} + \tilde{I}_\eta - \tilde{I}_\theta).
\end{aligned}$$

Нулевое приближение есть стационарное распределение вида

$$\begin{aligned}
& \Lambda(\vec{k}, \vec{l}) P(\vec{k}, \vec{l}) = \\
& = \sum_{\alpha, \beta, \gamma, \theta, \eta=0}^{3n} \Phi_{\alpha\beta\gamma\theta\eta}(\vec{k}, \vec{l}) P(\vec{k} + \tilde{I}_\alpha + \tilde{I}_\beta - \tilde{I}_\gamma, \vec{l} + \tilde{I}_\eta - \tilde{I}_\theta).
\end{aligned}$$

Заключение

В статье предложена стохастическая модель КС с АПО с возможностью заявок переходить в более быстрые сервера без отклика, если он осуществляется медленно. Получена система РДУ для вероятностей состояний такой сети и решена с помощью МПП. Показано, что вероятности состояний сети могут быть представлены в виде сходящихся степенных рядов. Данные результаты могут быть использованы для прогнозирования моментов пиковой нагрузки сети и необходимости её перераспределения, что позволит повысить пропускную способность АПО.

СПИСОК ЛИТЕРАТУРЫ

1. Науменко, В. В. Исследование и применение G-сети с доходами и сигналами со случайным временем активизации // Вестник ГрГУ. 2014. № 3. С. 142–152.

2. *Малинковский Ю.В., Кравченко С.В.* Сети массового обслуживания с отрицательными заявками, обходами узлов заявками и ограниченным временем ожидания в очередях // Известия Национальной академии наук Беларуси. Сер. физ.-мат. наук. 2001. № 4. С. 118–122.
3. *Ивницкий, В. А., Ивницкий О.В.* Нахождение нестационарных математических ожиданий количества «нетерпеливых» требований в узлах замкнутых марковских сетей массового обслуживания // Обзорение прикладной и промышленной математики. 2008. Т. 15, №2. С. 304–305.
4. *Малинковский Ю. В., Бородин Н. Н.* Сети массового обслуживания с конечным числом потоков отрицательных заявок и с ограниченным временем пребывания // ПФМТ. 2018. Т. 1. № 1. С. 64–68.
5. *Letunovich, Yu., Yakubovich, O.* Open Markov queuing networks with control queues and quarantine node. // Tomsk State University Journal of Control and Computer Science. 2017. Т. 41. № 3. С. 32–38.
6. *Копать, Д. Я.* Асимптотический анализ G-сети с многолинейными системами с контрольными и карантинными очередями // ПФМТ, 2023. № 3. С. 48–55.
7. *Harrison. P. G.* Transient behaviour of queueing networks // Journal of Applied Probability. 1981. V. 18. № 2. P. 482–490.
8. *Matalytski, M., Kopat, D.* Finding nonstationary probabilities of open Markov networks with multiple classes of customers and various features // PEIS. 2021. Т. 34. № 1. С. 157–179.

Копать Дмитрий Ярославович — канд. физ.-мат, наук, доцент кафедры
ЛиМУ ГрГУ им. Янки Купалы. E-mail: dk80395@mail.ru

АСИМПТОТИЧЕСКИЙ АНАЛИЗ МНОГОПОТОЧНОЙ НЕОДНОРОДНОЙ СМО В УСЛОВИИ ПРЕДЕЛЬНО РЕДКИХ ИЗМЕНЕНИЙ СОСТОЯНИЙ СЛУЧАЙНОЙ МАРКОВСКОЙ СРЕДЫ

Е. А. Невенченко, С. С. Шепилов, С. П. Моисеева

*Национальный исследовательский Томский государственный университет,
г. Томск, Россия*

В работе рассматривается система массового обслуживания, функционирующая в случайной среде. На вход системы поступают два ММРР-потока разнотипных заявок, обслуживание осуществляется в соответствии с типом заявки на двух блоках, каждый из которых содержит неограниченное число приборов. Для рассматриваемой системы в условии предельно редких изменений состояний управляющей цепи Маркова получены распределения вероятностей и вероятностные характеристики числа заявок каждого типа. Проведен сравнительный анализ точных характеристик и асимптотических. **Ключевые слова:** *случайная среда, бесконечнолинейная СМО, ММРР-поток, асимптотический анализ, предельно редкие изменения состояний цепи Маркова.*

Введение

Интерес к исследованию бесконечнолинейных гетерогенных систем массового обслуживания растет с распространением беспроводной передачи данных. Одной из важнейших задач при передаче данных является оптимизация, как с точки зрения пропускной способности, так и требуемого трафика. Данным разных типов не требуется одинаковое количество ресурса.

Для отражения адекватности реальных систем будем строить математическую модель, учитывая разные типы данных [1], а также изменения во внешней среде, которые могут влиять на интенсивность поступления заявок, например, ремонт или отказ оборудования, природные явления, сбои связи при передаче данных и другие. Системы массового

обслуживания, отображающие реальные процессы, связанные с возмущениями внешней среды, называют функционирующими в случайной среде [2, 3, 4, 5, 6].

1. Постановка задачи и математическая модель

Рассмотрим математическую модель с потоком разнотипных заявок и гетерогенным обслуживанием в виде СМО с двумя узлами, отличающимися характеристиками обслуживания, каждый из которых содержит достаточное количество необходимых ресурсов. В данной работе случайная среда влияет только на входящие потоки, поэтому можно говорить, что на вход системы поступает два марковски модулированных пуассоновских потока, управляемые одной цепью Маркова с конечным числом состояний $k(t) = 1, \dots, K$, заданной матрицей инфинитезимальных характеристик $\mathbf{Q} = \|q_{ij}\|$, $i, j = 1, \dots, K$, и диагональными матрицами условных интенсивностей $\mathbf{\Lambda}^1$ и $\mathbf{\Lambda}^2$ с элементами $\lambda_k^1, \lambda_k^2 \geq 0$, $k = 1, \dots, K$, на главной диагонали.

Поступившие заявки обслуживаются в блоке i -ого типа в течение случайного времени τ_i , экспоненциально распределенного с параметром μ_i , $i = 1, 2$.

Введём трехмерный случайный процесс $\{k(t), i_1(t), i_2(t)\}$, характеризующий состояние управляющей цепи Маркова и число занятых приборов в соответствующих блоках.

Ставится задача исследования трехмерного марковского процесса, для его стационарного распределения вероятностей

$$\pi(k, i_1, i_2) = P\{k(t) = k, i_1(t) = i_1, i_2(t) = i_2\}$$

запишем систему уравнений Колмогорова:

$$\begin{aligned} 0 = & -(\lambda_k^1 + \lambda_k^2 + i_1\mu_1 + i_2\mu_2)\pi(k, i_1, i_2) + \\ & + \lambda_k^1\pi(k, i_1 - 1, i_2) + \lambda_k^2\pi(k, i_1, i_2 - 1) + (i_1 + 1)\mu_1\pi(k, i_1 + 1, i_2) + \\ & + (i_2 + 1)\mu_2\pi(k, i_1, i_2 + 1) + \sum_{\nu} q_{\nu k}\pi(\nu, i_1, i_2). \end{aligned} \quad (1)$$

Введем частичные характеристические функции трехмерного стационарного распределения $\pi(k, i_1, i_2)$ в виде:

$$H(k, u_1, u_2) = \sum_{i_1=0}^{\infty} \sum_{i_2=0}^{\infty} e^{ju_1 i_1} e^{ju_2 i_2} \pi(k, i_1, i_2),$$

где $j = \sqrt{-1}$.

Тогда (1) перепишем в виде дифференциальных уравнений для частичных характеристических функций $H(k, u_1, u_2)$:

$$\begin{aligned} & j\mu_1 (e^{-ju_1} - 1) \frac{\partial H(k, u_1, u_2)}{\partial u_1} + j\mu_2 (e^{-ju_2} - 1) \frac{\partial H(k, u_1, u_2)}{\partial u_2} = \\ & = \{\lambda_k^1 (e^{ju_1} - 1) + \lambda_k^2 (e^{ju_2} - 1)\} H(k, u_1, u_2) + \sum_{\nu=1}^K H(\nu, u_1, u_2) q_{\nu k} \end{aligned} \quad (2)$$

с начальными условиями

$$H(k, 0, 0) = r(k), \quad k = \overline{1, K}.$$

Здесь $\mathbf{r} = [r(1), \dots, r(K)]$ — вектор стационарного распределения вероятностей управляющей цепи Маркова, определяемый системой линейных уравнений

$$\begin{cases} \mathbf{r}\mathbf{Q} = 0, \\ \mathbf{r}\mathbf{e} = 1, \end{cases} \quad (3)$$

где \mathbf{e} — единичный вектор-столбец размерности $K \times 1$.

Для получения выражений для вероятностных характеристик воспользуемся свойствами характеристических функций. Тогда среднее число занятых приборов каждого типа определяется по формуле:

$$m_1^{(i)}(k) = \frac{r(k)\lambda_k^i}{\mu_i}, \quad i = 1, 2, \quad (4)$$

а вторые начальные моменты

$$m_2^{(i)}(k) = \frac{m_1^{(i)}(k)(\lambda_k^i - \mu_i) + r(k)\lambda_k^i}{2\mu_i}, \quad i = 1, 2. \quad (5)$$

Тогда, используя (4) и (5), запишем корреляционный момент

$$m_{12}(k) = \frac{m_1^{(1)}(k)\lambda_k^2 + m_1^{(2)}(k)\lambda_k^1}{\mu_1 + \mu_2}. \quad (6)$$

2. Асимптотический анализ

Будем искать решение (2) в асимптотическом условии предельно редких изменений состояний случайной среды.

Обозначим $T_k = -\frac{1}{q_{kk}}$, $k = 1, \dots, K$ — среднее время пребывания управляющего процесса в k -ом состоянии, тогда под условием предельно редких изменений состояний среды будем понимать

$$q_{\nu k} = \varepsilon q_{\nu k}^* \quad (7)$$

при $\varepsilon \rightarrow 0$. Решение системы (2) с учетом (7) будем искать в виде $H(k, u_1, u_2) = F(k, u_1, u_2, \varepsilon)$ с начальным условием

$$F(k, 0, 0, \varepsilon) = H(k, 0, 0) = r(k), \tag{8}$$

запишем уравнения в новых обозначениях

$$\begin{aligned} j\mu_1(e^{-ju_1} - 1) \frac{\partial F(k, u_1, u_2, \varepsilon)}{\partial u_1} + j\mu_2(e^{-ju_2} - 1) \frac{\partial F(k, u_1, u_2, \varepsilon)}{\partial u_2} = \\ = \{\lambda_k^1(e^{ju_1} - 1) + \lambda_k^2(e^{ju_2} - 1)\}F(k, u_1, u_2, \varepsilon) + \\ + \sum_{\nu} \varepsilon q_{\nu k}^* F(\nu, u_1, u_2, \varepsilon). \end{aligned} \tag{9}$$

В условии $\varepsilon \rightarrow 0$ в силу теоремы о непрерывной зависимости решения системы дифференциальных уравнений от параметра: $\lim_{\varepsilon \rightarrow 0} F(k, u_1, u_2, \varepsilon) = F_1(k, u_1, u_2)$, система уравнений (9) примет вид

$$\begin{aligned} j\mu_1(e^{-ju_1} - 1) \frac{\partial F_1(k, u_1, u_2)}{\partial u_1} + j\mu_2(e^{-ju_2} - 1) \frac{\partial F_1(k, u_1, u_2)}{\partial u_2} = \\ = \{\lambda_k^1(e^{ju_1} - 1) + \lambda_k^2(e^{ju_2} - 1)\}F_1(k, u_1, u_2). \end{aligned} \tag{10}$$

Для системы уравнений в частных производных (10) запишем соответствующую систему обыкновенных дифференциальных уравнений:

$$\begin{aligned} \frac{\partial u_1}{j\mu_1(e^{-ju_1} - 1)} = \frac{\partial u_2}{j\mu_2(e^{-ju_2} - 1)} = \\ = \frac{\partial F_1(k, u_1, u_2)}{\{\lambda_k^1(e^{ju_1} - 1) + \lambda_k^2(e^{ju_2} - 1)\}F_1(k, u_1, u_2)}. \end{aligned} \tag{11}$$

Решая (11), с учетом начальных условий (8) получаем

$$\begin{aligned} H(k, u_1, u_2) = F(k, u_1, u_2, \varepsilon) \approx F_1(k, u_1, u_2) = \\ = r(k) \exp \left\{ \frac{\lambda_k^1(e^{ju_1} - 1)}{\mu_1} + \frac{\lambda_k^2(e^{ju_2} - 1)}{\mu_2} \right\}. \end{aligned} \tag{12}$$

Просуммировав (12) по всем $k = 1, \dots, K$, найдем характеристическую функцию совместного распределения вероятностей числа заявок каждого типа в системе, которая определяется формулой:

$$H(u_1, u_2) \approx \sum_{k=1}^K r(k) \exp \left\{ \frac{\lambda_k^1(e^{ju_1} - 1)}{\mu_1} + \frac{\lambda_k^2(e^{ju_2} - 1)}{\mu_2} \right\}. \tag{13}$$

3. Численный анализ

Для того, чтобы определить точность аппроксимации (13), применяя свойства характеристических функций, вычислим асимптотические вероятностные характеристики числа заявок каждого типа.

Асимптотическое математическое ожидание:

$$m_1^{(i)} as = \sum_{k=1}^K r(k) \frac{\lambda_k^i}{\mu_i}, i = 1, 2. \quad (14)$$

Асимптотический второй начальный момент:

$$m_2^{(i)} as = \sum_{k=1}^K r(k) \left\{ \frac{\lambda_k^i}{\mu_i} + \left(\frac{\lambda_k^i}{\mu_i} \right)^2 \right\}, i = 1, 2. \quad (15)$$

Асимптотический корреляционный момент:

$$m_{12} as = \sum_{k=1}^K r(k) \frac{\lambda_k^1 \lambda_k^2}{\mu_1 \mu_2}. \quad (16)$$

Для численного примера рассмотрим систему с $K = 4$ состояниями, заданную параметрами $\mu_1 = 0, 1, \mu_2 = 0, 2$ и матрицами

$$\mathbf{\Lambda}^1 = \begin{pmatrix} 4 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 8 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \mathbf{\Lambda}^2 = \begin{pmatrix} 10 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 13 \end{pmatrix},$$

$$\mathbf{Q} = \begin{pmatrix} -4 & 1 & 2 & 1 \\ 0,6 & -2 & 0,4 & 1 \\ 3 & 1 & -9 & 5 \\ 2 & 0 & 8 & -10 \end{pmatrix}.$$

Поскольку асимптотические и точные математические ожидания определяются одинаковыми формулами, будем оценивать относительную погрешность сравнивая моменты второго порядка (5) и (15) по формуле:

$$\Delta_i = \frac{|m_2^{(i)} - m_2^{(i)} as|}{m_2^{(i)}}, i = 1, 2. \quad (17)$$

В таблице 2 представлены значения относительных погрешностей между точным и аппроксимационным значениями вторых начальных моментов.

Из таблицы 1 видно, что при уменьшении ε значения относительной погрешности для каждого типа заявок также уменьшаются. Следовательно, значения момента второго порядка, полученные асимптотическим путем, стремятся к значениям, полученным аналитически. При $\varepsilon \leq 0,005$ относительная погрешность составляет менее 10%.

Таблица 1

Относительная погрешность аппроксимации

	$\varepsilon = 0,1$	$\varepsilon = 0,01$	$\varepsilon = 0,005$
Δ_1	0,29	0,16	0,11
Δ_2	0,35	0,11	0,06

Заключение

В работе построена математическая модель бесконечнолинейной многопоточной неоднородной СМО, функционирующей в случайной среде. В асимптотическом условии предельно редких изменений состояний управляющей цепи Маркова получены выражения для вероятностных характеристик, проведен численный анализ для сравнения точных и асимптотических характеристик. Полученные результаты позволяют решать задачи оптимизации передачи данных.

СПИСОК ЛИТЕРАТУРЫ

1. *Наумов В. А., Самуйлов К. Е.* О моделировании систем массового обслуживания с множественными ресурсами // Вестник РУДН. Серия: математика, информатика, физика. 2014. № 3. С. 60–64.
2. *Eisen M., Taimiter M.* Stochastic variations in queueing processes // Operations Research. 1963. V. 11. P. 922–927.
3. *Naor P., Yechiali U.* Queueing problems with heterogeneous arrivals and service // Operations Research. 1971. V. 19. № 3. P. 722–734.
4. *Yechiali U.* A queueing type birth and death process defined as a continuous time markov chain // Operations Research. 1973. V. 21. № 2. P. 604–629.
5. *Neuts M. F.* A queue subject to extraneous phase changes // Advances in applied probability. 1971. V. 3. P. 78–119.
6. *Neuts M. F.* Matrix-geometric solutions in stochastic models // Baltimore and London: The John Hopkins University Press. 1981.

Невенченко Екатерина Алексеевна — старший преподаватель кафедры теории вероятностей и математической статистики ИПМКН НИ ТГУ. E-mail: *pravkovakaty_a_2010@mail.ru*

Шепилов Степан Сергеевич — студент ИПМКН НИ ТГУ. E-mail: *stepanshepilovwork@gmail.com*

Моисеева Светлана Петровна — доктор физико-математических наук, профессор, заведующая кафедрой теории вероятностей и математической статистики ИПМКН НИ ТГУ. E-mail: *sмоiseeva@mail.ru*

АНАЛИЗ ПИКОВОГО ВОЗРАСТА ИНФОРМАЦИИ В СЕТИ ИНТЕГРИРОВАННОГО ДОСТУПА И ТРАНЗИТА

Е. А. Гайдамака¹, А. А. Живцова¹, А. А. Милехин¹, К. Е. Самуйлов^{1,2}

¹ *Кафедра теории вероятностей и кибербезопасности, Российский университет дружбы народов имени Патриса Лумумбы, Россия, Москва*

² *Федеральный исследовательский центр «Информатика и управление» Российской академии наук (ФИЦ ИУ РАН), Россия, Москва*

В работе построена модель сети интегрированного доступа и транзита в виде сети Джексона. Каждая базовая станция сети представлена системой массового обслуживания типа $M/M/1/\text{inf}$, где обслуживание имеет физический смысл передачи по исходящему звену. Для модели получены формулы преобразования Лапласа-Стилтьеса и функции распределения пикового возраста информации на абонентских устройствах. Для проверки аналитических результатов и проведения численного анализа на языке Python написана программа имитационного моделирования. **Ключевые слова:** *возраст информации, сеть интегрированного доступа и транзита, сети массового обслуживания, имитационное моделирование.*

Введение

Возраст информации, концепция которого начала развиваться в научной литературе в начале 2010-х годов, является важным показателем качества, характеризующим актуальность передаваемых данных [1]. Его величина определяет время, прошедшее с момента генерации данных до их получения и обработки. Этот показатель представляет интерес в системах реального времени, например, в беспилотных транспортных аппаратах, сетях интернета вещей и многих других областях. В настоящей работе исследуется возраст информации в сети интегрированного доступа и транзита (Integrated Access and Backhaul, IAB) [2].

1. Математическая модель

Рассмотрим сеть IAB, состоящую из сервера, генерирующего пакеты информации, M базовых станций (БС) и P абонентских устройств

Исследование выполнено за счет гранта Российского научного фонда № 24-19-00804, <https://rscf.ru/project/24-19-00804/>. Публикация выполнена в рамках проекта № 021937-2-000 Системы грантовой поддержки научных проектов РУДН.

(АУ). Пакеты, сгенерированные сервером, моментально передаются к БС-донору, а от нее передаются АУ по каналам связи, проходящим через другие БС. Рассматривается только нисходящее направление передачи. БС, АУ и каналы связи формируют древовидный граф, в корневой вершине которого находится БС-донор, а в листовых вершинах – АУ [3].

Представим описанную выше сеть в виде сети Джексона [4]. Множество M узлов сети соответствует множеству БС, где первый узел, $i = 1$ соответствует БС-донору. Исток $\{a\}$ – сервер, стоки $\{b_p, p = \overline{1, P}\}$ – АУ. Пример сети IAB, представленной в виде сети Джексона, приведен на рис. 1.

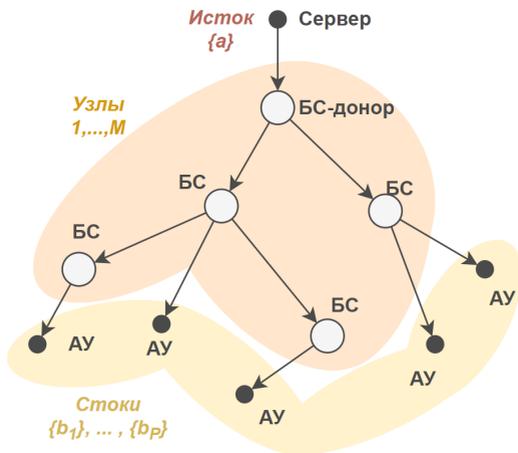


Рис. 1. Пример сети IAB. Красным цветом указан исток, оранжевым – множество БС, желтым – множество стоков.

В силу того, что сеть представлена древовидным графом, существует единственный маршрут от сервера до некоторого АУ p . Будем называть этот маршрут p -путь. Обозначим множество узлов, составляющих p -путь, через M_p , $M_p \subseteq M$, а множество путей – через \mathcal{P} .

Обозначим интенсивность поступления пакетов на p -путь через Λ_p , $p = \overline{1, P}$, а интенсивность поступления пакетов из узла i в узел j – через λ_{ij} . Суммарная интенсивность входящего потока и выходящего потока пакетов для узла i равны и обозначены через $\lambda_{\bullet i} = \lambda_{i \bullet} = \lambda_i$. Значения λ_{ij} для рассмотренного примера сети представлены на рис. 2. При нумерации узлов так, чтобы из двух любых узлов тот, кто находится ближе к донору, всегда имел меньший номер, матрица стационарных

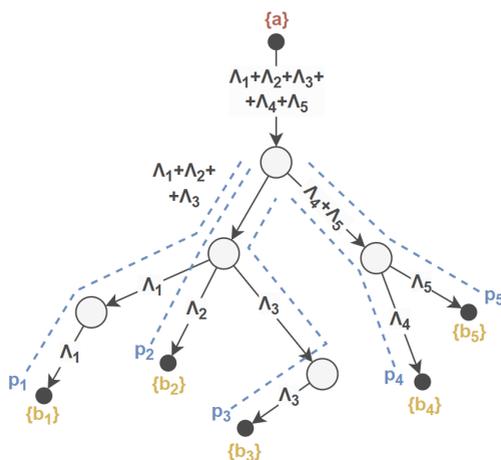


Рис. 2. Схема потоков сети. Красным цветом указан исток $\{a\}$, желтым – стоки $\{b_1\}, \dots, \{b_P\}$. Синими пунктирными линиями изображены пути сети p_1, \dots, p_P .

интенсивностей $\Lambda = (\lambda_{ij})$ принимает вид (1), где все столбцы кроме первого содержат ровно один ненулевой элемент. Элементы маршрутной матрицы Θ равны $\theta_{i,j} = \lambda_j / \lambda_i, \forall i, j \in \mathcal{M}_p \mid \lambda_i > 0$.

$$\Lambda = \begin{matrix} & \{a\} & 1 & 2 & \dots & M & \{b_1\} & \dots & \{b_P\} \\ \begin{matrix} \{a\} \\ 1 \\ 2 \\ \vdots \\ M \\ \{b_1\} \\ \vdots \\ \{b_P\} \end{matrix} & \left[\begin{array}{cccccccc} 0 & \lambda_{01} & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & \lambda_{12} & \dots & \lambda_{1M} & \lambda_{1M+1} & \dots & \dots & \lambda_{1M+P} \\ 0 & 0 & 0 & \dots & \lambda_{2M} & \lambda_{2M+1} & \dots & \dots & \lambda_{2M+P} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \lambda_{MM+1} & \dots & \dots & \lambda_{MM+P} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \end{array} \right] \cdot \end{matrix} \quad (1)$$

Будем моделировать передачу пакета между двумя БС в сети IAB как обслуживание пакета передающим узлом в сети Джексона. Для этого введем C_{ij} – емкость канала между БС i и j , тогда C_{ij} – интенсивность обслуживания узлом i пакетов, передающихся узлу j , в сети Джексона, $i \in \mathcal{M}, j \in \mathcal{M} \cup \{b_p\}_{p=1, \dots, P}$. Для простоты будем считать $C_{ij} = C$ для всех каналов сети IAB.

Случайная величина V_i задержки на узле i вычисляется как время пребывания в системе $M/M/1/\infty$ с интенсивностью входящего потока λ_i и интенсивностью обслуживания C . Обозначим $\gamma_i = C - \lambda_i$, тогда $V_i \sim \exp(\gamma_i)$. Обозначим через U_p случайную величину времени пребывания пакетов на p -пути:

$$U_p = \sum_{i \in \mathcal{M}_p} V_i.$$

Таким образом, U_p имеет обобщенное распределение Эрланга с $|\mathcal{M}_p| = M_p$ этапами, время пребывания на каждом из которых распределено экспоненциально с параметрами γ_i , $i \in \mathcal{M}_p$.

2. Возраст информации

Рассмотрим некоторую пару сервер-АУ. Обозначим через t_k момент создания сервером (и поступления на БС-донор) пакета k и через t'_k – момент получения АУ пакета k . Согласно [1], возраст информации $\Delta(t)$ в момент времени t равен времени, прошедшему с момента создания последнего доставленного пакета:

$$\Delta(t) = t - t_{N(t)}, \quad (2)$$

где $N(t)$ – номер последнего пакета, полученного АУ к моменту времени t . График зависимости возраста информации $\Delta(t)$ от времени представлен на рис. 3. Обозначим через Y_k время между генерациями пакетов $k-1$ и k , T_k – время доставки пакета k от сервера до АУ:

$$Y_k = t_k - t_{k-1}, \quad T_k = t'_k - t_k. \quad (3)$$

Пиковый возраст информации A_k пакета k определен следующим образом

$$A_k = t'_k - t_{k-1} = Y_k + T_k. \quad (4)$$

Пусть Q_k – площадь под графиком $\Delta(t)$, указанная на рис. 3. Средний возраст информации Δ_τ за время τ определяется формулой

$$\Delta_\tau = \frac{1}{\tau} \int_0^\tau \Delta(t) dt = \frac{1}{\tau} \sum_{k=1}^{N(\tau)+1} Q_k, \quad (5)$$

Введем Y – случайная величина времени между созданиями сервером пакетов, T – случайная величина времени доставки пакета, Q – случайная величина площади под графиком $\Delta(t)$ в промежуток между получениями АУ пакетов. Геометрически, Q можно выразить как

сумму площадей прямоугольного треугольника и параллелограмма, $Q = \frac{1}{2}Y^2 + YT$. Таким образом среднее значение возраста информации $\Delta = \lim_{\tau \rightarrow \infty} \Delta_\tau$ выражается как

$$\Delta = \frac{MQ}{MY} = \frac{1/2 M[Y^2] + M[YT]}{MY}, \tag{6}$$

а случайная величина пикового возраста информации A имеет вид

$$A = Y + T. \tag{7}$$

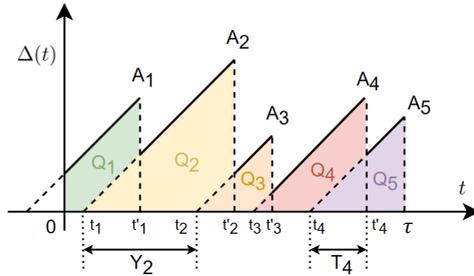


Рис. 3. График зависимости возраста информации $\Delta(t)$ от времени.

3. Расчет возраста информации в сети IAB

Из описания математической модели следует, что случайная величина времени между созданиями сервером пакетов для p -пути имеет вид $Y_p \sim \exp(\Lambda_p)$. Тогда случайная величина пикового возраста информации для пакетов на p -пути

$$A_p = Y_p + \sum_{i \in \mathcal{M}_p} V_i. \tag{8}$$

имеет обобщенное распределение Эрланга с $M_p + 1$ этапами, время пребывания на каждом из которых распределено экспоненциально с параметрами $\Lambda_p, \gamma_i, i \in \mathcal{M}_p$. Для обобщенного распределения Эрланга существует матричная форма записи преобразования Лапласа-Стилтьеса $\alpha(s)$ и функции распределения $F_{A_p}(x)$:

$$\alpha(s) = \beta^T (s\mathbf{I} - \mathbf{N}_p)^{-1} \mathbf{n}_p, \quad F_{A_p}(x) = 1 - \beta^T e^{\mathbf{N}_p x} \mathbf{1}, \tag{9}$$

где

$$\beta = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \mathbf{N}_p = \begin{pmatrix} -\Lambda_p & \Lambda_p & 0 & 0 \dots & \\ 0 & -\gamma_1 & \gamma_1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & -\gamma_{M_p} \end{pmatrix}, \quad \mathbf{n}_p = (\gamma_1 \dots \gamma_{M_p}).$$

Средний пиковый возраст информации может быть вычислен по формуле

$$MA = \frac{1}{\Lambda_p} + \sum_{i \in \mathcal{M}_p} \frac{1}{\gamma_i}. \quad (10)$$

4. Численный анализ

Для проверки аналитических результатов на языке Python написана программа дискретно-событийного имитационного моделирования. Эксперимент проведен для примера сети на рис. 1, принято $\Lambda_p = \Lambda \forall p \in \mathcal{P}$ и $C = 10 \text{ мс}^{-1}$. Отметим, что для существования стационарного режима необходимо, чтобы интенсивность входящего потока не превышала интенсивность обслуживания $\lambda_i \leq C, \forall i \in \mathcal{M}$. В частности для узла БС-донор необходимо выполнение $P\Lambda \leq C$ или $\Lambda \leq C/P = 10/5 = 2$. Отсюда принято $\Lambda = 1, \dots, 2, \forall p \in \mathcal{P}$.

На рис. 4 представлен график зависимости среднего пикового возраста информации от нагрузки на сеть. При низкой интенсивности Λ пиковый возраст информации принимает высокие значения в силу того, что интервал между генерациями пакетов велик. С ростом интенсивности поступления пакетов возраст информации убывает, но при высокой нагрузке снова растет, так как увеличивается время ожидания обслуживания на узлах. На правом графике представлена относительная погрешность симулятора. Погрешность не превышает 4%.

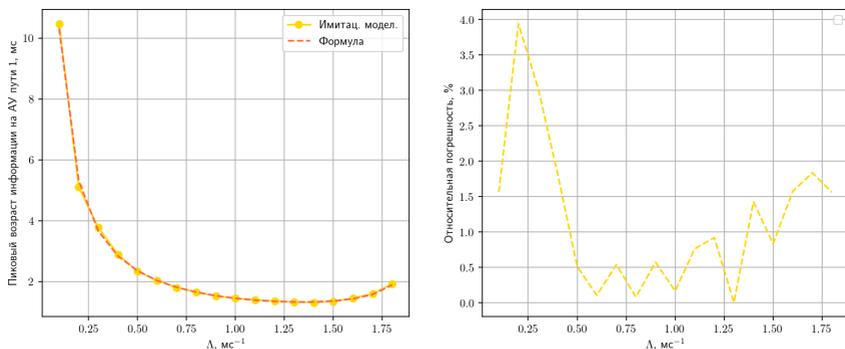


Рис. 4. График зависимости среднего пикового возраста информации на первом пути $A_{p=1}$ от нагрузки на сеть Λ/C при $C = 10 \text{ мс}^{-1}$.

Заключение

В ходе работы построена модель сети IAB в виде сети Джексона, для которой получены формулы преобразования Лапласа-Стилтьеса и функции распределения пикового возраста информации на абонентских устройствах. Формулы подтверждены результатами работы программы имитационного моделирования, написанной на языке Python.

СПИСОК ЛИТЕРАТУРЫ

1. *Antzela Kosta, Nikolaos Pappas and Vangelis Angelakis* Age of Information: A New Concept, Metric, and Tool // *Foundations and Trends in Networking*. 2017. Vol. 12. № 3. P. 162–259.
2. <https://portal.3gpp.org> – 3GPP TR 38.801 v14.0.0 Study on new radio access technology: Radio access architecture and interfaces. 2017.
3. *Manan Gupta et al* System-Level Analysis of Full-Duplex Self-Backhauled Millimeter Wave Networks // *IEEE Transactions on Wireless Communications* 22, 2021. P. 1130–1144.
4. *Башарин Г. П.* Лекции по математической теории телеграфика: Учеб. пособие. Изд. 2-е испр. и доп. Москва: РУДН, 2007. 268 с.

Гайдамака Елизавета Александровна — студент, кафедра теории вероятностей и кибербезопасности. E-mail: 1032216434@rudn.ru

Живцова Анна Александровна — студент, кафедра теории вероятностей и кибербезопасности. E-mail: zhivtsova-aa@rudn.ru

Милехин Александр Андреевич — аспирант, кафедра теории вероятностей и кибербезопасности. E-mail: *1142230064@rudn.ru*

Самуйлов Константин Евгеньевич — д.т.н., проф., зав. каф. теории вероятностей и кибербезопасности, директор Института компьютерных наук и телекоммуникаций, Российский университет дружбы народов им. П. Лумумбы; вед.н.с., Федеральный исследовательский центр «Информатика и управление» Российской академии наук (ФИЦ ИУ РАН). E-mail: *samuylov-ke@rudn.ru*

НАТУРНОЕ МОДЕЛИРОВАНИЕ НАРЕЗКИ СЕТИ 5G НА БАЗЕ КОМПЛЕКСНОЙ ИНТЕГРАЦИИ SDN И SDR

И. Е. З. Гебриал¹, К. А. Леонтьева¹, А. А. Хакимов¹,
А. С. А. Мутханна¹, И. А. Кочеткова^{1,2}

¹ *Российский университет дружбы народов, г. Москва, Россия*

² *Федеральный исследовательский центр «Информатика и управление»
Российской академии наук, г. Москва, Россия*

Беспроводные сети пятого поколения, или 5G, предоставляют широкий спектр услуг и приложений. Эти услуги можно отделить от сетевой и физической инфраструктуры, что позволяет создавать несколько виртуальных сетей на одной физической. Для управления ресурсами необходимо эффективное распределение. В работе представлена реализация натурального эксперимента с нарезкой сети. Распределение абонентов происходит между двумя операторами: основным и виртуальным. Абоненты разделены на три группы в зависимости от типа услуг: виртуальная реальность, просмотр видео и видеоигры. Для построения сети используют программно-определяемые радиосистема и сеть, а также коммутаторы и контроллер. Построены функции распределения задержек передачи данных.

Ключевые слова: *5G, нарезка сети, контроллер, SDN, SDR.*

Введение

Сети будущего – это значительный прорыв в сфере телекоммуникаций, который решает многие текущие проблемы и открывает новые возможности для пользователей. Эти сети обладают высокой степенью масштабируемости и программируемости, позволяя легко увеличивать пропускную способность между различными точками или центрами обработки данных. При этом нет необходимости обновлять оборудование или программное обеспечение.

Эффективное управление сетями 5G и ресурсами достигается благодаря нарезке сети на основе таких технологий, как SDN (software-defined networking) и виртуализация сетевых функций. SDN позволяет централизованно управлять сетью с помощью унифицированных программных средств, разделять процессы управления и передачи данных,

а также виртуализировать физические сетевые ресурсы. Эти технологии активно применяются в различных исследованиях [2, 3]. Анализ нарезки сети также осуществляется с помощью математических методов, например, управляемой системы массового обслуживания [4].

Интерес представляет не только теоретический анализ, но и практическая реализация сети 5G [5] и ее нарезки на конкретном оборудовании. В данной работе построена сеть, в которой распределение абонентов происходит между основным и виртуальным операторами, а также по требуемым услугам. Реализация тестового стенда проводилась на базе оборудования научного центра моделирования беспроводных сетей новых поколений РУДН.

1. Архитектура тестового стенда

Рассмотрим сеть основного оператора (mobile network operator, MNO), который владеет базовой станцией (radio access network, RAN), спектром радиочастот и ядром сети (рис. 1). Коммутаторы обеспечивают связь по физическим кабелям между ядром сети, RAN и серверами приложений. Коммутаторы 1 и 2 (switch_1, switch_2) предназначены для пропуска трафика абонентов, а к коммутаторам 3, 4 и 5 (switch_3, switch_4, switch_5) подключены серверы приложений.

Виртуальный оператор (virtual network operator, VNO) арендует ресурсы у основного оператора. Он использует ту же физическую инфраструктуру – все коммутаторы и RAN, для него создано виртуальное ядро сети.

Услуги можно разделить на две группы. Первая группа – это услуги, которые предоставляются от RAN по определенным маршрутам до коммутаторов, имеющих прямое или близкое подключение к серверам приложений. К ним относятся виртуальная реальность (VR), видеоигры (gaming), просмотр видео (video watching). Вторая группа – это прочие услуги, связанные непосредственно с ядром сети.

Нарезка сети достигается за счет разделения каналов L2 между RAN и ядром MNO и VNO. Дополнительно происходит разделение радиоресурсов по услугам для каждого оператора. В результате разделения каналов L2 создаются виртуальные каналы для различных услуг. Канал абонент – приложение управляется удаленно с ядра своего оператора.

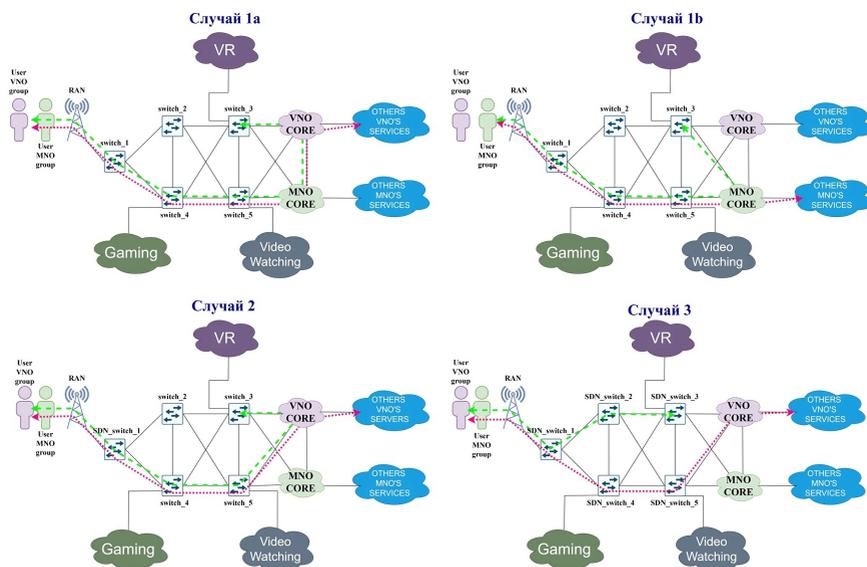


Рис. 1. Архитектура тестового стенда

Рассмотрим три возможных сценария работы данной сети, которые также представлены на рис. 1:

- 1) Коммутаторы $switch_i, i = 1, \dots, 5$ не относятся к архитектуре SDN.
- 2) Коммутатор $switch_1$ выполняет роль SDN-контроллера, который определяет принадлежность абонента конкретному оператору.
- 3) Коммутатор $switch_1$ играет роль SDN-контроллера, а коммутаторы $switch_i, i = 2, 3, 4, 5$ являются SDN-коммутаторами.

В первом сценарии (случай 1a), когда абонент виртуального оператора запрашивает услугу VR и подключается к RAN, трафик проходит через коммутаторы $switch_1, switch_4$ и $switch_5$, а затем достигает MNO. MNO определяет, что абонент принадлежит к виртуальному оператору VNO, основываясь на IP-адресе, и направляет его запрос туда. VNO затем распределяет абонента на коммутатор $switch_3$ для получения нужной услуги. В случае 1b абонент принадлежит MNO, поэтому запрос направляется непосредственно на $switch_3$.

Во втором сценарии SDN-контроллер SDN_switch_1 играет ключевую роль в определении принадлежности абонента к конкретному оператору. После этого запрос отправляется оператору через ядро VNO

Таблица 1

Параметры тестового стенда

Параметр	Характеристика
Software Defined Radio (SDR)	
Модель	NI USRP-2954R
Количество каналов	2 (2 Rx и 2 Tx)
Диапазон частот	10 МГц – 6 ГГц
Диапазон усиления	0 дБ – 20 дБ для Tx, 0 дБ – 15 дБ для Rx
Пропускная способность	160 МГц
Контроллер	
Модель	Intel(R) Xeon(R) CPU E5-2620 v4 @ 2.10GHz
Память	32 Гб
Коммутаторы	
Модель	Intel(R) Xeon(R) CPU E5-2650 v4 @ 2.20GHz
Память	32 Гб

напрямую на коммутатор `switch_3`, соответствующий требуемой услуге, минуя MNO.

В третьем сценарии, благодаря архитектуре SDN, определение оператора и типа услуги происходит на коммутаторе `SDN_switch_1`. Затем запрос абонента сразу отправляется на коммутатор `SDN_switch_3` без необходимости посещения VNO.

2. Натурный эксперимент

Ключевые характеристики оборудования, на котором была реализована сеть, представлены в табл. 1. В ходе эксперимента была использована радиокommunikационная система SDR (Software Defined Radio), которая использует программное обеспечение для обработки сигналов вместо обычных аппаратных компонентов.

Входной поток состоял из подключения 3000 абонентов в течение часа, каждый из которых пользовался определенной услугой одного из операторов. Общий канал имел пропускную способность 1 Гбит/с и был равномерно разделен между двумя операторами по 500 Мбит/с для каждого.

Из графиков на рис. 3 можно сделать следующие выводы:

- В сценарии 3, когда все коммутаторы относятся к архитектуре SDN, задержки передачи данных для различных типов услуг (gaming, video, VR) были наименьшими. Это говорит об эффективности данного подхода.
- Наибольшие задержки по передаче данных были зафиксированы для услуг gaming во всех трех случаях. Это подчеркивает высокие

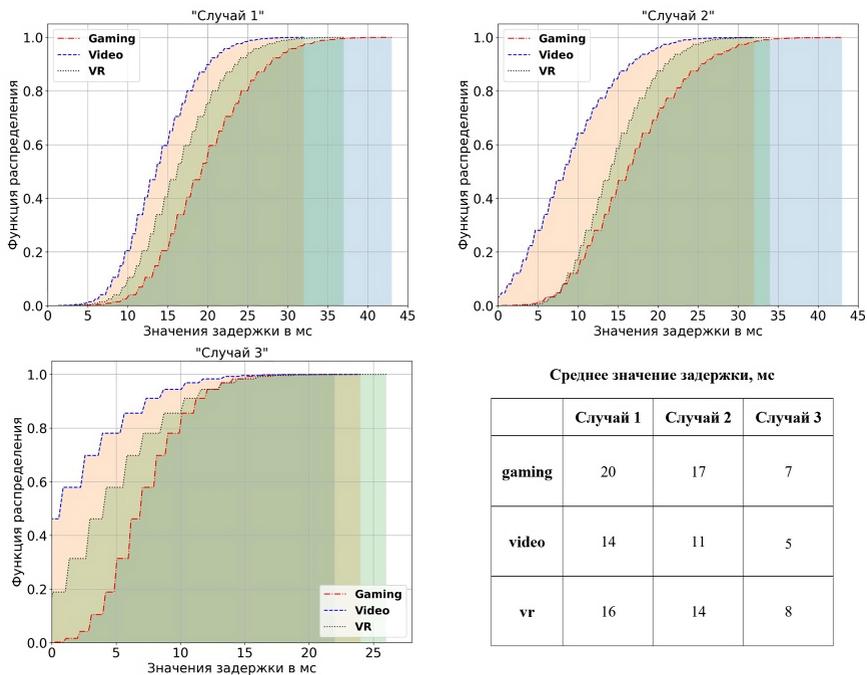


Рис. 2. Функция распределения времени задержки

требования к сети для данного типа услуг. Особенно заметно это в сценариях 1 и 2, где задержки для gaming оказались существенно выше, чем для других услуг.

- Услуги video и VR имеют более схожие распределения задержек. Это указывает на немного меньшие требования к сети по сравнению с Gaming.

Заключение

Исследование показало, что использование SDN позволяет эффективно управлять ресурсами. В дальнейших исследованиях планируется провести тестирование сети при большем количестве услуг и абонентов. Планируется провести дополнительные исследования по оптимизации нарезки сети с учетом динамически изменяющихся условий нагрузки на операторов.

СПИСОК ЛИТЕРАТУРЫ

1. *Maule M., Mekikis, P.-V., Ramantas K., Vardakas, J., and Verikoukis C.* Real-Time Dynamic Network Slicing for the 5G Radio Access Network // GLOBECOM 2019 – 2019 IEEE Global Communications Conference, 2019. P. 1–6.
2. *D'oro S., Restuccia F., Talamonti, A., and Melodia T.* The Slice Is Served: Enforcing Radio Access Network Slicing in Virtualized 5G Systems // IEEE International Conference on Computer Communications (INFOCOM), 2019.
3. *Kochetkova I., Leonteva K., Ghebrial, I., Vlaskina A., Burtseva S., Kushchazli A., and Samouylov K.* Controllable Queuing System with Elastic Traffic and Signals for Resource Capacity Planning in 5G Network Slicing // Future Internet, 2024.
4. *Khakimov A., Poluektov D., Mokrov, E., Koucheryavy Y., and Samouylov K.* Prototyping 5G Network in Laboratory Testbed with Virtual Deployment Environment // International Scientific and Technical Conference Modern Computer Network Technologies (MoNeTeC), 2020.

Гебриал Ибрам Есам Зекри — аспирант кафедры теории вероятностей и кибербезопасности РУДН. E-mail: *gebrial-ie@rudn.ru*

Леонтьева Ксения Андреевна — студент кафедры теории вероятностей и кибербезопасности РУДН. E-mail: *leontyeva-ka@rudn.ru*

Хакимов Абдукодир Абдукаримович — к.т.н, старший научный сотрудник научного центра моделирования беспроводных сетей новых поколений РУДН. E-mail: *khakimov-aa@rudn.ru*

Мутханна Аммар Салех Али — д.т.н, директор научного центра моделирования беспроводных сетей новых поколений РУДН. E-mail: *mutkhanna-as@rudn.ru*

Кочеткова Ирина Андреевна — к.ф.-м.н., доцент, доцент кафедры теории вероятностей и кибербезопасности РУДН, старший научный сотрудник ФИЦ ИУ РАН. E-mail: *kochetkova-ia@rudn.ru*

О ПРИМЕНЕНИИ СЕТЕЙ МАССОВОГО ОБСЛУЖИВАНИЯ ДЛЯ МОДЕЛИРОВАНИЯ РАБОТЫ МОРСКОГО КОНТЕЙНЕРНОГО ТЕРМИНАЛА

Х. З. Ву¹, М. Л. Жарков², А. Л. Казаков^{1,2}, А. А. Лемперт²

¹*Иркутский национальный исследовательский технический университет, г. Иркутск, Россия*

²*Институт динамики систем и теории управления имени В.М. Матросова СО РАН, г. Иркутск, Россия*

В статье предлагается и исследуется математическая модель работы морского контейнерного терминала, который обеспечивает перевалку контейнеров между наземным и морским видами транспорта, имеющая вид сети массового обслуживания. Поток контейнеров, прибывающих на грузовых автомобилях, описывается простейшим потоком заявок, а поступление контейнеров на морских судах – неординарным пуассоновским; работа ворот и склада моделируется системами массового обслуживания (СМО) с ограниченным ожиданием, причала – с неординарным обслуживанием и ограниченным ожиданием, якорной стоянки – СМО с групповым обслуживанием и отказами. Конкретным объектом исследования является крупный морской терминал на севере Вьетнама. На основе результатов численного исследования его модели определены текущий уровень загрузки и узкие места системы. **Ключевые слова:** математическое моделирование, сеть массового обслуживания, численный эксперимент, морской контейнерный терминал.

Введение

Методы теории массового обслуживания (ТМО) эффективны при исследовании технических систем, функционирование которых подвержено существенному влиянию случайных факторов [1, 2]. В области транспорта эти методы активно применяются еще с 80-х годов [3], как правило, для анализа наземных систем. Так, в [4, 5] предложены модели движения пассажиров через терминалы транспортно-пересадочных узлов, в [6, 7] – работы различных типов железнодорожных станций.

В области морского транспорта ТМО применяется реже. Большая часть исследований посвящена контейнерным терминалам. При этом из-за сложности их структуры обычно изучается работа отдельных подсистем. В частности, в [9] моделируется разгрузка прибывающих в морской порт грузовиков, в [10] оценивается эффективность различных конфигураций склада, в [11] описывается совместная работа склада и береговых кранов.

В данной статье ТМО используется для комплексного моделирования работы морского контейнерного терминала. В разделе 1 приводится описание объекта исследования. В разделе 2 строится модель его работы в виде сети массового обслуживания (СеМО) с двумя входящими потоками заявок. В разделе 3 проводится численное исследование модели, и на основе полученных результатов делается вывод об уровне загрузки и определяются узкие места в структуре системы.

1. Объект исследования

Морские контейнерные терминалы представляют собой механизированные склады, предназначенные для перевалки контейнеров между наземным и морским видами транспорта. Типовая схема таких объектов включает три основные подсистемы: грузовой фронт для наземного транспорта; контейнерную площадку (склад); причал с береговыми кранами. Морские терминалы являются узловыми элементами транспортной сети, поэтому их исследование с целью увеличения пропускной способности является актуальной задачей.

В качестве конкретного объекта исследования рассматривается международный контейнерный терминал Хайфон (Haiphong International Container Terminal), который является одним из крупнейших подобных объектов на севере Вьетнама и обладает типовой конфигурацией: наличествуют ворота для грузовиков, везущих контейнеры, склад и два причала; в акватории порта имеется якорная стоянка, где суда ждут разрешения на швартовку. Вследствие укрепления экономических связей между Россией и Вьетнамом наблюдается увеличение товарооборота на линии Владивосток – Хайфон (см. <https://www.interfax.ru/business/942097>), что влечет нагрузку на морские терминалы. Так, Хайфон в 2023 году обработал 636 тыс. контейнеров, что на 16% выше уровня ожиданий, и придает проведенному исследованию особую значимость в контексте развития двусторонних экономических контактов.

2. Математическая модель

Модель работы терминала Хайфон строится в виде СеМО в два этапа. На первом описываются поступающие потоки контейнеров, на втором – процесс их обработки в подсистемах.

Описание входящих потоков. В среднем в Хайфон с суши поступает ежедневно 871 грузовик, каждый доставляет один контейнер, их поступление описывается простейшим потоком M_1 с интенсивностью $\lambda_1 = 36,3$ заявок в час; с моря прибывает 10,3 контейнеровозов в неделю, каждый из них перевозит от 320 до 960 контейнеров (см. <https://hict.net.vn/thu-vien/Pages/thu-vien.aspx>), которые образуют неординарный пуассоновский поток M_2 с интенсивностью $\lambda_2 = 0,057$ групп заявок в час. Путем анализа статистических данных установлено, что размер группы является случайной величиной, которая подчиняется дискретному равномерному распределению $U(320; 960)$.

Работа подсистем терминала моделируется следующим образом: количество узлов равно числу направлений движения грузов (один или два). Распределения размера обслуживаемой группы заявок и времени ее обслуживания в узле определяются на основе результатов статистической обработки данных из открытых источников (см. <https://hict.net.vn>).

Подсистемы терминала Хайфон описываются 12 узлами, формальное описание которых в терминах ТМО имеет следующий вид:

- **Узел 1** – $M_1/M/9/0$, $\lambda = 80$ в час – въездные ворота;
- **Узел 2** – $*/M/5/0$, $\lambda = 80$ в час – выездные ворота;
- **Узлы 3 и 4** – $*/M/26/0$, $\lambda = 15$ в час – грузовики, перевозящие контейнеры от ворот к складу и обратно соответственно;
- **Узел 5** – $*/M/13/0$, $\lambda = 50$ в час – козловые краны, перегружающие контейнеры на склад;
- **Узел 6** – $*/M/13/20160$, $\lambda = 50$ в час – склад;
- **Узлы 7 и 8** – $*/M/15/0$, $\lambda = 12$ в час – тягачи, перевозящие контейнеры от склада к причалу и обратно соответственно;
- **Узел 9** – $*/M/3/0$, $\lambda = 30$ в час – береговые краны, которые загружают суда;
- **Узел 10** – $*/M^{X_{10}}/1/X$, $\lambda = 0,057$ в час, $X_{10} = U(420; 1060)$ – отправление загруженного судна с причала;
- **Узел 11** – $*/M/3/960$, $\lambda = 30$ в час – причал для разгружаемых судов;
- **Узел 12** – $M_2/M^{X_{12}}/10/0$, $\lambda = 0,44$ в час, $X_{12} = 960$ – якорная стоянка.

К этим узлам необходимо добавить еще два фиктивных, которые являются источниками потоков заявок [1]: **узел 0** для M_1 , **узел 13** для M_2 . Вероятности переходов групп заявок между узлами в описанной СеМО представлены в виде весов на рис. 1.

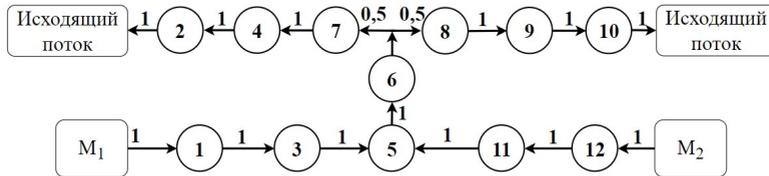


Рис. 1. Схема СеМО

Таким образом, модель работы терминала Хайфон имеет вид СеМО с двумя входящими потоками заявок и 14 узлами, из которых два являются фиктивными. **Узлы 1 и 2** описывают работу ворот, **Узлы 5 и 6** — склада, а **Узлы 9-12** — функционирование причалов и якорной стоянки. Перевозка контейнеров между подсистемами моделируется **Узлами 3, 4, 7 и 8**.

3. Вычислительный эксперимент

Исследование модели выполнено численно с помощью разработанного авторами программного обеспечения. Виртуальное время моделирования равно одному году. В табл. 1 представлены средние результаты за 10 пусков имитационной модели: P_{loss} — вероятность отказа; $T_{СМО}$ — среднее время пребывания в системе отдельной заявки; K_i — коэффициент занятости каналов в узле i , T_i — среднее время пребывания заявки (час), b_i — общее время блокировки одного канала (час) и L_i — средняя длина очереди в узле i .

Можно видеть, что в настоящее время Хайфон справляется с текущим объемом грузоперевозок. Однако наблюдается высокий уровень загрузки системы. В частности, в среднем 0,11% прибывающих грузовиков получают отказ в обслуживании из-за переполнения склада, которое происходит при пиковых нагрузках. Береговые краны являются наиболее загруженными элементами морского терминала, и при увеличении потоков контейнеров их пропускной способности будет недостаточно для своевременной обработки судов. В частности, при возрастании среднего размера поступающей партии контейнеров на 10% среднее время ожидания на якорной стоянке вырастет на 33%. Более подробная информация о полученных результатах будет представлена в докладе.

Таблица 1

Принято	Показатели	K_i	T_i	b_i	L_i
Заявок М1 312765,3	Узел 1	0,05	0,01	32,8	-
	Узел 2	0,09	0,01	0	-
Заявок М2 313560,6	Узел 3	0,09	0,07	33,8	-
	Узел 4	0,09	0,07	1,2	-
Отклонено	Узел 5	0,11	0,02	35,7	-
Заявок М1 349,0	Узел 6	0,74	52,18	9036,6	3724,8
	Узел 7	0,20	0,08	29,5	-
Заявок М2 0	Узел 8	0,77	0,33	8248,9	-
	Узел 9	0,82	0,07	6017,6	-
$P_L = 0,00056$	Узел 10	0,80	32,18	0	538,6
$T_{СМО} = 94,1$	Узел 11	0,40	4,90	38,3	174,0
	Узел 12	0,02	4,13	120,9	-

Выводы

В статье ТМО применяется для комплексного описания работы международного контейнерного терминала Хайфон, который расположен во Вьетнаме. Модель имеет вид сети массового обслуживания с несколькими входящими потоками заявок, в том числе неординарными. На основе результатов численного исследования модели оценена текущая загрузка системы и установлено, что пропускной способности береговых кранов при увеличении потока контейнеров будет недостаточно.

Дальнейшие исследования могут быть направлены на детализацию модели, в частности, учет нестационарности транспортных потоков. Также интерес представляет моделирование работы системы при поступлении сверхбольших контейнеровозов, перевозящих до 10000 контейнеров, поскольку для их своевременного обслуживания потребуется изменение технологии разгрузки/погрузки контейнеров на причале и введение дополнительных бригад.

СПИСОК ЛИТЕРАТУРЫ

1. *Dudin A., Klimenok V., Vishnevsky V.* The Theory of Queuing Systems with Correlated Flows. Cham: Springer, 2019. 410 p.
2. *Назаров А. А., Терпугов А. Ф.* Теория массового обслуживания. Томск: Изд-во НТЛ, 2010. 228 с.
3. *Клейнрок Л.* Теория массового обслуживания. М. : Машиностроение, 1979. 432 с.

4. *Майоров Н. Н.* Построение функциональной схемы обработки пассажиропотоков в морском пассажирском терминале // Системный анализ и логистика. 2020. № 2(24). С. 102–110.
5. *Журавская М. А., Казаков А. Л., Жарков М. Л., Парсюрова П. А.* Моделирование работы транспортно-пересадочного узла мегаполиса как трехфазной системы массового обслуживания // Транспорт Урала. 2015. № 3(46). С. 17–22.
6. *Бычков И. В., Казаков А. Л., Лемперт А. А., Бухаров Д. С., Столбов А. Б.* Интеллектуальная система управления развитием транспортно-логистической инфраструктурой региона // Проблемы управления. 2014. № 1. С. 27–35.
7. *Zharkov M., Lempert A., Pavidis M.* Simulation of Railway Marshalling Yards Based on Four-Phase Queuing Systems // In: Dudin A., Nazarov A., Moiseev A. (eds) ITMM. Queueing Theory and Applications. ITMM 2020, Cham: Springer, 2020. № 1391. P. 143–154.
8. *Vychkov I., Kazakov A., Lempert A., Zharkov M.* Modeling of railway stations based on queuing networks // Applied Sciences. 2021. № 11(5). P. 2425.
9. *Янченко А. А., Маликова Т. Е.* Методика анализа технологического процесса обработки груза на контейнерном терминале // Эксплуатация морского транспорта. 2020. № 2(95). С. 20–26.
10. *Roy D., de Koster R.* Optimal Stack Layout Configurations at Automated Container Terminals Using Queuing Network Models // In: Bose J.W. (eds) Handbook of Terminal Planning, Verlag: Springer, 2020. P. 437–461.
11. *Canonaco P., Legato P., Mazza R. M., Musmanno R.* A queuing network model for the management of berth crane operations // Computers & Operations Research. 2008. № 35(8), P. 243–2446.

Хоанг Занг Ву — аспирант. E-mail: giangtrang14022016@gmail.com

Жарков Максим Леонидович — к.т.н., н.с., Лаб. 6.2. Логических и оптимизационных методов анализа сложных систем. E-mail: zharkm@mail.ru

Казаков Александр Леонидович — д.ф.-м.н., г.н.с., Отд. 2. Прикладных проблем математической физики и теории поля. E-mail: kazakov@icc.ru

Лемперт Анна Ананьевна — к.ф.-м.н., зав.лаб., Лаб. 6.2. Логических и оптимизационных методов анализа сложных систем. E-mail: lempert@icc.ru

МОДЕЛИРОВАНИЕ СИСТЕМЫ ХРАНЕНИЯ И ТРАНСПОРТИРОВКИ ГРУЗОВ СЕТЬЮ МАССОВОГО ОБСЛУЖИВАНИЯ С ГРУППОВЫМ ОБСЛУЖИВАНИЕМ ТРЕБОВАНИЙ

Н. В. Сергеева, И. Е. Тананко, Е. П. Станкевич

*Саратовский национальный исследовательский государственный
университет имени Н. Г. Чернышевского, г. Саратов, Россия*

В качестве математической модели системы грузоперевозок рассматривается открытая сеть массового обслуживания, состоящая из систем обслуживания двух типов. Системы первого типа отображают распределительные пункты, а системы второго типа - транспортные средства. Предполагается, что длительности обслуживания в системах сети являются независимыми случайными величинами с экспоненциальным распределением. В системы первого типа из источника поступает пуассоновский поток требований. Из систем первого типа требования по одному могут переходить в источник или системы второго типа. Из систем второго типа требования переходят группами только в системы первого типа. Переходы требований в сети определяются маршрутной матрицей. Для данной сети получен инфинитезимальный оператор и основные стационарные характеристики.

Ключевые слова: *сеть массового обслуживания, групповое обслуживание, стационарные характеристики.*

Введение

Анализ и оптимизация сложных стохастических систем с сетевой структурой диктует необходимость создания, в частности, моделей сетей массового обслуживания с групповым обслуживанием требований. Обзор методов анализа систем массового обслуживания с групповым обслуживанием можно найти в [1]. Важнейшая область применения систем с групповым входящим потоком требований и групповым обслуживанием – транспортные системы. Так в работе [2] разработана и проанализирована модель транспортной системы модульного типа (подъемники, канатные дороги, метро и пр.). Целью моделирования являлось определение размера входящей группы в кабину на нижних станциях, чтобы уменьшить время ожидания пассажиров на верхних станциях.

С помощью построенной модели были разработаны рекомендации по загрузке гондольного подъемника горнолыжного курорта.

В работе [3] описана модель, которая используется при планировании производства полупроводникового завода. Модель представляет собой сеть массового обслуживания, состоящую из систем типа $GI^x/G^{(b,b)}/c$. Путем реконфигурирования входящего потока данные системы сводятся к системам типа $GI/G/c$, для которых получены приближенные формулы для вычисления средних характеристик. Основываясь на таком же допущении, авторы [4] посредством декомпозиции сети массового обслуживания и применения генетического алгоритма показывают, как можно решить задачу управления запасами, а именно задачу определения оптимальных размеров партий в системе производства нескольких продуктов с целью минимизации общего времени цикла. В работе [5] с помощью сети массового обслуживания с блокировкой моделируется работа погрузочно-разгрузочной тандемной системы. Предложен приближенный метод вычисления характеристик сети, разработан алгоритм поэтапного нахождения оптимального размера буферов.

С развитием интернет-торговли высокая загрузка машин и сокращение сроков доставки являются противоречивыми целями и нуждается в построении моделей для анализа и планирования работы интернет-магазинов. В данной работе построена математическая модель компаний, использующих систему грузоперевозок, в качестве которой выбрана открытая сеть массового обслуживания с групповым обслуживанием требований.

1. Описание модели

Рассматривается маркетплейс, состоящий из пунктов хранения и выдачи товаров, а также транспортных средств, которые перевозят товар между этими пунктами. Будем исследовать загруженность такой системы и время доставки товаров покупателям.

В качестве математической модели данной системы рассмотрим открытую сеть массового обслуживания. В сеть из источника поступает пуассоновский поток требований с интенсивностью λ_0 . Переходы требований в сети определяются маршрутной матрицей $\Theta = (\theta_{ij})$, $i, j = 0, 1, \dots, L$. Предполагается, что сеть обслуживания состоит из L систем массового обслуживания S_i , $i \in I$, $I = \{1, \dots, L\}$, двух типов. Системы первого типа обозначим S'_j , $j \in J$, $J = \{j : j \in I, \theta_{j0} \neq 0, \theta_{0j} \neq 0\}$. В системах S'_j содержится κ_j обслуживающих приборов. Системы вто-

рого типа обозначим S''_k , $k \in K$, $K = I \setminus J$. Они содержат один прибор, обслуживающий требования группой размера b_k .

Системы первого типа S'_j отображают в модели пункты хранения и выдачи. Число приборов соответствует числу работников на распределительном пункте. Полагаем, что длительность обслуживания в системе S'_j является экспоненциально распределенной случайной величиной с параметром μ_j , $j \in J$. Требования в системы этого типа поступают по одному из источника или группой размера b_k из систем S''_k , $k \in K$. Системы второго типа S''_k отображают в модели транспортные средства. Считаем, что длительность обслуживания в системе S''_k – экспоненциально распределенная случайная величина с параметром μ_k , $k \in K$. Системы S''_k связаны только с системами S'_j и не связаны с источником.

2. Метод анализа

Обозначим состояние сети вектором $n = (n_i)$, $i = 1, \dots, L$, где n_i – число требований, находящихся в системе S_i . Рассмотрим все возможные переходы между состояниями и запишем их интенсивности переходов. Изменение состояния сети может происходить в следующих случаях:

- 1) требование поступило в сеть из источника. Интенсивность перехода в этом случае имеет вид:

$$q(n, n + e_j) = \lambda_0 \theta_{0j}, \quad j \in J,$$

где e_j – это вектор, j -я компонента которого равна 1, а все остальные компоненты равны 0;

- 2) требование обслужилось в системе S'_j и ушло в источник, тогда

$$q(n, n - e_j) = \mu_j \min(n_j, \varkappa_j) \theta_{j0}, \quad j \in J;$$

- 3) требование обслужилось в системе S'_j и перешло в систему S''_k , тогда

$$q(n, n - e_j + e_k) = \mu_j \min(n_j, \varkappa_j) \theta_{jk} \mathbf{1}(n_k < b_k),$$

где $\mathbf{1}(A) = 1$, если A – истинно, иначе $\mathbf{1}(A) = 0$;

- 4) требование обслужилось в системе S''_k и перешло в систему S'_j , тогда

$$q(n, n - b_k e_k + b_k e_j) = \mu_k \theta_{kj} \mathbf{1}(n_k = b_k);$$

- 5) в остальных случаях

$$q(n, n') = 0, \quad n \neq n'.$$

Построив инфинитезимальный оператор $Q = (q(n, n'))$, стационарное распределение можно найти численно решив систему $\pi Q = 0$.

Интенсивности входящих потоков в системы S_i , $i \in I$, будем вычислять по формулам [6]

$$\lambda_i = \frac{\omega_i}{\omega_0} \lambda_0, \quad i = 1, \dots, L,$$

где относительные интенсивности потоков ω_i являются решением уравнения $\omega \Theta = \omega$ с условием $\sum_{i=0}^L \omega_i = 1$.

Сеть массового обслуживания функционирует в стационарном режиме, когда

$$\frac{\lambda_j}{\varkappa_j \mu_j} < 1, \quad j \in J.$$

Для систем S_k'' всегда существует стационарный режим.

Математическое ожидание (м. о.) числа требований в системах S_i , $i \in I$ [7],

$$\bar{n}_i = \sum_{m=0}^{\infty} m P_i(m),$$

где стационарная вероятность пребывания системы S_i в состоянии m

$$P_i(m) = \sum_{\forall n: n_i=m} \pi(n).$$

Используя закон Литтла, м. о. длительности пребывания требований в системе S_i

$$\bar{u}_i = \bar{n}_i / \lambda_i.$$

Время реакции сети массового обслуживания

$$\bar{\tau} = \frac{1}{\lambda_0} \sum_{i=1}^L \lambda_i \bar{u}_i.$$

3. Численный пример

Рассмотрим открытую сеть массового обслуживания, состоящую из $L = 5$ систем массового обслуживания. Системы S_1, S_2, S_5 относятся к системам типа S' , а системы S_3, S_4 к системам S'' . Для системы S_3 размер обслуживаемой группы требований $b_3 = 2$, а для системы S_4 размер $b_4 = 3$. Вектор интенсивностей обслуживания $\mu = (1.1; 1.2; 1.7; 2.3; 2)$,

вектор числа обслуживающих приборов $\varkappa = (3; 3; 1; 1; 2)$, маршрутная матрица имеет вид

$$\Theta = \begin{pmatrix} 0 & 0.4 & 0.3 & 0 & 0 & 0.3 \\ 0.6 & 0 & 0 & 0.2 & 0.2 & 0 \\ 0.7 & 0 & 0 & 0 & 0.3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0.3 & 0 & 0 & 0 & 0.7 \\ 0.5 & 0 & 0 & 0 & 0.5 & 0 \end{pmatrix}.$$

Результаты моделирования для некоторых интенсивностей входящего потока представлены в таблице 1.

Таблица 1

Стационарные характеристики сети массового обслуживания

	$\lambda_0 = 0.1$	$\lambda_0 = 1$
λ	(0.1; 0.06; 0.03; 0.02; 0.07; 0.09)	(1; 0.59; 0.3; 0.12; 0.64; 0.87)
\bar{n}	(0.05; 0.03; 0.52; 1.42; 0.05)	(0.55; 0.26; 0.57; 1.43; 0.57)
\bar{u}	(0.91; 0.84; 43.61; 22.02; 0.55)	(0.93; 0.86; 4.77; 2.23; 0.60)
$\bar{\tau}$	20.59	3.33

Из результатов видно, что увеличение λ_0 в 10 раз привело к значительному уменьшению времени реакции сети обслуживания. Изменение произошло за счёт уменьшения времени формирования групп требований в системах S_3 и S_4 . М. о. длительностей пребывания требований в системах S_1 , S_2 и S_5 практически не изменилось. Таким образом, использование модели позволяет планировать и оптимизировать работу пунктов распределения и хранения товаров.

Заключение

В данной работе построена математическая модель процессов хранения и перевозки товаров в компаниях, занимающихся интернет-продажей. Для этой модели получены основные стационарные характеристики. Приведены численные примеры, демонстрирующие загруженность моделируемой системы и времени доставки товаров покупателям. Результаты, полученные в работе, могут использоваться при моделировании процессов: производства и сборки изделий из многих элементов; сортировки и комплектования материалов; хранения в автоматизированных складских системах; пассажирских и грузовых перевозок транспортных систем. Математические особенности метода анализа открытых сетей массового обслуживания с групповым обслуживанием позво-

ляют решать задачи крупных компаний по проектированию, анализу и оптимизации систем, связанных с производством, переработкой, транспортировкой, продажей, сервисом.

СПИСОК ЛИТЕРАТУРЫ

1. *Першаков Н. В., Самуйлов К. Е.* Система M/G/1 с групповым обслуживанием и их применение к анализу модели протокола управления потоковой передачей. Часть I // Вестник РУДН. Серия Математика. Информатика. Физика. 2009. № 1. С. 34–44.
2. *Grippa P., Schilcher U., Bettstetter Ch.* On Access Control in Cabin-Based Transport Systems // IEEE Transactions on Intelligent Transportation Systems. 2019. Vol. 20. no. 6. P. 2149–2156.
3. *Hanschke Th., Zisgen H.* Queueing networks with batch service // European J. Industrial Engineering. 2011. Vol. 5. no. 3. P. 313–326.
4. *Rabta B., Reiner G.* Batch sizes optimisation by means of queueing network decomposition and genetic algorithm // International Journal of Production Research. 2012. Vol. 50. no. 10. P. 2720–2731.
5. *Yu A.-L., Zhang H.-Y., Chen Q.-X., Mao N., Xi Sh.-H.* Buffer allocation in a flow shop with capacitated batch transports // Journal of the Operational Research Society. 2022. Vol. 73. no. 4. P. 888–904.
6. *Stankevich E., Tananko I., Pagano M.* Optimization of Open Queuing Networks with Batch Services // Mathematics. 2022. Vol. 10. no. 16. 3027.
7. *Винневский В. М.* Теоретические основы проектирования компьютерных сетей. Москва : Техносфера, 2003. 506 с.

Сергеева Надежда Викторовна — ст. преподаватель, кафедра системного анализа и автоматического управления. E-mail: *SergeevaNV@info.sgu.ru*

Тананко Игорь Евстафьевич — к. ф.-м. н., доцент, заведующий кафедрой, кафедра системного анализа и автоматического управления. E-mail: *TanankoIE.sgu@gmail.com*

Станкевич Елена Петровна — к. ф.-м. н., доцент, кафедра системного анализа и автоматического управления. E-mail: *Stankevichelena@mail.ru*

АЛГОРИТМ ВЫЧИСЛЕНИЯ СТАЦИОНАРНОГО РАСПРЕДЕЛЕНИЯ ВЕРОЯТНОСТЕЙ СИСТЕМЫ С ПРИОРИТЕТАМИ ДЛЯ ПРОИЗВОЛЬНОГО ЧИСЛА МЕСТ В ОЧЕРЕДИ

Н. Г. Хаустова¹, С. П. Моисеева¹, Е. А. Пакулова²

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

² *Южный федеральный университет, г. Таганрог, Россия*

Рассматривается модель обработки информации с двумя потоками заявок — приоритетных и неприоритетных. Время обслуживания заявок случайное с экспоненциальным распределением вероятностей, причем его параметры соответствуют типу заявки. Также у каждого типа заявок имеется свой бункер ожидания, ограниченный по объему. Информация, находящаяся в очереди, обладает сроком жизни, после истечения которого ее передача может стать неактуальной. Срок жизни — случайная величина, также имеющая экспоненциальное распределение. Для указанной системы реализован алгоритм вычисления стационарного распределения вероятностей, позволяющий автоматизировать одну из наиболее трудоемких частей работы с приоритетными системами вручную — построение матриц коэффициентов. Данный алгоритм открывает возможность исследования систем с приоритетами на сколь угодно больших объемах бункеров, чего добиться без автоматизации крайне затруднительно.

Ключевые слова: *системы массового обслуживания с приоритетами, разнотипные данные, построение матриц коэффициентов, распределение вероятностей системы, срок жизни информации.*

Введение

Системы массового обслуживания как модель бизнес-процессов, происходящих в обществе, плотно вошли в жизнь человека, а их непрерывное усовершенствование и наращивание функционала позволяет удовлетворять все больше потребностей. Одним из наиболее актуальных направлений в теории телеграфика [1] является изучение систем с приоритетами [2], благодаря чему появляется возможность грамотно

перераспределять ресурсы рассматриваемой сети — к примеру, колл-центр компании, обслуживающей в первую очередь новых клиентов [4].

Однако нередко в ходе рассмотрения той или иной системы возникают проблемы, требующие применения методов оптимизации и алгоритмизации. К таким относится и исследование системы с приоритетами для произвольного числа мест в очереди. Эта проблема была выявлена в ходе рассмотрения систем с приоритетами на малых объемах бункеров ожидания и построения матрицы коэффициентов с последующим получением стационарного распределения вероятностей вручную, чего сделать на больших объемах бункеров не представляется физически возможным. К тому же заявки, поступающие на обслуживание прибору, в отличие от рассматриваемых в статье [4], являются нетерпеливыми — то есть информация имеет срок жизни, по истечении которого ее передача становится неактуальной. В настоящей статье приводится подробный алгоритм вычисления стационарного распределения вероятностей, благодаря которому исследование систем с приоритетами при нетерпеливости заявок и произвольном числе мест в очереди становится не просто возможным, но и наиболее удобным для пользователя.

1. Математическая модель системы с приоритетами при нетерпеливости заявок и произвольном числом мест в очереди

Сформулируем постановку задачи. Пусть $i(t)$ — число приоритетных заявок в бункере в момент времени t , $j(t)$ — число неприоритетных заявок в бункере, $k(t)$ — состояние прибора в момент времени t : $k(t) = 0$ — прибор свободен, $k(t) = 1$ — занят приоритетной заявкой, $k(t) = 2$ — занят неприоритетной заявкой.

Ставится задача нахождения стационарного распределения вероятностей трехмерного марковского процесса $\{k(t), i(t), j(t)\}$: $P_k(i, j) = \lim_{t \rightarrow \infty} P\{k(t) = k, i(t) = i, j(t) = j\}$ при следующих параметрах системы:

λ_p — интенсивности потоков, μ_p — интенсивности обслуживания, α_p — интенсивности выхода из бункеров, N_p — количество мест для ожидания в бункере, $p = 1, 2$.

2. Формализация событий в системе

Необходимо начать с формализации ключевых событий системы, которые при переводе с человеческих понятийных образов в формальный математический вид позволят сформировать корректный алгоритм.

События в системе для приоритетного потока:

— Пришла заявка приоритетного потока (λ_1)

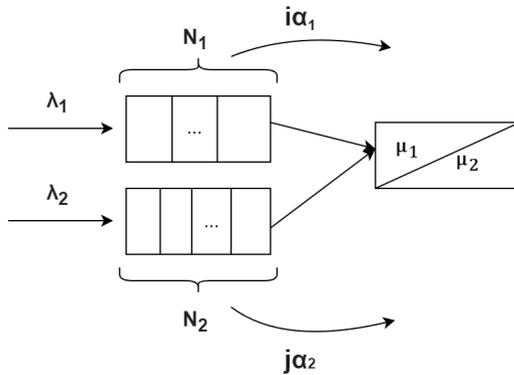


Рис. 1. Математическая модель обработки информации с приоритетами и нетерпеливыми заявками

1. Если прибор свободен, то состояние прибора равно 1.
 2. Если прибор занят, и в 1-ой очереди есть место, то число заявок в 1-ой очереди увеличивается на 1.
- Закончилось время ожидания заявки приоритетного потока ($i\alpha_1$)
 - 3. Если в 1-ой очереди есть заявки, то число заявок в 1-ой очереди уменьшается на 1.
 - Завершила обслуживание заявка приоритетного потока (μ_1)
 - 4. Если в 1-ой очереди есть заявка, то число заявок в 1-ой очереди уменьшается на 1.
 - 5. Если в 1-ой очереди нет заявок, во 2-ой очереди есть заявка, то состояние прибора равно 2, а число заявок во 2-ой очереди уменьшается на 1.
 - 6. Если в 1-ой очереди нет заявок и во 2-ой очереди нет заявок, то состояние прибора равно 0.
- События в системе для неприоритетного потока:**
- Пришла заявка неприоритетного потока (λ_2)
 - 7. Если прибор свободен, то состояние прибора равно 2.
 - 8. Если прибор занят, и во 2-ой очереди есть место, то число заявок во 2-ой очереди увеличивается на 1.
 - Закончилось время ожидания заявки неприоритетного потока ($j\alpha_2$)

- 9. Если во 2-ой очереди есть заявки, то число заявок во 2-ой очереди уменьшается на 1.
- **Завершила обслуживание заявка неприоритетного потока (μ_2)**
- 10. Если в 1-ой очереди нет заявок, во 2-ой очереди есть заявка, то число заявок во 2-ой очереди уменьшается на 1.
- 11. Если в 1-ой очереди есть заявка, то состояние прибора равно 1, а число заявок в 1-ой очереди уменьшается на 1.
- 12. Если в 1-ой очереди нет заявок, и во 2-ой очереди нет заявок, то состояние прибора равно 0.

На рисунках 2 и 3 представлены соответствующие графы переходов состояний рассматриваемого трехмерного марковского процесса.

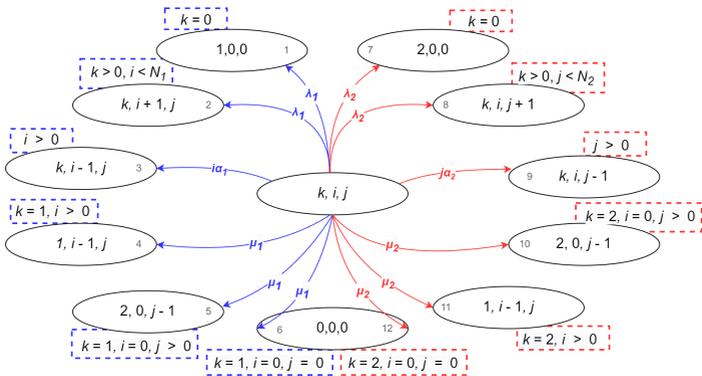


Рис. 2. Граф переходов для диагональных элементов матрицы процесса $\{k(t), i(t), j(t)\}$

3. Алгоритм вычисления стационарного распределения вероятностей

Наиболее трудоемкой частью исследования системы вручную является составление матрицы коэффициентов, ведь при увеличении количества мест в бункерах ожидания всего на единицу наблюдается резкое увеличение числа уравнений, а значит, и размерности матрицы. К тому же, важно сопоставить индекс состояния с самим состоянием $\{k(t), i(t), j(t)\}$. Именно эти два момента представляют интерес при создании алгоритма, описанного в настоящей статье. Алгоритм реализован на языке программирования Python.

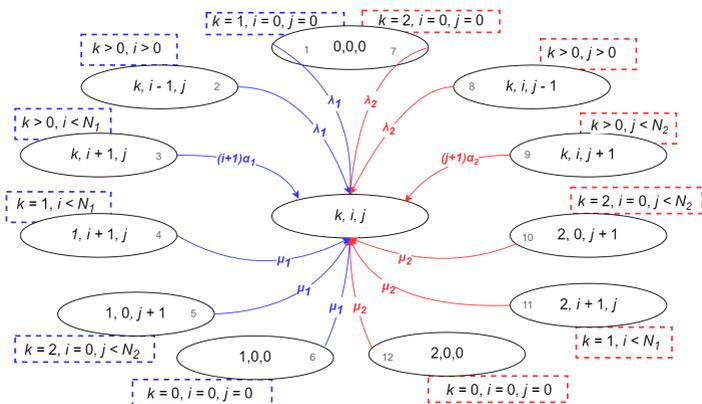


Рис. 3. Граф переходов для определенных состояний системы $\{k(t), i(t), j(t)\}$

Алгоритм вычисления стационарного распределения вероятностей:

1. Установить взаимно-однозначное соответствие между индексом состояния $n \in [0, \dots, 2 * (N_1 + 1) * (N_2 + 1)]$ и состоянием (k, i, j) при помощи структуры dict (словарь).
2. Определить нулевую матрицу **A** в структуре numpy.array (двумерный массив) размерности $|U| * |U|$.
3. Заполнить матрицу **A** в соответствии с таблицей 1.
4. Заполнить последнюю строку матрицы **A** единицами.
5. Заполнить вектор свободных членов **b** в структуре numpy.array (одномерный массив), у которого все элементы равны нулю, кроме последнего — он равен единице.
6. Решить систему линейных алгебраических уравнений $\mathbf{Ax}=\mathbf{b}$ с помощью метода numpy.linalg.solve(), аргументами которого являются матрица **A** и вектор **b**.

Ключевым результатом работы алгоритма будет являться вывод распределения вероятностей по всем возможным состояниям системы. Например, для параметров $N_1 = 1, N_2 = 2, \lambda_1 = 0.2, \lambda_2 = 0.5, \mu_1 = 0.2, \mu_2 = 0.3, \alpha_1 = 0.02, \alpha_2 = 0.03$ программа, использующая описанный алгоритм, выдаст следующее распределение вероятностей:

(0, 0, 0) : 0.03	(2, 0, 0) : 0.06
(1, 0, 0) : 0.015	(2, 0, 1) : 0.108
(1, 0, 1) : 0.044	(2, 0, 2) : 0.1
(1, 0, 2) : 0.222	(2, 1, 0) : 0.016
(1, 1, 0) : 0.005	(2, 1, 1) : 0.042
(1, 1, 1) : 0.033	(2, 1, 2) : 0.108
(1, 1, 2) : 0.217	

Таблица 1

Алгоритм заполнения матрицы коэффициентов

Условие	Действие
$k = 0, i = 0, j = 0$	$A[n, n]^- = \lambda_1 + \lambda_2; A[n, (1, 0, 0)]^+ = \mu_1;$ $A[n, (2, 0, 0)]^+ = \mu_2$
$k = 1, i > 0$	$A[n, n]^- = \mu_1$
$k = 2, i > 0$	$A[n, n]^- = \mu_2$
$k = 1, i = 0, j = 0$	$A[n, n]^- = \mu_1; A[n, (0, 0, 0)]^+ = \lambda_1$
$k = 2, i = 0, j = 0$	$A[n, n]^- = \mu_2; A[n, (0, 0, 0)]^+ = \lambda_2$
$k = 1, i = 0, j > 0$	$A[n, n]^- = \mu_1$
$k = 2, i = 0, j > 0$	$A[n, n]^- = \mu_2$
$i > 0$	$A[n, n]^- = i\alpha_1$
$j > 0$	$A[n, n]^- = j\alpha_2$
$k > 0, i > 0$	$A[n, (k, i - 1, j)]^+ = \lambda_1$
$k > 0, j > 0$	$A[n, (k, i, j - 1)]^+ = \lambda_2$
$k > 0, i < N_1$	$A[n, n]^- = \lambda_1; A[n, (k, i + 1, j)]^+ = (i + 1)\alpha_1$
$k > 0, j < N_2$	$A[n, n]^- = \lambda_2; A[n, (k, i, j + 1)]^+ = (j + 1)\alpha_2$
$k = 1, i < N_1$	$A[n, (1, i + 1, j)]^+ = \mu_1; A[n, (2, i + 1, j)]^+ = \mu_2$
$k = 2, i = 0, j < N_2$	$A[n, (1, 0, j + 1)]^+ = \mu_1; A[n, (2, 0, j + 1)]^+ = \mu_2$

Заключение

В настоящем материале описан алгоритм вычисления стационарного распределения вероятностей, позволяющий рассматривать системы с приоритетами на сколь угодно больших объемах бункеров ожидания. Представленный алгоритм не только расширяет возможности исследования систем с приоритетами, но и оптимизирует вычислительную работу исследователей, автоматизируя заполнение матрицы коэффициен-

тов и, как следствие, построение системы уравнений для вычисления распределения вероятностей.

СПИСОК ЛИТЕРАТУРЫ

1. *Степанов С. Н.* Теория телеграфика: концепции, методы, приложения. М: Горячая линия — Телеком, 2015. 868 с.
2. *Бронштейн О. И., Духовный И. М.* Модели приоритетного обслуживания в информационно-вычислительных системах. М: Наука, 1976. 220 с.
3. *Назаров А. А., Змеев О. А.* Ограничение нагрузки в телекоммуникационных сетях // Научное творчество молодежи. Математика. Информатика. Материалы XVIII Всерос.научн.-практ. конфер. 2014. С. 11–15.
4. *Moiseeva S. P., Pakulova E. A, Haustova N. G.* Mathematical Model of Information Processing with Relative Priority and Different Service Intensity. //2023 IEEE 19 th International Asian School-Seminar Optimization Problems of Complex Systems (OPCS), 14-22 august 2023. P. 65-69.

Хаустова Наталья Глебовна — студент Национального исследовательского Томского государственного университета. E-mail: *khaustovan@inbox.ru*

Моисеева Светлана Петровна — д.ф.-м.н., профессор Национального исследовательского Томского государственного университета. E-mail: *smoiseeva@mail.ru*

Пакулова Екатерина Анатольевна — к.т.н., доцент, Южный федеральный университет. E-mail: *epakulova@sfnedu.ru*

АСИМПТОТИЧЕСКИЙ АНАЛИЗ RQ-СИСТЕМЫ С КАТАСТРОФАМИ В УСЛОВИИ ВЫСОКОЙ ИНТЕНСИВНОСТИ ВХОДЯЩЕГО ПОТОКА

Н. П. Мелошникова

*Национальный исследовательский Томский государственный университет,
г. Томск, Россия*

В работе представлено исследование однолинейной RQ-системы с катастрофами. На вход поступает простейший поток положительных заявок, время обслуживания прибора распределено по экспоненциальному закону. Заявки, не получившие обслуживание, отправляются на орбиту, где осуществляют случайную задержку. Катастрофы происходят в моменты наступления событий простейшего потока отрицательных заявок, которые в момент прихода уничтожают все «положительные» заявки, находящиеся в системе: как на приборе, так и на орбите. В работе найдено стационарное распределение вероятностей числа заявок на орбите методом асимптотического анализа в условии высокой интенсивности входящего потока. **Ключевые слова:** RQ-система, отрицательные заявки, катастрофы, высокая интенсивность входящего потока.

Введение

Системы массового обслуживания с повторными вызовами (Retrial queueing systems или RQ-системы) это современные математические модели теории массового обслуживания (ТМО). Такие модели широко используются для построения, анализа и оптимизации различных инфо-коммуникационных систем, в которых используются механизмы для осуществления повторных попыток получить обслуживание спустя некоторое случайное время. Примерами таких систем являются сети связи, использующие протоколы множественного случайного доступа, call-центры, центры распределенных вычислений и другие системы [1, 2].

Также важным направлением ТМО являются G-системы и G-сети [3, 4]. Данные модели включают в себя понятие отрицательных заявок, которые могут оказывать различные виды негативных воздействий на систему: от уничтожения одного обрабатывающегося запроса до катастроф, приводящих к обнулению всей системы [5]. Актуальность таких моделей обусловлена различными негативными воздействиями на

инфо-коммуникационные системы, такими как вирусные поражения, хакерские атаки, системные сбои и другие угрозы.

В работе предлагается асимптотический метод исследования RQ-системы M/M/1 с катастрофами в условии высокой интенсивности входящего потока [6].

1. Математическая модель

В работе рассматривается однолинейная RQ-система M/M/1 (Рисунок 1). На вход системы поступает простейший поток положительных заявок с параметром λ . Время обслуживания каждой заявки распределено по экспоненциальному закону с параметром μ . Когда прибор оказывается занят, то входящая заявка идет на орбиту, где осуществляет случайную задержку, продолжительность которой имеет экспоненциальное распределение с параметром σ . С орбиты после случайного времени задержки заявка вновь обращается к обслуживающему прибору с повторной попыткой.

Также на вход поступает простейший поток отрицательных заявок с параметром γ . Отрицательная заявка не нуждается в обслуживании, при поступлении, она «обнуляет» всю систему, то есть все заявки, находящиеся на обслуживании и орбите покидают систему. Такие модели называют системами с катастрофами.

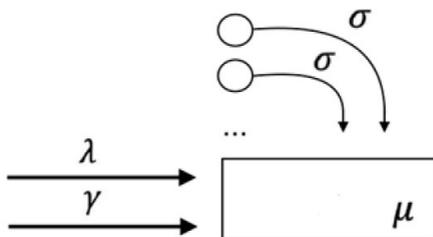


Рис. 1. RQ-система M|M|1 с катастрофами

Пусть $i(t)$ – число заявок на орбите, а $k(t)$ – определяет состояние прибора следующим образом:

$$k(t) = \begin{cases} 0, & \text{прибор свободен,} \\ 1, & \text{прибор занят.} \end{cases}$$

Ставится задача нахождения стационарного распределения вероятностей числа заявок на орбите. Обозначим $P\{k(t) = k, i(i) = i\} = P(k, i)$ – стационарные вероятности того, что прибор находится в состоянии k , а на орбите i заявок. Двумерный процесс $\{k(t), i(t)\}$ является Марковским. Система уравнений Колмогорова в стационарном виде будет иметь вид:

при $i = 0$:

$$\begin{cases} -(\lambda + \gamma)P(0, 0) + \mu P(1, 0) + \gamma = 0, \\ -(\lambda + \mu + \gamma)P(1, 0) + \lambda P(0, 0) + \sigma P(0, 1) = 0, \end{cases}$$

при $i > 0$:

(1)

$$\begin{cases} -(\lambda + i\sigma + \gamma)P(0, i) + \mu P(1, i) = 0, \\ -(\lambda + \mu + \gamma)P(1, i) + \lambda P(0, i) + \sigma(i + 1)P(0, i + 1) + \lambda P(1, i - 1) = 0. \end{cases}$$

Составим систему уравнений, определяющих частичные характеристические функции $H(k, u) = \sum_{i=0}^{\infty} e^{ju i} P(k, i)$. Система (1) переписется в виде:

$$\begin{cases} -j\sigma \frac{\partial H(0, u)}{\partial u} = -(\lambda + \gamma)H(0, u) + \mu H(1, u) + \gamma, \\ j\sigma e^{-ju} \frac{\partial H(0, u)}{\partial u} = -(\lambda + \mu + \gamma)H(1, u) + \lambda H(0, u) + \lambda e^{ju} H(1, u). \end{cases} \quad (2)$$

Сложив уравнения системы (2), получим:

$$\begin{aligned} j\sigma(e^{-ju} - 1) \frac{\partial H(0, u)}{\partial u} = \\ = \gamma - \gamma(H(0, u) + H(1, u)) + \lambda(e^{ju} - 1)H(1, u). \end{aligned} \quad (3)$$

Для решения системы (2)–(3) в работе предлагается метод асимптотического анализа в условии высокой интенсивности входящего потока ($\lambda \rightarrow \infty$).

2. Асимптотический анализ

Введем обозначения:

$$\lambda = \frac{\hat{\lambda}}{\varepsilon}, \quad \text{где } \varepsilon \rightarrow 0, \quad u = \varepsilon w, \quad H(1, u) = F(1, w, \varepsilon), \quad H(0, u) = \varepsilon F(0, w, \varepsilon).$$

Тогда из уравнений (2)–(3) получим:

$$\left\{ \begin{array}{l} -j\sigma \frac{\partial F(0, w, \varepsilon)}{\partial w} = -\left(\frac{\hat{\lambda}}{\varepsilon} + \gamma\right) \varepsilon F(0, w, \varepsilon) + \mu F(1, w, \varepsilon) + \gamma, \\ j\sigma e^{-j\varepsilon w} \frac{\partial F(0, w, \varepsilon)}{\partial w} = -\left(\frac{\hat{\lambda}}{\varepsilon} + \mu + \gamma\right) F(1, w, \varepsilon) + \hat{\lambda} F(0, w, \varepsilon) + \\ \quad + \frac{\hat{\lambda}}{\varepsilon} e^{j\varepsilon w} F(1, w, \varepsilon), \\ j\sigma (e^{-j\varepsilon w} - 1) \frac{\partial F(0, w, \varepsilon)}{\partial w} = \gamma - \gamma(\varepsilon F(0, w, \varepsilon) + F(1, w, \varepsilon)) + \\ \quad + \frac{\hat{\lambda}}{\varepsilon} (e^{j\varepsilon w} - 1) F(1, w, \varepsilon). \end{array} \right. \quad (4)$$

Выполним предельный переход $\varepsilon \rightarrow 0$ в (4):

$$\left\{ \begin{array}{l} -j\sigma \frac{\partial F(0, w)}{\partial w} = -\hat{\lambda} F(0, w) + \mu F(1, w) + \gamma, \\ j\sigma \frac{\partial F(0, w)}{\partial w} = -(\mu + \gamma) F(1, w) + \hat{\lambda} F(0, w) + \hat{\lambda}(jw)F(1, w), \\ 0 = \gamma - \gamma F(1, w) + \hat{\lambda} jw F(1, w). \end{array} \right. \quad (5)$$

Из последнего уравнения системы (5) выразим:

$$F(1, w) = \frac{\gamma}{\gamma - \hat{\lambda} jw}.$$

Тогда асимптотическая характеристическая функция числа заявок на орбите в условии большой загрузки может быть представлена функцией:

$$h(u) \approx F\left(1, \frac{u}{\varepsilon}\right) = \frac{\gamma}{\gamma - ju\lambda}.$$

Таким образом распределение вероятности числа заявок на орбите является экспоненциальным с параметром $\frac{\gamma}{\lambda}$.

Заключение

В работе был проведен асимптотический анализ многолинейной RQ-системы M/M/1 с катастрофами в условии высокой интенсивности входящего потока. В ходе исследования было получено, что распределение вероятностей числа заявок на орбите в рассмотренной модели экспоненциальное с параметром $\frac{\gamma}{\lambda}$.

СПИСОК ЛИТЕРАТУРЫ

1. *Falin G. I., Templeton J. G. C.* Retrial queues // London: Chapman and Hall. 1997. V. 7. № 1. P. 127–167.
2. *Artalejo J. R., Gomez-Corral A.* Retrial Queueing Systems. Berlin: Springer, 2008. 267 p.
3. *Gelenbe E.* Product-form queueing networks with negative and positive customers // Journal of Applied Probability. 1991. № 28. P. 656–663.
4. *Do T. V.* Bibliography on G-networks, negative customers and applications // Mathematical and Computer Modelling. 2011. V. 53. № 1-2. P. 205–212.
5. *Shin Y. W.* Multi-server retrial queue with negative customers and disasters // Queueing Syst. 2007. V. 55. P. 223–337.
6. *Moiseev A., Nazarov A. A.* Queueing network MAP-(GI/∞)K with high-rate arrivals // European Journal of Operational Research. 2016. V. 254. № 2. P. 161–168.

Мелошникова Наталья Павловна — аспирант кафедры Теории вероятностей и математической статистики Института прикладной математики и компьютерных наук Томского государственного университета. E-mail: *meloshnikovana@gmail.com*

ОБЛАСТЬ ПРИМЕНИМОСТИ АСИМПТОТИЧЕСКИХ РЕЗУЛЬТАТОВ ПРИ ИССЛЕДОВАНИИ МАРКИРОВАННОГО ММРР

С. В. Пауль, А. А. Назаров, И. Л. Лапатин, Х. А. Ву

*Национальный исследовательский Томский государственный университет,
г. Томск, Россия*

В работе рассматривается модель двумерного маркированного ММРР и решается задача построения аппроксимаций числа событий, наступивших в потоке за определенное время, которые строятся на основе реализации метода асимптотического анализа потока в предельном условии растущего времени наблюдения за потоком. Качество предельных распределений подтверждается путем их сравнения с допредельным распределением.

Ключевые слова: *маркированный ММРР, метод асимптотического анализа, аппроксимация распределения вероятностей*

Введение

Проектируя современные многомодальные интерфейсы, в которых одновременно поступают запросы разной природы, необходимо учитывать особенности процессов при их совместном взаимодействии. При этом, моделируя потоки разрозненной информации, важно выделять необходимые входящие сообщения. Для описания таких неоднородных процессов разработаны маркированные потоки [2], [3], [4], [5]. Ввиду сложной структуры таких потоков исследователи применяют методы имитационного и численного моделирования, реализуя которые часто сталкиваются с проблемами увеличения размерности решаемых систем. В работе описывается асимптотический метод, реализация которого для одно из классов маркированных потоков (ММРР) позволяет получить достаточно простые выражения для распределения вероятностей числа событий, наступивших в исследуемом потоке за определенное время. Область применимости асимптотических результатов определяется путем сравнения с допредельными распределениями, которые удается записать в интегральном виде.

1. Математическая модель

Двумерный маркированный ММРР задан генератором \mathbf{Q} инфинитезимальных характеристик управляющей цепи Маркова $m(t)$ с непрерывным временем; матрицами $\mathbf{\Lambda}^{(1)}$ и $\mathbf{\Lambda}^{(2)}$ интенсивностей $\lambda_m^{(1)}$ и $\lambda_m^{(2)}$ наступления событий первого и второго типов в m -ом состоянии. Состояния двумерного ММРР совпадают со значениями, которые принимает управляющая цепь Маркова $m(t)$ в момент времени t . Когда двумерный маркированный ММРР перейдет в состояние v в момент времени t_m , он пробудет в нем до момента t_{m+1} согласно генератору инфинитезимальных характеристик \mathbf{Q} . В течение этого времени будут наступать события первого и второго типов согласно диагональным матрицам интенсивностей $\mathbf{\Lambda}^{(1)}$ и $\mathbf{\Lambda}^{(2)}$ соответственно. Далее в момент времени t_{m+1} маркированный ММРР перейдет в некоторое другое состояние и процедура повторится. Обозначим процессы $n_1(t)$ и $n_2(t)$ – число событий первого и второго типов, наступивших в исследуемом потоке за время t .

Определим трехмерный марковский процесс $\{n_1(t), n_2(t), m(t)\}$, для распределения вероятностей

$$P_m(n_1, n_2, t) = P\{n_1(t) = n_1, n_2(t) = n_2, m(t) = m\} \quad (1)$$

составим систему дифференциальных уравнений Колмогорова

$$\begin{aligned} \frac{\partial P_m(n_1, n_2, t)}{\partial t} = & - \left(\lambda_m^{(1)} + \lambda_m^{(2)} \right) P_m(n_1, n_2, t) + \\ & + \lambda_m^{(1)} P_m(n_1 - 1, n_2, t) + \lambda_m^{(2)} P_m(n_1, n_2 - 1, t) + \sum_{v=0}^M P_v(n_1, n_2, t) q_{vm}. \end{aligned} \quad (2)$$

Решая данную систему, получим интегральные формулы для вычисления маргинальных распределений вероятностей числа событий первого и второго типов, наступивших в ММРР.

Теорема 1. Распределения вероятностей числа событий первого и второго типов, наступивших в ММРР за время t имеют вид

$$\begin{aligned} P(n_l, t) = & \frac{1}{2\pi} \times \\ & \times \int_{-\infty}^{\infty} e^{-j\alpha t} \mathbf{r} \left(\left(\mathbf{\Lambda}^{(l)} - \mathbf{Q} - j\alpha \mathbf{I} \right)^{-1} \mathbf{\Lambda}^{(l)} \right)^{n_l} \left(\mathbf{\Lambda}^{(l)} - \mathbf{Q} - j\alpha \mathbf{I} \right)^{-1} \mathbf{e} d\alpha, \end{aligned} \quad (3)$$

где $l = 1, 2$, $j = \sqrt{-1}$, $\mathbf{r} = [r_m]$, $m = 0, \dots, M$ – вектор стационарных вероятностей состояний цепи Маркова $m(t)$, определяемый системой линейных алгебраических уравнений и условием нормировки $\mathbf{r}\mathbf{Q} = 0$, $\mathbf{r}\mathbf{e} = 1$.

Нахождение значений полученных распределений (3) требует затрат машинного времени, которое не всегда приводит к определенным результатам. Поэтому получим простые выражения для предельных распределений вероятностей числа событий, наступивших за время t в ММРР. Реализуя метод асимптотического анализа, систему (2) перепишем для частичных характеристических функций

$$H_m(u_1, u_2, t) = \sum_{n_1=0}^{\infty} \sum_{n_2=0}^{\infty} e^{ju_1 n_1} e^{ju_2 n_2} P_m(n_1, n_2, t), \quad (4)$$

Введем следующие обозначения

$$\begin{aligned} \mathbf{H}(u_1, u_2, t) &= \{H_1(n_1, n_2, u, t), \dots, H_M(n_1, n_2, u, t)\}, \\ \frac{\partial \mathbf{H}(u_1, u_2, t)}{\partial t} &= \frac{\partial H_1(u_1, u_2, t)}{\partial t}, \dots, \frac{\partial H_M(u_1, u_2, t)}{\partial t}, \end{aligned} \quad (5)$$

запишем задачу Коши в матричном виде

$$\begin{aligned} \frac{\partial \mathbf{H}(u_1, u_2, t)}{\partial t} &= \mathbf{H}(u_1, u_2, t) \{ \mathbf{Q} + (e^{ju_1} - 1) \mathbf{\Lambda}^{(1)} + (e^{ju_2} - 1) \mathbf{\Lambda}^{(2)} \}, \\ \mathbf{H}(u_1, u_2, 0) &= \mathbf{r}. \end{aligned} \quad (6)$$

Последняя система будет основной в дальнейших исследованиях.

2. Асимптотический анализ числа заявок, наступивших в ММРР в предельном условии растущего времени наблюдения

Рассмотрим систему (8) в предельном условии растущего времени наблюдения за потоком, которое характеризуется заменой $t = \tau T$, где T – неограниченно возрастающая положительная величина. Это равенство определяет зависимость времени t от «медленного времени» τ , так как для любого фиксированного значения $\tau > 0$ значение t неограниченно возрастает.

Метод асимптотического анализа реализуется в два этапа. На первом этапе вводится обозначение $1/T = \varepsilon$, где ε – малый положительный параметр, в задаче (8) выполняются замены $t\varepsilon = \tau$, $u_1 = \varepsilon w_1$, $u_2 = \varepsilon w_2$, $\mathbf{H}(u_1, u_2, t) = \mathbf{F}(w_1, w_2, \tau, \varepsilon)$. Первая замена определяет асимптотическое условие растущего времени наблюдения за двумерным маркированным ММРР, при выполнении которого число событий, наступивших за время t , также неограниченно возрастает. Для компенсации этого неограниченного возрастания выполняется вторая замена, которая эквивалентна рассмотрению класса процессов $\varepsilon n_1(t) = \varepsilon n_1\left(\frac{\tau}{\varepsilon}\right)$, $\varepsilon n_2(t) =$

$\varepsilon n_2\left(\frac{\tau}{\varepsilon}\right)$, для которых можно ожидать сходимости при $\varepsilon \rightarrow 0$. Реализуя метод асимптотического анализа на первом этапе строится предельная характеристическая функция, которую мы будем называть асимптотической первого порядка числа событий первого и второго типов, наступивших в двумерном маркированном ММРР.

Теорема 2. Предельная характеристическая функция $h_1(u_1, u_2, t)$ числа событий первого и второго типа сообщений, наступивших в двумерном маркированном ММРР за время t имеет вид

$$h_1(u_1, u_2, t) = \exp \left\{ ju_1 \varkappa_1^{(1)} Tt + ju_2 \varkappa_1^{(2)} Tt \right\}. \quad (7)$$

Здесь величины $\varkappa_1^{(1)}$ и $\varkappa_1^{(2)}$ определяются равенствами

$$\varkappa_1^{(1)} = \mathbf{r} \mathbf{\Lambda}^{(1)} \mathbf{e}, \quad \varkappa_1^{(2)} = \mathbf{r} \mathbf{\Lambda}^{(2)} \mathbf{e}. \quad (8)$$

Очевидно, асимптотика первого порядка $h_1(u_1, u_2, t)$ определяет средние значения $\varkappa_1^{(1)} t$ и $\varkappa_1^{(2)} t$ числа событий первого и второго типов, наступивших в потоке за время t .

Для построения предельного распределения вероятностей реализуется второй этап метода асимптотического анализа. Систему (8) перепишем с учетом замены

$$\mathbf{H}(u_1, u_2, t) = \exp \left\{ ju_1 \varkappa_1^{(1)} Tt + ju_2 \varkappa_1^{(2)} Tt \right\} \mathbf{H}_2(u_1, u_2, t).$$

Введем обозначения $1/T = \varepsilon^2$, и сделаем замены $t\varepsilon^2 = \tau$, $u_1 = \varepsilon w_1$, $u_2 = \varepsilon w_2$, $\mathbf{H}_2(u_1, u_2, t) = \mathbf{F}_2(w_1, w_2, \tau, \varepsilon)$. Аналогично асимптотике первого порядка первая замена определяет асимптотическое условие растущего времени наблюдения за потоком, при выполнении которого возрастают не только значения процессов $n_1(t)$ и $n_2(t)$, но также возрастают и значения центрированных процессов $n_1(t) - \varkappa_1^{(1)} t$ и $n_2(t) - \varkappa_1^{(2)} t$. Для компенсации этого возрастания выполняется вторая замена, которая эквивалентна рассмотрению класса процессов

$$\varepsilon \left(n_l(t) - \varkappa_1^{(l)} t \right) = \varepsilon \left(n_l \left(\frac{\tau}{\varepsilon^2} \right) - \varkappa_1^{(l)} \frac{\tau}{\varepsilon^2} \right), \quad l = 1, 2,$$

для которых можно ожидать сходимость при $\varepsilon \rightarrow 0$. Результатом применения метода асимптотического анализа на втором этапе будет получена предельная характеристическая функция числа событий, наступивших в исследуемом потоке, которую будем называть асимптотикой второго порядка. Отметим, что асимптотики первого и второго порядков

рассматриваемого метода асимптотического анализа аналогичны закону больших чисел и центральной предельной теореме теории вероятностей.

Теорема 3. Предельная характеристическая функция $h_2(u_1, u_2, t)$ числа событий первого и второго типа сообщений, наступивших в двумерном маркированном ММРР за время t имеет вид

$$h_2(u_1, u_2, t) = \exp \left\{ ju_1 \varkappa_1^{(1)} Tt + ju_2 \varkappa_1^{(2)} Tt + \frac{(ju_1)^2}{2} \varkappa_2^{(1)} Tt + \frac{(ju_2)^2}{2} \varkappa_2^{(2)} Tt + ju_1 \cdot ju_2 \cdot \varkappa_{12} Tt \right\}. \quad (9)$$

Здесь величины $\varkappa_1^{(1)}$ и $\varkappa_1^{(2)}$ определяются равенствами (10), величины $\varkappa_2^{(1)}$, $\varkappa_2^{(2)}$ и \varkappa_{12} – равенствами

$$\begin{aligned} \varkappa_2^{(1)} &= 2\mathbf{g}_1 \mathbf{\Lambda}^{(1)} \mathbf{e} + \varkappa_1^{(1)}, \\ \varkappa_2^{(2)} &= 2\mathbf{g}_2 \mathbf{\Lambda}^{(2)} \mathbf{e} + \varkappa_1^{(2)}, \\ \varkappa_{12} &= \mathbf{g}_1 \mathbf{\Lambda}^{(2)} \mathbf{e} + \mathbf{g}_2 \mathbf{\Lambda}^{(1)} \mathbf{e}, \end{aligned} \quad (10)$$

векторы \mathbf{g}_1 и \mathbf{g}_2 являются решениями систем линейных алгебраических уравнений

$$\mathbf{g}_l \mathbf{Q} = \mathbf{r} \left(\varkappa_1^{(l)} \mathbf{I} - \mathbf{\Lambda}^{(l)} \right), \quad \mathbf{g}_l \mathbf{e} = 0, \quad l = 1, 2. \quad (11)$$

Получили, что предельное распределение вероятностей числа событий, наступивших в двумерном маркированном ММРР, можно аппроксимировать двумерным гауссовским распределением с асимптотическими средними $\varkappa_1^{(1)} Tt$ и $\varkappa_1^{(2)} Tt$, дисперсиями $\varkappa_2^{(1)} Tt$ и $\varkappa_2^{(2)} Tt$ и ковариацией $\varkappa_{12} Tt$ числа событий первого и второго типов, которые насупили в потоке за время t .

На основе двумерного распределения не сложно получить маргинальные распределения вероятностей числа событий первого и второго типов, аппроксимации которых строятся по формуле

$$P(n_l) = \frac{p(n_l)}{\sum_{n_l=0}^{\infty} p(n_l)}, \quad l = 1, 2, \quad (12)$$

где $p(x_1, x_2)$ – предельная одномерная гауссовская плотность распределения вероятностей числа событий l -го типа.

3. Численный пример

Для модели двумерного маркированного ММРР зададим матрицы

$$\mathbf{Q} = \begin{pmatrix} -1 & 0,4 & 0,6 \\ 0,2 & -0,5 & 0,3 \\ 0,3 & 0,4 & -0,7 \end{pmatrix}, \quad T = 30, \quad t = 1,$$

$$\mathbf{\Lambda}^{(1)} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}, \quad \mathbf{\Lambda}^{(2)} = \begin{pmatrix} 0,7 & 0 & 0 \\ 0 & 0,8 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

Асимптотические характеристики принимают следующие значения $\varkappa_1^{(1)}Tt = 64,87$, $\varkappa_1^{(2)}Tt = 36,33$, $\varkappa_2^{(1)}Tt = 90,04$, $\varkappa_2^{(2)}Tt = 55,71$ и $\varkappa_{12}Tt = 16,27$. Графики, иллюстрирующие сравнение построенных аппроксимаций $P1as(n)$, $P2as(n)$ распределения вероятностей числа заявок двух типов с допредельными распределениями $P1dop(n)$, $P2dop(n)$ представлены на рисунке 1.

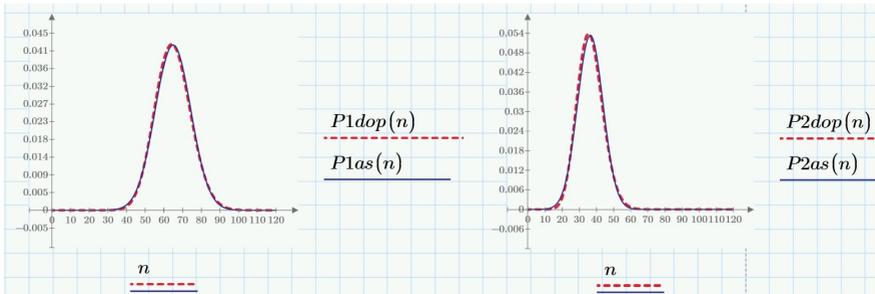


Рис. 1. Сравнение аппроксимаций распределений вероятностей числа событий двух типов с допредельными распределениями в двумерном маркированном ММРР

Полученные численные результаты говорят о достаточно высокой точности предложенной гауссовской аппроксимации. С увеличением значений времени t точность гауссовской аппроксимации естественно повышается в силу предельного условия. При малых параметрах t рекомендуется пользоваться допредельными интегральными формулами для нахождения маргинальных распределений вероятностей числа событий в двумерном маркированном ММРР в виде (3). При достаточно больших значениях времени t реализация численных алгоритмов нахождения распределения вероятностей становится практически невозможным, и в этом случае целесообразно применение асимптотических результатов.

Заключение

В работе выполнено исследование двумерного маркированного ММРР. Найдены распределения вероятностей числа событий двух типов, наступивших в потоке за время t . Эти распределения найдены как в допредельной ситуации в виде достаточно сложных выражений, численная реализация которых требует значительных затрат машинного времени, так и в предельном условии растущего времени наблюдения за потоком. Показано, что асимптотические распределения вероятностей являются гауссовскими и определяются несложными выражениями.

СПИСОК ЛИТЕРАТУРЫ

1. Вишнеvский В. М., Дудин А. Н., Клименок В. И. Стохастические системы с коррелированными потоками. Теория и применение в телекоммуникационных сетях. – М.: ТЕХНОСФЕРА. 2018. 564 с.
2. HE Q. M. Queues with marked customers // Adv. Appl. Probab. 1996. № 28. P. 567–587.
3. Dudin A. N., Dudin S. A.; Dudina O. S. Randomized Threshold Strategy for Providing Flexible Priority in Multi-Server Queueing System with a Marked Markov Arrival Process and Phase-Type Distribution of Service Time // Mathematics. 2023. № 11. P. 2669.
4. Naumov V., Gaidamaka Y., Yarkina N., Samouylov K. Matrix and Analytical Methods for Performance Analysis of Telecommunication Systems. Springer: Berlin/Heidelberg, Germany. 2021. 305 p.

Пауль Светлана Владимировна — доктор физико-математических наук, доцент, профессор, кафедра теории вероятностей и математической статистики. E-mail: paulsv82@mail.ru

Назаров Анатолий Андреевич — доктор технических наук, профессор, профессор, кафедра теории вероятностей и математической статистики. E-mail: nazarov.tsu@gmail.com

Лапатин Иван Леонидович — кандидат физико-математических наук, доцент, кафедра прикладной информатики. E-mail: ilapatin@mail.ru

Ву Хоанг Ань — студент бакалавриата института прикладной математики и компьютерных наук.

G-СЕТЬ С ДОХОДАМИ КАК МОДЕЛЬ КЛАСТЕРНОЙ СИСТЕМЫ

Т. В. Русилко, Д. А. Сальников

*Гродненский государственный университет имени Янки Купалы, г. Гродно,
Беларусь*

Статья посвящена исследованию и аналитическому моделированию кластерной системы. Замкнутая по структуре G-сеть массового обслуживания с запросами положительного и отрицательного классов применяется для построения математической модели кластерной системы при учете негативного воздействия вредоносного кода и ошибок на ее узлы. Генерируется множество доходов в зависимости от направления пакетов между узлами кластера. Получено дифференциальное уравнение для расчета дохода кластерной системы за время t при известном начальном состоянии. Целью работы является расчет дохода, генерируемого кластерной моделью в асимптотическом случае большого числа обрабатываемых запросов. **Ключевые слова:** *сеть массового обслуживания, G-сеть, сеть с доходами, метод асимптотического анализа, кластерная система.*

Введение

Системы, предназначенные для параллельной обработки данных, имеют широкое распространение. В основе таких систем лежит обработка поступающих задач. Эти задачи поступают на узлы системы, запрашивая ресурсы для обработки. Одним из средств моделирования таких систем являются сети массового обслуживания. В данной статье построена модель кластерной системы. Кластер — это связанный набор вычислительных систем (компьютеров), используемый в качестве объединенного вычислительного ресурса.

Марковский процесс служит математической моделью при исследовании сложных систем. Р. Ховард исследовал марковский процесс с непрерывным временем в предположении, что он получает вознаграждение (доход) в R_{ij} условных единиц (у.е.), когда система совершает переход из состояния i в состояние j , $i \neq j$, а также получает доход в размере R_i у.е. за единицу времени пребывания в состоянии i . Ховард назвал такой процесс «марковским процессом с вознаграждениями (доходами)» [1]. Концепция марковских процессов с доходами была

использована М. А. Маталыцким для определения сети массового обслуживания (СеМО) с доходами — НМ(Howard—Matalytski)-сети [2, 3].

G-сети — это обобщенные СеМО с несколькими типами запросов (заявок), в частности положительными и отрицательными заявками. G-сети впервые были представлены и изучены Э. Геленбе [4]. Область их применения — моделирование вычислительных систем и сетей, оценка их производительности, моделирование нейронных сетей и другие [5].

Целью данной работы является анализ доходов, генерируемых кластерной моделью. С этой целью производится асимптотический анализ замкнутой экспоненциальной G-сети с вознаграждениями с использованием аппроксимационного метода исследования СеМО при критическом предположении о большом, но ограниченном числе запросов [6, 7].

1. Описание модели. Постановка проблемы

Рассмотрим кластерную систему, которая включает в себя вычислительные устройства, выступающие в качестве отправителей, получателей и обработчиков данных. Устройства объединяются в сетевую структуру с помощью линий связи, таким образом в основе кластера лежит сеть передачи данных. В сети передача информации происходит посредством отдельных пакетов данных, при этом скорость передачи данных по каналам, процессинг и направление пакетов имеют свои ограничения. В кластерной системе могут передаваться полезные данные, потерявшие актуальность или поврежденные запросы, а также вредоносное программное обеспечение (ПО). Предполагается, что прием и обработка полезной информации приносит некоторый доход, а выполнение неактуальных задач и вредоносного кода приводит к потерям. В качестве модели кластерной системы применяется G-сеть с доходами.

В основе математической модели лежит замкнутая экспоненциальная G-сеть из конечного числа узлов S_0, S_1, \dots, S_n . Между узлами циркулирует фиксированное число K запросов, соответствующих передаваемым на обработку данным. Узел S_0 представляет собой IS-узел (Infinite Server). IS-узел играет роль «внешней среды» или конечного источника K запросов. Узел S_0 генерирует пуассоновский поток запросов с интенсивностью $\lambda_0 k_0$, где λ_0 — интенсивность потока, k_0 — количество запросов в узле S_0 . Входящий поток делится на поток положительных и отрицательных запросов. Обычные запросы относятся к положительному классу, а поврежденные, устаревшие или содержащие вредоносное ПО принадлежат к отрицательному классу. Вероятность прихода положительного запроса за короткий промежуток времени Δt равна

$\lambda_0 k_0 p_{0i}^+ \Delta t + o(\Delta t)$, в то время как вероятность прихода отрицательного запроса составляет $\lambda_0 k_0 p_{0i}^- \Delta t + o(\Delta t)$, $i = \overline{1, n}$, $\sum_{i=1}^n (p_{0i}^+ + p_{0i}^-) = 1$.

Каждый узел кластерной системы считается системой массового обслуживания S_i , $i = \overline{1, n}$, имеющей m_i линий обслуживания и бесконечный буфер. Вероятность завершения обслуживания положительного запроса в узле S_i за короткий интервал времени Δt равна $\mu_i \min(m_i, k_i) \Delta t + o(\Delta t)$, где k_i — число запросов в узле S_i ; вероятность завершения обслуживания двух и более запросов равна $o(\Delta t)$. События завершения обслуживания запросов в разных узлах сети являются взаимно независимыми. Запросы обслуживаются по правилу FIFO. Когда запрос завершает обслуживание в узле S_i , он мгновенно передается в узел S_j как положительный с вероятностью p_{ij}^+ или отрицательный с вероятностью p_{ij}^- , иначе он передается в IS-узел S_0 с вероятностью $p_{i0}^+ = 1 - \sum_{j=1}^n (p_{ij}^+ + p_{ij}^-)$, $i \neq j$, $i, j = \overline{1, n}$. Отрицательные запросы, прибывающие в узел, не обслуживаются. Отрицательный запрос, прибывший в узел S_i , $i = \overline{1, n}$, удаляет один положительный запрос из того же узла и оба они немедленно передаются в IS-узел S_0 .

Состояние исследуемой G-сети описывается n -мерным марковским процессом с непрерывным временем и конечным множеством состояний $\mathbf{k}(t) = (k_1(t), k_2(t), \dots, k_n(t))$, где $k_i(t)$ — количество запросов в узле S_i в момент t , $0 \leq k_i(t) \leq K$, $i = \overline{1, n}$, $t \in [0, +\infty)$, $k_0(t) = K - \sum_{i=1}^n k_i(t)$.

Предположим, что описанная выше G-сеть зарабатывает R_{ij}^+ у.е., когда положительный запрос переходит из узла S_i в узел S_j , а также R_{ij}^- у.е., когда отрицательный запрос совершает такой же переход, $i \neq j$, $i, j = \overline{0, n}$. Мы называем R_{ij}^+ и R_{ij}^- «доходами», связанными с переходом положительного и отрицательного запроса, соответственно, из S_i в S_j , $R(\mathbf{k})$ — доход в единицу времени пребывания процесса в состоянии \mathbf{k} . Интересует вопрос: каков будет ожидаемый доход G-сети через время t , $t \in T$, если сейчас она находится в состоянии \mathbf{k} .

2. Асимптотический анализ G-сети с доходами

Общая цель асимптотических методов в теории массового обслуживания — изучение процессов обслуживания в системах и сетях массового обслуживания путем нахождения для них подходящих приближений при некотором критическом (предельном) предположении.

Используемая в качестве модели замкнутая экспоненциальная G-сеть с доходами исследуется в асимптотическом случае большого числа

запросов K . Осуществлен предельный переход от марковской цепи $\mathbf{k}(t)$ к непрерывному марковскому процессу $\mathbf{s}(t)$. Математический подход, используемый в данной работе основан на дискретной модели непрерывного марковского процесса [8].

Пусть $V(\mathbf{k}, t)$ — общий ожидаемый доход G-сети, который сеть получит за время t при начальном состоянии \mathbf{k} . Для плотности дохода в случае асимптотической аппроксимации используется обозначение $v(\mathbf{x}, t)$, \mathbf{x} — начальное состояние, t — оставшееся время. Пусть $\theta(x)$ — функция Хевисайда, доопределенная при $x = 0$ значением 0. В дальнейшем используется методика исследования СеМО при $K \rightarrow \infty$. В таком случае процесс $\mathbf{s}(t) = \left(\frac{k_1(t)}{K}, \frac{k_2(t)}{K}, \dots, \frac{k_n(t)}{K} \right)$ задает состояние сети. Пусть r_{ij}^+ , r_{ij}^- и $r(\mathbf{x})$ — параметры дохода на единицу пространства состояний $\mathbf{s}(t)$

Теорема 1. В асимптотическом случае большого числа запросов K функция плотности вознаграждения $v(\mathbf{x}, t)$ при условии, что она дифференцируема по t и дважды непрерывно дифференцируема по x_i , $i = \overline{1, n}$, удовлетворяет с точностью $o(\varepsilon^2) = o(1/K^2)$, где $\varepsilon = 1/K$, обобщенному многомерному обратному уравнению Колмогорова:

$$\frac{\partial v(\mathbf{x}, t)}{\partial t} = - \sum_{i=1}^n A_i(\mathbf{x}, t) \frac{\partial v(\mathbf{x}, t)}{\partial x_i} + \frac{\varepsilon}{2} \sum_{i,j=1}^n B_{ij}(\mathbf{x}, t) \frac{\partial^2 v(\mathbf{x}, t)}{\partial x_i \partial x_j} + q(\mathbf{x}), \quad (1)$$

$$A_i(\mathbf{x}, t) = \sum_{j=1}^n \mu_j \min(l_j, x_j) (p_{ji}^- - p_{ji}^+ + \delta_{ji}) +$$

$$+ \mu_i \min(l_i, x_i) \sum_{j=1}^n p_{ij}^- (1 - \theta(x_j)) - \lambda \left(1 - \sum_{i=1}^n x_i \right) (p_{0i}^+ - p_{0i}^-),$$

$$B_{ii}(\mathbf{x}, t) = \sum_{j=1}^n \mu_j \min(l_j, x_j) (p_{ji}^+ + p_{ji}^- + \delta_{ji}) +$$

$$+ \mu_i \min(l_i, x_i) \sum_{j=1}^n p_{ij}^- (1 - \theta(x_j)) + \lambda \left(1 - \sum_{i=1}^n x_i \right) (p_{0i}^+ - p_{0i}^-),$$

$$B_{ij}(\mathbf{x}, t) = \mu_i \min(l_i, x_i) (p_{ij}^- - p_{ij}^+), \quad i \neq j,$$

$$q(\mathbf{x}) = r(\mathbf{x}) + K \left(\sum_{i,j=1}^n \mu_i \min(l_i, x_i) (p_{ij}^+ r_{ij}^+ + p_{ij}^- r_{ij}^- (2 - \theta(x_j))) + \right.$$

$$\left. + \sum_{i=1}^n \mu_i \min(l_i, x_i) p_{i0}^+ r_{i0}^+ + \sum_{i=1}^n \lambda \left(1 - \sum_{i=1}^n x_i \right) (p_{0i}^+ r_{0i}^+ + p_{0i}^- r_{0i}^-) \right),$$

где δ_{ji} — дельта Кронекера.

3. Математическая модель ожидаемого дохода

Решение уравнения (1) затруднительно. В связи с этим требуется ряд последующих приближений. Очевидно, что коэффициенты диффузии $B_{ij}(\mathbf{x}, t)$ уравнения (1) имеют порядок ε , тогда член $\frac{\varepsilon}{2} \sum_{i,j=1}^n B_{ij}(\mathbf{x}, t) \frac{\partial^2 v(\mathbf{x}, t)}{\partial x_i \partial x_j}$ в правой части (1) является $O(\varepsilon^2)$. Значит, с точностью до членов порядка $O(\varepsilon^2)$ плотность дохода задается уравнением:

$$\frac{\partial v(\mathbf{x}, t)}{\partial t} = - \sum_{i=1}^n A_i(\mathbf{x}, t) \frac{\partial v(\mathbf{x}, t)}{\partial x_i} + q(\mathbf{x}). \quad (2)$$

Интегрируя плотность $v(\mathbf{x}, t)$ в пределах n -мерной области D , $D \subseteq X$, получаем общий ожидаемый доход, который сеть заработает за время t при начальном состоянии $\mathbf{x} \in D$:

$$V_D(t) = \int \int \dots \int_D v(\mathbf{x}, t) d\mathbf{x}.$$

Применяя это преобразование к обеим частям (2), используя правила интегрирования, интегральное правило Лейбница, линейность коэффициентов $A_i(\mathbf{x}, t)$ по \mathbf{x} и граничное условие $A(\mathbf{x}, t)v(\mathbf{x}, t) = 0$, $\mathbf{x} \in \Gamma(D)$, где $\Gamma(D)$ — отражающая граница области D [8], мы получаем обыкновенное линейное дифференциальное уравнение первого порядка для ожидаемого вознаграждения $V_D(t)$:

$$\frac{d}{dt} V_D(t) = \sum_{i=1}^n \frac{\partial A_i(\mathbf{x}, t)}{\partial x_i} \cdot V_D(t) + \int \int \dots \int_D q(\mathbf{x}) d\mathbf{x}. \quad (3)$$

Уравнение (3) — математическая модель для расчета ожидаемого дохода кластерной системы $V_D(t)$ при известном $V_D(0)$.

По аналогии с центром тяжести материального тела D в физике мы можем найти точку равновесия ожидаемого дохода, когда $\mathbf{x} \in D$:

$$E_i^D(t) = \frac{1}{V_D(t)} \int \int \dots \int_D x_i v(\mathbf{x}, t) d\mathbf{x}, \quad i = \overline{1, n}.$$

Заключение

В данной работе представлена аналитическая модель кластерной системы. Анализ ее доходности произведен на основе замкнутой экспо-

ненциальной G-сети с доходами. Последовательность доходов, создаваемая сетью, возникает за счет передачи запросов между узлами. Представлены результаты математического моделирования, которые были использованы для прогнозирования общего ожидаемого дохода G-сети как функции оставшегося времени t , когда известно начальное состояние сети.

СПИСОК ЛИТЕРАТУРЫ

1. *Howard R. A.* Dynamic programming and Markov processes. Cambridge: Technology Press of Massachusetts Institute of Technology, 1960. 136 p.
2. *Маталыцкий М. А., Паньков А. В.* Вероятностный анализ доходов в банковских сетях // Вестник БГУ. Серия 1. Физика, математика, информатика. 2004. № 9. С. 79–92.
3. *Matalytski M., Pankov A.* Analysis of the stochastic model of changing of incomes in the open banking network // Computer Science. 2003. Vol. 3. № 5. P. 19–29.
4. *Gelenbe E* Product form queueing networks with negative and positive customers // Journal of Applied Probability. 1991. Vol. 28. P. 656–663.
5. *Caglayan M. U.* G-networks and their applications to machine learning, energy packet networks and routing: introduction to the special issue // Probability in the Engineering and Informational Sciences. 2017. Vol. 31. № 4. P. 381–395.
6. *Rusilko T.* Asymptotic analysis of a closed G-network of unreliable nodes // Journal of Applied Mathematics and Computational Mechanics. 2022. Vol. 21. № 2. P. 91–102.
7. *Rusilko T. V.* The G-network as a stochastic data network model // Journal of the Belarusian State University. Mathematics and Informatics. 2023. № 2. P. 45–54.
8. *Тихонов В. А., Миронов М. А.* Марковские процессы. Москва: Советское радио, 1977. 488 с.

Русилко Татьяна Владимировна — кандидат физико-математических наук, доцент, кафедра фундаментальной и прикладной математики, Гродненский государственный университет имени Янки Купалы. E-mail: tatiana.rusilko@gmail.com

Сальников Дмитрий Александрович — магистрант, кафедра математического и информационного обеспечения экономических систем, Гродненский государственный университет имени Янки Купалы. E-mail: dima.saln.gr@gmail.com

АСИМПТОТИЧЕСКИЙ АНАЛИЗ СИСТЕМЫ $M(T)|M(T)|\infty$

В. А. Вавилов, А. А. Назаров

Томский государственный университет, Томск, Российская Федерация

В работе предлагается математическая модель системы массового обслуживания с неограниченным числом приборов, высокоинтенсивным нестационарным входящим потоком и параметрами обслуживания, зависящими от времени. Для распределения вероятностей числа заявок в системе составлена система дифференциальных уравнений Колмогорова. Выполнен асимптотический анализ полученной системы в условиях высокой интенсивности входящего потока. Показано, что предельное распределение числа занятых приборов является нестационарным гауссовским. В допредельном исследовании аналитически получено, что распределение числа занятых приборов является пуассоновским. Расстояние Колмогорова между гауссовской аппроксимацией и пуассоновским распределением является вполне приемлемым. **Keywords:** *система массового обслуживания с неограниченным числом приборов, высокоинтенсивный входящий поток, гауссовская аппроксимация, пуассоновское распределение.*

Введение

В настоящее время в связи с развитием цифровых и нейросетевых технологий актуальным становится вопрос исследования систем обслуживания с параметрами, изменяющимися под воздействием тех или иных факторов. Так, например, в работе [1] изучены системы обслуживания, параметры которых зависят от значений управляющих случайных процессов. В работах [2-3] представлены системы, параметры функционирования которых - есть периодические функции, зависящие от времени. В работе [4] можно найти исследование бесконечнолинейной системы массового обслуживания с нестационарным пуассоновским потоком, когда интенсивность входящего потока зависит от времени. В работе [5] представлено исследование различных моделей бесконечнолинейных систем обслуживания асимптотическим методом. В данной работе рассмотрим систему обслуживания с неограниченным числом приборов, нестационарным высокоинтенсивным пуассоновским входя-

щим потоком и параметрами обслуживания, описываемыми периодическими функциями.

1. Описание математической модели

Рассмотрим систему массового обслуживания с неограниченным числом приборов, на вход которой поступает нестационарный пуассоновский [6] с параметром $N\lambda(t)$ поток заявок, где N – некоторый параметр, принимающий достаточно большие значения. Заявка, поступившая в систему, занимает свободный прибор и начинает немедленно обслуживаться с интенсивностью $\mu(t)$. По завершении обслуживания заявка покидает систему. Число заявок в системе обозначим $i(t)$.

Обозначим $P\{i(t) = i\} = P(i, t)$. В любой момент времени должно выполняться условие нормировки $\sum_{i=0}^{\infty} P(i, t) = 1$. Для распределения вероятностей $P(i, t)$ можно составить систему дифференциальных уравнений Колмогорова:

$$\frac{\partial P(0, t)}{\partial t} = -N\lambda(t)P(0, t) + \mu(t)P(1, t),$$

$$\begin{aligned} \frac{\partial P(i, t)}{\partial t} = & -(N\lambda(t) + i\mu(t))P(i, t) + N\lambda(t)P(i-1, t) + \\ & + (i+1)\mu(t)P(i+1, t). \end{aligned} \quad (1)$$

Для исследования системы (1) перейдём от вероятностей $P(i, t)$ к характеристической функции:

$$H(u, t) = \sum_{i=0}^{\infty} e^{jui} P(i, t),$$

где $j = \sqrt{-1}$. Затем поделим левую и правую части на N , система (1) примет вид:

$$\frac{1}{N} \frac{\partial H(u, t)}{\partial t} = (e^{ju} - 1) \left(\lambda(t)H(u, t) + je^{-ju} \frac{\mu(t)}{N} \frac{\partial H(u, t)}{\partial u} \right). \quad (2)$$

2. Асимптотика первого порядка

В системе вида (2) выполним следующие замены: $1/N = \varepsilon, u = \varepsilon w, H(u, t) = F_1(w, t, \varepsilon)$, в результате имеем:

$$\varepsilon \frac{\partial F_1(w, t, \varepsilon)}{\partial t} = (e^{j\varepsilon w} - 1) \left(\lambda(t)F_1(w, t, \varepsilon) + je^{-j\varepsilon w} \mu(t) \frac{\partial F_1(w, t, \varepsilon)}{\partial w} \right). \quad (3)$$

Обозначим: $\lim_{\varepsilon \rightarrow 0} F_1(w, t, \varepsilon) = F_1(w, t)$.

Используем разложение функции $e^{j\varepsilon w}$ в ряд Тейлора с точностью до $O(\varepsilon^2)$, поделим левую и правую части (3) на ε , выполним предельный переход при $\varepsilon \rightarrow 0$, получим:

$$\frac{\partial F_1(w, t)}{\partial t} = jw \left(\lambda(t)F_1(w, t) + j\mu(t) \frac{\partial F_1(w, t)}{\partial w} \right). \quad (4)$$

В (4) выполним замену $F_1(w, t) = e^{jwx(t)}$, будем иметь:

$$x'(t) = \lambda(t) - \mu(t)x(t). \quad (5)$$

Получили линейное неоднородное дифференциальное уравнение первого порядка [7]. Полагаем, что в начальный момент времени система свободна, поэтому начальное условие имеет вид $x(0) = 0$, а решение (5) - следующий вид:

$$x(t) = \exp\left(-\int_0^t \mu(s)ds\right) \left(x(0) + \int_0^t \lambda(y) \exp\left(\int_0^y \mu(s)ds\right) dy \right). \quad (6)$$

3. Асимптотика второго порядка

Обозначим правую часть дифференциального уравнения (5) как $a(t)$. Вернёмся к переменной u , тогда $e^{jwx(t)} = e^{jNux(t)}$. В уравнении (2) выполним замену $H(u, t) = H_2(u, t)e^{jNux(t)}$ и запишем его в виде:

$$\begin{aligned} & \frac{1}{N} \frac{\partial H_2(u, t)}{\partial t} + jua(t)H_2(u, t) = (e^{ju} - 1) \times \\ & \times \left((\lambda(t) - e^{-ju}x(t)\mu(t))H_2(u, t) + je^{-ju} \frac{\mu(t)}{N} \frac{\partial H_2(u, t)}{\partial u} \right). \end{aligned} \quad (7)$$

В (7) выполним замены: $1/N = \varepsilon^2$, $u = \varepsilon w$, $H_2(u, t) = F_2(w, t, \varepsilon)$, получим:

$$\begin{aligned} & \varepsilon^2 \frac{\partial F_2(w, t, \varepsilon)}{\partial t} + j\varepsilon wa(t)F_2(w, t, \varepsilon) = (e^{j\varepsilon w} - 1) \times \\ & \times \left((\lambda(t) - e^{-j\varepsilon w}x(t)\mu(t))F_2(w, t, \varepsilon) + j\varepsilon e^{-j\varepsilon w} \mu(t) \frac{\partial F_2(w, t, \varepsilon)}{\partial w} \right). \end{aligned} \quad (8)$$

Разложим $e^{j\varepsilon w}$ и $e^{-j\varepsilon w}$ в ряды Тейлора с точностью до $O(\varepsilon^3)$, подставим эти разложения в (8), выполним преобразования, обозначим $\lim_{\varepsilon \rightarrow 0} F_2(w, t, \varepsilon) = F_2(w, t)$, получим:

$$\frac{\partial F_2(w, t)}{\partial t} = -w\mu(t) \frac{\partial F_2(w, t)}{\partial w} - \frac{w^2}{2} (\lambda(t) + \mu(t)x(t))F_2(w, t). \quad (9)$$

Решение (9), то есть функцию $F_2(w, t)$ будем искать в виде:

$$F_2(w, t) = \exp\left(\frac{(jw)^2}{2}K(t)\right). \quad (10)$$

Функция вида (10) - есть характеристическая функция нормально-го распределения с нулевым математическим ожиданием и дисперсией $K(t)$. Подставим (10) в (9), выполним преобразования, получим:

$$K'(t) = -2\mu(t)K(t) + \lambda(t) + \mu(t)x(t). \quad (11)$$

Решив обыкновенное неоднородное дифференциальное уравнение (11), получим выражение для функции $K(t)$ в явном виде:

$$K(t) = \exp\left(-2 \int_0^t \mu(s)ds\right) \times \\ \times \left(K(0) + \int_0^t (\lambda(y) + \mu(y)x(y))\exp\left(2 \int_0^y \mu(s)ds\right) dy\right). \quad (12)$$

Таким образом, предельное распределение числа занятых приборов является нестационарным гауссовским с математическим ожиданием $Nx(t)$ и дисперсией $NK(t)$.

4. Допредельное исследование

Решение уравнения (2) будем искать в виде $H(u, t) = \exp(e^{ju-1}\rho(t))$. После несложных преобразований имеем линейное неоднородное дифференциальное уравнение первого порядка:

$$\rho'(t) = N\lambda(t) - \mu(t)\rho(t). \quad (13)$$

Решение (13) при условии $\rho(0) = 0$ имеет вид:

$$\rho(t) = \exp\left(-\int_0^t \mu(s)ds\right) \times \\ \times \left(\rho(0) + N \int_0^t \lambda(y)\exp\left(\int_0^y \mu(s)ds\right) dy\right). \quad (14)$$

5. Численный анализ

Пусть $\lambda(t) = 5 + \sin(2\pi t)$, $\mu(t) = 2 - \cos(2\pi t)$, $N = 10$, $\rho(0) = 0$, $x(0) = 0$, $K(0) = 0$. На рис. 1 изображены графики математического ожидания $\rho(t)$, полученного в допредельном случае, а также асимптотического среднего $Nx(t)$ и дисперсии $NK(t)$ числа заявок в системе.

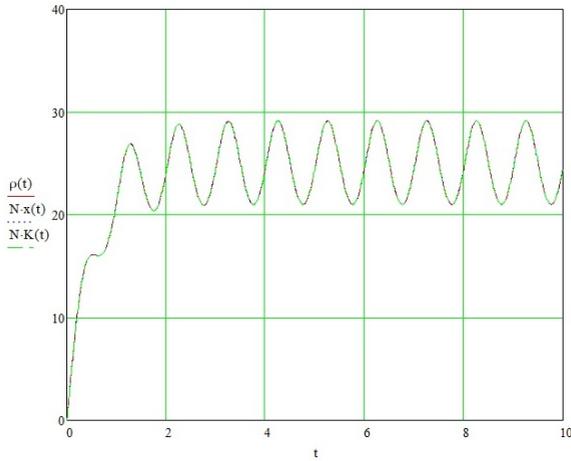


Рис. 1. Математическое ожидание, асимптотическое среднее и дисперсия

Пусть $t = 10$, $n = 50$. Распределение Пуассона $P_i, i = 1, \dots, n$ с параметром $\rho(10) = 24,461$ изображено на рис. 2 сплошной линией. Гауссовская аппроксимация $PA_i, i = 1, \dots, n$ изображена пунктиром. Расстояние Колмогорова составляет $\Delta = 0.013$.

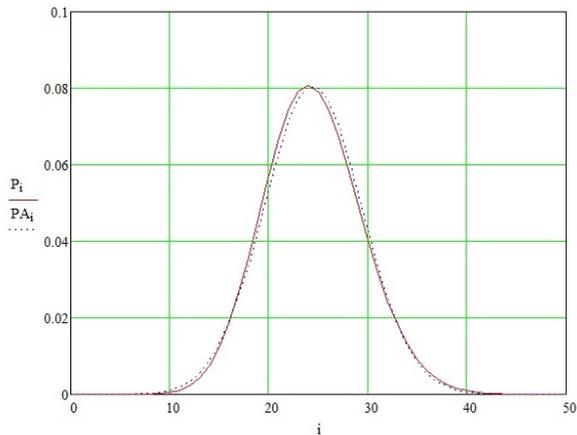


Рис. 2. Допредельное и асимптотическое распределения вероятностей числа заявок в системе

6. Заключение

В работе представлено исследование математической модели бесконечнолинейной системы с нестационарным входящим потоком и параметрами обслуживания, зависящими от времени. Показано, что в условиях высокой интенсивности $N \rightarrow \infty$ входящего потока предельное распределение числа занятых приборов является нестационарным гауссовским с математическим ожиданием $Nx(t)$ и дисперсией $NK(t)$, где $x(t)$ и $K(t)$ определяются выражениями (6) и (12) соответственно. В допредельном исследовании аналитически получено, что распределение числа заявок в системе является пуассоновским с параметром $\rho(t)$, определяемым выражением (14). По результатам численного анализа расстояние Колмогорова между допредельным и асимптотическим распределениями не превышает величины 0.05 при $N \geq 2$.

СПИСОК ЛИТЕРАТУРЫ

1. *Коротаев И. А.* Системы массового обслуживания с переменными параметрами. Томск: Изд-во Том. ун-та, 1991. 167 с.
2. *Granovsky B. L., Zeifman A. I.* Nonstationary Markovian queues // *Math. Sci.* 99, 2000. С. 1415–1438.
3. *Ковалёв И. А., Сатин Я. А., Ситникова А. В., Зейфман А. И.* Об одном подходе к оцениванию скорости сходимости нестационарных марковских моделей систем обслуживания // *Информатика и ее применения.* 2022. Т. 16. № 3. С. 75–82.
4. *Назаров А. А., Терпугов А. Ф.* Теория массового обслуживания: учебное пособие. 2-е изд., испр. Томск: Изд-во НТЛ, 2010. 228 с.
5. *Моисеев А. Н., Назаров А. А.* Бесконечнолинейные системы и сети массового обслуживания. Томск: Изд-во НТЛ, 2015. 240 с.
6. *Миллер Б. М., Панков А. Р.* Теория случайных процессов. М.: Физматлит, 2007. 317 с.
7. *Эльсгольц Л. Е.* Дифференциальные уравнения и вариационное исчисление. М.: Наука, 1969. 424 с.

Вавилов Вячеслав Анатольевич — кандидат физико-математических наук, доцент, доцент кафедры программной инженерии, институт прикладной математики и компьютерных наук. E-mail: vavilovv@yandex.ru

Назаров Анатолий Андреевич — доктор технических наук, профессор, профессор кафедры теории вероятностей и математической статистики, институт прикладной математики и компьютерных наук. E-mail: nazarov.tsu@gmail.com

ОБ УСЛОВИЯХ СУЩЕСТВОВАНИЯ СТАЦИОНАРНОГО РАСПРЕДЕЛЕНИЯ В СМО С ЦИКЛИЧЕСКИМ УПРАВЛЕНИЕМ ПЕРИОДИЧЕСКИМИ КОНФЛИКТНЫМИ ПОТОКАМИ

В. Л. Цодиков, А. В. Зорин

Национальный исследовательский Нижегородский государственный университет им. Н. И. Лобачевского, г. Нижний Новгород, Россия

В работе изучается система массового обслуживания с периодической интенсивностью конфликтных входных потоков, отражающая, в частности, функционирование системы управления транспортными потоками в классе циклических алгоритмов с фиксированным ритмом переключения и суточными колебаниями интенсивности транспортного потока. Рассматривается сужение цепи Маркова на класс существенных сообщаемых состояний, устанавливаются рекуррентные соотношения для производящих функций для распределения длины очереди, состояния обслуживающего устройства и времени суток. Использование итеративно-мажорантного подхода изучаются необходимые и достаточные условия существования стационарного распределения для длин очередей.

Ключевые слова: конфликтные потоки, нестационарный поток Пуассона, циклическое управление с фиксированным ритмом, счётная цепь Маркова, условия существования стационарного распределения.

Введение

В настоящее время разными исследованиями разрабатываются как детерминированные, так и стохастические модели образования транспортных потоков [1, 2, 3, 4, 5]. В этих и многих других работах по управлению дорожным движением не учитывается периодический характер изменения характеристик потоков, связанный с суточными, недельными и другими причинами. В то же время, в теории массового обслуживания известное число работ посвящено системам обслуживания с зависящими от времени и, в частности, с периодически изменяющимися законами для входного потока и промежутков обслуживания [6, 7, 8, 9, 10].

В докладе продолжается исследование системы обслуживания конфликтных потоков, начатое в [11, 12]. Кратко напомним постановку

задачи. В систему обслуживания поступают $m < \infty$ входных потоков Π_1, \dots, Π_m . Требования по потоку Π_j образуют нестационарный поток Пуассона, $\lambda_j(t)$ — мгновенная интенсивность потока Π_j в момент t . Требования потока Π_j помещаются в накопитель O_j неограниченной вместимости. Функция $\lambda_j(t)$ — периодическая с периодом $T_D > 0$. Величина может иметь смысл продолжительности суток. Обслуживающее устройство имеет $2m$ внутренних состояний $\Gamma^{(1)}, \dots, \Gamma^{(2m)}$, сменяющихся циклически. В состоянии вида $\Gamma^{(2j-1)}$ обслуживаются только требования потока Π_j . В состоянии вида $\Gamma^{(2j)}$ требования не обслуживаются. Время пребывания обслуживающего устройства в состоянии $\Gamma^{(r)}$ неслучайно и равно T_r . За время T_{2j-1} из очереди Π_j может обслужиться не более, чем ℓ_j требований. Следуя работе [11], обозначим через $\tau_0, \tau_1, \tau_2, \dots$ моменты смены состояния обслуживающего устройства, через Γ_i — состояние обслуживающего устройства на промежутке $[\tau_{i-1}, \tau_i)$, $i = 1, 2, \dots$; $\eta_{j,i}$ — число требований потока Π_j , поступивших на промежутке $[\tau_i, \tau_{i+1})$, $\varkappa_{j,i}$ — число требований в очереди O_j в момент τ_i ; $\xi_{j,i}$ — максимально возможное число требований из очереди O_j , которые можно обслужить на промежутке $[\tau_i, \tau_{i+1})$, а $\bar{\xi}_{j,i}$ — число фактически обслуженных требований из очереди O_j на промежутке $[\tau_i, \tau_{i+1})$. Для натурального числа r определим операцию $r \oplus 1$ равенствами $r \oplus 1 = r + 1$ для $r < 2m$ и $(2m) \oplus 1 = 1$. Введём величины τ'_i как остаток от деления τ_i на T_D . Величина τ'_i имеет смысл времени суток. Величины T_1, T_2, \dots, T_{2m} и T_D предполагаются соизмеримыми, кратными некоторому $\Delta > 0$. Тогда $\tau'_i \in \{0, \Delta, 2\Delta, \dots, T_D - \Delta\}$. Из соизмеримости следует, что существуют такие целые числа p, q , что

$$pT = qT_D, \tag{1}$$

где $T = T_1 + T_2 + \dots + T_{2m}$. Так же считаем известной функцию

$$\Lambda_j(x) = \int_0^x \lambda_j(t) dt, \quad j = 1, 2, \dots, m.$$

Введём m -мерную неотрицательную решётку

$$X = \{0, 1, \dots\} \times \{0, 1, \dots\} \times \dots \times \{0, 1, \dots\}.$$

В работе [11] была введена функция $\text{rem}(x, d)$, дающая остаток от деления x на d . Были определены для $j = 1, 2, \dots, 2m$ множества

$$G_j = \bigcup_{s=0}^{p-1} \{(\text{rem}(sT, T_D), \Gamma^{(j)}), (\text{rem}(sT + T_{j\oplus 1}, T_D), \Gamma^{(j\oplus 1)}), (\text{rem}(sT + T_{j\oplus 1} + T_{j\oplus 2}, T_D), \Gamma^{(j\oplus 2)}), \dots, (\text{rem}(sT + T_{j\oplus 1} + T_{j\oplus 2} + \dots + T_{j\oplus (2m-1)}, T_D), \Gamma^{(j\oplus (2m-1))})\}.$$

Каждое из данных множеств G_j , $j = 1, 2, \dots, 2m$ содержит одинаковое число $2mp$ элементов. Было доказано, что при заданном распределении вектора $(\tau'_0, \Gamma_0, \varkappa_0)$ последовательность

$$\{(\tau'_i, \Gamma_i, \varkappa_i); i = 0, 1, \dots\} \quad (2)$$

является многомерной однородной марковской цепью с пространством состояний вида $(G_1 \cup G_2 \cup \dots \cup G_{2m}) \times X$. Далее было установлено, что если все множества G_1, G_2, \dots, G_{2m} совпадают, то все состояния принадлежат одному классу существенных сообщающихся состояний. А если среди указанных множеств есть не совпадающие, то могут быть несколько классов существенных сообщающихся состояний. Кроме того, последовательность $\{(\tau'_i, \Gamma_i, \varkappa_{j,i}); i = 0, 1, \dots\}$ также является однородной цепью Маркова.

В работе [12] была приведена лемма, показывающая, что переходные вероятности в существенной части зависят только от разности между длинами очередей. Был применён метод цензурирования цепей и построен алгоритм вычисления стационарного распределения вероятностей этой цепи. Однако, в указанной работе отсутствовали условия существования стационарного распределения. Поэтому в настоящей работе мы уделяем внимание именно этому вопросу.

1. Основные результаты

Не уменьшая общности, будем рассматривать цепь Маркова (2) на множестве $G_1 \times X$, на котором она является неразложимой. Введём производящие функции:

$$\Psi_{j,i}(z; r, t) = \mathbf{M} \left(I(\Gamma_i = \Gamma^{(r)}, \tau'_i = t) \cdot z^{\varkappa_{j,i}} \right),$$

где $(\Gamma^{(r)}, t) \in G_1$, а также функцию $q_r(z, t)$:

$$q_r(z, t) = e^{(\Lambda_1(t+T_r) - \Lambda_1(t))(z-1)},$$

где r, t такие, что $(\Gamma^{(r)}, t) \in G_1$.

Лемма 1. Справедливы следующие рекуррентные соотношения для первого потока:

$$\Psi_{1,i+1}(z; r \oplus 1, \text{rem}(t + v(\Gamma_r), T_D)) = q_{r+1}(z, t) \cdot \Psi_{1,i}(z; r, t),$$

для $(\Gamma^{(r)}, t) \in G_1, r \neq 2m,$

$$\begin{aligned} \Psi_{1,i+1}(z; 1, \text{rem}(t + v(\Gamma_{2m}), T_D)) &= \sum_{x=0}^{\ell_1-1} \mathbf{P}(\Gamma_i = \Gamma^{(r)}, \varkappa_{1,i} = x, \tau'_i = t) \times \\ &\times \sum_{s=0}^{\ell_1-x-1} \frac{(\Lambda_1(t + T_1) - \Lambda_1(t))^s}{s!} e^{-(\Lambda_1(t+T_1) - \Lambda_1(t))} (1 - z^{x+s-\ell_1})_+ \\ &+ z^{-\ell_1} q_1(z, t) \Psi_{j,i}(z; r, t), \text{ для } (\Gamma^{(r)}, t) \in G_1, r = 2m. \end{aligned}$$

Аналогичные рекуррентные соотношения выводятся и для остальных очередей с номером $j = 2, \dots, m$.

Основные результаты работы приводятся в следующих теоремах. Для доказательства используется итеративно-мажорантный метод, см. [13, 14].

Теорема 1. Для существования стационарного распределения цепи Маркова $\{(\tau'_i, \Gamma_i, \varkappa_i); i = 0, 1, \dots\}$ необходимо выполнение неравенств

$$\Lambda_j(T_D) \cdot q < p \cdot \ell_j, j = 1, 2, \dots, m, \tag{3}$$

где участвуют параметры p и q из соотношения (1).

Неравенству (3) можно дать простую физическую интерпретацию: математическое ожидание числа поступивших требований за p полных циклов работы обслуживающего устройства должно быть строго меньше максимально возможного числа обслуженных требований по каждому из потоков.

Лемма 2. Если стационарного распределения не существует, то последовательность математических ожиданий

$$\{\mathbf{M}(\varkappa_i); i = 0, 1, \dots\}$$

неограниченно возрастает.

Теорема 2. Для существования стационарного распределения цепи Маркова $\{(\tau'_i, \Gamma_i, \varkappa_i); i = 0, 1, \dots\}$ достаточно выполнения неравенств

$$\Lambda_j(t + T) - \Lambda_j(t) < \ell_j, j = 1, \dots, m, \tag{4}$$

где $(\Gamma^{(r)}, t) \in G_1$.

Неравенству (4) тоже можно можно дать физическую интерпретацию: для каждого из потоков за любой полный цикл работы обслуживающего устройства должно поступить строго меньше требований, чем максимально возможное число обслуженных требований за цикл. Данное условие, по-видимому, является сильно ограничительным. В дальнейшем попробуем его ослабить.

Заключение

Таким образом, были получены рекуррентные соотношения для производящих функций, необходимое и достаточное условия существования стационарного распределения. Как видно, они не совпадают. Рассмотрение данной цепи Маркова будет продолжено.

СПИСОК ЛИТЕРАТУРЫ

1. *Гасников А. В. и др.* Введение в математическое моделирование транспортных потоков: Учебное пособие Под ред. А. В. Гасникова. Издание 2-е, испр. и доп. М.: МЦНМО, 2013. 429 с.
2. *Afanasyeva L., Bulinskaya E.* Stochastic Models of Transport Flows // Communications in Statistics — Theory and Methods. Vol. 40. P. 2830–2846.
3. *Федоткин М. А., Федоткин А. М., Кудрявцев Е. В.* Динамические модели неоднородного потока транспорта на магистралях // Автоматика и телемеханика. 2020. № 8. С. 149–164.
4. *Piccoli B. et al.* Vehicular Traffic: A Review of Continuum Mathematical Models // Encyclopedia of complexity and systems science. 2009. Vol. 22. P. 9727–9749.
5. *Jabari S. E., Liu H. X.* A stochastic model of traffic flow: Theoretical foundations Author links open overlay panel // Transportation Research Part B: Methodological. 2012. Vol. 46, Issue 1. P. 156–174.
6. *Afanasyeva L. G.* On periodic distribution of waiting-time process // Stability Problems for Stochastic Models. Lecture Notes in Mathematics. Vol. 1155. P. 1–20.
7. *Зейфман А.И., Бенцинг В.Е., Соколов И.А.* Марковские цепи и модели с непрерывным временем М.: Элекс-КМ, 2008. 167 с.
8. *Петрова О.В., Ушаков В.Г.* Асимптотический анализ системы массового обслуживания $E_r(t)|G|1$ // Информатика и ее применения. 2009. Т. 3, вып. 4. С. 35–40.
9. *Schwarz J. A., Selinka G., Stolletz R.* Performance analysis of time-dependent queueing systems: Survey and classification // Omega. 2016. Vol. 63. P. 170–189.

10. *Vishnevsky V., Vytovtov K., Varabanova E., Semenova O.* Analysis of a MAP/M/1/N Queue with Periodic and Non-Periodic Piecewise Constant Input Rate // *Mathematics*. 2022. Vol. 10, Iss. 10. P. 1684.
11. *Цодиков В. Л., Зорин А. В.* Построение математической модели процесса циклического управления периодическими конфликтными потоками // XXI Международная конференция имени А.Ф.Терпугова, Томск: Издательство Томского государственного университета, 2023. С. 215–219.
12. *Цодиков В. Л., Зорин А. В.* Метод вычисления стационарного распределения СМО с циклическим управлением периодическими конфликтными потоками // XXII Международная конференция имени А.Ф.Терпугова, Томск: Издательство Томского государственного университета, 2023. С. 113–118.
13. *Зорин А. В.* Случайные колебания длины приоритетной очереди при обслуживании конфликтных потоков по алгоритму с продлением в случайной среде. // *Вестник Нижегородского университета им. Н.И. Лобачевского*. 2009. № 1. С. 112–118.
14. *Зорин А. В.* Устойчивость тандема систем обслуживания с бернуллиевским мгновенным перемещением требований // *Теория вероятностей и математическая статистика*. 2011. Т. 84. С. 163–176.

Цодиков Владимир Леонидович — студент 1 курса магистратуры по направлению «Фундаментальная информатика и информационные технологии» Института информационных технологий, математики и механики. E-mail: vovatsodikov@gmail.com

Зорин Андрей Владимирович — доктор физико-математических наук, доцент, заведующий кафедрой Теории вероятностей и анализа данных, Институт информационных технологий, математики и механики. E-mail: andrei.zorine@itmm.unn.ru

ТЕСТИРОВАНИЕ ДОЛГОЙ ПАМЯТИ В НАГРУЗКЕ ОДНОСЕРВЕРНОЙ СИСТЕМЫ

В. А. Иголкин¹, А. С. Румянцев²

¹ *Петрозаводский государственный университет, г. Петрозаводск, Россия*

² *ИПМИ КарНЦ РАН, г. Петрозаводск, Россия*

В работе исследована применимость методов статистического обнаружения свойств долгой памяти (долговременной зависимости) для процесса нагрузки в односерверной системе обслуживания, в которой времена обслуживания имеют распределение с тяжелым хвостом. **Ключевые слова:** *долгая память, долговременная зависимость, проверка гипотез, односерверная система обслуживания.*

Введение

Процессы с долгой памятью играют важную роль в исследовании сложных систем, например в сетях передачи данных [1]. Наличие долгой памяти может затруднять прогнозирование и управление системами, увеличивая задержки в системах обслуживания [2]. В этой связи важно найти удобный статистический инструмент для обнаружения долгой памяти и исследования её характеристик.

Существует множество эмпирических методов оценки параметров долгой памяти случайных процессов и временных рядов, такие как RS-анализ, спектральные методы, которые, однако, в основном используются в качестве описательных и не пригодны для статистического вывода [3]. В то же время, статистические тесты для проверки наличия долгой памяти напрямую не применялись на моделях систем обслуживания.

В этой связи, целью данной работы является исследование применимости статистических тестов на наличие или отсутствие долгой памяти к процессам в системах обслуживания на примере нагрузки односерверной системы, в которой время обслуживания клиентов имеет распределение с правильно меняющимся (тяжелым) хвостом. Известно, что при определенных моментных свойствах распределения времени обслуживания такой процесс обладает свойством долгой памяти [6]. В данной работе указанные теоретические результаты проверяются эмпирически.

1. Свойство долгой памяти

Стационарный в широком смысле случайный процесс $\mathcal{X} = \{X_n\}_{n \in \mathbb{Z}}$ с автоковариационной функцией $\gamma_{\mathcal{X}}(k) = \text{Cov}(X_0, X_k)$, $k \in \mathbb{Z}$ и спектральной плотностью (с.п.)

$$f_{\mathcal{X}}(\lambda) = (2\pi)^{-1} \sum_{k \in \mathbb{Z}} \gamma_{\mathcal{X}}(k) e^{-ik\lambda}, \quad \lambda \in [-\pi, \pi], \quad (1)$$

при $|\lambda| \rightarrow 0$ имеет [3]

- долгую память, если $f_{\mathcal{X}}(\lambda) \rightarrow \infty$;
- короткую память, если $f_{\mathcal{X}}(\lambda)$ сходится к конечной константе;
- свойство антиперсистентности, если $f_{\mathcal{X}}(\lambda) \rightarrow 0$.

Используя (1), можно получить эквивалентное определение на основе сходимости ряда $\sum_{k \in \mathbb{Z}} \gamma_{\mathcal{X}}(k)$. Именно, процесс \mathcal{X} имеет [3]

- долгую память, если автоковариационный ряд расходится;
- короткую память, если сумма ряда конечна и отлична от нуля;
- свойство антиперсистентности, если сумма ряда нулевая.

Указанные определения имеют эмпирический характер и уточняются в классе процессов, у которых с.п. имеет вид

$$f_{\mathcal{X}}(\lambda) = L_f(\lambda) |\lambda|^{-2d}, \quad (2)$$

где $L_f(\lambda) \geq 0$ – симметричная функция, медленно меняющаяся в нуле, т.е. для любого $u > 0$, имеет место асимптотическая эквивалентность $L_f(u\lambda) \sim L_f(\lambda)$ при $\lambda \rightarrow 0$ (иногда вместо правильного изменения функции L_f предполагается четность, непрерывность и положительность [5]). Напомним, асимптотическая эквивалентность $a(x) \sim b(x)$ при $x \rightarrow x_0$ для некоторого x_0 означает

$$\lim_{x \rightarrow x_0} a(x)/b(x) = 1.$$

Определение 1. Процесс \mathcal{X} со с.п. вида (2) имеет

- долгую память, если $d \in (0, 1/2)$;
- короткую память, если $d = 0$ и $\lim_{\lambda \rightarrow 0} L_f(\lambda) = c_f \in (0, \infty)$;
- свойство антиперсистентности, если $d \in (-1/2, 0)$.

Следующая теорема позволяет связать спектральные свойства и сходимость ряда автоковариаций процесса \mathcal{X} .

Теорема 1. [3, Теорема 1.3] Пусть \mathcal{X} стационарен в широком смысле. Тогда

— если автоковариационная функция имеет вид

$$\gamma_{\mathcal{X}}(k) = L_{\gamma}(k)|k|^{2d-1},$$

где $L_{\gamma}(k)$ медленно меняется на бесконечности (т.е. $L_{\gamma}(k) \sim L_{\gamma}(uk)$ для любого $u > 0$ при $k \rightarrow \infty$), и выполнено одно из условий:

а) $d \in (0, 1/2)$, б) $d \in (-1/2, 0)$ и $\sum_{k \in \mathbb{Z}} \gamma_{\mathcal{X}}(k) = 0$, тогда

$$f_{\mathcal{X}}(\lambda) \sim |\lambda|^{-2d} L_{\gamma}(\lambda^{-1}) \pi^{-1} \Gamma(2d) \sin\left(\frac{\pi}{2} - \pi d\right), \quad \lambda \rightarrow 0;$$

— если с.п. имеет вид

$$f_{\mathcal{X}}(\lambda) = L_f(\lambda)|\lambda|^{-2d}, \quad 0 < \lambda < \pi,$$

где $d \in (-1/2, 0) \cup (0, 1/2)$, а $L_f(\lambda)$ медленно меняется в нуле и имеет ограниченную вариацию на (a, π) для любого $a > 0$, то

$$\gamma_{\mathcal{X}}(k) \sim |k|^{2d-1} 2L_f(k^{-1}) \Gamma(1-2d) \sin \pi d, \quad k \rightarrow \infty.$$

Параметр d , параметр памяти процесса [4], используется для проверки гипотез о наличии или отсутствии долгой памяти. Пример такого статистического теста приведен в следующем разделе.

2. Тест в частотной области

В работе [5] был предложен статистический тест в частотной области для проверки гипотезы H_0 о наличии короткой памяти у стационарного процесса \mathcal{X} со с.п. вида (2) при альтернативной гипотезе H_1 о присутствии свойства долгой памяти у указанного процесса. Тест основан на асимптотических свойствах так называемой периодограммы $I_{n,k}^{(l)}$, определяемой по наблюдениям $X_{(k-1)l+1}, \dots, X_{kl}$:

$$I_{n,k}^{(l)}(\lambda) = \frac{1}{2\pi l} \left| \sum_{j=1}^l X_{(k-1)l+j} e^{ij\lambda} \right|^2. \quad (3)$$

Именно, для любого фиксированного j при $\lambda_j = \frac{2\pi j}{n}$, имеет место сходимость по распределению [4]

$$\frac{I_n(\lambda_j)}{f_{\mathcal{X}}(\lambda_j)} \Rightarrow E, \quad n \rightarrow \infty,$$

где E имеет стандартное экспоненциальное распределение, а $I_n = I_{n,1}^{(n)}$ — периодограмма, построенная по исходной выборке. В этой связи, тестовая статистика для данного теста определена следующим образом [5]:

$$Q_{n,m}(s) = m \sum_{j=1}^s \frac{I_n(\lambda_j)}{\sum_{k=1}^m I_{n,k}^{(l)}(\lambda_j)}, \quad (4)$$

где $s \geq 1$ – некий фиксированный (настраиваемый) целочисленный параметр теста, $m = \frac{n}{l}$, l – длина блока, m – количество блоков, n – длина временного ряда. Отметим в качестве рекомендованных в [5] параметры $m = \sqrt{n}$ и $s \leq 10$ для небольших выборок порядка $n = 10^4$.

Процедура проверки гипотезы о наличии короткой памяти в работе [5] определена для процесса \mathcal{X} , который имеет вид скользящего среднего

$$X_n = \sum_{j=0}^{\infty} a_j \varepsilon_{n-j}, \quad (5)$$

где $\{\varepsilon_j\}_{j \in \mathbb{Z}}$ – н.о.р.с.в. с нулевым средним и $E\varepsilon^4 < \infty$ (без индекса обозначена типичная с.в.). Однако это условие не слишком ограничительное, поскольку любой стационарный в широком смысле процесс с нулевым средним может быть представлен в виде (5) с помощью так называемого разложения Вольда [4]. Именно, доказана следующая теорема (приведенная здесь с незначительной модификацией формулировки).

Теорема 2. [5, Теорема 1] Пусть процесс \mathcal{X} имеет вид (5), а тестовая статистика $Q_{n,m}(s)$ определена в (4). Пусть $m \rightarrow \infty$, при этом $m = o(n)$. Тогда

— если \mathcal{X} – процесс с короткой памятью в том смысле, что

$$\sum_{j=1}^{\infty} |a_j| < \infty,$$

и $f_{\mathcal{X}}$ положительна. Тогда имеет место сходимость по распределению

$$Q_{n,m}(s) \Rightarrow \Gamma(s, 1),$$

где $\Gamma(s, 1)$ имеет распределение гамма с параметрами $(s, 1)$, плотность которого имеет вид

$$f_{\Gamma}(x) = x^{s-1} e^{-x} \Gamma(s)^{-1},$$

где $\Gamma(s)$ есть гамма-функция;

— если \mathcal{X} – процесс с долгой памятью в смысле Определения 1, тогда

$$Q_{n,m}(s) \Rightarrow \infty.$$

Таким образом, гипотеза H_0 (о короткой памяти) отвергается в пользу H_1 (о наличии долгой памяти) на уровне значимости α при

$$Q_{n,m}(s) > \Gamma_{\alpha}(s),$$

где $\Gamma_{\alpha}(s)$ – квантиль порядка $1 - \alpha$ распределения $\Gamma(s, 1)$.

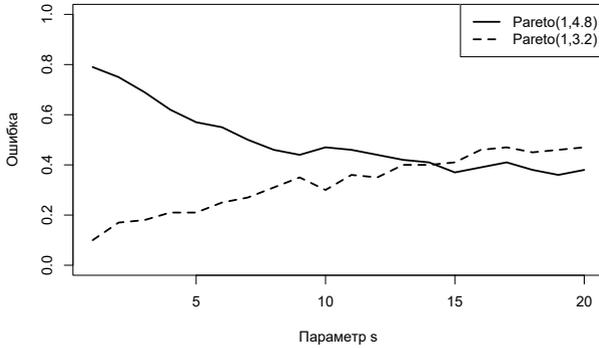


Рис. 1. Зависимость ошибки первого рода на уровне значимости $\alpha = 0.1$ от параметра s теста (4) на основе 100 траекторий (7) длиной 10^4 .

3. Долгая память процесса нагрузки системы GI/G/1

В работе [6] доказано наличие свойства долгой памяти в стационарной последовательности времен ожидания $\{W_n\}_{n \geq 1}$ односерверной системы GI/G/1, в которой распределение $B(x)$ н.о.р. времен обслуживания $\{S_n\}_{n \geq 1}$ имеет правильно меняющийся хвост. Именно, доказана следующая теорема.

Теорема 3. [6, Теорема 1] В системе GI/G/1 с конечным средним значением интервалов $\{T_n\}_{n \geq 1}$ между приходами клиентов, $ET < \infty$, и правильно меняющимся хвостом распределения времени обслуживания $\bar{B}(x) = 1 - B(x) = x^{-k_S} L_B(x)$ (где $L_B(x)$ медленно меняется на бесконечности), стационарная последовательность времен ожидания $\mathcal{W} = \{W_n\}_{n \geq 1}$ имеет коэффициент корреляции $\gamma_{\mathcal{W}}(n) = n^{3-k_S} L_{\mathcal{W}}(x)$ для некоторой медленно меняющейся функции $L_{\mathcal{W}}$.

С учетом Теоремы 1, Теорема 3 означает наличие долгой памяти в системе GI/G/1, в которой времена обслуживания имеют распределение Парето,

$$\bar{B}(x) \sim x^{-k_S}, \quad x \geq 1, \quad (6)$$

при $k_S \in (3, 4)$. Однако обнаружить это свойство с использованием имитационной модели (рекурсии Линдли)

$$W_{n+1} = (W_n + S_n - T_n)^+, \quad (7)$$

где $W_0 = 0$, достаточно сложно.

Для проверки указанного теоретического результата, был проведен численный эксперимент на модели GI/G/1 при относительно высокой

нагрузке, $\rho = ES/ET = 0.8$. Для этого к $M = 100$ независимым траекториям $\{W_n\}_{n \leq N}$, построенной в соответствии с рекурсией (7) (при длине траекторий $N = 10^4$) применялся тест (4) при $\bar{B}(x)$ вида (6) для $k_S = 3.2$ (теоретическое наличие долгой памяти) и $k_S = 4.8$ (отсутствие долгой памяти). Подсчитывалось число ошибок первого рода в зависимости от параметра s теста на уровне значимости $\alpha = 0.1$. Результаты представлены на графике 1. Видно, что значения $10 \leq s \leq 20$ дают умеренно высокую ошибку и могут применяться для грубого построения оценки.

4. Заключение

В ходе работы были получены эмпирические результаты, подтверждающие наличие долгой памяти в односерверной системе обслуживания с определенными моментными свойствами распределения времени обслуживания. Также стоит отметить, что проведенные эксперименты показали сильное влияние параметра нагрузки ρ на качество теста, что требует дальнейших исследований.

СПИСОК ЛИТЕРАТУРЫ

1. *Park C. et al.* Long-range dependence in a changing Internet traffic mix // *Computer Networks*. 2005. № 3 (48). С. 401–422.
2. *Dahl T. A., Willemain T. R.* The effect of long-memory arrivals on queue performance // *Operations Research Letters*. 2001. № 3 (29). С. 123–127.
3. *Beran J. et al.* Long-Memory Processes: Probabilistic Properties and Statistical Methods. Berlin: Springer, 2013. 884 с.
4. *Giraitis L., Koul H. L., Surgailis D.* Large Sample Inference for Long Memory Processes. Imperial College Press, 2012.
5. *Gromykov G., Ould Haye M., Philipp A.* A frequency-domain test for long range dependence // *Statistical Inference for Stochastic Processes*. 2018. № 21(3). С. 513–526.
6. *Carpio K. J. E.* Long-range dependence of stationary processes in single-server queues // *Queueing Systems*. 2007. № 2 (55). С. 123–130.

Иголкин Владислав Александрович — магистр, ИМИТ. E-mail: *easyufbin@gmail.com*

Румянцев Александр Сергеевич — д.ф.-м.н., вед.н.с., лаб. СМИТС. E-mail: *ar0@krc.karelia.ru*

ИМИТАЦИОННЫЙ АНАЛИЗ ХАРАКТЕРИСТИК ВЫХОДЯЩЕГО ПОТОКА БЕСКОНЕЧНОЛИНЕЙНОЙ СИСТЕМЫ С ДЕГРАДАЦИЕЙ СКОРОСТИ ОБСЛУЖИВАНИЯ

И. Л. Лапатин, В. В. Романов

*Национальный исследовательский Томский государственный университет,
г. Томск, Россия*

В работе рассматривается модель массового обслуживания с неограниченным числом приборов, простейшим входящим потоком и разными законами распределения объема поступающих требований. Интенсивность обслуживания поступающих заявок не является постоянной и зависит от числа заявок в системе. Эта зависимость выражается функцией деградации, которая определяет снижение скорости обслуживания каждой заявки при росте общего числа заявок в системе. Объектом исследования в данной работе является выходящий поток системы. Для исследования выходящего потока была построена имитационная модель. Получены характеристики длин интервалов выходящего потока системы при различных законах распределения объема поступающих заявок и разных видах функции деградации.

Ключевые слова: *выходящий поток, теория массового обслуживания, деградация скорости обслуживания, имитационное моделирование, системы с неограниченным числом приборов.*

Введение

Модели массового обслуживания с неограниченным числом приборов и наличием зависимости скорости обслуживания являются моделями работы вычислительных узлов и систем. Такие системы характеризуются тем, что ресурсы ограничены и при увеличении одновременно выполняемых задач может возникать конкуренция за ресурсы [1, 2], что приводит к снижению скорости выполнения каждого отдельного запроса. При построении модели и анализе ее характеристик этот эффект обязательно нужно учитывать, так как он вносит значительное влияние на значения полученных характеристик. Моделям с деградацией скорости обслуживания посвящены работы [3, 4]. Основным объектом исследования в них является распределение вероятностей числа заявок в системе, которая показывает какую долю времени система находится в разной степени загрузки.

Не менее важной является информация о выходящем потоке. Мотивация исследования выходящих потоков достаточно проста - выходящий поток одного этапа обслуживания является входящим для следующего. А именно характеристики потока значительно влияют на показатели качества обслуживания в системе. Так, например, чем больше коэффициент вариации длин интервалов во входящем потоке (при одинаковых средних значениях), тем большая пропускная способность системы нужна для обеспечения таких же показателей качества обслуживания. Еще в середине XX века для системы массового обслуживания с неограниченным числом приборов, простейшим входящим потоком и экспоненциальным законом распределения обслуживания поступающих заявок было показано [5], что выходящий поток также является простейшим. Через несколько лет этот результат был обобщен [6] на случай произвольного распределения времени обслуживания заявок. Это означает, что для системы с неограниченным числом приборов, на вход которой поступает простейший поток, выходящий поток также является простейшим. Причем этот результат инвариантен относительно закона распределения времени обслуживания. То есть коэффициент вариации длин интервалов выходящего потока в таких системах равен 1, а коэффициент корреляции длин соседних интервалов равен 0.

В нашей работе мы хотим проверить, как наличие зависимости скорости обслуживания от числа заявок в системе будет влиять на характеристики выходящего потока. Под характеристиками будем понимать как раз оценку коэффициента вариации длин интервалов в выходящем потоке и оценку коэффициента корреляции.

1. Математическая модель

Рассмотрим математическую модель системы массового обслуживания с деградацией скорости обслуживания заявок. Будем полагать, что моменты поступления запросов на узел образует простейший поток с параметром λ , а объем работы поступающих запросов являются независимыми и одинаково распределенными случайными величинами с функцией распределения $B(x)$ и средним значением $b = 1/\mu$. Скорость обслуживания поступающих заявок не является постоянной, она зависит от числа запросов в системе i и определяется значениями функции $f(i)$. Функцию $f(i)$ мы называем функцией деградации скорости обслуживания, так как в предлагаемой постановке задачи, при увеличении числа одновременно обрабатываемых запросов скорость выполнения каждого будет снижаться. Но предлагаемая математическая модель допускает использование более общего вида функции $f(i)$, в том числе возрастающей. Заметим, что если распределение вероятностей по-

ступающих запросов является экспоненциальным, то получаем систему (Рисунок 1) с переменной интенсивностью $\mu_i = \mu \cdot f(i)$.

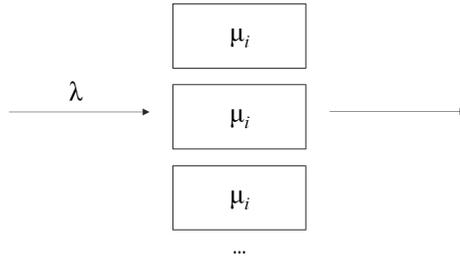


Рис. 1. Математическая модель бесконечнолинейной СМО с деградацией скорости обслуживания

Объектом исследования в данной работе является выходящий поток, события которого наступают при завершении обслуживания какой-либо заявки. При помощи имитационного моделирования будем искать характеристики длин интервалов между моментами наступления событий в выходящем потоке.

2. Имитационная модель

Выходящий поток существенно зависит от функционирования самой системы, и аналитическое исследование его для данной системы является нетривиальной задачей даже для случая экспоненциального закона распределения вероятностей объема поступающих задач. Мы хотим посмотреть характеристики выходящего потока при разных законах распределения объема поступающих требований. Для решения этой задачи была построена имитационная модель предлагаемой системы.

Для написания имитационной модели было принято решение использовать язык программирования C++. Такой выбор был сделан по нескольким основным причинам: множество удобных библиотек, которые могут быть полезными для работы со случайными величинами, в том числе генераторы случайных величин, а также наличие отладки, которая существенно сокращала поиск неочевидных ошибок [5]. Также стоит заметить, что при написании алгоритма для имитационной модели использовался дискретно-событийный подход [1].

Особенностью модели является то, что необходимо учитывать зависимость скорости обслуживания от числа заявок в системе. Для этого при поступлении заявки в систему моделируется плановое время обра-

ботки запроса, исходя из объема заявки и текущей скорости обработки, которая зависит от числа i запросов в системе. Как только приходит новый запрос или какой-то заканчивает обслуживание, остаточное время для всех активных заявок просто корректируется (растягивается или уменьшается) в зависимости от изменения интенсивности обслуживания. Это делается путём умножения остаточного времени на коэффициент, в числителе которого находится интенсивность до изменения состояния системы, а в знаменателе - новая интенсивность. Данный подход соответствует имитации текущей реализации поведения вычислительного узла, а также легко применим к случаям, когда распределение объема работы в заявке имеет неэкспоненциальное распределение.

Результатом работы имитационной модели является набор значений длин интервалов между моментами наступления событий в выходящем потоке. По этим данным уже можно строить оценки числовых характеристик.

3. Численные эксперименты

Используя результаты имитационного моделирования были построены оценки коэффициента автокорреляции и коэффициента вариации длин интервалов в выходящем потоке рассматриваемой системы при разных законах распределения объема поступающих требований.

Автокорреляция - корреляция между величиной и её запаздыванием в один или более периодов времени. Число периодов, по которым рассчитывается коэффициент автокорреляции, часто называется лагом или порядком автокорреляции. В нашем случае $\tau = 1$ - порядок автокорреляции [2]. Формула расчета оценки коэффициента автокорреляции:

$$r_{\tau} = \frac{\sum_{t=1+\tau}^n (y_t - \bar{y}_{1\tau}) \cdot (y_{t-\tau} - \bar{y}_{2\tau})}{\sqrt{\sum_{t=1+\tau}^n (y_t - \bar{y}_{1\tau})^2 \cdot \sum_{t=1+\tau}^n (y_{t-\tau} - \bar{y}_{2\tau})^2}}, \quad (1)$$

где:

$$\bar{y}_{1\tau} = \frac{\sum_{t=1+\tau}^n y_t}{n - \tau}; \quad \bar{y}_{2\tau} = \frac{\sum_{t=1+\tau}^n y_{t-\tau}}{n - \tau}, \quad (2)$$

Формула расчета оценки коэффициента вариации:

$$CV = \frac{\sigma}{\bar{y}}, \quad (3)$$

где σ - оценка среднеквадратического отклонения, \bar{y} - выборочное среднее.

Полученные оценки характеристик выходящего потока при $\lambda = 0.7, \mu = 1$ для соответствующих законов распределения времени обслуживания и функций деградации (изменения) скорости обслуживания приведены в Таблицах 1 и 2.

Таблица 1

Оценка коэффициента корреляции для длин соседних интервалов в выходящем потоке

Распределение объема заявок	$f(i) = 1 + i$	$f(i) = \frac{1}{\sqrt{1+i}}$	$f(i) = \frac{1}{i}$
Экспоненциальное ($\mu = 1$)	-0.004	-0.005	-0.019
Детерминированное ($d = 1$)	-0.016	0.074	0.113
Равномерное ($a = 0, b = 2$)	-0.014	0.003	-0.021
Гамма ($\alpha = \beta = 0.5$)	-0.0005	-0.007	-0.022
Гамма ($\alpha = \beta = 2$)	-0.008	-0.002	-0.033

Таблица 2

Оценка коэффициента вариации для длин соседних интервалов в выходящем потоке

Распределение объема заявок	$f(i) = 1 + i$	$f(i) = \frac{1}{\sqrt{1+i}}$	$f(i) = \frac{1}{i}$
Экспоненциальное ($\mu = 1$)	1.005	0.975	0.929
Детерминированное ($d = 1$)	1.02	0.871	0.704
Равномерное ($a = 0, b = 2$)	1.013	0.948	0.871
Гамма ($\alpha = \beta = 0.5$)	1.004	0.991	0.956
Гамма ($\alpha = \beta = 2$)	1.01	0.964	0.885

Из таблиц можно сделать вывод, что значения длин соседних интервалов мало зависимы друг от друга в независимости от наличия деградации и её функции, а также распределения объёма заявок.

4. Заключение

В ходе исследования характеристик выходящего потока бесконечнолинейной системы массового обслуживания с деградацией скорости обслуживания была построена имитационная модель и получены оценки коэффициента вариации и коэффициента корреляции длин интервалов для различных видов распределения объёма поступающих требова-

ний и различных видах функции деградации. Из результатов численных экспериментов можем заметить, что наличие деградации скорости обслуживания не приводит к появлению значительной корреляции в выходящем потоке. Для большинства распределений сохраняется тенденция, что чем сильнее деградация скорости обслуживания в системе, тем выше значение коэффициента корреляции и меньше значение коэффициента вариации длин интервалов. Наибольшее отличие значений числовых характеристик от классических результатов без деградации скорости обслуживания наблюдалось при детерминированном объеме поступающих требований.

СПИСОК ЛИТЕРАТУРЫ

1. *Canosa-Reyes R. M., Tchernykh A., Cortés-Mendoza J. M., and others.* Dynamic performance-Energy tradeoff consolidation with contention-aware resource provisioning in containerized clouds // PLoS ONE. 2022. Т. 17. Выпуск 1. Номер статьи e0261856.
2. *Goodell, D., Balaji, P., Buntinas, D., and others.* Minimizing MPI Resource Contention in Multithreaded Multicore Environments // IEEE International Conference on Cluster Computing, 2010. С. 1–8.
3. *А.К. Красников, Н.С. Щербачев* Многоканальные системы массового обслуживания с деградацией // Приборы и системы. Управление, контроль, диагностика. С. 24–28.
4. *А.А. Назаров, Е.А. Фёдорова, О.Д. Лизюра, С.В. Пауль [и др.]* Двухфазная СМО с неограниченным числом приборов и деградацией обслуживания // Информационные технологии и математическое моделирование : материалы XXI Международной конференции имени А. Ф. Терпугова, 2022 г. Томск: Изд-во Том. ун-та, 2023. С. 18–23.
5. *Burke, P.J.* The Output of a Queuing System // Operations Research. 1956. Vol 4. P. 699–704.
6. *P.J. Mirasol* On the Departure Process of Queues in Heavy Traffic // Journal of the Operations Research Society of America. 1960. Vol. 8. № 2, P. 246–255.

Лапатин Иван Леонидович — канд. физ.-мат. наук, доцент кафедры теории вероятностей и математической статистики Томского государственного университета. E-mail: *ilapatin@mail.ru*

Романов Вячеслав Васильевич — магистрант кафедры теории вероятностей и математической статистики Томского государственного университета. E-mail: *slaves2001@gmail.com*

ИССЛЕДОВАНИЕ БЕСКОНЕЧНОЛИНЕЙНОЙ СМО С ПРОВЕРКОЙ КАЧЕСТВА ОБРАБОТКИ ПРИ НЕСТАЦИОНАРНОМ РЕЖИМЕ РАБОТЫ

М. Д. Лавриненко, М. А. Шкленник

*Национальный исследовательский Томский государственный университет,
г. Томск, Россия*

В данной работе рассматривается система массового обслуживания (СМО) с неограниченным числом обслуживающих устройств, с наличием как мгновенной, так и отсроченной обратной связи. Процесс обслуживания включает в себя обработку заявок в течение случайного времени, распределение которого экспоненциально. После обработки каждая заявка подвергается проверке на наличие ошибок, которые могут быть мгновенно устранены либо требовать некоторого времени для их устранения. После ликвидации ошибок заявкам необходимо пройти повторную обработку, на которую требуется уже отличное от первичного обслуживания время. Таким образом в исследуемой системе реализуется одновременно мгновенная и отсроченная обратная связь. **Ключевые слова:** *бесконечнолинейная система массового обслуживания, проверка качества обработки, мгновенная обратная связь, отсроченная обратная связь, производящая функция*

Введение

Системы массового обслуживания (СМО) являются эффективными математическими моделями для анализа процессов обслуживания в различных областях, включая телекоммуникационные системы, компьютерные сети, и производственные процессы [2]. Важным аспектом в анализе СМО является наличие обратной связи, которая существенно влияет на характеристики производительности и качества обслуживания. Мгновенная обратная связь предполагает, что каждая обработанная заявка мгновенно возвращается в систему в случае необходимости дополнительного обслуживания [1]. В отличие от мгновенной, отсроченная обратная связь предусматривает возможность задержки возвращения заявки в систему после первичного обслуживания [3]. Такие системы представляют собой компромисс между мгновенной реакцией и возможностью эффективного управления ресурсами [4].

1. Математическая модель

Рассмотрим систему массового обслуживания с неограниченным числом обслуживающих устройств с мгновенной и отсроченной обратной связью (рис. 1). В систему поступает простейший поток заявок с параметром λ . Время обслуживания заявок входящего потока является случайной величиной, имеющей экспоненциальное распределение с параметром μ_1 . После обслуживания каждая заявка проверяется на наличие ошибок, возникающих в процессе обработки. Пусть существует два типа ошибок:

- первый тип устраняется мгновенно, но при этом требуется повторное обслуживание. Вероятность возникновения ошибок первого типа равна r_1 ,
- второй тип ошибок возникает с вероятностью r_2 и для их устранения требуется случайное время, распределенное экспоненциально с параметром σ . После чего заявке также требуется повторное обслуживание.

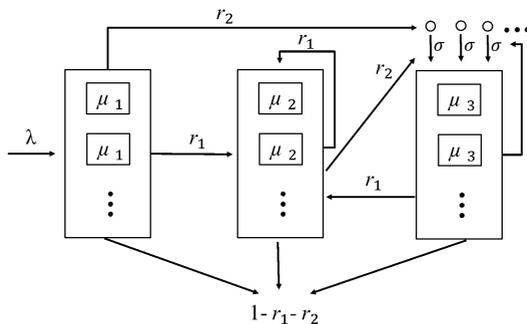


Рис. 1. – Система массового обслуживания с неограниченным числом обслуживающих устройств с мгновенной и отсроченной обратной связью

Таким образом, каждая заявка входящего потока после завершения обслуживания может с вероятностью r_1 поступить во второй блок, где время обслуживания распределено экспоненциально с параметром μ_2 , либо с вероятностью r_2 уйти на орбиту, где осуществляется задержка в течение случайного времени с параметром σ . С орбиты заявка поступает на третий блок обслуживания системы и обрабатывается там в течение экспоненциально распределенного времени с параметром μ_3 . После обработки на втором и третьем блоке проверка качества обслуживания

повторяется, и в случае обнаружения ошибок первого или второго типа схема их устранения и обработки повторяется.

Введем следующие обозначения:

- $i_1(t)$ – число заявок входящего потока, обслуживающихся в системе в момент времени t ;
- $i_2(t)$ – число заявок, которые обслуживаются после обнаружения ошибки первого типа, в момент времени t ;
- $i_3(t)$ – число заявок, которые обслуживаются после обнаружения ошибки второго типа, в момент времени t ;
- $i(t)$ – число заявок, находящихся в момент времени t на орбите.

Ставится задача исследования четырёхмерного процесса $\{i_1(t), i_2(t), i_3(t), i(t)\}$ в нестационарном режиме.

Пусть $P(i_1, i_2, i_3, i, t) = P\{i_1(t) = i_1, i_2(t) = i_2, i_3(t) = i_3, i(t) = i\}$ – распределение вероятностей исследуемого процесса. Тогда, используя Δt -метод, можем записать систему дифференциальных уравнений Колмогорова

$$\begin{aligned} \frac{\partial P(i_1, i_2, i_3, i, t)}{\partial t} = & -P(i_1, i_2, i_3, i, t)(\lambda + i_1\mu_1 + i_2\mu_2 + i_3\mu_3 + i\sigma) + \\ & + \lambda P(i_1 - 1, i_2, i_3, i, t) + (i_1 + 1)\mu_1 r_1 P(i_1 + 1, i_2 - 1, i_3, i, t) + \\ & + (i_1 + 1)\mu_1 r_2 P(i_1 + 1, i_2, i_3, i - 1, t) + (i_1 + 1)\mu_1(1 - r_1 - r_2)P(i_1 + 1, i_2, i_3, i, t) + \\ & + (i_2 + 1)\mu_2 r_2 P(i_1, i_2 + 1, i_3, i - 1, t) + i_2\mu_2 r_1 P(i_1, i_2, i_3, i, t) + \\ & + (i_2 + 1)\mu_2(1 - r_1 - r_2)P(i_1, i_2 + 1, i_3, i, t) + \\ & + (i_3 + 1)\mu_3 r_1 P(i_1, i_2 - 1, i_3 + 1, i, t) + (i_3 + 1)\mu_3 r_2 P(i_1, i_2, i_3 + 1, i - 1, t) + \\ & + (i_3 + 1)\mu_3(1 - r_1 - r_2)P(i_1, i_2, i_3 + 1, i, t) + \\ & + (i + 1)\sigma P(i_1, i_2, i_3 - 1, i + 1, t) \end{aligned} \quad (1)$$

с начальным условием

$$P(i_1, i_2, i_3, i, 0) = \begin{cases} 1, & \text{если } i_1 = i_2 = i_3 = i = 0, \\ 0, & \text{если } i_1 \neq 0, i_2 \neq 0, i_3 \neq 0, i \neq 0. \end{cases} \quad (2)$$

2. Метод производящих функций

Для получения искомого выражения для распределения будем использовать метод производящих функций [15]. Введем производящую функцию

$$G(x, y, z, v, t) = \sum_{i_1} \sum_{i_2} \sum_{i_3} \sum_i x^{i_1} y^{i_2} z^{i_3} v^i P(i_1, i_2, i_3, i, t).$$

Тогда система (1) равносильна дифференциальному уравнению в частных производных для производящей функции

$$\begin{aligned} \frac{\partial G(x, y, z, v, t)}{\partial t} + \mu_1(x - 1 - r_1(y - 1) - r_2(v - 1)) \frac{\partial G(x, y, z, v, t)}{\partial x} + \\ + \mu_2((y - 1)(1 - r_1) - r_2(v - 1)) \frac{\partial G(x, y, z, v, t)}{\partial y} + \\ + \mu_3(z - 1 - r_1(y - 1) - r_2(v - 1)) \frac{\partial G(x, y, z, v, t)}{\partial z} + \\ + \sigma(v - z) \frac{\partial G(x, y, z, v, t)}{\partial v} = \lambda(x - 1)G(x, y, z, v, t) \end{aligned} \quad (3)$$

с начальным условием

$$G(x, y, z, v, 0) = 1. \quad (4)$$

Решение задачи Коши (3) - (4) имеет вид

$$\begin{aligned} G(x, y, z, v, t) = \exp \left\{ \frac{\lambda}{\Omega} \left[\frac{\alpha_1}{\mu_1 \varkappa_1} (\gamma_3 - \gamma_2) (\varkappa_1(1 - e^{-\mu_1 t}) - \mu_1(1 - e^{-\varkappa_1 t})) - \right. \right. \\ \left. \left. - \frac{\alpha_2}{\mu_1 \varkappa_2} (\gamma_3 - \gamma_1) (\varkappa_2(1 - e^{-\mu_1 t}) - \mu_1(1 - e^{-\varkappa_2 t})) + \right. \right. \\ \left. \left. + \frac{\alpha_3}{\mu_1 \varkappa_3} (\gamma_2 - \gamma_1) (\varkappa_3(1 - e^{-\mu_1 t}) - \mu_1(1 - e^{-\varkappa_3 t})) \right] (y - 1) + \right. \\ \left. + \left[\frac{\alpha_1}{\mu_1 \varkappa_1} (\beta_3 - \beta_2) (\varkappa_1(1 - e^{-\mu_1 t}) - \mu_1(1 - e^{-\varkappa_1 t})) - \right. \right. \\ \left. \left. - \frac{\alpha_2}{\mu_1 \varkappa_2} (\beta_3 - \beta_1) (\varkappa_2(1 - e^{-\mu_1 t}) - \mu_1(1 - e^{-\varkappa_2 t})) + \right. \right. \\ \left. \left. + \frac{\alpha_3}{\mu_1 \varkappa_3} (\beta_2 - \beta_1) (\varkappa_3(1 - e^{-\mu_1 t}) - \mu_1(1 - e^{-\varkappa_3 t})) \right] (z - 1) + \right. \end{aligned}$$

$$\begin{aligned}
& + \left[\frac{\alpha_1}{\mu_1 \varkappa_1} (\beta_3 \gamma_2 - \beta_2 \gamma_3) (\varkappa_1 (1 - e^{-\mu_1 t}) - \mu_1 (1 - e^{-\varkappa_1 t})) - \right. \\
& - \frac{\alpha_2}{\mu_1 \varkappa_2} (\beta_3 \gamma_1 - \beta_1 \gamma_3) (\varkappa_2 (1 - e^{-\mu_1 t}) - \mu_1 (1 - e^{-\varkappa_2 t})) + \\
& \left. + \frac{\alpha_3}{\mu_1 \varkappa_3} (\beta_2 \gamma_1 - \beta_1 \gamma_2) (\varkappa_3 (1 - e^{-\mu_1 t}) - \mu_1 (1 - e^{-\varkappa_3 t})) \right] (v - 1) - \\
& - \frac{(x - 1)}{\mu_1} \Omega (1 - e^{-\mu_1 t}) \Bigg\}, \tag{5}
\end{aligned}$$

где $\varkappa_1, \varkappa_2, \varkappa_3$ – корни характеристического уравнения

$$\begin{aligned}
& \varkappa^3 - (\mu_3 + \sigma + \mu_2(1 - r_2))\varkappa^2 + (\mu_3\sigma(1 - r_2) + \\
& + \mu_2\sigma(1 - r_1) + \mu_2\mu_3(1 - r_1))\varkappa - \mu_2\mu_3\sigma(1 - r_1 - r_2) = 0,
\end{aligned}$$

$$\Omega = \gamma_1(\beta_2 - \beta_3) + \gamma_2(\beta_3 - \beta_1) + \gamma_3(\beta_1 - \beta_2),$$

$$\alpha_k = \frac{\mu_1}{\varkappa_k - \mu_1} \left(r_1 + r_2 - \varkappa_k \left(\frac{1}{\sigma} + \frac{1}{\mu_3} \right) + \frac{\varkappa_k^2}{\sigma\mu_3} \right),$$

$$\beta_k = 1 - \frac{\varkappa_k}{r_1} \left(\frac{1}{\sigma} + \frac{1}{\mu_3} \right) + \frac{\varkappa_k^2}{\sigma\mu_3 r_1}, \gamma_k = 1 - \frac{\varkappa_k}{\sigma}, k = 1, 2, 3.$$

По виду выражения (5) можно сделать вывод о том, что числа заявок в каждом блоке системы являются величинами независимыми, причем каждая компонента исследуемого многомерного процесса представляет собой нестационарный пуассоновский процесс.

При $t \rightarrow \infty$ несложно получить производящую функцию финального распределения вероятностей исследуемого процесса, которое совпадает со стационарным. Запишем его в виде

$$\begin{aligned}
G(x, y, z, v) = \exp \Bigg\{ \frac{\lambda}{\Omega} \Bigg[& \frac{\alpha_1}{\mu_1 \varkappa_1} (\gamma_3 - \gamma_2) (\varkappa_1 - \mu_1) - \\
& - \frac{\alpha_2}{\mu_1 \varkappa_2} (\gamma_3 - \gamma_1) (\varkappa_2 - \mu_1) + \frac{\alpha_3}{\mu_1 \varkappa_3} (\gamma_2 - \gamma_1) (\varkappa_3 - \mu_1) \Bigg] (y - 1) \\
& + \left[\frac{\alpha_1}{\mu_1 \varkappa_1} (\beta_3 - \beta_2) (\varkappa_1 - \mu_1) - \frac{\alpha_3}{\mu_1 \varkappa_3} (\beta_2 - \beta_1) (\varkappa_3 - \mu_1) \right] (z - 1) \\
& + \left[\frac{\alpha_1}{\mu_1 \varkappa_1} (\beta_3 \gamma_2 - \beta_2 \gamma_3) (\varkappa_1 - \mu_1) - \frac{\alpha_2}{\mu_1 \varkappa_2} (\beta_3 \gamma_1 - \beta_1 \gamma_3) (\varkappa_2 - \mu_1) + \right. \\
& \left. + \frac{\alpha_3}{\mu_1 \varkappa_3} (\beta_2 \gamma_1 - \beta_1 \gamma_2) (\varkappa_3 - \mu_1) \right] (v - 1) + \frac{\Omega(x - 1)}{\mu_1} \Bigg\}. \tag{6}
\end{aligned}$$

По виду производящей функции (6) можно сделать вывод о том, что стационарные распределения числа заявок на каждом из этапов обработки подчиняется закону Пуассона с параметрами, которые можно вычислить, используя (6).

Заключение

В работе представлена математическая модель системы массового обслуживания с проверкой качества обработки в виде реализации мгновенной и отсроченной обратной связи. В результате проведенного исследования получено выражение для производящей функции нестационарного распределения вероятностей четырехмерного случайного процесса, описывающего числа заявок каждого типа в исследуемой СМО. Показано, что числа заявок каждого типа, находящиеся в системе в момент времени t , являются величинами независимыми и являются нестационарными пуассоновскими процессами. Также записана производящая функция стационарного распределения исследуемого процесса, которая будет использована в дальнейших исследованиях потоков заявок в рассматриваемой системе.

СПИСОК ЛИТЕРАТУРЫ

1. Шкленник М. А., Моисеев А. Н., Задиранова Л. А. Исследование потока повторных обращений в двухфазной системе массового обслуживания GI/GI/ ∞ с мгновенной обратной связью методом марковского суммирования // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2021. № 56. С. 29–40.
2. Van Doorn E. A., Jagers A. A. Note on the GI/GI/infinity system with identical service and interarrival-time distributions // Queueing Systems. 2004. № 47. P. 45–52.
3. Кудрявцева Е. Н., Росляков А. В. Применение теории сетевого исчисления к исследованию систем массового обслуживания с обратной связью // Т-Сотм: Телекоммуникации и транспорт. 2015. NFI. С. 17–21.
4. Назаров А. А., Терпугов А. Ф. Теория вероятностей и случайных процессов. Томск: Издательство, 2010. 199 с.

Лавриненко Милена Дмитриевна — аспирантка ИПМКН ТГУ. E-mail: mila-guyy@mail.ru

Шкленник Мария Александровна — кандидат физ.-мат. наук, доцент кафедры ТВиМС, ИПМКН ТГУ. E-mail: shklennikm@yandex.ru

МАТЕМАТИЧЕСКАЯ МОДЕЛЬ КОММЕРЧЕСКОГО БАНКА, ИСПОЛЬЗУЮЩЕГО СТРАХОВАНИЕ КРЕДИТНЫХ ПРОДУКТОВ

Н. А. Кузнецов, М. А. Шкленник

*Национальный исследовательский Томский государственный университет,
г. Томск.*

В данной работе предложена математическая модель коммерческого банка, представленного как система массового обслуживания с обратной связью. Поток клиентов, обратившихся в банк впервые, задается простейшим пуассоновским потоком. Так как при оформлении кредитного продукта клиент может застраховать свою ответственность перед банком или же отказаться от этой услуги, то входящий поток заявок, поступающий в предложенную СМО, разделяется. Найдена производящая функция для нестационарного распределения вероятностей числа клиентов, имеющих действующие застрахованные и незастрахованные кредитные договоры, в любой момент времени. **Ключевые слова:** *математическая модель, обратная связь, коммерческий банк, страхование кредитных продуктов, система массового обслуживания.*

Введение

В современном мире экономических отношений сложно представить человека, который не сталкивался с очередями на кассе, не пользовался банкоматом или не брал билет в общественном транспорте. Эти ситуации можно описать с помощью систем массового обслуживания (СМО) — математических моделей, представляющих стохастические процессы, где заявки на обслуживание поступают случайным образом и требуют обработки в течение случайного времени [2, 7]. СМО играют ключевую роль в экономике благодаря оптимизации ресурсов, улучшению качества обслуживания, повышению эффективности процессов и способности к анализу и прогнозированию. Особенно заметно их применение в банковской сфере, где финансовые операции, такие как выдача кредитов и управление страховыми продуктами, требуют точного анализа и управления. В нашем исследовании мы создадим математическую модель для анализа работы коммерческого банка как системы массового

обслуживания с обратной связью [2, 4]. Цель работы заключается в получении выражений для основных вероятностных характеристик числа клиентов, имеющих действующие застрахованные и незастрахованный кредитные договора в данном банке в определенный момент времени, что важно для оптимизации процесса выдачи кредитов и повышения удовлетворенности клиентов.

1. Математическая модель

Рассмотрим коммерческий банк, в котором имеется возможность застраховать свой кредит, минимизируя риски как банка, так и самого заемщика от непредвиденных обстоятельств. Будем считать, что клиенты банка после заключения кредитного договора до погашения обязательств не открывают новых кредитов в данном банке. После полного гашения кредита клиент может сделать выбор в пользу другого банка. Также клиент может снова обратиться за финансовой помощью к данному банку, но при этом изменить условия кредитного договора, путем его страхования, или же взять новый кредит на тех же условиях, то есть без страховки. Процесс изменения числа клиентов с учетом механизма обратной связи имеет решающее значение для кредитного учреждения.

Поток клиентов, впервые обратившихся за финансовой помощью в банк, будем рассматривать как простейший поток с параметром λ . Клиент, придя в банк, с вероятностью r_0 оформит свой первый кредитный договор со страхованием и будет выплачивать его на протяжении некоторого случайного времени, которое распределено по экспоненциальному закону с параметром μ_1 . После закрытия долговых обязательств клиент может снова обратиться в банк для оформления нового кредитного договора. При этом у него есть возможность поменять условия кредитования по своим предпочтениям и оформить кредитный договор без использования страхования с вероятностью r_{12} , либо же заново оформить кредит с теми же условиями с вероятностью r_{11} . Также клиент может прекратить взаимоотношения с банком, перейдя в другой банк, с вероятностью $(1 - r_{11} - r_{12})$. Предположим, если клиент поменял условия кредитования и оформил кредит без страхования, то политика его выплат меняется, и время, за которое клиент полностью выплатит долговую сумму, будет случайной величиной, распределенной экспоненциально с параметром μ_2 .

Обозначим $i_1(t)$ – число кредитных договоров в любой момент времени t , оформленных со страхованием, а $i_2(t)$ – соответственно, число кредитных договоров, оформленных без страховки. Тогда случайный процесс $\{i_1(t), i_2(t)\}$ будет характеризовать число застрахованных и незастрахованных кредитных договоров. Найдем основные вероят-

ностные характеристики данного процесса в нестационарном и стационарном режимах, используя механизмы теории массового обслуживания.

2. Исследование системы массового обслуживания

Для математического моделирования нашей задачи будем использовать бесконечнолинейную систему массового обслуживания (СМО) с двумя обслуживающими блоками и обратной связью (рисунок 2). В каждом из двух блоков системы имеется неограниченное число обслуживающих устройств [3]. В систему поступает простейший поток с параметром λ , который на входе разделяется по биномиальной схеме, оставаясь при этом тем же простейшим потоком [5]. Заявка, попадая в систему, с вероятностью r_0 направляется в первый блок, занимая любой из доступных приборов, где её обслуживание происходит в течение некоторого случайного времени, распределенного по экспоненциальному закону с параметром μ_1 . После обслуживания заявка с вероятностью r_{11} может повторно обратиться к этому блоку, перейти в другой блок с вероятностью r_{12} , или с вероятностью $(1 - r_{11} - r_{12})$ покинуть систему. Аналогичная ситуация наблюдается и для второго блока, параметры системы отображены на рисунке 2.

Обозначим $i_1(t)$ – число занятых приборов в первом блоке обслуживания в момент времени t , а $i_2(t)$ – во втором соответственно. Ставится приоритет на вычисление распределения вероятностей двумерного случайного процесса $\{i_1(t), i_2(t)\}$.

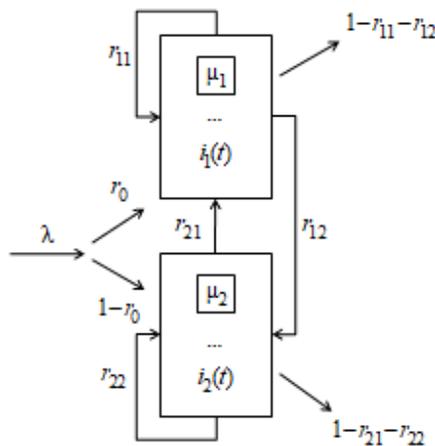


Рис. 1. СМО с двумя обслуживающими блоками и обратной связью

Описанный двумерный процесс является марковским [5]. Определим распределение вероятностей для данного процесса в виде $P(i_1, i_2, t) = P\{i_1(t) = i_1, i_2(t) = i_2\}$. Тогда, используя формулу полной вероятности и Δt -метод, можно получить систему дифференциальных уравнений Колмогорова для искомого распределения вероятностей.

$$\begin{aligned} \frac{dP(i_1, i_2, t)}{dt} = & r_0\lambda P(i_1 - 1, i_2, t) + r_{12}\mu_1(i_1 + 1)P(i_1 + 1, i_2 - 1, t) + \\ & + \mu_1(i_1 + 1)(1 - r_{11} - r_{12})P(i_1 + 1, i_2, t) + r_{21}\mu_2(i_2 + 1)P(i_1 - 1, i_2 + 1, t) + \\ & + \mu_2(i_2 + 1)(1 - r_{21} - r_{22})P(i_1, i_2 + 1, t) + (1 - r_0)\lambda P(i_1, i_2 - 1, t) - \\ & - (\lambda + \mu_1 i_1 + \mu_2 i_2 - r_{11} i_1 \mu_1 - r_{22} i_2 \mu_2)P(i_1, i_2, t). \end{aligned} \quad (1)$$

Начальное условие для системы (1):

$$P(i_1, i_2, 0) = \begin{cases} 1, & \text{если } i_1 = i_2 = 0, \\ 0, & \text{если } i_1 > 0 \text{ или } i_2 > 0. \end{cases}$$

Определим общий вид производящей функции двумерного распределения в виде:

$$G(x, y, t) = \sum_{i_1=0}^{\infty} \sum_{i_2=0}^{\infty} P(i_1, i_2, t) x^{i_1} y^{i_2}. \quad (2)$$

Из системы (1), принимая во внимание вид производящей функции (2), можно получить линейное дифференциальное уравнение первого порядка для функции $G(x, y, t)$ в частных производных:

$$\begin{aligned} \frac{\partial G(x, y, t)}{\partial t} + \mu_1[(1 - r_{11})x - r_{12}y - (1 - r_{11} - r_{12})] \frac{\partial G(x, y, t)}{\partial x} + \\ + \mu_2[(1 - r_{22})y - r_{21}x - (1 - r_{21} - r_{22})] \frac{\partial G(x, y, t)}{\partial y} + \\ = \lambda[r_0 x + (1 - r_0)y - 1]G(x, y, t). \end{aligned} \quad (3)$$

Начальное условие для уравнения (3):

$$G(x, y, 0) = 1.$$

Решив уравнение (3), получим выражение для производящей функции нестационарного распределения числа занятых приборов в каждом из блоков системы:

$$G(x, y, t) = \exp \left\{ \lambda \left[A(x-1) \left(1 - \frac{(m-n+\sqrt{D})e^{-k_1 t} - (m-n-\sqrt{D})e^{-k_2 t}}{2\sqrt{D}} \right) + \right. \right. \\ \left. \left. + B(y-1) \left(1 - \frac{(m-n+\sqrt{D})e^{-k_1 t} + (m-n-\sqrt{D})e^{-k_2 t}}{2\sqrt{D}} \right) + \right. \right. \\ \left. \left. + \frac{B(x-1)}{\sqrt{D}} \mu_2 r_{21} (e^{-k_1 t} - e^{-k_2 t}) + \frac{A}{\sqrt{D}} \mu_1 r_{12} (y-1) (e^{-k_2 t} - e^{-k_1 t}) \right] \right\}, \quad (4)$$

где

$$A = \frac{r_0(1-r_{22}) + (1-r_0)r_{21}}{\mu_1((1-r_{22})(1-r_{11}) - r_{21}r_{12})}, \quad B = \frac{r_0r_{12} + (1-r_{11})(1-r_0)}{\mu_2((1-r_{22})(1-r_{11}) - r_{21}r_{12})}, \quad (5)$$

$$m = \mu_1(1-r_{11}), \quad n = \mu_2(1-r_{22}), \quad (6)$$

$$k_1 = \frac{1}{2} (m-n+\sqrt{D}), \quad k_2 = \frac{1}{2} (m-n-\sqrt{D}),$$

$$D = (m-n)^2 + 4\mu_1\mu_2r_{12}r_{21} > 0. \quad (7)$$

Применяя свойства производящей функции, вычислим математическое ожидание числа занятых приборов в блоках системы в момент времени t

$$M\{i_1(t)\} = \frac{\lambda}{\sqrt{D}} \left[A \left(\sqrt{D} - \frac{(m-n+\sqrt{D})e^{-k_1 t} - (m-n-\sqrt{D})e^{-k_2 t}}{2} \right) + \right. \\ \left. + B\mu_2r_{21} (e^{-k_1 t} - e^{-k_2 t}) \right], \quad (8)$$

$$M\{i_2(t)\} = \frac{\lambda}{\sqrt{D}} \left[B \left(\sqrt{D} - \frac{(m-n+\sqrt{D})e^{-k_1 t} - (m-n-\sqrt{D})e^{-k_2 t}}{2} \right) + \right. \\ \left. + A\mu_1r_{12} (e^{-k_1 t} - e^{-k_2 t}) \right], \quad (9)$$

где A и B определяются выражениями (10), m и n (6), а D - (7).

При $\Delta t \rightarrow \infty$ получаем вид производящей функции стационарного распределения числа заявок в первом и втором блоках:

$$G(x, y) = \exp\{\lambda[A(x-1) + B(y-1)]\}. \quad (10)$$

Заключение

На основе проведенного анализа найдена производящая функция (9) нестационарного распределения двумерного случайного процесса, описывающего число застрахованных и незастрахованных кредитных продуктов банка в момент времени t . При задании входящих параметров можно оценить среднее количество застрахованных и незастрахованных кредитов в любой момент времени t , используя формулы (8) и (9).

Можно также сделать вывод, что случайные процессы $\{i_1(t)\}$, $\{i_2(t)\}$, отражающие количество занятых приборов в обслуживающих блоках системы в стационарном режиме, независимы и следуют пуассоновскому распределению с параметрами A и B (10) для первого и второго блоков соответственно.

СПИСОК ЛИТЕРАТУРЫ

1. Жидкова Л. А., Моисеева С. П. Математическая модель потоков покупателей двухпродуктовой торговой компании в виде системы массового обслуживания с повторными обращениями к блокам // Известия Томского политехнического университета. 2013. Т. 322. № 6. С. 5–9.
2. Ивановская И. А., Моисеева С. П. Исследование математической модели параллельного обслуживания заявок смешанного типа // Известия Томского политехнического университета. 2010. Т. 317. № 5. С. 32–34.
3. Моисеева С. П., Морозова А. С. Исследование потока обращений в бесконечнолинейной СМО с повторным обслуживанием // Обработка данных и управление в сложных системах. Томск: Изд-во Том. ун-та, 2005. Вып. 7.
4. Назаров А. А. Теория вероятностей и случайных процессов: Учебное пособие. Томск: Изд-во НТЛ, 2006. 204 с.
5. Чегодаева Е. Г., Моисеева С. П., Полин Е. П. Математическая модель числа клиентов страховой компании в виде системы массового обслуживания с повторными обращениями // Математическое и компьютерное моделирование в экономике, страховании и управлении рисками. 2022. № 7. С. 136–139.

Кузнецов Никита Андреевич — студент ИМПКН. E-mail: nikita_kuznecov2004@mail.ru

Шкленник Мария Александровна — Канд. физ.-мат. наук, доцент кафедры ТВиМС, ИППКН. E-mail: shklennikm@yandex.ru

АНАЛИЗ СИСТЕМЫ СБОРКИ ФРАГМЕНТОВ ТРЕБОВАНИЙ, ОБСЛУЖЕННЫХ В АЦИКЛИЧЕСКОЙ СЕТИ МАССОВОГО ОБСЛУЖИВАНИЯ

О. С. Постнова, И. Е. Тананко, Е. С. Рогачко

*Саратовский национальный исследовательский государственный
университет имени Н. Г. Чернышевского, г. Саратов, Россия*

Рассматривается модель системы сборки требований из фрагментов, поступающих в систему, после завершения их обслуживания в ациклической открытой экспоненциальной сети массового обслуживания. Каждое требование разбивается на два фрагмента при поступлении в сеть. Система сборки представляется системой массового обслуживания с бесконечным числом приборов. Предполагается, что входящий в систему сборка поток фрагментов требований является пуассоновским, длительности сборки требований из фрагментов являются экспоненциально распределенными случайными величинами. Вычисляется математическое ожидание длительности сборки требований. Для проверки адекватности предложенной модели системы сборки используется метод имитационного моделирования.

Ключевые слова: *ациклическая сеть массового обслуживания, деление и слияние требований, фрагменты требований.*

Введение

Системы и сети массового обслуживания часто используются в качестве математических моделей дискретных стохастических систем с сетевой структурой. Одной из актуальных задач в настоящее время является построение и исследование имитационных моделей [1, 2], а также систем и сетей массового обслуживания с делением и слиянием требований [3, 5, 4, 6] для проектирования и анализа систем с многопутевой маршрутизацией, систем параллельной обработки информации, распределенных вычислительных систем.

В работе [3] были получены асимптотические выражения для стационарных вероятностей состояний двух параллельных систем массового обслуживания типа $M/M/1$ с делением требований на два фрагмента. На основе результатов работы [3] в работе [4] получено точное выражение для времени реакции. Для сети массового обслуживания, состоя-

щей из m систем типа $M/M/s$ с делением требований на m фрагментов, получены приближенные выражения для вычисления времени реакции [5]. В работе [6] рассматривается класс открытых сетей массового обслуживания с произвольной топологией, которые являются развитием fork-join сетей, состоящих из параллельных систем обслуживания.

В данной работе описывается модель системы сборки требований, каждое из которых состоит из двух фрагментов. Источником фрагментов для модели является открытая сеть обслуживания, в которую поступает пуассоновский поток разделяемых на фрагменты требований.

1. Описание модели

Рассматривается открытая экспоненциальная сеть массового обслуживания, состоящая из L систем S_i типа $M/M/1$ с интенсивностями обслуживания μ_i , $i = 1, \dots, L$. Из источника S_0 в сеть обслуживания с интенсивностью λ_0 поступает пуассоновский поток требований одного класса. Предполагается, что сеть — ациклическая, то есть структура сети представляется ориентированным ациклическим графом.

Каждое из поступающих требований состоит из двух фрагментов. В момент поступления требования в сеть это требование разделяется на два фрагмента, называемых родственными, и оба фрагмента независимо друг от друга поступают в одну из систем сети, с которыми связан источник. Фрагменты одного требования независимо друг от друга и независимо от фрагментов других требований переходят между системами сети до тех пор, пока не завершится их обслуживание в сети. После этого фрагменты требований поступают в систему S_{L+1} , называемую системой сборки. Родственные фрагменты, ранее принадлежавшие одному требованию, собираются в единое требование, которое покидает систему сборки. Переходы фрагментов между системами сети, источником и системой сборки осуществляются в соответствии с маршрутной матрицей $\Theta = (\theta_{ij})$, $i = 0, 1, \dots, L$, $j = 1, \dots, L+1$, где θ_{ij} — вероятность того, что после пребывания в системе S_i фрагмент перейдет в S_j .

Сеть массового обслуживания с введенными ранее параметрами систем, в которой обслуживаются фрагменты требований, без системы сборки обозначим через N . Элементы маршрутной матрицы $\bar{\Theta} = (\bar{\theta}_{ij})$, $i, j = 0, 1, \dots, L$, сети N определяются как $\bar{\theta}_{ij} = \theta_{ij}$, если $j \neq 0$, и $\bar{\theta}_{i,0} = \theta_{i,L+1}$. Предполагается, что в сети N число смежных для S_i , $i = 0, 1, \dots, L$ (кроме S_i , для которых $\bar{\theta}_{i,0} = 1$), систем обслуживания, в которые возможен переход фрагментов из S_i , намного больше двух (числа фрагментов требования), и вероятности $\bar{\theta}_{ij} > 0$, $j = 0, 1, \dots, L$, сравнимы. В этом случае выходящий из сети N поток фрагментов требований можно считать пуассоновским [7] с интенсивностью $2\lambda_0$.

Предполагается, что для сети N выполняется необходимое условие существования стационарного режима

$$\lambda_0 < \frac{1}{2} \min_{i=1, \dots, L} \frac{\mu_i \omega_0}{\omega_i} = a, \quad (1)$$

где вектор $\omega = (\omega_j)$, $j = 0, 1, \dots, L$, находится как решение уравнения $\omega = \omega \Theta$ с условием нормировки $\sum \omega_j = 1$.

Система сборки S_{L+1} состоит из бесконечного числа обслуживающих приборов. Назначение этой системы — сбор требований из родственных фрагментов, поступающих из сети N . При поступлении в систему сборки одного из родственных фрагментов — первого по времени поступления в систему, называемого «первым» фрагментом требования, — он занимает свободный прибор. В момент поступления «второго» из родственных фрагментов мгновенно происходит объединение фрагментов в единое требование, которое покидает систему сборки, и освобождение прибора. Таким образом, длительность сборки требования из родственных фрагментов совпадает с длительностью пребывания в системе «первого» фрагмента или длительностью интервала времени между моментами выхода родственных фрагментов из сети N .

Задачей данной работы является нахождение математического ожидания длительности сборки требований и стационарных вероятностей состояний системы сборки.

2. Приближенный анализ системы сборки

Предположим, что система сборки является системой массового обслуживания типа $M/M/\infty$. В систему поступает пуассоновский поток фрагментов с интенсивностью λ_0 , так как обслуживаются в системе только «первые» фрагменты.

Результаты имитационного моделирования показывают [2], что длительность сборки требования из фрагментов можно представить экспоненциально распределенной случайной величиной с математическим ожиданием $1/\mu$. При этом изменение параметра μ приближенно опишем убывающей линейной функцией интенсивности λ_0 , $\mu = \mu(\lambda_0)$, $\lambda_0 \in (0, a)$. Графиком функции $\mu(\lambda_0)$ является прямая, проходящая через две точки $(0, \nu)$ и $(a, 0)$, где ν — значение μ , получаемое при $\lambda_0 = \varepsilon$, где $\varepsilon > 0$ — малое число. Таким образом,

$$\mu = -\frac{\nu}{a} \lambda_0 + \nu, \quad \lambda_0 \in (0, a). \quad (2)$$

Найдем ν . При $\lambda_0 = \varepsilon$ математическое ожидание длительности сборки требования из фрагментов или, иначе, математическое ожидание

длительности интервала времени между родственными фрагментами в выходящем из сети N потоке фрагментов совпадает со среднеквадратическим отклонением σ_0 длительности реакции τ_0 сети N . Тогда

$$\nu = 1/\sigma_0, \quad \sigma_0 = \sqrt{E\tau_0^2 - (E\tau_0)^2}, \quad (3)$$

где $E\tau_0$ и $E\tau_0^2$ — первый и второй начальные моменты величины τ_0 .

Известно [8], что

$$E\tau_0 = \frac{1}{2\lambda_0} \sum_{i=1}^L \frac{\psi_i}{1 - \psi_i}, \quad (4)$$

где $\psi_i = \lambda_i/\mu_i$, $\lambda_i = 2\lambda_0\omega_i/\omega_0$ — интенсивность потока фрагментов требований в систему S_i , $i = 1, \dots, L$.

Обозначим τ_i — длительность пребывания в сети N фрагментов требований, поступивших в сеть через систему S_i , $i = 1, \dots, L$. Математические ожидания $E\tau_i$ могут быть найдены из системы уравнений:

$$E\tau_i = \frac{1}{\mu_i(1 - \psi_i)} + \sum_{j=1}^L \bar{\theta}_{ij} E\tau_j, \quad i = 1, \dots, L.$$

Вторые моменты $E\tau_i^2$ удовлетворяют следующей системе уравнений:

$$E\tau_i^2 = \frac{2E\tau_i}{\mu_i(1 - \psi_i)} + \sum_{j=1}^L \bar{\theta}_{ij} E\tau_j^2, \quad i = 1, \dots, L. \quad (5)$$

Из (5) получим

$$2\lambda_0 \sum_{i=1}^L \bar{\theta}_{0i} E\tau_i^2 = 2 \sum_{i=1}^L \frac{\lambda_i}{\mu_i(1 - \psi_i)} E\tau_i$$

или

$$E\tau_0^2 = \frac{1}{\lambda_0} \sum_{i=1}^L \frac{\psi_i}{1 - \psi_i} E\tau_i. \quad (6)$$

Подставив найденные по формулам (4) и (6) значения в формулу (3), найдем значение ν , используемое в формуле (2).

Обозначим P_k — стационарная вероятность пребывания системы сборки в состоянии k , $k = 0, 1, 2, \dots$, где k — число «первых» фрагментов в системе сборки. Очевидно, что

$$P_k = \frac{(\lambda_0/\mu)^k}{k!} e^{-\lambda_0/\mu}, \quad k = 0, 1, 2, \dots,$$

где μ вычисляется по формуле (2).

3. Численные результаты

Для проверки адекватности предложенной модели системы сборки была построена имитационная модель рассматриваемой сети массового обслуживания с системой сборки [2]. Результаты аналитического вычисления стационарного распределения системы сборки сравнивались с результатами имитационного моделирования. В качестве критерия близости распределений использовалось расстояние Колмогорова:

$$\Delta = \max_{i \geq 0} \left| \sum_{k=0}^i P_k - \sum_{k=0}^i \tilde{P}_k \right|,$$

где P_k и \tilde{P}_k — вероятности состояния k системы сборки, вычисленные с помощью аналитической и имитационной модели соответственно.

В табл. 1 приводятся расстояния Колмогорова для случая $L = 20$, $\mu^1 = (\mu_i^1) = (4.5, 3.8, 3.1, 4.8, 3.2, 3.0, 2.9, 2.8, 4.7, 3.8, 4.1, 4.2, 4.6, 4.8, 3.9, 4.2, 4.3, 3.3, 4.0, 5.0)$ и $\mu^2 = (\mu_i^2) = (4.1, 12.0, 3.9, 8.0, 3.2, 4.0, 10.0, 9.0, 4.3, 10.0, 11.0, 8.0, 6.0, 4.0, 4.5, 10.0, 3.8, 5.0, 7.0, 4.0)$, $\lambda_0^1 = 0.1a$ и $\lambda_0^2 = 0.5a$, удовлетворяющих условию (1), при двух видах сети N : I — «двухслойная» сеть параллельных систем и II — произвольная древовидная сеть. Из таблицы видно, что аппроксимация стационарного распределения системы сборки улучшается при уменьшении λ_0 .

Таблица 1

Значения расстояний Колмогорова

Вид сети	μ^1		μ^2	
	λ_0^1	λ_0^2	λ_0^1	λ_0^2
I	0.0075	0.1506	0.0102	0.1755
II	0.0471	0.0793	0.0265	0.1352

Заключение

Был предложен приближенный метод вычисления математического ожидания длительности сборки требований из фрагментов, обслуженных в ациклической сети массового обслуживания, и стационарного распределения системы сборки. Зная данные характеристики, можно вычислить другие стационарные характеристики рассматриваемой сети массового обслуживания.

СПИСОК ЛИТЕРАТУРЫ

1. Кельтон В., Лоу А. Имитационное моделирование. Классика CS. СПб.: Питер, 2004. 847 с.
2. Карпенко О. С., Тананко И. Е., Рогачко Е. С. Исследование имитационной модели открытой сети массового обслуживания с делением и слиянием требований // Системы управления, информационные технологии и математическое моделирование. Омск: Омский государственный технический университет, 2023. С. 209–214.
3. Flatto L., Hahn S. Two parallel queues created by arrivals with two demands I // SIAM Journal on Applied Mathematics. 1984. Vol. 44. № 5. P. 1041–1053.
4. Nelson R., Tantawi A. N. Approximate analysis of fork/join synchronization in parallel queues // IEEE Transactions on Computers. 1988. Vol. 37. № 6. P. 739–743.
5. Ko S. S., Serfozo R. Response times in $M/M/s$ fork-join networks // Advances in Applied Probability. 2004. Vol. 36. № 3. P. 854–871.
6. Осипов О. А., Тананко И. Е. Сети массового обслуживания произвольной топологии с делением и слиянием требований: случай бесконечноприборных систем обслуживания // Вестник Тверского государственного университета. Серия: Прикладная математика. 2017. № 4. С. 43–58.
7. Цициашвили Г. Ш., Осипова М. А. Стационарные потоки в ациклических сетях массового обслуживания // Дальневосточный математический журнал. 2016. № 2. С. 223–228.
8. Lemoine A. J. Networks of queues — a survey of equilibrium analysis // Management Science. 1977. Vol. 24. № 4. P. 464–481.

Постнова Оксана Сергеевна — магистрант, кафедра системного анализа и автоматического управления. E-mail: popova.np0062@gmail.com

Тананко Игорь Евстафьевич — кандидат физико-математических наук, доцент, заведующий кафедрой, кафедра системного анализа и автоматического управления. E-mail: TanancoIE@info.sgu.ru

Рогачко Екатерина Сергеевна — кандидат физико-математических наук, доцент, кафедра системного анализа и автоматического управления. E-mail: RogachkoES@info.sgu.ru

КОНЦЕПЦИЯ АРХИТЕКТУРЫ ПРИЛОЖЕНИЯ ДЛЯ ПОСТРОЕНИЯ АНАЛИТИЧЕСКОЙ АППРОКСИМАЦИИ РАСПРЕДЕЛЕНИЯ ВЕРОЯТНОСТЕЙ ЧИСЛА ЗАЯВОК В СМО НА ОСНОВЕ ИМИТАЦИОННОГО МОДЕЛИРОВАНИЯ

Ю. А. Прокудина, А. Н. Моисеев

*Национальный исследовательский Томский государственный университет,
г. Томск, Россия*

В статье представлена идея и концепция общей архитектуры приложения для построения аналитической аппроксимации распределения вероятностей числа заявок в системах массового обслуживания, которое выполняется на основе результатов серии экспериментов имитационного моделирования.

Ключевые слова: *имитационное моделирование, система массового обслуживания, аппроксимация распределения вероятностей*

Введение

При исследовании систем массового обслуживания (СМО) [1] одной из главных задач является получение в аналитическом виде распределения вероятностей числа заявок в системе, обычно в стационарном режиме работы. При этом из-за трудностей математического характера только в редких случаях удается получить точный аналитический результат. В некоторых случаях удается аналитически получить аппроксимации распределений, используя асимптотические методы [2], [3] или аппроксимации исходных распределений [4]. В остальных случаях возможно получение рекуррентных алгоритмов [5] либо чаще всего – провести численное и имитационное моделирование [6].

Имитационное моделирование позволяет получить численный результат практически для любой конфигурации СМО. Поэтому возникает идея – провести серию экспериментов по имитационному моделированию одной и той же СМО с разными значениями параметров, получить наиболее подходящие в каждом случае распределения вероятностей числа заявок в системе, а затем систематизировать эти результаты и попробовать оценить влияние каждого входного параметра на результирующую функцию распределения. Таким образом, в качестве

результата будет предложена аппроксимация функции распределения вероятностей числа заявок в системе в виде какой-то известной функции распределения, параметры которой определяются в терминах исходных параметров модели. Предполагается, что в приложении будет некоторый банк распределений, среди которых производится поиск.

В настоящее время в ИПМКН ТГУ ведется разработка программного комплекса SimQ [7], предназначенного для имитационного моделирования систем массового обслуживания. Предполагается использовать ядро этой системы для выполнения непосредственно процесса имитационного моделирования и получения эмпирических рядов распределения, а затем с помощью надстроенного каскада модулей построить аналитическую аппроксимацию числа заявок в системе. В данной работе предложена концепция архитектуры программного комплекса, который будет выполнять данную задачу.

1. Концепция архитектуры приложения

Чтобы реализовать описанный процесс, предлагается следующая концепция архитектуры разрабатываемого программного комплекса (Рис. 1). Мы выделяем пять основных частей системы: UI, Controller, Estimator, Ядро моделирования SimQ, Analytic.

Пакет UI предназначен для элементов, формирующих интерфейс пользователя, в котором он выбирает конфигурацию системы, задает некоторые параметры модели и параметры процесса построения аналитической аппроксимации. Здесь же находятся объекты, с помощью которых производится вывод результатов по окончании процесса построения.

Пакет Controller реализует функции управления процессом построения аппроксимации, многократно запускает имитационное моделирование выбранной системы (конфигурации СМО) для различных значений варьируемых параметров модели, получает от модуля Estimator соответствующие оценки распределений и передает их в модуль Analytic.

Работа пакета Estimator подробно описана в [8]. Его задачей является поиск наиболее близкого распределения вероятностей из имеющейся коллекции распределений и оценка его параметров. На самом деле, оценки строятся для каждого распределения из коллекции, если это возможно, при этом вычисляется метрика близости построенной оценки к эмпирическому распределению (в настоящий момент для этого используется расстояние Колмогорова).

Ядро моделирования SimQ [7] занимается непосредственно имитационным моделированием заданной СМО и выдает в качестве результата эмпирическое распределение вероятностей числа заявок в системе.

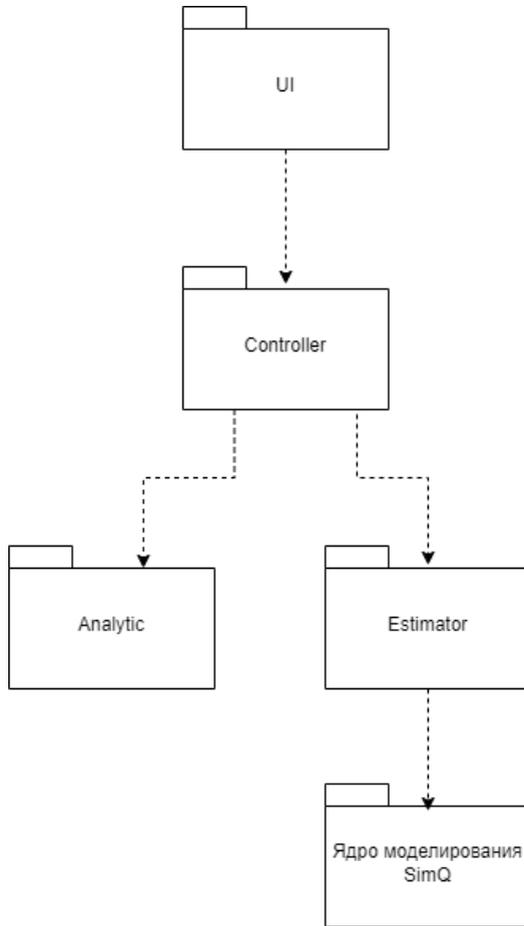


Рис. 1. Концепция архитектуры приложения для построения аналитической аппроксимации в SimQ

Наконец, задачей пакета Analytic является построение аппроксимации распределения вероятностей числа заявок в выбранной СМО в аналитической форме. Для этого используются выражения для распределений из имеющейся коллекции, а также построенные с помощью Estimator оценки распределений и их метрики близости к результатам имитационного моделирования.

В результате пользователь получает в аналитическом виде предполагаемое (аппроксимацию) распределение вероятностей числа заявок в

системе для выбранной модели СМО. В принципе, приложение может выдавать пользователю все аппроксимации, полученные для каждого элемента коллекции распределений, сортированные по значению некоторой агрегированной метрики близости.

2. Общий контур процесса построения аппроксимации

Рассмотрим общий ход предлагаемого процесса построения аналитической аппроксимации распределения вероятностей в разрабатываемой системе (Рис. 2).

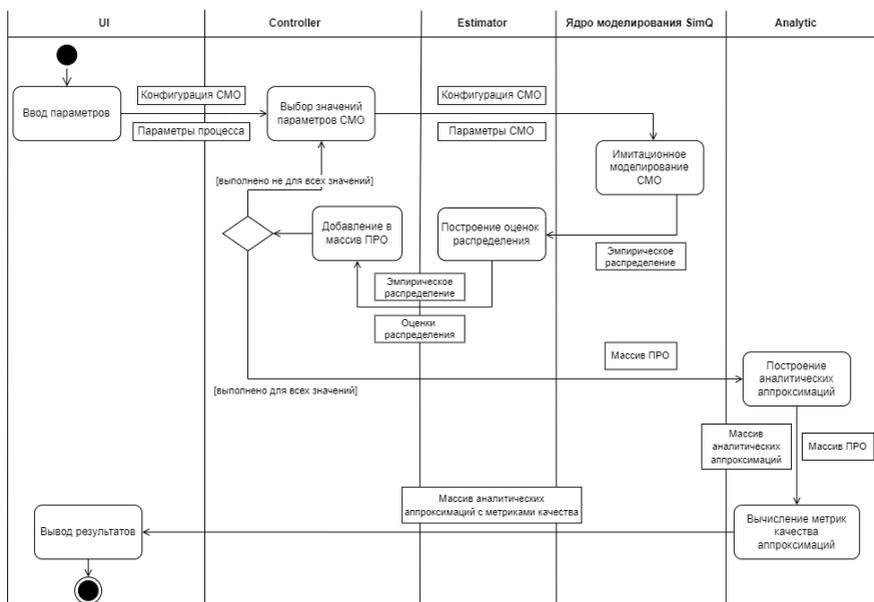


Рис. 2. Процесс построения аналитической аппроксимации

Пользователь выбирает конфигурацию (тип) СМО, в некоторых случаях задает значения неизменяемых параметров модели, а также параметры самого процесса построения. Эти данные поступают в Controller, который в цикле перебирает все значения изменяемых параметров модели, необходимые для построения аналитических выражений, запускает для них имитационное моделирование соответствующей СМО в Ядре моделирования SimQ, передает получаемые при этом эмпирические распределения модулю Estimator, который строит для них оценки распределения и вычисляет значения метрики качества для каждой

оценки. Эти оценки и соответствующие значения метрики записываются в Массив ПРО («Параметры-Распределение-Оценки»), который индексируется по выбранным значениям параметров СМО. Этот процесс продолжается до тех пор, пока не будут выбраны все необходимые значения варьируемых параметров модели.

Далее сформированный Массив ПРО передается модулю Analytic, который на его основе строит аналитические аппроксимации распределения вероятностей числа заявок в СМО выбранной пользователем конфигурации, а также для каждой такой аппроксимации вычисляет некоторую агрегированную метрику качества, например, используя значения метрики качества оценок распределений, полученных ранее и хранящихся в Массиве ПРО.

Результат выдается пользователю в виде аналитических выражений аппроксимаций для распределения вероятностей, выполненных в терминах исходных (варьируемых) параметров СМО.

Заключение

Таким образом, в результате работы сформулированы общая концепция архитектуры приложения и основной контур построения аналитической аппроксимации функции распределения числа заявок в СМО, конфигурацию которой определяет пользователь. Далее планируется реализовать рабочие прототипы модулей данной системы и выполнить испытание их работы в составе программного комплекса (прототипы модулей Ядро моделирования SimQ и Estimator уже реализованы).

СПИСОК ЛИТЕРАТУРЫ

1. Назаров А. А. Теория массового обслуживания: учебное пособие. – 2-е изд., испр. Томск: Изд-во НТЛ, 2010. 228 с.
2. Назаров А. А., Моисеева С. П. Метод асимптотического анализа в теории массового обслуживания. Томск: Изд-во НТЛ, 2006. 112 с.
3. Моисеев А. Н., Назаров А. А. Бесконечнолинейные системы и сети массового обслуживания. Томск: Изд-во НТЛ, 2015. 240 с.
4. Moiseev A. N., Nazarov A. A., Paul S. V. Asymptotic Diffusion Analysis of Multi-Server Retrial Queue with Hyper-Exponential Service // Mathematics. 2020. V. 8. № 4. P. 531. 16 p.
5. Полховская А. В., Данилюк Е. Ю., Моисеева С. П., Бобкова О. С. Вероятностная модель системы совместного доступа с коллизиями, Н-настойчивостью и отказами // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2022. № 58. С. 35–46.

6. *Приступа И. О., Сущенко И. О., Михеев И. О., Моисеев И. О.* Анализ эффективности прямой коррекции ошибок на уровне транспортного протокола // Автоматика и телемеханика. 2022. № 7. С. 59–81.
7. *Прокудина Ю. А., Оруджов Э. А., Моисеев А. Н.* Концепция архитектуры программного комплекса SimQ // Материалы X-й Международной молодежной научной конференции "Математическое и программное обеспечение информационных, технических и экономических систем Томск, 26-29 мая 2023 г., Томск: Труды Томского государственного университета; т. 308 : Серия физико-математическая., 2023. С. 155–159.
8. *Прокудина Ю. А., Моисеев А. Н.* Прототип приложения для поиска распределения, наиболее близкого к заданному эмпирическому // Материалы XI-й Международной молодежной научной конференции "Математическое и программное обеспечение информационных, технических и экономических систем Томск, 24-27 мая 2024 г., Томск: Труды Томского государственного университета. (в печати)

Прокудина Юлия Андреевна — аспирант, НИ ТГУ, ИПМКН, кафедра программной инженерии. E-mail: *prokudina_ua@mail.ru*

Моисеев Александр Николаевич — доктор физико-математических наук, доцент, НИ ТГУ, ИПМКН, кафедра программной инженерии. E-mail: *moiseev.tsu@gmail.com*

Математическое моделирование

PURSUIT PROBLEM IN A DIFFERENTIAL GAME WITH INERTIAL PLAYERS UNDER INTEGRAL CONSTRAINTS ON CONTROLS

B. Samatov¹, M. Turgunboeva²

¹ *V. I. Romanovskiy Institute of Mathematics, Uzbekistan Academy of Sciences, Tashkent, Uzbekistan*

² *Namangan State University, Namangan, Uzbekistan*

We investigate a pursuit problem in a differential game with two inertial players, where the controls of both players are subject to integral constraints. The pursuit problem is addressed by implementing a parallel pursuit strategy and sufficient conditions for the solvability of the pursuit problem are derived. Furthermore, in the case of the same dynamics of the players, an explicit formula for a meeting domain of the players is provided by the main lemma characterizing its monotonicity property. **Keywords:** *differential game, Π -strategy, pursuit, acceleration, integral constraint, meeting domain.*

Introduction

Pursuit-evasion problems are particularly important in the field of differential games due to several qualities. This quality is demonstrably effective in real-world applications. A prime example is “the game with a Life-line”, which was pioneered by R. Isaacs (see [1], problem 9.5.1), featuring players’ simple dynamics. L.A. Petrosjan analyzed this game and established an exclusive strategy named “parallel pursuit strategy” (Π -strategy) that was proved highly effective in work [2]. This success led to the Π -strategy’s applications in solving broader classes of differential pursuit games (see, [2–5]).

Belousov [6] introduced the resolving functions method for differential games with integral constraints on the control functions, which gives a sufficient condition for ending the pursuit. This method was later improved for problems of group pursuit in differential games with integral constraints imposed on players’ controls, such as the Pontryagin control example and group pursuit with a simple motion for the case of ℓ -catch [7–8].

1. Main results

We analyze a differential game with two conflict-controlled inertial players moving in the Euclidian space \mathbf{R}^n . One of them, called the Pursuer

(denoted by \mathbb{P}), chases another one, called the Evader (denoted by \mathbb{E}). We represent their positions using vectors x and y in \mathbf{R}^n . The motions of dynamic system of the players is characterized by the following second-order differential equations, which are accompanied by their respective initial conditions:

$$\mathbb{P} : \quad \ddot{x} = u, \quad x(0) = x_{10}, \quad \dot{x}(0) = x_{11}, \quad (1)$$

$$\mathbb{E} : \quad \ddot{y} = v, \quad y(0) = y_{10}, \quad \dot{y}(0) = y_{11}, \quad (2)$$

where $x, y, u, v \in \mathbf{R}^n, n \geq 2$; x_{10}, y_{10} are the initial positions of the players, and x_{11}, y_{11} are their initial velocity vectors, respectively. It is assumed that $x_{10} \neq y_{10}$ at the time $t = 0$. The parameters u and v are the acceleration vectors which are utilized as controls of the players.

Definition 1. Measurable functions $u(\cdot) = (u(t), t \geq 0)$ and $v(\cdot) = (v(t), t \geq 0)$ are called the admissible controls of Pursuer \mathbb{P} and Evader \mathbb{E} if they meet the following integral constraints

$$\int_0^t (t-s)|u(s)|^2 ds \leq \rho_0^2, \quad t \geq 0, \quad (3)$$

$$\int_0^t (t-s)|v(s)|^2 ds \leq \sigma_0^2, \quad t \geq 0, \quad (4)$$

respectively, where ρ_0 and σ_0 are given positive parametric numbers that represent the maximum resource values of the players.

We denote the set of all admissible controls $u(\cdot)$ of the Pursuer by \mathbb{U}_I and the set of all admissible controls $v(\cdot)$ of the Evader by \mathbb{V}_I .

Definition 2. For each pair $(\rho_0, u(\cdot)), u(\cdot) \in \mathbb{U}_I$ and $(\sigma_0, v(\cdot)), v(\cdot) \in \mathbb{V}_I$, it is said that the quantities

$$\rho(t) = \rho_0^2 - \int_0^t (t-s)|u(s)|^2 ds, \quad \rho(0) = \rho_0^2,$$

$$\sigma(t) = \sigma_0^2 - \int_0^t (t-s)|v(s)|^2 ds, \quad \sigma(0) = \sigma_0^2$$

are the residual resources of the players at the current time $t, t \geq 0$.

Definition 3. If $u(\cdot) \in \mathbb{U}_I$ and $v(\cdot) \in \mathbb{V}_I$, then by virtue of equations (1) and (2), the triplets $(x_{10}, x_{11}, u(\cdot))$ and $(y_{10}, y_{11}, v(\cdot))$ generate the trajectories of the players as follows:

$$x(t) = x_{10} + x_{11}t + \int_0^t (t-s)u(s)ds,$$

$$y(t) = y_{10} + y_{11}t + \int_0^t (t-s)v(s)ds,$$

respectively.

The primary objective of Pursuer \mathbb{P} is to catch Evader \mathbb{E} , i.e., to specifically achieve the equality $x(t_*) = y(t_*)$ at some $t_* > 0$. For Evader \mathbb{E} , the main goal is to maintain the relation $x(t) \neq y(t)$ for each $t \in [0, +\infty)$, and if this is impossible, then put back the occurrence of capture. Additionally, we suppose that the initial positions x_{10}, y_{10} satisfy the conditions $x_{10} \neq y_{10}$ at the start of the game.

Let's introduce new denotations:

$$z = x - y, \quad z_{10} = x_{10} - y_{10}, \quad z_{11} = x_{11} - y_{11}.$$

First of all, we consider the differential game (1)–(4) in the case of the same dynamic of the players, i.e., $x_{11} = y_{11}$

Definition 4. In the differential game (1)–(4), the function

$$\mathbf{u}_1(v) = v - \gamma(v)\xi_{10} \tag{5}$$

is called the Π -strategy of the Pursuer for the case of $x_{11} = y_{11}$, where

$$\gamma(v) = \max\{0, \delta + 2\langle v, \xi_{10} \rangle\}, \quad \delta = (\rho_0^2 - \sigma_0^2)/|z_{10}|, \quad \xi_{10} = z_{10}/|z_{10}|.$$

and $\mathbf{u}_1(v(t))$ ($t \geq 0$) - its realization for any $v(\cdot) \in \mathbb{V}_I$.

Theorem 1. If $\rho_0 > \sigma_0$, then the Pursuer wins by using Π -strategy (5) on the time interval $[0, T_1]$, where $T_1 = \sqrt{2}|z_{10}|/(\rho_0 - \sigma_0)$.

Let the quadruplets (x, y, ρ, σ) describe the current state of the game at some moment in time t , where $x \neq y$. The set $\Omega(x, y, \rho, \sigma)$ is defined as the set of all points ω where the Pursuer moving from the position x and consuming the resource ρ should encounter the Evader moving from the position y and consuming the resource σ .

Referring to Theorem 1, if $\rho_0 > \sigma_0$, then by virtue of the Π -strategy (5) the Evader is captured at some point in the space \mathbf{R}^n . The players \mathbb{P} and \mathbb{E} will meet at various points according to the choice of the control $v(\cdot) \in \mathbb{V}_I$.

So, we will construct a set of meeting points of the players for the case $\rho_0 > \sigma_0$.

We define the following multi-valued mapping

$$\Omega(x(t), y(t), \rho(t), \sigma(t)) = \left\{ \omega : |\omega - x(t)| \geq \sqrt{\rho(t)/\sigma(t)}|\omega - y(t)| \right\}, \quad (6)$$

as long as $\sigma(t) > 0$ on $[0, t_*]$, $t_* \in [0, T_1]$, where t_* is the encounter time of the players, which the equality $x(t_*) = y(t_*)$ holds and, $\rho(t)$ and $\sigma(t)$ are the residual resources of the players as defined in Definition 2. Note that

$$\Omega(x_{10}, y_{10}, \rho_0, \sigma_0) = \{ \omega : |\omega - x_{10}| \geq \rho_0/\sigma_0|\omega - y_{10}| \}. \quad (7)$$

Lemma 1. The multi-valued mapping $\Omega(x(t), y(t), \rho(t), \sigma(t)) - tx_{11}$, $t \in [0, t_*]$, is monotonically decreasing with respect to the inclusion, i.e. if $t_1, t_2 \in [0, t_*]$ and $t_1 < t_2$, then

$$\Omega(x(t_1), y(t_1), \rho(t_1), \sigma(t_1)) - t_1x_{11} \supset \Omega(x(t_2), y(t_2), \rho(t_2), \sigma(t_2)) - t_2x_{11}.$$

Corollary 1. From Lemma 1, it can be concluded that:

- a) $\Omega(x(t), y(t), \rho(t), \sigma(t)) \subset \Omega(x_{10}, y_{10}, \rho_0, \sigma_0) + tx_{11}$ at each $t \in [0, t_*]$;
- b) $y(t) \in \Omega(x_{10}, y_{10}, \rho_0, \sigma_0) + tx_{11}$ for all $t \in [0, t_*]$ are met.

Definition 5. The set

$$\Omega^*(x_{10}, y_{10}, \rho_0, \sigma_0, T_1) = \bigcup_{t=0}^{T_1} \{ \Omega(x_{10}, y_{10}, \rho_0, \sigma_0) + tx_{11} \} \quad (8)$$

is called the meeting domain of the players in the pursuit problem, where T_1 is the guaranteed capture time (see Theorem 1).

Let W be the closed subset of \mathbf{R}^n , and let the differential game (1)–(4) with the “Life-line” W be studied.

Theorem 2. If the meeting domain of the players doesn’t intersect with the set W , i.e. $\Omega^*(x_{10}, y_{10}, \rho_0, \sigma_0, T_1) \cap W = \emptyset$ in the “Life-line” game, then the Pursuer wins by using Π -strategy (5) on the time interval $[0, T_1]$.

In the next step, we consider the differential game (1)–(4) for the case where the vectors z_{10} and z_{11} are collinear, which means there is a finite number k such that

$$z_{11} = kz_{10}, \quad k \in \mathbf{R} \setminus \{0\}.$$

Lemma 2. Let $\rho_0 \geq \sigma_0$ and $k < 0$. Then there exists at least a positive root of the following equation

$$\Lambda_*(t) = 0$$

and we denote by T_* the smallest positive root, where

$$\Lambda_*(t) = 1 + kt - \frac{t}{|z_{10}|} \left[\frac{\delta}{4}t - \frac{\sigma_0}{\sqrt{2}} + \sqrt{\left(\frac{\delta}{4}t - \frac{\sigma_0}{\sqrt{2}}\right)^2 - \frac{1}{2}k\delta|z_{10}|t} \right].$$

Lemma 3. Let $\rho_0 > \sigma_0 + \sqrt{2}k|z_{10}|$ and $k > 0$. Then there exists at least positive root of the following equation

$$\Lambda^*(t) = 0$$

and we denote by T^* the smallest positive root, where

$$\Lambda^*(t) = 1 + kt - \frac{1}{|z_{10}|} \max \left\{ 0, \frac{\delta}{2(1+kt)}t^2 - \sqrt{2}\sigma_0t \right\},$$

$$T^* = \frac{\sqrt{2}|z_{10}|}{\rho_0 - \sigma_0 - \sqrt{2}k|z_{10}|}.$$

Next, we will define the Π -strategy based on the existing works [3–5]. Let us introduce the following denotation:

$$\lambda(v) = \begin{cases} \lambda_*(v) & \text{if } k < 0, \rho_0 \geq \sigma_0, \\ \lambda^*(v) & \text{if } k > 0, \rho_0 > \sigma_0 + \sqrt{2}k|z_{10}|, \end{cases} \tag{9}$$

where

$$\lambda_*(v) = \frac{\delta}{2} + \langle v, \xi_{10} \rangle + \sqrt{\left(\frac{\delta}{2} + \langle v, \xi_{10} \rangle\right)^2 - \frac{2}{T_*}k\delta|z_{10}|},$$

$$\lambda^*(v) = \max \left\{ 0, \frac{\delta}{1+kT^*} + 2\langle v, \xi_{10} \rangle \right\},$$

$\delta = (\rho_0^2 - \sigma_0^2)/|z_{10}|$ and $\xi_{10} = z_{10}/|z_{10}|$, T_* and T^* are defined in Lemma 2 and Lemma 3, respectively.

Definition 6. The control function

$$\mathbf{u}_2(v) = v - \lambda(v)\xi_{10}, \tag{10}$$

is called the Π -strategy of the Pursuer in the differential game (1)–(4) for the case of $z_{11} = kz_{10}$.

Theorem 3. Let Lemma 2 and Lemma 3 be met. Then the Pursuer wins by employing the Π -strategy (10) during the time interval $[0, T_2]$, where

$$T_2 = \begin{cases} T_* & \text{if } k < 0, \rho_0 \geq \sigma_0, \\ T^* & \text{if } k > 0, \rho_0 > \sigma_0 + \sqrt{2}k|z_{10}|. \end{cases}$$

Conclusion

In conclusion, we can say that the pursuit problem is solved in a differential game with inertial players under integral constraints on controls by applying the II-strategy, and the meeting domain of the players is constructed for the case of the same dynamics of the players. The proposed method can be generalized to more general linear and nonlinear conflict-controlled systems with various constraints on controls in the future.

REFERENCES

1. *Isaacs R.* Differential games. New York: John Wiley and Sons, 1965.
2. *Pshenichnii B.* Simple pursuit by several objects // Cybernetics and System Analysis. 1976. Vol. 12. № 5. P. 484–485.
3. *Petrosjan L.* Differential games of pursuit. Series on optimization. Singapore: World Scientific Publishing, 1993.
4. *Azamov A.* About the quality problem for games pursuit with the restriction // Serdica Bulgariacae math. Publ. Sofia. 1986. Vol. 2. № 12. P. 38–43.
5. *Azamov A., Samatov B.* The II-Strategy: Analogies and Applications // The Fourth International Conference on Game Theory and Management, Petersburg, 2010. Vol. 4. P. 33–47.
6. *Belousov A.* Method of resolving functions for differential games with integral constraints // Theory of Optimal Solution. 2010. Vol. 9. P. 10–16.
7. *Samatov B.* Problems of group pursuit with integral constraints on controls of the players I // Cybernetics and Systems Analysis. 2013. Vol. 49. № 5 P. 786–767.
8. *Samatov B.* Problems of group pursuit with integral constraints on controls of the players II // Cybernetics and Systems Analysis. 2013. Vol. 49. № 6 P. 907–921.

Samatov Bahrom — DSc, professor, senior researcher, Department of Dynamic and Controllable Systems. E-mail: *samatov57@gmail.com*

Turgunboeva Mokhisanam — PhD student, teacher, Mathematics. E-mail: *turgunboevamokhisanam95@gmail.com*

ON AUTOMORPHISMS OF A GENETIC ALGEBRA ASSOCIATED BY A NON-VOLTERRA OPERATOR

M. Ibragimov

Andijan State University, Andijan, Uzbekistan

The aim of this paper is to examine the automorphisms of genetic algebras associated with permuted Volterra operators defined on three-dimensional simplex. Furthermore, three dimensional permuted Volterra algebras are considered and the description of their automorphisms are given as well. **Keywords:** *genetic algebra, non-Volterra, quadratic stochastic operator, automorphisms.*

Introduction

The work of Bernstein [1] has pioneered the application of quadratic stochastic operators (QSO), and in general, these operators are employed within the domain of biology especially in the presentation of time evolution of countless species. There are various established classes of non-associative algebras and so forth, and all these have significantly contributed to the theoretical population genetics (see e.g. [3]). One class is named by genetic algebra which is concerned with the mathematical modeling of genetic processes, inheritance patterns, and genetic diversity using algebraic structures and techniques. The application of a type of genetic algebra called evolution algebra was demonstrated in Tian [4].

In genetic algebra, each mapping determines the character of the next generation. These are cases where parental traits and characteristics are retained or are not rediscovered in a new generation. The mappings we study are different, and automorphisms occupy a special place among them.

Automorphisms are one-to-one mappings of an algebra into itself and play an important role in genetic algebra. Understanding automorphisms is crucial in algebraic structures as they provide insight into the symmetries and invariant properties of these structures under transformations that preserve their algebraic operations. In [2] automorphisms in Volterra algebras are studied. It is natural question: what about automorphism in genetic algebras associated by non-Volterra operators? Below we shall particularly answer to this question and present preliminary results.

1. Main part

Recall the definition of an evolution algebra of a free population following [3]. Let $E = \{1, \dots, m\}$. By \mathbf{e}_i , $i = 1, \dots, m$ we denote the standard basis

in \mathbb{R}^m . The convex hull of these basis elements is $(m - 1)$ - dimensional simplex, that is

$$S^{m-1} = \left\{ \mathbf{x} \in \mathbb{R}^m : x_i \geq 0, \forall i \in E, \sum_{i \in E} x_i = 1 \right\}.$$

A *quadratic stochastic operator* (QSO) is a mapping $V : S^{m-1} \rightarrow S^{m-1}$ of the form

$$V : x'_k = \sum_{i,j \in E} p_{ij,k} x_i x_j, \quad k \in E \tag{1}$$

where $p_{ij,k}$ are heredity coefficients, which satisfy the following conditions:

$$p_{ij,k} \geq 0, \quad p_{ij,k} = p_{ji,k}, \quad \sum_{k=1}^m p_{ij,k} = 1, \quad i, j, k \in E \tag{2}$$

Let V be a QSO and let $\mathbf{x} = (x_1, \dots, x_m)$, $\mathbf{y} = (y_1, \dots, y_m) \in \mathbb{R}^m$ are arbitrary vectors, we introduce a multiplication rule (see [2]) on \mathbb{R}^m by

$$(\mathbf{x} \circ_V \mathbf{y})_k = \sum_{i,j \in E} p_{ij,k} x_i y_j, \quad k \in E. \tag{3}$$

Using (2), (3) it is easy to see that $\mathbf{x} \circ_V \mathbf{y} = \mathbf{y} \circ_V \mathbf{x}$, i.e. the multiplication is commutative. In general, the genetic algebra is not necessarily to be associative.

Definition 1. A linear transformation \mathbf{A} of the algebra \mathcal{A}_V is an *endomorphism* if $\mathbf{A}(\mathbf{x} \circ_V \mathbf{y}) = \mathbf{A}(\mathbf{x}) \circ_V \mathbf{A}(\mathbf{y})$ for all $\mathbf{x}, \mathbf{y} \in \mathcal{A}_V$. We say that a bijective endomorphism \mathbf{A} of the algebra \mathcal{A}_V is an *automorphism* of the algebra \mathcal{A}_V .

Let us consider the following Volterra QSO defined on the S^2

$$V : \begin{cases} x'_1 = x_1^2 + (1 - a)x_1x_2 + (1 + c)x_1x_3, \\ x'_2 = x_2^2 + (1 + a)x_1x_2 + (1 - b)x_2x_3, \\ x'_3 = x_3^2 + (1 - c)x_1x_3 + (1 + b)x_2x_3, \end{cases} \tag{4}$$

where $-1 \leq a, b, c \leq 1$. Let us fix the following permutation of the set $E = \{1, 2, 3\}$

$$\pi = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}.$$

From [5] we know that the operator $V_\pi = T_\pi V$ is a quadratic homeomorphisms of the two-dimensional simplex, where $T_\pi(\mathbf{x}) = (x_{\pi(1)}, x_{\pi(2)}, x_{\pi(3)})$.

Obviously that if we take $\pi = Id$, then V_π will be a Volterra quadratic operator. The following theorem is valid in the algebra defined by the Volterra quadratic operator.

Theorem 1. [2] Let $\mathcal{A}_V(0, 0, 0)$ be the Lotka-Volterra algebra associated by (4). Then $Aut(\mathcal{A}_V(0, 0, 0))$ is the set of all linear automorphisms σ of \mathcal{A}_V such that $\sigma(H) = H$, where $H = \{\mathbf{x} \in \mathcal{A} : x_1 + x_2 + x_3 = 1\}$.

Another cases the operator V_π will be non-Volterra operator, which is called permuted Volterra operator.

Now, by (4) the permuted Volterra operator V_π is defined as follows:

$$V_\pi : \begin{cases} x'_1 = x_2^2 + (1 + a)x_1x_2 + (1 - b)x_2x_3, \\ x'_2 = x_3^2 + (1 - c)x_1x_3 + (1 + b)x_2x_3, \\ x'_3 = x_1^2 + (1 - a)x_1x_2 + (1 + c)x_1x_3, \end{cases} \quad (5)$$

where $-1 \leq a, b, c \leq 1$.

Then, by (5) the corresponding algebra $\mathcal{A}_{V_\pi} := \mathcal{A}_{V_\pi}(a, b, c)$ has the following multiplication table:

$$\begin{aligned} \mathbf{e}_1 \circ_{V_\pi} \mathbf{e}_1 &= \mathbf{e}_3 & \mathbf{e}_1 \circ_{V_\pi} \mathbf{e}_2 &= \mathbf{e}_2 \circ_{V_\pi} \mathbf{e}_1 = \frac{1+a}{2}\mathbf{e}_1 + \frac{1-a}{2}\mathbf{e}_3 \\ \mathbf{e}_2 \circ_{V_\pi} \mathbf{e}_2 &= \mathbf{e}_1 & \mathbf{e}_1 \circ_{V_\pi} \mathbf{e}_3 &= \mathbf{e}_3 \circ_{V_\pi} \mathbf{e}_1 = \frac{1-c}{2}\mathbf{e}_2 + \frac{1+c}{2}\mathbf{e}_3 \\ \mathbf{e}_3 \circ_{V_\pi} \mathbf{e}_3 &= \mathbf{e}_2 & \mathbf{e}_2 \circ_{V_\pi} \mathbf{e}_3 &= \mathbf{e}_2 \circ \mathbf{e}_3 = \frac{1-b}{2}\mathbf{e}_1 + \frac{1+b}{2}\mathbf{e}_2 \end{aligned} \quad (6)$$

Now we consider the algebra $\mathcal{A}_{V_\pi}(0, 0, 0)$ and other algebras will be studied in our subsequent paper.

Theorem 2. Let $\mathbf{A} : \mathcal{A}_{V_\pi}(0, 0, 0) \rightarrow \mathcal{A}_{V_\pi}(0, 0, 0)$ be a non-zero linear map defined as

$$\begin{aligned} \mathbf{A}(\mathbf{e}_1) &= \alpha\mathbf{e}_1 + \beta\mathbf{e}_2 + \gamma\mathbf{e}_3, \\ \mathbf{A}(\mathbf{e}_2) &= \gamma\mathbf{e}_1 + \alpha\mathbf{e}_2 + \beta\mathbf{e}_3, \\ \mathbf{A}(\mathbf{e}_3) &= \beta\mathbf{e}_1 + \gamma\mathbf{e}_2 + \alpha\mathbf{e}_3, \end{aligned}$$

where $\alpha, \beta, \gamma \in \mathbb{R}$. Then \mathbf{A} is an automorphism on the algebra \mathcal{A}_{V_π} if and only if $\alpha + \beta + \gamma = 1$ and $\alpha \neq \beta \neq \gamma$.

Proof. Let a linear map \mathbf{A} be an automorphism of the algebra \mathcal{A}_{V_π} and $\mathbf{A}(\mathbf{e}_i) = a_{i1}\mathbf{e}_1 + a_{i2}\mathbf{e}_2 + a_{i3}\mathbf{e}_3$, $i = 1, 2, 3$, that is A is the corresponding matrix to linear map \mathbf{A} . From the equation

$$\mathbf{A}(\mathbf{e}_i \circ_{V_\pi} \mathbf{e}_j) = \mathbf{A}(\mathbf{e}_i) \circ_{V_\pi} \mathbf{A}(\mathbf{e}_j), \quad i, j = 1, 2, 3$$

we have the following system

$$\left\{ \begin{array}{l} a_{12}(a_{11} + a_{12} + a_{13}) = a_{31}, \\ a_{13}(a_{11} + a_{12} + a_{13}) = a_{32}, \\ a_{11}(a_{11} + a_{12} + a_{13}) = a_{33}, \\ a_{22}(a_{21} + a_{22} + a_{23}) = a_{11}, \\ a_{23}(a_{21} + a_{22} + a_{23}) = a_{12}, \\ a_{21}(a_{21} + a_{22} + a_{23}) = a_{13}, \\ a_{32}(a_{31} + a_{32} + a_{33}) = a_{21}, \\ a_{33}(a_{31} + a_{32} + a_{33}) = a_{22}, \\ a_{31}(a_{31} + a_{32} + a_{33}) = a_{23}, \\ a_{11} + a_{31} = a_{12}(a_{21} + a_{22} + a_{23}) + a_{22}(a_{11} + a_{12} + a_{13}), \\ a_{12} + a_{32} = a_{13}(a_{21} + a_{22} + a_{23}) + a_{23}(a_{11} + a_{12} + a_{13}), \\ a_{13} + a_{33} = a_{11}(a_{21} + a_{22} + a_{23}) + a_{21}(a_{11} + a_{12} + a_{13}), \\ a_{21} + a_{31} = a_{12}(a_{31} + a_{32} + a_{33}) + a_{32}(a_{11} + a_{12} + a_{13}), \\ a_{22} + a_{32} = a_{13}(a_{31} + a_{32} + a_{33}) + a_{33}(a_{11} + a_{12} + a_{13}), \\ a_{23} + a_{33} = a_{11}(a_{31} + a_{32} + a_{33}) + a_{31}(a_{11} + a_{12} + a_{13}), \\ a_{11} + a_{21} = a_{22}(a_{31} + a_{32} + a_{33}) + a_{32}(a_{21} + a_{22} + a_{23}), \\ a_{12} + a_{22} = a_{23}(a_{31} + a_{32} + a_{33}) + a_{33}(a_{21} + a_{22} + a_{23}), \\ a_{13} + a_{23} = a_{21}(a_{31} + a_{32} + a_{33}) + a_{31}(a_{21} + a_{22} + a_{23}). \end{array} \right. \quad (7)$$

Adding the first (resp. second and third) three equations of the system (7) we get

$$\begin{aligned} (a_{11} + a_{12} + a_{13})^2 &= a_{31} + a_{32} + a_{33}, \quad (a_{21} + a_{22} + a_{23})^2 = a_{11} + a_{12} + a_{13} \\ (a_{31} + a_{32} + a_{33})^2 &= a_{21} + a_{22} + a_{23} \Rightarrow \\ (a_{11} + a_{12} + a_{13})(1 - (a_{11} + a_{12} + a_{13})^7) &= 0 \Rightarrow \\ a_{11} + a_{12} + a_{13} &= 0 \quad \text{or} \quad a_{11} + a_{12} + a_{13} = 1, \\ a_{21} + a_{22} + a_{23} &= 0 \quad \text{or} \quad a_{21} + a_{22} + a_{23} = 1, \\ a_{31} + a_{32} + a_{33} &= 0 \quad \text{or} \quad a_{31} + a_{32} + a_{33} = 1. \end{aligned}$$

From the system (7) one has that if $a_{i1} + a_{i2} + a_{i3} = 0$ is true for a $i = 1, 2, 3$ then it holds $a_{ij} = 0$ for all $i, j = 1, 2, 3$.

Also if $a_{i1} + a_{i2} + a_{i3} = 1$ holds for all $i = 1, 2, 3$ then we have $a_{11} = a_{22} = a_{33} = \alpha$, $a_{12} = a_{23} = a_{31} = \beta$ and $a_{13} = a_{32} = a_{21} = \gamma$, where

$\alpha, \beta, \gamma \in \mathbb{R}$ such that $\alpha + \beta + \gamma = 1$. Further the determinant of the matrix of \mathbf{A} is equal

$$\det(A) = \alpha^3 + \beta^3 + \gamma^3 - 3\alpha\beta\gamma = (\alpha + \beta + \gamma)(\alpha^2 + \beta^2 + \gamma^2 - \alpha\beta - \alpha\gamma - \beta\gamma)$$

$$\alpha^2 + \beta^2 + \gamma^2 - \alpha\beta - \alpha\gamma - \beta\gamma \geq 0,$$

where we have used well known rearrangement inequality and the equality case holds when $\alpha = \beta = \gamma$. Using $\alpha + \beta + \gamma = 1$ we obtain that $\det(A) = 0$ if and only if $\alpha = \beta = \gamma = \frac{1}{3}$. It is easy to verify that these solutions also satisfy the remaind nine equations of the system (7). ■

2. Conclusion

We considered automorphisms of the Lotka-Volterra algebra $\mathcal{A}_V(0, 0, 0)$ and a genetic algebra $\mathcal{A}_{V_\pi}(0, 0, 0)$. From the statements of Theorem 1 and Theorem 2 we conclude that $\text{Aut}(\mathcal{A}_V(0, 0, 0))$ is richer than $\text{Aut}(\mathcal{A}_{V_\pi}(0, 0, 0))$.

REFERENCES

1. *Bernstein S.N.* Solution of a mathematical problem connected with the theory of heredity // Ann. Sci. de Ukraine. 1924. P. 83–114.
2. *Fernandez J.S.G., Garcia C.I.* On Lotka-Volterra algebras // Jour. Alg. its Appl. 2019. Art. num. 1950187.
3. *Lyubich Y.I.* Mathematical structures in population genetics // Lect. Notes of Biomathematics. Springer-Verlag, Berlin, 1992. Vol. 22. 373 p.
4. *Tian J.P.* Evolution algebras and their applications // Lecture Notes in Math., Springer-Verlag, Berlin, 2008. Vol. 1921. 133 p.
5. *Ganikhodzhaev R.N., Eshmamatova D.B.* Quadratic automorphisms of a simplex and the asymptotic behavior of their trajectories // Vladikavkaz. Mat. Zh., 2006. Vol. 8. P. 12–28. (in Russian)

Ibragimov Mahammadjon — PhD student, senior teacher, teacher, department of Mathematics. E-mail: *muhammadibragimov1972@gmail.com*

STOCHASTIC TWO-CRITERIA MODEL FOR OPTIMISING THE LOCATION OF AGRICULTURAL PRODUCTION UNDER CONDITIONS OF SOIL SATURATION

N.U. Uteuliev¹, A.J. Seytov^{2,3}, Yu.Q. Qutlimuratov¹, B.N. Begilov¹,
Z.B. Madreymova⁴, M. Esonturdiyev⁵

¹*Nukus Branch of Tashkent University of Information Technologies, Nukus,
Uzbekistan*

²*National University of Uzbekistan, Tashkent, Uzbekistan*

³*Central Asia University, Tashkent, Uzbekistan*

⁴*Karakalpak State University, Nukus, Uzbekistan*

⁵*Chirchiq state pedagogical universitety, Tashkent, Uzbekistan*

The paper develops a two-criteria stochastic ecological and economic model of agricultural production, taking into account the level of land salinisation. The distribution of areas for farms under agricultural crops is determined for agricultural production under given resource constraints: land, labor, fodder and ecological. Taking into account the limitations, it is required that the expected total volume of crop and livestock production in certain proportions was maximized and the total volume of water consumption was minimized. For the question of specialization and placement of developed agricultural production based on maximisation of crop and livestock production and minimisation of water use, in the two-criteria ecological-economic model taking into account the level from land salinisation λ ($\lambda = 1, 2, 3, 4$) and application of chemical fertilizers, the indicators of the amount of harmful substances stored in the composition are taken into account. Theorems 1 and 2 of dual relations for two-criteria ecological-economic model are established and qualitative analysis is carried out. Methods of calculation of water use volumes and total water consumption volumes per unit volume of crop and livestock production from dual relations are given. **Keywords:** *Stochastic model, dual relations, qualitative analysis, salinity parameters, Lagrangian function.*

Introduction

Today, in the Aral Sea region, questions about agricultural production indicators are of great importance for the location and specialization of agriculture. When determining the economic efficiency of location and specialization of agricultural production, it is advisable to use a system of in-

terrelated indicators that could determine the degree of use of land, labour, machinery, etc [2]. Assessment of economic efficiency of agricultural production and specialization of agriculture is associated with the use of such indicators as gross and marketable output, gross and net income, yield productivity, cost price, labour inputs, profitability and others [3].

It should be noted that the experience of some scientists shows that improving the structure of sown areas, more rational placement of agricultural production and its optimal connection with all links of the agro-industrial complex can give a significant economic effect. At more rational placement of agricultural crops (e.g. cotton, wheat, rice, vegetables, etc.), the gross harvest of these crops increases in the proportions determined by the planned tasks for the harvest of each crop with unknown resources [4].

1. MATERIALS AND METHODS

Similar ecological and economic models of agricultural production were considered in [5, 5, 6]. Formation of constraints and target functions of the stochastic model of location and specialisation of agricultural production were considered in [7].

Problems of water resource losses in main canals during settlement of agricultural production are presented in works [8, 9], and solutions to such a class of problems are considered in works [10, 11].

The stochastic nature of the considered tasks in this area is explained by various reasons. Therefore, the question arises about the need to consider the stochastic variant of the task [12].

The following two - criteria stochastic problem of placement and specialisation of agricultural production is considered. The problem is written in x-form, which is convenient for numerical solution, it can be written in $(x, y(\theta))$ -form, according to the terminology of [13], which has the following form:

$$G_1(x, y, z, \bar{h}_1(\theta), \bar{h}_2) = \alpha_1 E \bar{h}_1(\theta) - \alpha_2 \bar{h}_2 \rightarrow \max, \tag{1}$$

$$\sum_{i \in I} \sum_{k \in K_i} \sum_{\lambda \in \Lambda} a_{ijk}^\lambda(\theta) x_{ijk}^\lambda - \sum_{i \in I} y_{ij} \geq A_j h_1, p.n. j \in J_1^i, \tag{2}$$

$$\sum_{i \in I} z_{iv} \geq B_v h_1, v \in J_2^i, \tag{3}$$

$$\sum_{j \in J_1^i} \sum_{k \in K_i} \sum_{\lambda \in \Lambda} \beta_{ijk}^\lambda x_{ijk}^\lambda + \sum_{v \in J_2^i} \bar{\beta}_{iv} z_{iv} \leq Q_i h_2, i \in I, \tag{4}$$

$$\sum_{j \in J_1^i} \chi_{jr} y_{ij} - \sum_{v \in J_2^i} b_{ivr} z_{iv} \geq 0, i \in I, r \in R, \tag{5}$$

$$\sum_{j \in J_1^i} \sum_{k \in K_i} \sum_{\lambda \in \Lambda} t_{ijk}^\lambda x_{ijk}^\lambda + \sum_{v \in J_2^i} \tau_{iv} z_{iv} \leq T_i, i \in I, \tag{6}$$

$$\sum_{j \in J_1^i} \sum_{k \in K_i} \sum_{\lambda \in \Lambda} (\gamma_{mjk}^{1\lambda} + \gamma_{mjk}^{2\lambda}) x_{ijk}^\lambda + \sum_{v \in J_2^i} \bar{\gamma}_{mv} z_{iv} \leq D_{mi}, i \in I, m \in M, \tag{7}$$

$$\sum_{j \in J_1^i} \sum_{k \in K_i} \sum_{\lambda \in \Lambda} q_{ijk}^\lambda x_{ijk}^\lambda + \sum_{v \in J_2^i} \bar{q}_{iv} z_{iv} \leq M_{il}, i \in I, l \in L_i, \tag{8}$$

$$\sum_{j \in J_1^i} \sum_{k \in K_i} x_{ijk}^\lambda \leq S_i^\lambda, i \in I, \lambda \in \Lambda. \tag{9}$$

Here, $a_{ijk}^\lambda(\theta)$ -yield of the j -th crop in the i -th farm on the k -th type of land and λ -planted soil, depending on the random variable θ , which is a state of nature. E-mathematical expectation of the random variable θ . Other parameters of model (1) - (9) are introduced in [14] deterministic version of models (1) - (9).

Let us consider the study of qualitative properties of the two-criteria stochastic problem of placement and specialisation of agricultural production. For this purpose, we construct a dual problem on the problems (1) - (9), which has the following form:

$$G_2(u(\theta), \vartheta, \bar{u}, d, \nu, \rho, \mu, \omega) = \sum_{i \in I} T_i \Lambda_i + \sum_{i \in I} S_i^\lambda \omega_i^\lambda + \sum_{m \in M} \sum_{i \in I} D_{mi} \rho_{mi} + \sum_{l \in L} \sum_{i \in I} M_{il} \mu_{il} \rightarrow \min, \tag{10}$$

$$t_{ijk}^\lambda \Lambda_i + \beta_{ijk}^\lambda \bar{u}_i + \omega_i^\lambda + \sum_{l \in L} q_{ijk}^\lambda \mu_{il} + \sum_{m \in M} (\gamma_{mjk}^{1\lambda} + \gamma_{mjk}^{2\lambda}) \rho_{mj} - E(a_{ijk}^\lambda(\theta) u_j(\theta)) \geq 0, \tag{11}$$

$$u_j(\theta) - \sum_{r \in R} \chi_{jr} d_{ir} \geq 0, \tag{12}$$

$$\sum_{r \in R} b_{ivr} d_{ir} + \tau_{iv} \Lambda_i + \bar{\beta} \bar{u}_i + \sum_{l \in L} \bar{q}_{iv} \mu_{iv} + \sum_{m \in M} \bar{\gamma}_{mv} \rho_{mi} - \vartheta_v \geq 0, \tag{13}$$

$$\sum_{j \in J_1^i} A_j u_j(\theta) + \sum_{v \in J_2^i} B_v \vartheta_v \geq \alpha_1, \tag{14}$$

$$\sum_{j \in J_1^i} Q_i \bar{u}_i \leq \alpha_2, \tag{15}$$

where $u_j(\theta)$ - means an the random variable θ .

Other parameters of model (10)-(15) are introduced in [14] deterministic version of models (1)-(9). Problem (1)-(9) and (10)-(15) are mutually dual problems of parametric linear programming. The corresponding first and second duality theorems [15] are valid for them, which are formulated in the following form.

Theorem 1. If one of the problems of the dual pair (1)-(9) and (10) - (15) has a solution at any values of $\alpha_1 \geq 0, \alpha_2 \geq 0$, then the other problem is also solvable [16]. The equality holds for any optimal solutions $X^* = \{x_{ijk}^{\lambda*}, y_{ij}^*, z_{iv}^*, \bar{h}_1^*(\theta), \bar{h}_2^*\}$ and $Y^* = \{u_j^*(\theta), \vartheta_v^*, \bar{u}_i^*, d_{ir}^*, \Lambda_i^*, \rho_{mi}^*, \mu_{il}^*, \omega_i^{\lambda*}\}$:

$$G_1(X^*) = G_2(Y^*). \quad (16)$$

Theorem 2. For the admissible solution of $\{x_{ijk}^{\lambda*}, y_{ij}^*, z_{iv}^*, \bar{h}_1^*(\theta), \bar{h}_2^*\}$ of the problem (1) - (9) to be optimal, it is necessary and sufficient that there exist such values $\{u_j^*(\theta), \vartheta_v^*, \bar{u}_i^*, d_{ir}^*, \Lambda_i^*, \rho_{mi}^*, \mu_{il}^*, \omega_i^{\lambda*}\}$, satisfying conditions (10) - (15), so that the relations [16] are satisfied.

2. Results

Based on the study of doicity and optimality condition for this two-criteria stochastic model and the established 1st and 2nd duality theorem, the norm of water resources utilisation is formed. From the 2nd theorem, we find the total payment for water consumption for crop and livestock production and the specific volumes of the water consumption balance [19]. Analyzing the results of the work, the following conclusions can be approved. The relationship of duality and the optimality conditions of the two-criterion ecological model are investigated, the 1st and 2nd duality theorems are established, and water use formulas can also be derived.

3. Conclusion

In this regard, a two-criteria ecological and economic model of agricultural production has been developed, taking into account the level of land salinity, which helps to determine the types of crops to be planted on each land plot and allows obtaining optimal alternative options for the placement of crop species in accordance with reclamation conditions. Also, the issue of specialisation and location of agricultural production developed under water resources allocation leads to maximisation of crop and livestock production and minimisation of water use, and thus leads to saving of water resources compared to the actual situation, and secondly, to ensure satisfaction of demand for crops. An additional characteristic of chemical fertiliser use is to

ensure that the amount of harmful substances contained in the composition does not exceed the established normative level.

REFERENCES

1. *J.V. Belovoe* Pokazateli effektivnosti razmesheniya selskoxozyaystvennogo proizvodstva // <https://cyberleninka.ru/article/n/pokazateli-effektivnosti-razmesheniya-selskoxozyaystvennogo-proizvodstva/viewer>.
2. *Akmarov P B, Tretyakova E S and Kharisov R G* Ecological and economic aspects of increasing the energy productivity of agricultural land in models of optimal farming // Bulletin of the Perm National Research Polytechnic University. 2012. Vol. 15. P. 66–73.
3. *Ivanyo Ya M, Kovaleva E A and Petrova S. A.* Algorithm of realization of ecological mathematical problems of optimization of agricultural production in conditions of uncertainty // Proceedings of VSU, Series: systems analysis and information technologies. 2020. Vol. 2. P. 79–91. DOI 10.17308/sait.2020.2/2918.
4. *Ermoliev, Y.M., Mikhalevich, M.V., Uteuliev, N.U.* Economic Modeling of International Water Use (The Case of the Aral Sea Basin) // IIASA Research Report (Reprint). IIASA, Laxenburg, Austria: RR-95-005. Reprinted from Cybernetics and Systems Analysis. 1994. Vol. 30(4).
5. *Lyashenko I.N., Mikhalevich M.V., Uteuliev N.U.* Metodi ekologo-ekonomicheskogo modelirovaniya (Methods of ecological and economic modeling). Nukus: Bilim, 1994. 211 p.
6. *Podinovskiy V.V., Nogin V.D.* Pareto-optimalne resheniya mnogokriterialnix zadach (Pareto-optimal solutions to multi-criteria problems). Moscow: Nauka, 2007. 255 p.
7. *Uteuliev N.U., Qutlimuratov Yu.Q., Yadgarov Sh.A.* Using the methods and optimizing criteria in making the managing decisions of agricultural production // 2019 International Conference on Information Science and Communications Technologies (ICISCT)- 2019. Tashkent. P. 1–4. DOI:10.1109/AICT50176.2020.9368798
8. *Tkachyev, A.A., Olgarenko, I.V.* Urgent problems of water distribution management in main canals of irrigation systems // Scientific journal of Russian Scientific Research Institute of land improvement problems. 2021. Vol. 11, no. 2. P. 1–23.
9. *Jiang, Y., Xiong, L., Yao, F., Xu, Z.* Optimizing regional irrigation water allocation for multi-stage pumping-water irrigation system based on multi-level optimization-coordination model. // Journal of Hydrology. 2019. № 4. Vol. 100038. P. 1–13.
10. *Seytov, S.J., Sayfullayev, B.Sh., Anorbayev, M.M.* Separating a Finite System of Points from each Other in Real Euclidean Space // AIP Conference Proceedingsthis link is disabled. 2023. Nukus. Vol. 2781. P. 020043.

11. *A.Seytov, L.Varlamova, S.Bahramov, Yu.Qutlimuratov, B.Begilov, D.Seytimbetov.* Optimal management of water resources of large main canals with cascades of pumping stations. // AIP Conference Proceedingsthis link is disabled. 2024. Vol. 3147. P. 030020.
12. *Ajjaz Maqbool Dar, Mushtaq Ahmad Lone.* Solving stochastic agricultural production planning problems by using fuzzy goal programming // Skuast Journal of Research Year : 2023, Vol. 25(3). P. 1-8. DOI : 10.5958/2349-297X.2023.00049.1
13. *Ermolev Yu.M., Yastremskiy A.I.* Stoxasticheskie modeli i metodi v ekonomicheskom planirovanii. Moscow: Nauka, 1988. 253 p.
14. *Yu. Qutlimuratov, B. Jarilkanov, Z. Madreymova.* Ecological and economic model of plant placement based on soil salinity monitoring // AIP Conference Proceedingsthis link is disabled. 2024. Vol. 3147. P. 030020.
15. *Golshteyn E.G.* Vipukloe programmirovaniye (elementi teorii). Moscow: FIZ-MATLIT, 2005. 368 p.
16. *Yudin D.B., Golshteyn E.G.* Lineynoe programmirovaniye (teoriya, metodi i prilozheniya). Moscow: Nauka, 1969. 424 p.

Uteuliev Nietbay Uteulievich — Doctor of physical and mathematical sciences, Professor, Head of department of Software engineering and mathematical modeling. E-mail: *uteuliev@mail.ru*

Seytov Aybek Jumabaevich — DSc, Assistant professor, Professor of department of Computational mathematics and information systems. E-mail: *saybek868@gmail.com*

Qutlimuratov Yusup Qulbaevich — PhD, Assistant professor, Head of the Department of Computer Systems. E-mail: *q-yusup@mail.ru*

Begilov Bakhadir Nazarbaevich — PhD, Senior Lecturer of department of Software injiniring and mathematical modeling. E-mail: *bbegilov@gmail.com*

Madreymova Zulfiya Bazarbaevna — PhD student of department of Applied Mathematics and Informatics. E-mail: *madreimovazulfiya@gmail.com*

Esonturdiyev Mamatqobil — Deputy dean for educational affairs of Chirchik State Pedagogical University.

IMPLEMENTATION OF GENETIC ALGORITHM OPTIMIZATION APPROACH FOR SIGMA-DELTA MEMS ACCELEROMETER DESIGN

R. Starodubov

Peter the Great St.Petersburg Polytechnic University, St.Petersburg, Russian Federation

This paper presents the novel approach to development and optimization of the signal processing for capacitive MEMS accelerometer based on electromechanical sigma-delta modulator technique (EM-SDM). The EM-SDM closed-loop architecture leads to high sensor characteristics at the cost of increasing design complexity. Also, due to interrelation of parameters and parasitic couplings there are a lot of difficulties in the design of such structures. Therefore, the optimization technique for the third order EM-SDM loop based on a genetic algorithm has developed. System level simulations show the achieved signal to quantization noise ratio (SQNR) of 105 dB. **Keywords:** *MEMS, capacitive accelerometer, electromechanical sigma-delta modulator, genetic algorithm, simulation, optimization.*

Introduction

Capacitive micro-electromechanical accelerometers have comparatively high sensitivity and simple fabrication [1]. The main parts of basic interface circuit are capacitance-to-voltage converter (C/V) and analog-to-digital converter (ADC). Systems in open-loop configuration (without a feedback) are relatively simple but have a low performance which heavily depends on parameters of the sensing element; such as bandwidth or eigenfrequencies. Furthermore, the displacement-to-capacitance conversion (x/C) is not linear and pull-in effect limits the maximum effective mass displacement.

The embedding of sensing element in a negative feedback closed-loop control system allows to neglect linearity requirements for each constituent block, lower the dependencies on sensing element parameters and increase the dynamic range. One of the effective solutions for implementation of the closed-loop system based on a $\Sigma\Delta$ -modulator working principle. It preserves all advantages of closed-loop operation and provides easier feedback due to linearity of voltage to force conversion.

The present work is organized as follows. Section 1 describes the basics of electromechanical $\Sigma\Delta$ -modulator. Section 2 dedicated to genetic

algorithm. The features of EM-SDM structure optimization and modeling results are in the Section 3. Finally, the conclusion is discussed in Section 4.

1. Electromechanical sigma-delta modulator

The sensing element of the capacitive accelerometer can be described in Laplace terms as follows:

$$M(s) = \frac{1}{m \cdot s^2 + b \cdot s + k} \tag{1}$$

where m - effective proof mass, b - damping factor, k - spring stiffness.

Thus, the transfer function of the sensing element is an equivalent of the second order integrator. The sensing element model is presented in Figure 1.

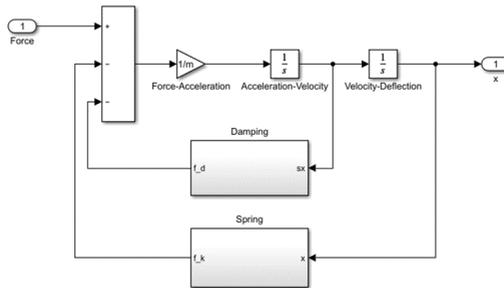
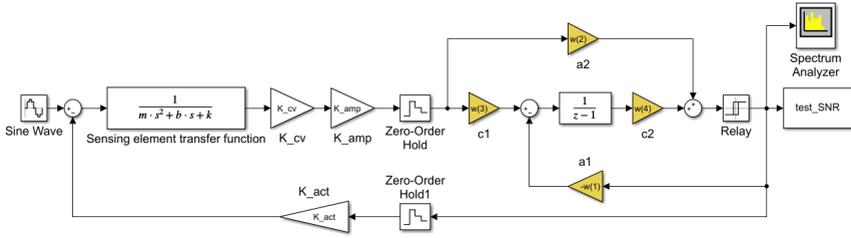


Figure 1. Sensing element model

The EM-SDM consists of one-bit or multi-bits quantizer, electronic filters, accelerometer, capacitive interface, C/V converter and V/F converter. Electronic filters can be implemented using field programmable gate array (FPGA). That offers extra flexibility in optimizing the EM-SDM structure and allows reconfiguration after the fabrication. However, multi-bits ADC is needed and it should have quantization noise level at least one order less than quantizer’s to neglect it’s contribution in resulting noise level of the system [2].

Simulink model of the EM-SDM is shown on Figure 2. Zero-Order Hold is used as idealized ADC and DAC, relay is an equivalent to one-bit quantizer.

There is no formalized methods to synthesize coefficients for electronic filters in EM-SDM [3]. As a basis, it is proposed to use a tool for synthesizing

Figure 2. Electromechanical $\Sigma\Delta$ -modulator model

$\Sigma\Delta$ -modulator coefficients (e.g. Delta Sigma Toolbox). That allows to get prototype transfer function with specified order and H_∞ - out-of-band gain of noise transfer function (NTF). Signal and noise transfer functions can be expressed through transfer functions from modulator input and output to quantizer input [4]:

$$STF(z) = \frac{L_0(z)}{1 + L_1(z)} \quad (2)$$

$$NTF(z) = \frac{1}{1 + L_1(z)} \quad (3)$$

Assume that $L_0(z) = L_1(z) = L(z)$. Now the analytic formula for 3rd electromechanical $\Sigma\Delta$ -modulator $L_{EM}(z)$ should be found:

$$L_{EM}(z) = \frac{N2 \cdot z^2 + N1 \cdot z + N0}{D3 \cdot z^3 + D2 \cdot z^2 + D1 \cdot z + D0} \quad (4)$$

Now the corresponding numerator coefficients of the $L_{EM}(z)$ and prototype transfer functions should be combined into a system of equations and extra coefficients should be fixed to achieve the only solution to the system of equations. The comparison of $L_{EM}(z)$ and prototype transfer function is presented on Figure 3. The resulting $NTF(z)$ of the electromechanical $\Sigma\Delta$ -modulator is equivalent to prototype by stability factor although it has one extra zero in-band.

2. Genetic algorithm

Genetic algorithm is a stochastic multidimensional optimization method inspired by the process biological evolution. This method can be used for finding the global minimum, however the convergence is not guaranteed. One of the main advantages of this method - it can be applied without an analytic formula of the fitness function.

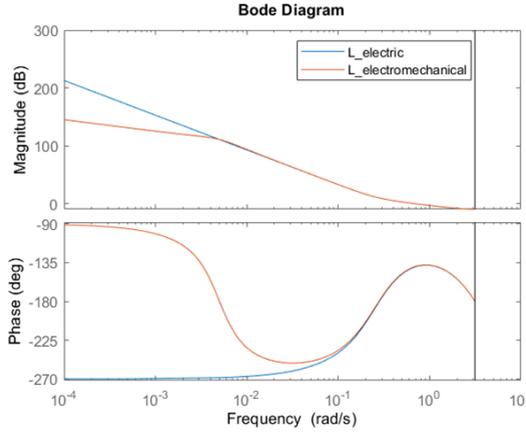


Figure 3. Amplitude-frequency responses $L_{EM}(z)$ and $L_E(z)$

Initially there is a uniform distribution of points inside the initial range defined by variables boundaries. The fitness function evaluates in every point and than after the ascending sort every function value is scaled and gets rating:

1. Rank scaling. Every function value is normed by $\frac{1}{\sqrt{n}}$. Where n is the number of the point in sorted list. Rating is higher for lower function values.
2. Top scaling. Some percent of the set of points with lowest function values get the same equal rating and the others get rating equals to zero.

The difference between these two scaling methods is in the points diversity on every following iteration. Rank scaling is more efficient for the functions with many local minimums.

There is three different mechanisms to obtain new set of points for the next iteration:

1. Elite. These points with highest rating are getting to the next set without any changes.
2. Crossover. Two or more points shuffle variables values between each other to form a new point. Points for a crossover choose randomly based on the rating.
3. Mutation. Point gets a random change to one or more variables.

3. Optimization of electromechanical sigma-delta modulator

Existing tools for synthesizing $\Sigma\Delta$ -modulators change the placement of transfer function poles simultaneously to match H_∞ . The research showed that it is possible to achieve higher SQNR by moving poles independently.

Genetic algorithm has a slow speed of convergence and it heavily depends on the initial range. Electronic filters coefficients have specific order, for instance, $a_1 \approx 10^{-3}$ and $c_2 \approx 10^4$.

Boundaries for coefficients can be found based on Lee's stability criteria:

$$|H_\infty| < 1.5 \quad (5)$$

This criteria actually depends on the exact structure - the more order of the modulator the stricter it becomes. When modulator's work becomes unstable SQNR sharply decreases. Because GA is a minimization method lets assume that fitness function is $SQNR' = -SQNR$. That guarantees that method does not fall into unstable work area.

The method for finding boundaries based on Delta-Sigma Toolbox and method introduced in Section 1. This method measures characteristics $SQNR$ vs H_∞ while varying the normalized amplitude of the input signal (Figure 4). Therefore, the rough evaluation of the peak SQNR can be found and used to set the boundaries for optimization method.

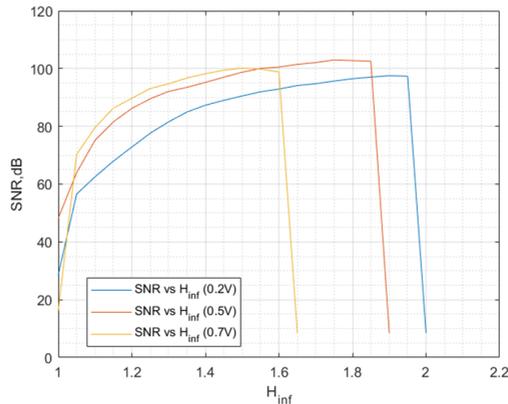


Figure 4. SQNR depending on H_∞ for different normalized amplitudes

The progress of genetic algorithm applied to electromechanical $\Sigma\Delta$ -modulator with usage of previously mentioned methods is shown on Figure 5.

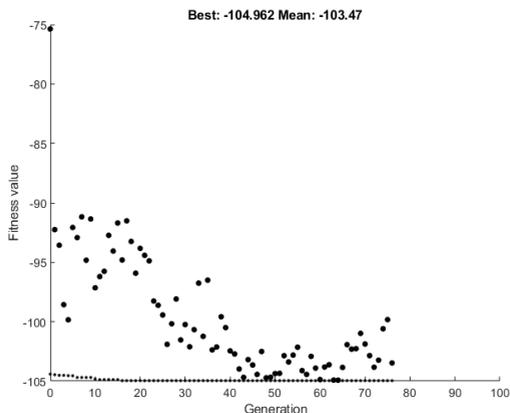


Figure 5. Best function values on each iteration of GA

Conclusion

Methods for synthesizing coefficients of the electromechanical $\Sigma\Delta$ -modulator and defining boundaries for the further optimization and have been developed. The results showed that optimization allows to achieve 3% more SQNR compared to conventional methods. System model can be expanded which can allow to have more degrees of freedom for optimization and formulate the requirements for each individual block of the structure.

REFERENCES

1. *Chen F., Li X., Kraft M.* Electromechanical Sigma-Delta Modulators ($\Sigma\Delta M$) Force Feedback Interfaces for Capacitive MEMS Inertial Sensors: A Review // *IEEE Sensors Journal*, 2016. Vol. 16 № 17. P. 6476–6495.
2. *Dong Y., Zwahlen P., Nguyen A.-M., Rudolf F., Stauffer J.-M.* High performance inertial navigation grade sigma-delta MEMS accelerometer // *IEEE/ION Position, Location and Navigation Symposium*. 2010. P. 32–36.
3. *Gad M., Elshennawy A., Ismail A.A.* Design Method for Delta Sigma Force-Feedback Accelerometer Interface Systems // *18th IEEE International New Circuits and Systems Conference (NEWCAS)*. 2020. P. 5–8.
4. *Pavan S., Schreier R., Temes G.C.* Understanding delta-sigma data converters. Hoboken: John Wiley & Sons, Inc., 2017. 561 p.

THE DEVELOPMENT OF RESONANT MEMS ACCELEROMETER BASED ON SYSTEM LEVEL APPROACH

V. Ivannikov

Peter the Great St. Petersburg Polytechnic University, St. Petersburg, Russian Federation

This paper deals with the system level simulation of a resonant MEMS accelerometer. System level design utilizes the advantages of top-down approach for design and verification in terms of increased design accuracy and reduced development time. The first step is the development of an analytical model of sensor. Secondly, FEM-based simulation of sensor structure and model reduction is done. This model is then used to construct the MEMS-oscillator circuit. Finally, a comprehensive system model including both mechanical and electronic parts is developed. This system model allows early stage optimization and sensor performance prediction. The simulations show that the scale factor is 400 Hz/g and the resonant frequency is 59 kHz. **Keywords:** *MEMS sensors, MEMS accelerometer, system modeling, resonant sensor, frequency modulation.*

Introduction

The continuous advancements in microelectromechanical systems have significantly influenced various fields, including consumer electronics, automotive safety, and biomedical applications. Unlike conventional capacitive accelerometers, resonant accelerometers use oscillating beams whose resonant frequency varies with the applied acceleration. This design offers a number of advantages, including higher sensitivity, resistance to environmental fluctuations, and quasi-digital output.

A systems design approach is used to address the problems of performance prediction and design optimization of resonant accelerometers without the need for extensive physical prototypes. Such a model serves as a powerful tool for understanding the complex interactions between the mechanical, electrical, and control components of an accelerometer. This paper is organized as follows. Section 1 presents the working principle of the device and the calculation of the mechanical model. Section 2 presents the equivalent electrical model of the sensing element and the oscillating circuit. Section 3 presents the process of building a system model of the accelerom-

eter. The final part with a discussion of the results of the work done and the main conclusions is presented in Section 4.

1. Device Design

The resonant accelerometer consists of a square proof mass, two levers, two symmetrically arranged resonators, and sensing and excitation electrodes. The design is shown in Figure 1.

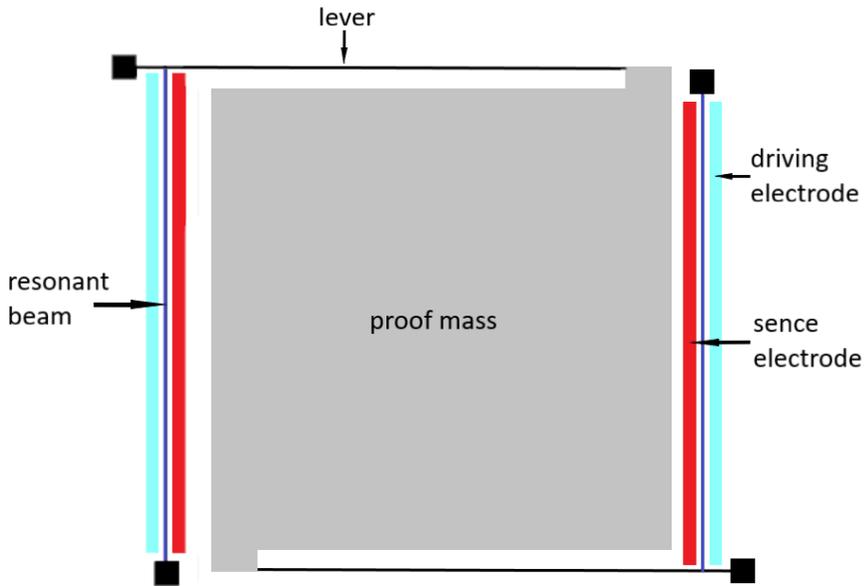


Figure 1. Scheme of the resonant accelerometer

In the absence of acceleration, the resonators have an oscillation natural frequency (f_0) of 59298 Hz. When acceleration is applied, the proof mass is displaced, and the force is transmitted to the resonators by means of levers. One resonator will experience stretching - its resonant frequency increases accordingly, and the second resonator will experience compression - its resonant frequency decreases. The excitation electrode sets it in motion, and the frequency-modulated signal is picked up by the sensing electrode.

Next, the development of a mathematical model of the sensing element is presented, which allows to calculate its characteristics. Basic geometric parameters of the accelerometer [1]: length of proof mass, levers and resonating beams $L = 400 \mu m$, resonant beam width and lever width

$w_{balc} = 1.2 \mu m$, structure thickness $h = 15 \mu m$, distance between the sensing and driving electrodes and the beam $d = 2.1 \mu m$, distance between the lever mounting point and the resonator $d_{load} = 6.6 \mu m$, beam cross-sectional area $S_{blk} = 1.8 * 10^{-11} m^2$, coefficients of boundary conditions: $c = 4.73$, $\alpha = 0.0246$.

Consider a resonant beam, one end of which is rigidly fixed and the other end is attached to a lever. The fundamental frequency of the beam without an applied axial load is calculated as [1]-[5]:

$$f_0 = \frac{c^2}{2\pi L^2} \sqrt{\frac{EI}{\rho S_{blk}}}, \quad (1)$$

where ρ is the density of polysilicon, $E = 160 GPa$ is Young's modulus, I - moment of inertia of the beam.

The resonant frequency at applied load F_{load} is calculated as [1]-[4]:

$$f_{new} = f_0 \sqrt{1 + \alpha F_{load} \frac{L^2}{EI}}. \quad (2)$$

Equation (1) and (2) is satisfied if the transverse vibrations can be considered small with respect to the beam thickness in the direction.

When acceleration is applied to the proof mass, the force exerted by one arm on the resonator is equal to [5], [6]:

$$F_{res} = \frac{0.5gL^2h\rho L}{d_{load}}. \quad (3)$$

This force causes a differential change in the resonant frequencies of the two beams, which is used to determine the magnitude of the applied acceleration.

To investigate the mechanical performance of the accelerometer, a finite element model was created using COMSOL Multiphysics. Different accelerations in the range [-1g, 1g] were applied to the model and the changes in the resonant frequencies of the beams were observed. The simulation results showed that the sensitivity of the accelerometer, is 400 Hz/g.

2. Design Electrical Circuit

This section presents the equivalent circuit diagram of the sensing element, which is further used to construct the oscillating circuit.

The micromechanical resonator can be represented as an equivalent series *RLC* circuit. The electromechanical transduction factor will be calculated as [7]:

$$n = \frac{0.825V_p\varepsilon_0S_{el}}{d^2}, \quad (4)$$

where $V_p = 6 V$ is the DC voltage applied to the beam, $v_a = 180mV$ is the AC voltage applied to the excitation electrode.

According to the study [7], the equivalent electrical parameters R_{eq} , C_{eq} , L_{eq} are calculated as:

$$R_{eq} = \frac{2\pi f_0 M}{Qn^2}, \quad (5)$$

$$C_{eq} = \frac{n^2}{M(2\pi f_0)^2}, \quad (6)$$

$$L_{eq} = \frac{M}{n^2}, \quad (7)$$

where f is the natural frequency of the resonator, M is the equivalent mass of the resonating beam.

In this equivalent substitution scheme, R_{eq} accounts for the losses in the resonator, L_{eq} accounts for the inertia of the resonator, and C_{eq} accounts for the resonance and elastic properties of the resonator. For the case in the absence of acceleration ($f = 59 kHz$) the elements are $R_{eq} = 8.76 M\Omega$, $L_{eq} = 4702 H$, $C_{eq} = 1.532 fF$.

The next step is to build an oscillating circuit, it is shown in Figure 2.

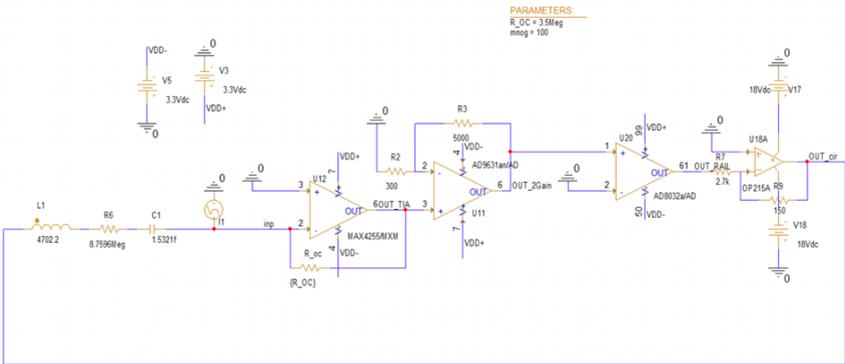


Figure 2. Scheme of the oscillating circuit

The circuit works as follows: the current from the sensing electrode is fed to the MAX4255 transimpedance amplifier, then the signal is fed to the

AD9631an inverting amplifier. The AD8032a operational amplifier is used for amplitude limiting, its output generates a rectangular waveform signal, which is fed to the processing circuitry. OP215A is used to reduce the signal to an optimum amplitude value of 180 mV before it is fed to the driving electrode. The occurrence of auto oscillations is ensured by selecting the gain of the second stage and selecting the inverting configuration for the second amplifying stage.

3. System Model Development

In this section, a system model is presented which includes the mechanical part of the sensing element and the electrical part of the oscillating circuit. The schematic of the system model is presented in Figure 3.

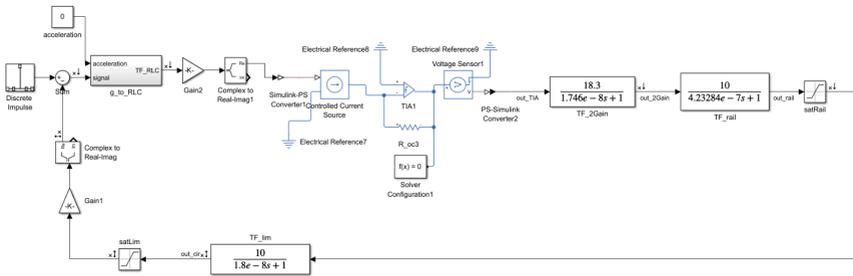


Figure 3. Schematic of the system model

First in the system model is the mechanical block ($g.to_RLC$). Its structure is shown in detail in the Figure 4 and consists of:

- 1) A block to convert the applied acceleration into a force acting on the lever ($acceleration_to_Force$),
- 2) Block for calculating the force applied to the resonator ($Lever$),
- 3) A block including calculation of the resonant frequency as a function of the force applied to it ($F.to_freq$),
- 4) Block for calculating the equivalent RLC circuit ($freq.to_RLC$),
- 5) Block forming the voltage-to-current transfer function of the equivalent RLC circuit (RLC_equiv).

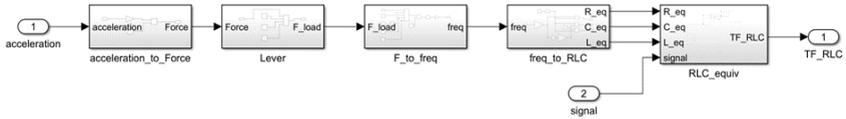


Figure 4. Mechanical system model of the sensing element unit

All presented blocks use equations (1)-(7) to calculate values. The signal from the block g_to_RLC output is input to a transimpedance amplifier. The signal then passes sequentially through the transfer function blocks in the following order: the second voltage gain stage block (TF_2Gain), the Rail-to-Rail operational amplifier block (TF_rail), and the amplitude limiter block (TF_lim). From the amplitude limiter output, the signal goes to the 180° phase rotation block, since the second gain stage is an inverting amplifier.

4. Conclusion

This paper presents a comprehensive approach to the system-level design of a resonant MEMS accelerometer. The study covers several key steps: creation of a mechanical model, computation of the equivalent substitution circuit of the sensing element, construction of an oscillatory circuit, and integration of the results into a single system model. The system-level model fully describes the electrical and mechanical properties of the device and allows predicting its characteristics at early stages. The paper considers a bottom-up design approach to demonstrate the truth of the model, but using the previously described theoretical formulas, a top-down approach can be effectively employed for rapid system-level design and optimization.

REFERENCES

1. *Comi C., Corigliano A., Langfelder G., Longoni A., Tocchio A., Simoni B.* A Resonant Microaccelerometer With High Sensitivity Operating in an Oscillating Circuit. // Journal of Microelectromechanical Systems. 2010. Vol. 19. № 5. P. 1140–1152.
2. *Comi C., Corigliano A., Langfelder G., Longoni A., Tocchio A., Simoni B.* A high sensitivity uniaxial resonant accelerometer. // IEEE 23rd International Conference on Micro Electro Mechanical Systems (MEMS), City: Hong Kong, 2010, P. 260–263.

3. *Comi C., Corigliano A., Langfelder G., Longoni A., Tocchio A., Simoni B.* A new two-beam differential resonant micro accelerometer. // *SENSORS*. City: Christchurch, 2009, P. 158–163.
4. *Mizumoto T., Hirai Y., Banerjee A., Tsuchiya T.* Mems Reservoir Computing Using Frequency Modulated Accelerometer. // *IEEE 35th International Conference on Micro Electro Mechanical Systems Conference (MEMS)*, City: Tokyo, 2022. P. 487–490.
5. *Stokey W.* Vibration of systems having distributed mass and elasticity. // *Shock and Vibration Handbook*. City: New York, 1988. P. 1–50.
6. *Wallin M., Ivarsson N., Tortorelli D.* Stiffness optimization of non-linear elastic structures. // *Computer Methods in Applied Mechanics and Engineering*. City: United States, 2018, P. 292–307.
7. *Kaajakari V., Mattila T., Oja A., Kiihamäki J., Seppä H.* Square-extensional mode single-crystal silicon micromechanical resonator for low phase noise oscillator applications. // *IEEE Electron Device Letters*. 2004. Vol. 25. № 4. P. 173–175.

Ivannikov Vladimir — Bachelor of Engineering, Department of Integrated Electronics. E-mail: *ivannikov.va@edu.spbstu.ru*

ПОСТРОЕНИЕ СРАВНИТЕЛЬНЫХ РЯДОВ ДЛЯ ЗАДАЧИ ОЛЬФАКТОРНОГО АНАЛИЗА

А. А. Солдатенко, Д. В. Семенова

Сибирский федеральный университет, г. Красноярск, Россия

Рассматривается комбинаторная задача построения сравнительных рядов ольфакторного исследования. Сравнительные ряды используются при проведении экспертиз судебными экспертами органов внутренних дел России, а также при тренировке собак-детекторов. Предлагается рассматривать каждый ряд как выбор слова из некоторого языка. В работе представлены языки для каждого из четырех сравнительных рядов, приведены верхние оценки количества слов в языках. **Ключевые слова:** ольфакторный анализ, этограмма, теория формальных языков.

Введение

Задача построения сравнительных рядов ольфакторного анализа является комбинаторной задачей. Требуется четыре различных ряда, каждый из которых должен удовлетворять регламентированным требованиям. Для составления рядов используются два вида запаховых проб: проба, взятая с вещи; проба, взятая из крови. В данной работе рассматривается задача построения сравнительных рядов для идентификационной экспертизы для одного объекта и одного проверяемого лица [1]. В этом случае в рядах используются шесть проб, взятых с вещей, и четыре пробы, взятых из крови. В табл. 1 содержатся названия проб и их идентификаторы, используемые в работе. Отметим, что в первых двух рядах не используется проба замены (номер 8), а в третьем и четвертом рядах убирается нейтральный эталон (проба Н.Э.). Каждый ряд должен состоять из десяти проб. На практике пробы выставляются на станциях по кругу и собака-детектор обходит их начиная с некоторой позиции и должна остановиться на станции с определенной пробой, где получает пищевое подкрепление. Такой обход проб будем называть проводкой. В первых двух рядах проводка должна закончиться на станции с пробой Н.Э., а в остальных рядах выполняется две остановки, на станциях с пробами И. и Э. Если собака не останавливается на предполагаемой станции, то экспертиза считается неудавшейся.

Работа поддержана Красноярским математическим центром, финансируемым Минобрнауки РФ (Соглашение 075-02-2024-1429). Авторы благодарят Н. В. Коршуну за постановку задачи.

Таблица 1

Виды исследуемых проб

Название пробы	Тип пробы	Идентификатор пробы
Исследуемый объект	Вещь	И.
Лицо контрольной группы	Вещь	1–5
Эталон	Кровь	Э.
Нейтральный Эталон	Кровь	Н.Э.
Лицо контрольной группы	Кровь	6–7
Лицо контрольной группы (замена)	Кровь	8

Сравнительные ряды должны удовлетворять ограничениям, представленным ниже.

- 01 Длина проводок должна быть разной в каждом ряду.
- 02 Движение должно быть в различном направлении (по часовой стрелки или против) и с различной станцией в каждом ряду.
- 03 Первые две проводки должны заканчиваться на станции с пробой Н.Э., а третья и четвертая проводки должны заканчиваться на станции с пробой Э.
- 04 Первая проводка должна содержать пробу со схожей ольфакторной характеристикой, взятой с вещи или из крови, перед пробами И. и Э., соответственно.
- 05 Все пробы, кроме пробы номер 8, должны присутствовать в первых двух проводках.
- 06 Пробы не должны встречаться на тех же станциях, что и в предыдущих рядах. Все пробы расположенные на станциях вне проводки не влияют на выполнение этого требования.
- 07 Пробы И., Э., Н.Э. не должны располагаться на соседних станциях.
- 08 В каждой из двух первых проводок должны содержаться пробы И., Э., Н.Э.
- 09 Недопустимо размещать более двух проб в третьей и четвертой проводках с одинаковой ольфакторной характеристикой, и более трех проб в первых двух проводках.
- 010 Недопустимо размещать пробу, взятую из крови, перед пробой И.
- 011 Недопустимо размещать пробы И., Э., Н.Э. на станциях, где собака получала пищевое подкрепление.

Дополнительно предъявляются необязательные рекомендации к сравнительным рядам.

- P1 Одинаковые пробы не должны размещаться перед пробами И., Э., Н.Э. в соседних рядах.
- P2 В третьей и четвертой проводках должно быть различное число станций между пробами И. и Э.
- P3 Проба номер 8 должна находиться в начале проводки.

В табл. 2 представлено влияние требований и рекомендаций на каждый сравнительный ряд. Символ «+» означает, что требование вляет непосредственно на шаблон ряда, а символ «*» означает, что требование зависит от конкретных реализаций предыдущих рядов.

Таблица 2

Виды исследуемых проб

	O1	O2	O3	O4	O5	O6	O7	O8	O9	O10	O11	P1	P2	P3
Ряд 1			+	+			+	+	+	+				
Ряд 2	*	*	+		*	*	+	+	+	+	*	*		
Ряд 3	*	*	+			*	+		+	+	*	*		+
Ряд 4	*	*	+			*	+		+	+	*	*	*	+

Согласно представленной таблицы, легко видеть, что начиная с первого ряда, пространство допустимых последующих рядов сокращается.

1. Постановка задачи

Пусть заданы два множества проб. Пробы, взятые с вещей, обозначим как $V = \{v_1, \dots, v_5, v_0\}$, где v_0 соответствует исследуемому объекту. Пробы, взятые из крови, обозначим как $K = \{k_6, k_7, k_8, k_e, k_n\}$, где k_e – эталон, k_n – нейтральный эталон. Тогда общий алфавит для задачи построения сравнительных рядов, обозначим как $\Sigma = V \cup K$. Здесь и далее, будем применять терминологию традиционную для теории формальных языков [2]. Под рядом будем понимать слово из языка L над алфавитом Σ . Общая схема построения языка L выглядит как

$$L = \{\alpha: \alpha = a_1 \dots a_{10}, a_i \in \Sigma, a_i \neq a_j, i \neq j\}.$$

Каждому ряду соответствует свой собственный язык с множеством ограничений, при этом справедливо $L_1, L_2, L_3, L_4 \subset L$. Поскольку каждый ряд состоит из проводки и остаточных элементов, то всякое слово из соответствующего языка можно представить как конкатенацию слов

$\alpha = \beta \cdot \gamma$. Таким образом, ограничения, влияющие на шаблон ряда, влияют только на слово β .

Таким образом, задача построения сравнительных рядов ольфакторного анализа формулируется следующим образом:

- Дано:** алфавит Σ , языки L_1, L_2, L_3 и L_4 .
Требуется: найти слова $\alpha_1 \in L_1, \alpha_2 \in L_2, \alpha_3 \in L_3$ и $\alpha_4 \in L_4$, удовлетворяющие требованиям О1–О11 и рекомендациям Р1–Р3.

В работах [3, 4] представлены определения языков L_1 и L_2 для первых двух рядов, соответственно. Также предложен рекурсивный алгоритм построения дерева шаблонов первых двух рядов, и метод построения слов α_1 и α_2 , по этим шаблонам.

Согласно табл. 2 третий и четвертый ряды имеют одинаковые ограничения на шаблон, а именно О3, О7, О9, О10, Р3. Рекомендация Р3 не имеет четкой формулировки, поэтому в работе предполагается, что проба номер 8 должна находиться в одной из первых трех станций проводки. В данных ограничениях, количество различных шаблонов слова β из $\alpha = \beta \cdot \gamma$ составляет всего 41 шаблон. Верхняя оценка числа возможных слов β составляет 1562 слова, тогда верхняя оценка возможных слов α – 58320 слов. Вид дерева шаблонов представлен на рис. 1.

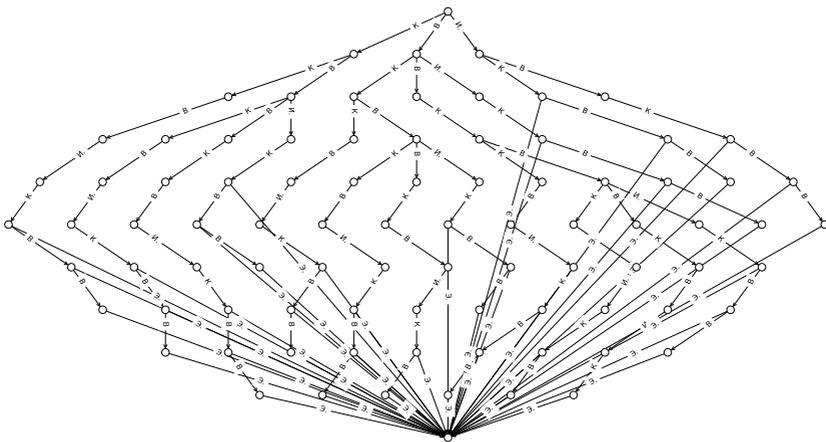


Рис. 1. Дерево шаблонов слова β для третьего и четвертого рядов, где символ К – место для пробы, взятой из крови, символ В – место для пробы, взятой с вези

2. Вычислительные эксперименты

Использование реализованных деревьев шаблонов позволяет выбирать шаблоны для проводки каждого ряда. Разработана программа, которая выбирает случайным образом шаблоны для каждого ряда и заполняет их таким образом, чтобы удовлетворять всем требованиям О1–О11 и рекомендациям Р1–Р3. В табл. 3 и табл. 4 представлены результаты работы программы в виде шаблонов рядов и рядов, заполненных пробами.

Таблица 3

Таблица шаблонов для каждого ряда

	1	2	3	4	5	6	7	8	9	10
Ряд 1	К →	В	В	К	Э.	В	И.	В	Н.Э. *П	О
Ряд 2	К	И.	В	Э.	В ←	Н.Э. *П	В	В	К	В
Ряд 3	В	К	В	И. *	К	В	Э. *П	О	О	В →
Ряд 4	К	В	И. ←*П	О *	О	О	О	Э. *П	В	В

Таблица 4

Таблица заполненных сравнительных рядов для ольфакторного анализа

	1	2	3	4	5	6	7	8	9	10
Ряд 1	6 →	3	5	7	Э.	2	И.	1	Н.Э. *П	4
Ряд 2	7	И.	1	Э.	4 ←	Н.Э. *П	5	2	6	3
Ряд 3	5	8	2	И. *	6	3	Э. *П	7	4	1 →
Ряд 4	6	4	И. ←*П	3	2	8	7	Э. *П	1	5

В табл. 3 символом «К» обозначена станция, где должна быть размещена проба, взятая из крови, символом «В» обозначена станция, где

должна быть размещена, проба, взятая с вещи, символом «О» обозначена станция, которая не будет пронохана собакой-детектором, символ «*» обозначает предполагаемое место остановки собаки-детектора, а символом «п» место пищевого подкрепления собаки-детектора.

В табл. 4 представлены сравнительные ряды, полностью удовлетворяющие требованиям и рекомендациям, построенные согласно шаблонам из табл. 3.

Заключение

В работе рассмотрена задача построения сравнительных рядов для ольфакторного анализа. Задача построения сформулирована как задача выбора четырех слов из соответствующих языков. Выполнен анализ третьего и четвертого рядов, и соответствующих им деревьев шаблонов. Получены верхние оценки числа возможных слов для этих рядов. Проведены вычислительные эксперименты, в ходе которых построены шаблоны сравнительных рядов, а также их непосредственная реализация.

СПИСОК ЛИТЕРАТУРЫ

1. *Образцов В. А.* Криминалистика: учебник. М.: Юрист, 1997. 760 с.
2. *Ахо А., Ульман Дж.* Теория синтаксического анализа, перевода и компиляции, том 1. М.: Мир, 1978. 612 с.
3. *Soldatenko A., Semenova D., Goldenok E., Kulikov V., Korshunova N.* Research on the Problem of Constructing Scent Lineups of Olfactory Analysis // 2023 IEEE 17th International Conference on Application of Information and Communication Technologies (AICT), Baku, Azerbaijan, 2023. P. 1–4,
4. *Солдатенко А. А., Семенова Д. В., Коршунова Н. В.* Исследование задачи ольфакторного анализа методами теории формальных языков // Системы управления, информационные технологии и математическое моделирование : Материалы V Всероссийской научно-практической конференции с международным участием, Омск, 25–26 апреля 2023 года / Отв. редактор В.А. Бадрызов, Омск: Омский государственный технический университет, 2023. С. 502–508.

Солдатенко Александр Александрович — кандидат физико-математических наук, доцент, СФУ. E-mail: ASoldatenko@sfu-kras.ru

Семенова Дарья Владиславовна — кандидат физико-математических наук, доцент, доцент, СФУ. E-mail: DVSeменова@sfu-kras.ru

ВЛИЯНИЕ МЕТАОКРУЖЕНИЙ НА ПОДБОР ГИПЕРПАРАМЕТРОВ СИСТЕМЫ ОБУЧЕНИЯ С ПОДКРЕПЛЕНИЕМ

А. С. Героев, О. М. Гергет

Институт проблем управления им. В. А. Трапезникова Российской академии наук, г. Москва, Россия

В работе представлен анализ влияния различных метаокружений на выбор и настройку гиперпараметров системы управления шагающего робота. Была проведена серия экспериментов по управлению роботом на разных ландшафтах с одинаковыми архитектурами агентов и с отличающимся набором гиперпараметров. На основании результатов экспериментов был сделан вывод о том, что метаокружение агента может напрямую влиять на изменение его Q-функции, если оно допускает нахождение агента в состояниях стагнации. Состояния стагнации — состояние агента, при котором всевозможные действия агента не приводят к изменению значения функции вознаграждения. **Ключевые слова:** обучение с подкреплением, метаокружение, DDPG.

Введение

Шагающие роботы - это роботы, которые могут передвигаться на ногах, подобно человеку или животным. Управление такими роботами - комплексная задача, требующая согласование множества движений, с целью достижения устойчивости и эффективности движения. В настоящее время наиболее перспективным направлением в управлении шагающими роботами является применение технологий, основанных на обучении с подкреплением. Обучение с подкреплением представляет собой концепцию взаимодействия агента и окружения. В рамках данного исследования в качестве агента выступает управляющая плата робота а средой является сам робот. Робот перемещается по определённому ландшафту, который предопределяет его возможные положения, определяя — **метаокружение робота**. Метаокружение представляет собой ландшафт, по которому перемещается робот. Метаокружение определяет предельные значения показаний окружения. В данной работе проведен анализ гиперпараметров обучения, по результатам которого сделаны рекомендации для выбора ландшафта обучения агентов при решении задач локомоции. **Средства моделирования:** Matlab, Simulink, Simscape Multibody, Reinforcement Learning Designer for MATLAB

1. Основные компоненты системы обучения с подкреплением для шагающего робота

1. Политика (policy) – это стратегия, определяющая, как агент выбирает действия в зависимости от текущего состояния среды. Политика может быть детерминированной или стохастической. [7]

2. Среда (environment) – это физическое или виртуальное пространство, в котором действует агент. Она включает в себя все элементы и условия, с которыми робот может взаимодействовать. В контексте управления шагающим роботом средой выступает сам робот.

3. Состояние (state) – это информация о текущем состоянии среды, которую агент получает в качестве входных данных для принятия решений. Состояние формируется из вектора измерений (наблюдений) за обучающей средой в данный момент времени. [9]

Наблюдения: среда робота предоставляет агенту 56 наблюдений, каждое из которых нормализовано между -1 и 1. Этими наблюдениями являются:

1. Z (вертикальное) и Y (боковое) положение центра масс туловища
2. Кватернион, представляющий ориентацию туловища
3. X (поступательная), Z (вертикальная) и Y (поперечная) скорости туловища в центре масс
4. Скорости крена, тангажа и рыскания туловища
5. Угловые положения и скорости тазобедренных и коленных шарниров для каждой ноги
6. Нормальная сила и сила трения из-за контакта с землёй для каждой ноги
7. Значения действия (крутящий момент для каждого сустава) с предыдущей временной итерации измерения показаний

Действие (action): Это выбор, сделанный агентом на основе своей политики. Действие воздействует на среду и вызывает изменения в её состоянии. **Вознаграждение** (reward): Это числовая оценка, которую агент получает от среды в ответ на совершенные действия. Цель агента - максимизировать суммарное вознаграждение, которое он получает в течение процесса взаимодействия со средой. Значение вознаграждения r определяется функцией (1).

$$r_t = v_x + 25 \frac{T_s}{T_f} - 40\dot{y}^2 - 30\theta^2 - 0.02 \sum_i (u_{t-1}^i)^2 \quad (1)$$

— v_x - скорость центра масс туловища в направлении x (м/с).

- T_s и T_f - это время выборки и время окончательного моделирования окружающей среды соответственно (с).
- \hat{y}^2 - квадратичная погрешность измерения высоты центра масс туловища с требуемой высоты в 0,75 метров (м).
- θ - угол крена туловища (рад).
- u_{t-1}^i - значение действия для сустава i с предыдущего временного шага (Н·м)

В ходе выполнения своей работы я использовал алгоритм DDPG [3] для обновления политики агента и его обучения. Применение алгоритма DDPG в управлении шагающим роботом основывается на его способности эффективно решать проблему управления непрерывными пространствами действий и состояний. Алгоритм DDPG использует нейронные сети для аппроксимации функции оценки Q и определения оптимальной стратегии действий. При этом используются две нейронные сети: критик (critic) и деятель (actor). Критик оценивает функцию Q, а деятель выбирает оптимальные действия на основе текущего состояния. Эти две сети взаимодействуют друг с другом в процессе обучения, позволяя получить более точное представление о ценности каждого состояния и выбрать наилучшее действие. В качестве метаокружений были выбраны 2 типа дорог: узкая сплошная и широкая сетчатая



Рис. 1. Узкая дорога

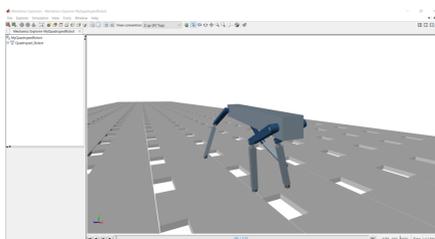


Рис. 2. Сетчатая дорога

2. Результаты исследования

Описание проводимых экспериментов проиллюстрируем на 3-х видах метаокружения (см. табл. 2)

Таблица 1

Условия экспериментов

метаокружение	число слоёв actor	число слоёв critic	число действий
узкая дорога	3	3	8
узкая дорога	3	3	12
широкая дорога	3	4	12

В первой серии экспериментов (серия содержит по 24 эксперимента) у робота отсутствовали шарниры латеральной подвижности, соответственно метаокружение не позволяло роботу перейти в состояние стагнации. В статье под состоянием стагнации понимается состояние агента, при котором всевозможные действия агента не приводят к изменению значения функции вознаграждения. Во второй серии экспериментов была реализована латеральная подвижность агента, путём добавления соответствующих шарниров, а соответственно появилась возможность стагнации. В 3-ей серии экспериментов добавление слоя dropout позволило не просто решить проблему, связанную с переобучением, а способствовало формированию у агента универсального паттерна походки, таким образом, добавление слоя dropout позволило агенту не просто обучиться «видеть» в каких местах содержатся ямы (отверстия) и обходить их, а выработать новое решение. Анализируя полученные в процессе обучения агента результаты, следует обратить внимание на рисунки 4,5, на которых явно видно, что функция вознаграждения и Q-функция на протяжении нескольких сотен эпох совпадают друг с другом. В это время агент находится в состоянии стагнации, так как его метаокружение содержит препятствия, дающие ему такую возможность. В первой серии экспериментов (рисунок 3) напротив – отсутствие подвижностей не позволило агенту войти в стагнирующее состояние и как следствие получить самые высокие значения Q-функции и функции вознаграждения. Таким образом, агенты, у которых присутствуют латеральные подвижности, должны обучаться в метаокружении без наличия препятствий, позволяющих агенту стагнировать

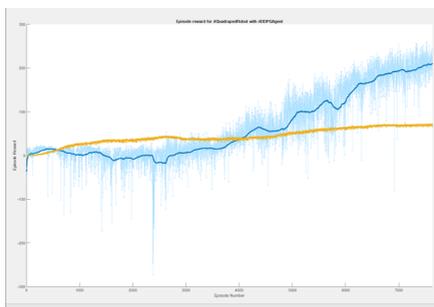


Рис. 3. Иллюстрация процесса обучения шагающего робота в отсутствии шарниров латеральной подвижности

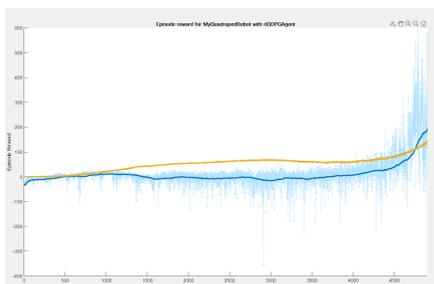


Рис. 4. Иллюстрация процесса обучения шагающего робота с шарнирами латеральной подвижности

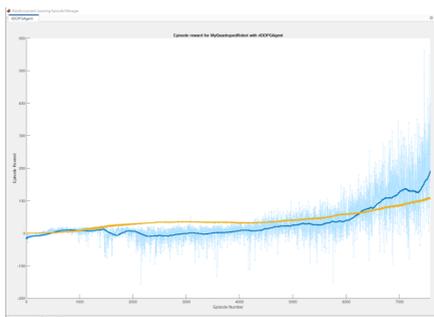


Рис. 5. Иллюстрация процесса обучения шагающего робота в отсутствии шарниров латеральной подвижности

Заключение

В ходе проделанной работы было выявлено, что в полностью детерминированных политиках необходимо составлять метаокружения агентов таким образом, чтобы избежать состояний стагнации. Если же политика имеет долю стохастичности, то метаокружение и действия агента, приводящие к стагнации могут в конечном итоге обучить агента наиболее гибкому алгоритму движения.

СПИСОК ЛИТЕРАТУРЫ

1. Reinforcement Learning Algorithms with Python. Birmingham: Packt, 2020. P. 123–143.
2. Decision Making Under Uncertainty and Reinforcement Learning Theory and Algorithms. Cham: Springer, 2022. P. 115–118.
3. https://www.researchgate.net/publication/332859622_Improving-DDPG_via_Prioritized_Experience_Replay_RL_course_report — ResearchGate 2024.
4. <https://uk.mathworks.com/help/reinforcement-learning/ug/quadruped-robot-locomotion-using-ddpg-gent.html#d124e31226> — mathworks.com 2024.

Героев Александр Сергеевич — инженер, лаборатория 80 ИПУ РАН. E-mail: *geroew_sasha@mail.ru*

Гергет Ольга Михайловна — доктор технических наук, доцент, ведущий научный сотрудник, лаборатория 80 ИПУ РАН. E-mail: *olgagerget@mail.ru*

СТАТИСТИЧЕСКИЕ ХАРАКТЕРИСТИКИ УГЛОВОГО РАСПРЕДЕЛЕНИЯ ЯРКОСТИ ДЛЯ РАЗЛИЧНЫХ МОДЕЛЕЙ АТМОСФЕРЫ

А. Р. Халиуллина

*Национальный исследовательский Томский государственный университет,
г. Томск, Россия*

Угловое распределение яркости - это функция, описывающая яркость излучения, фиксируемого приемным оптическим устройством в различных направлениях. Целью данной работы является расчет локальных оценок для углового распределения яркости точечного источника света в трехмерном пространстве для разных оптических моделей и исследование статистических характеристик данного распределения. В работе подробно описана методика моделирования движения фотонов. Также в работе представлены вычисления локальной оценки углового распределения яркости и приведены данные численных экспериментов для различных исходных условий. Кроме того, в работе представлены результаты статистической обработки имитационных моделей. **Ключевые слова:** *угловое распределение яркости, метод Монте-Карло, локальная оценка углового распределения яркости.*

Введение

Существует множество задач теории переноса излучения в атмосфере, требующих точного расчета. Самыми распространенными и актуальными являются задачи моделирования климата. При их решении важно учитывать множественные изменения светового потока при его распространении в атмосфере Земли. Именно поэтому важно строить качественные имитационные модели, учитывающие неоднородность атмосферы, локальные и временные характеристики поля излучения в ней. Для их получения в работе использован один из методов Монте-Карло - метод локальной оценки углового распределения яркости.

1. Реализация схемы «прямого моделирования» методом Монте-Карло

Взаимодействие рассматриваемой среды с излучением точечного источника моделируется как процесс случайных столкновений фотонов с аэрозольными частицами, которые приводят либо к рассеянию, либо

к поглощению [1]. Модель среды предполагает задание коэффициента поглощения, коэффициента рассеяния и коэффициента ослабления, который определяется суммой двух предыдущих. Положение фотонов в среде задается точкой, характеризующейся тремя декартовыми координатами для определения положения. Схема «прямого моделирования» методом Монте Карло представлена на рисунке 1 [3]:

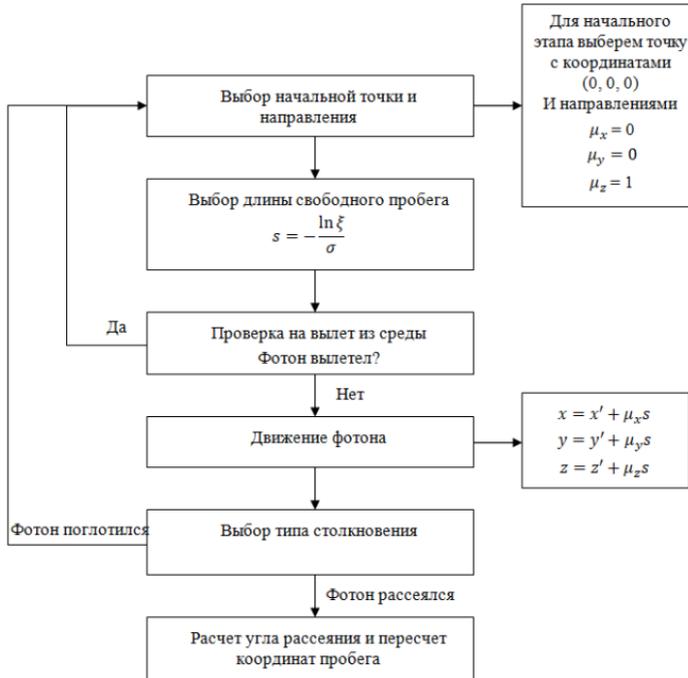


Рис. 1. Схема прямого моделирования

Пусть среда ограничена плоскостью $z = 0$, которая соответствует подстилающей поверхности Земли, и плоскостью $z = h$, при этом h является толщиной среды и выше данной границы процесс рассеяния отсутствует. Точечный приемник определяется координатами $(0, 0, h)$. Точечный источник излучения фотонов располагается на подстилающей поверхности в точке $(0, 0, 0)$. Определим начальное направление фотонов с помощью направляющих косинусов, установив их следующим образом (1):

$$\mu_x = 0; \mu_y = 0; \mu_z = 1. \quad (1)$$

Моделируем длину свободного пробега с помощью формулы (2):

$$s = -\frac{\ln \xi}{\sigma}, \quad (2)$$

где ξ - случайное число, распределенное между 0 и 1.

Формула (2) генерирует случайные длины свободного пробега s .

Для определения дальнейшего направления распространения мы аналогично будем использовать три направляющих косинуса, путем обновления координат по следующим формулам:

$$\begin{aligned} x &= x' + \mu_x s; \\ y &= y' + \mu_y s; \\ z &= z' + \mu_z s. \end{aligned}$$

В точке с обновленными координатами происходит столкновение, это столкновение приводит либо к поглощению, либо к рассеянию. Если в точке столкновения происходит поглощение, то мы возвращаемся к выбору начальной точки и направления (рис. 1), если же происходит рассеяние, то рассчитывается новое направление фотона. Чтобы уменьшить дисперсию и улучшить качество модели, применяем модификацию "без поглощения"[2]. Среднее значение косинуса угла рассеяния фотонов известно как анизотропия рассеяния g . Чтобы определить новое направление фотонов, нам нужно знать фазовую функцию рассеяния. В работе использована фазовая функция Хеньи-Гринштейна. Затем угол рассеяния θ определяется по следующей формуле (3):

$$\cos \theta = \left\{ \frac{1}{2g} \left[\frac{1 - g^2}{1 - g + 2g\xi} \right] \right\}, \text{ if } g \neq 0; \quad (3)$$

$$\cos \theta = 1 - 2\xi, \text{ if } g = 0. \quad (4)$$

На основе косинусов текущего направления мы можем найти новый набор направляющих косинусов. Новое направление распространения может быть представлено в декартовой системе координат с помощью формул (5), (6) и (7):

$$\mu'_x = \frac{\sin \theta (\mu_x \mu_z \cos \varphi - \mu_y \sin \varphi)}{\sqrt{1 - \mu_z^2}} + \mu_x \cos \theta; \quad (5)$$

$$\mu'_y = \frac{\sin \theta (\mu_y \mu_z \cos \varphi - \mu_x \sin \varphi)}{\sqrt{1 - \mu_z^2}} + \mu_y \cos \theta; \quad (6)$$

$$\mu'_z = -\sqrt{1 - \frac{z^2}{z^2}} \sin \theta \cos \varphi + \mu_z \cos \theta. \quad (7)$$

Предполагается, что полярный угол φ равномерно распределен между 0 и 2π .

Схема прямого моделирования была реализована в виде программного кода на языке python для количества фотонов, равного 10000, и геометрической толщины среды, равной 100. Рисунок 2 демонстрирует распространение фотонов в среде при анизотропиях рассеяния: 0,85; 0,4; 0,1.

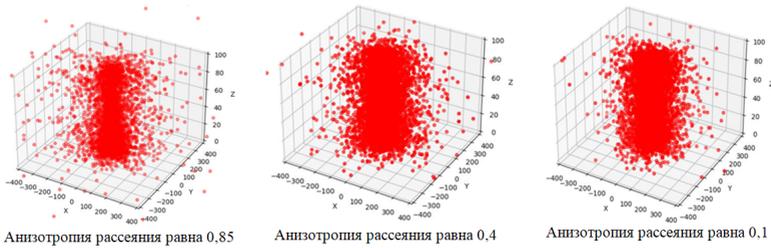


Рис. 2. Распределение фотонов в пространстве при разных значениях анизотропии

При увеличении оптической анизотропии и приближении ее значения к 1, рассеяние происходит преимущественно в прямом направлении, основной вес фотонов концентрируется вдоль оси OZ. Если же значение оптической анизотропии приближается к значению 0, то рассеяние происходит равновероятно по всем направлениям.

2. Реализация метода локальной оценки углового распределения яркости

Локальная оценка позволит проанализировать, как яркость распределяется в различных углах относительно точечного приемника и различных областей пространства. Получение углового распределения яркости предполагает вычисление следующего функционала (8):

$$\oint_{\Omega_i} \Phi(\vec{r}^*, \vec{\omega}^*) d\omega = \oint_X l_i(\vec{x}^j, \vec{x}^*) d\vec{x} = M \sum_{n=0}^N Q_n l_i(\vec{x}_n, \vec{x}^*); \quad (8)$$

$$l_i(\vec{x}, \vec{x}^*) = \frac{\exp[-\tau(\vec{r}^j, \vec{r}^*)] p(\mu^*)}{2\pi |\vec{r}^* - \vec{r}^j|^2} \Delta_i(\vec{s}^*); \quad (9)$$

$$\vec{s}^* = \frac{\vec{r}^* - \vec{r}}{|\vec{r}^* - \vec{r}|}, \mu^* = (\vec{\omega}^*, \vec{s}^*); \quad (10)$$

$$p(\mu^*) = \frac{1 - g^2}{(1 + g^2 - 2g\mu^*)^{\frac{3}{2}}}. \quad (11)$$

Здесь:

$\Delta_i(\vec{s})$ – индикатор исследуемой области;

$f(\vec{x})$ – плотность столкновений;

Φ – поток частиц в заданной точке \vec{x}^* ;

Q_n – вспомогательный вес.

Формула (8) определяет интеграл по исследуемой области пространства [1].

Формула (9) имеет смысл плотности вероятности того, что фотон, который находится в точке x , которая характеризуется положением и направлением, после столкновения рассеется по направлению к точке x^* . Точка x^* – точечный приемник.

По формуле (11) вычисляется индикатриса Хензи-Гринштейна, то есть вероятность, того что фотон рассеется под определенным углом.

Для определения областей пространства используем телесные углы. Разделим пространство на 89 телесных углов, представленных в виде круговых конусов с высотой, равной h и постоянно увеличивающимся радиусом.

Представленный алгоритм применяется для построения углового распределения яркости точечного источника, которое находится через отношение плотности вероятности (10) в области Ω_i к величине телесного угла. Мы исследовали угловое распределение яркости для коэффициента ослабления σ , численно равного 0,1 (рис. 3).

Анализ результатов позволил выявить следующие закономерности:

1) Значения плотности вероятности уменьшаются при отдалении от начальных углов приема. При малых углах отклонения от начального направления распространения фотоны испытывают меньше рассеяний в среде и, следовательно, теряют меньше энергии. По мере увеличения угла отклонения от начального направления, фотоны проходят большее расстояние в среде, что увеличивает вероятность их рассеяния. 2) По мере уменьшения показателя анизотропии g , средняя яркость точечного источника и ее вариация уменьшаются. Это указывает на то, что меньшая анизотропия приводит к более равномерному и менее интенсивному распределению яркости.

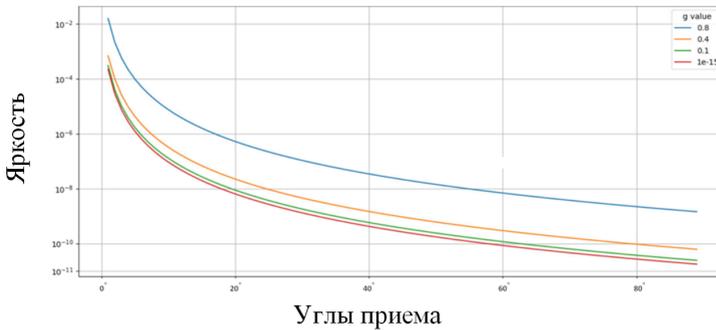


Рис. 3. Угловое распределение яркости точечного источника при $\sigma = 0.1$

Заклучение

Таким образом, результаты исследования позволяют сделать следующие выводы:

- 1) Если оптическая анизотропия увеличивается и приближается к значению 1, то рассеяние происходит преимущественно в прямом направлении, основной вес фотонов концентрируется вдоль оси OZ .
- 2) Увеличение угла отклонения фотонов от начального направления приводит к увеличению пути их прохождения в среде и, соответственно, к большей вероятности рассеяния. При уменьшении значений g происходит постепенное смещение распределения яркости к более низким значениям.

СПИСОК ЛИТЕРАТУРЫ

1. Марчук Г. И. Метод Монте-Карло в атмосферной оптике. Новосибирск: Наука, 1976. 284 с.
2. Зуев В. Е., Креков Г. М. Оптические модели атмосферы. Ленинград: Гидрометеиздат, 1986. 256 с.
3. https://en.wikipedia.org/wiki/Monte_Carlo_method — Википедия:свободная энциклопедия. 2013.

Алина Рафаиловна Халиуллина — студент ИПМКН НИ ТГУ. E-mail: alinalailai12347@mail.ru

ОБ АППРОКСИМАЦИИ МНОЖЕСТВА ТОЧЕК ПОВЕРХНОСТЬЮ НА ОСНОВЕ ПРИБЛИЖЕННОГО РАССТОЯНИЯ ИЗ РЯДА ТЕЙЛОРА

А. М. Кирилин, В. Р. Куликов

Сибирский федеральный университет, г. Красноярск, Россия

В работе исследуется вопрос аппроксимации множества точек поверхностью с точки зрения минимизации расстояния от точек до поверхности, возникший из задачи подбора новых имплантатов для краниопластики. Приведен новый подход на основе примерного расстояния, полученного из разложения неявной функции поверхности в ряд Тейлора. Произведено сравнение с подходом, базирующемся на приближении неявной функции к нулю. Новый подход показал большую устойчивость к шуму, при сохранении низких требований к функциям, задающим поверхность. **Ключевые слова:** *краниопластика, подгонка поверхности, аппроксимация поверхности, расстояние до поверхности, ряд Тейлора.*

Введение

На практике часто встречается задача «подгонки», или аппроксимации гиперповерхности к множеству точек A , иначе говоря, отыскании поверхности F^* из семейства поверхностей \mathcal{F} , проходящей максимально близко к точкам множества:

$$F^* = \arg \min_{F \in \mathcal{F}} \sum_{\mathbf{x} \in A} \rho(\mathbf{x}, F), \quad (1)$$

где $\rho(\mathbf{x}, F)$ — расстояние от точки \mathbf{x} до поверхности F . Причем, желательно, чтобы та часть поверхности, которая проходит через множество, была связанной. Именно эта задача возникла в процессе подбора новых имплантатов для краниопластики [1].

В общем случае решение такой задачи нетривиально. В первую очередь потому, что нахождение аналитической формулы $\rho(\mathbf{x}, F)$ для произвольной поверхности — это отдельный класс сложных задач. По этой

Работа поддержана Красноярским математическим центром, финансируемым Минобрнауки РФ (Соглашение 075-02-2024-1429).

причине мы вынуждены прибегать к решению эквивалентных задач или к использованию приближенного расстояния.

Одним из примеров такого является задача регрессии [2]. В ней мы минимизируем ошибку $|x_n - f(x_1, \dots, x_{n-1})|$ между экспериментальными данными x_n и значением модели $f(x_1, \dots, x_{n-1})$. Иначе говоря, мы минимизируем примерное расстояние между поверхностью F и точкой:

$$\rho(\mathbf{x}, F) \approx \rho_{reg}(\mathbf{x}, F) = |x_n - f(x_1, \dots, x_{n-1})|, \quad (2)$$

где $x_n = f(x_1, \dots, x_{n-1})$ — явный вид поверхности F . Этот подход широко используется и показывает хорошие результаты, но, очевидно, работает только с поверхностями, представимыми в явном виде.

Другим вариантом является приближение к 0 неявной функции поверхности $F(\mathbf{x})$. Иначе говоря, мы считаем её эквивалентной расстоянию в контексте этой задачи:

$$\rho(\mathbf{x}, F) \sim \rho_F(\mathbf{x}, F) = |F(\mathbf{x})|. \quad (3)$$

Действительно, и $\rho(\mathbf{x}, F) \rightarrow 0$, и $\rho_F(\mathbf{x}, F) \rightarrow 0$ при приближении \mathbf{x} к поверхности. Этот вариант потенциально применим к любой поверхности, однако эквивалентность выполняется далеко не всегда, как будет показано далее.

В этой статье мы предлагаем подход решения задачи, который объединяет надежность первого варианта и потенциальную универсальность второго. Он основан на получении приближенного расстояния из разложения неявной функции в ряд Тейлора.

1. Общие положения

Разложим задающую поверхность неявную функцию $F : \mathbb{R}^n \rightarrow \mathbb{R}$ в ряд Тейлора [3] в точке \mathbf{a} :

$$\begin{aligned} F(\mathbf{x}) = & F(\mathbf{a}) + \sum_{i=1}^n \frac{\partial F(\mathbf{a})}{\partial x_i} (x_i - a_i) + \\ & + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2 F(\mathbf{a})}{\partial x_i \partial x_j} (x_i - a_i)(x_j - a_j) + \dots + \\ & + \frac{1}{k!} \underbrace{\sum_{i_1=1}^n \dots \sum_{i_k=1}^n}_{k} \frac{\partial^k F(\mathbf{a})}{\partial x_{i_1} \dots \partial x_{i_k}} (x_{i_1} - a_{i_1}) \dots (x_{i_k} - a_{i_k}) + \dots \quad (4) \end{aligned}$$

Предположим, что \mathbf{x} — ближайшая к \mathbf{a} точка на поверхности. Тогда расстояние от \mathbf{a} до поверхности выражается через расстояние между этими точками:

$$\rho(\mathbf{a}, F) = \|\mathbf{a} - \mathbf{x}\|. \quad (5)$$

Также, $F(\mathbf{x}) \equiv 0$ в силу того, что \mathbf{x} лежит на поверхности.

Мы хотим выразить $\|\mathbf{a} - \mathbf{x}\|$ из (4). Очевидно, что решить задачу в общем случае невозможно, поэтому мы будем выражать из ряда Тейлора с остаточным членом. Пренебрегая остаточным членом, мы получим приблизительное расстояние до поверхности.

2. Ряд Тейлора с остаточным членом 1 степени

В простейшем случае мы возьмем ряд Тейлора с остаточным членом 1 степени $r_1(\mathbf{x})$:

$$0 = F(\mathbf{a}) + \sum_{i=1}^n \frac{\partial F(\mathbf{a})}{\partial x_i} (x_i - a_i) + r_1(\mathbf{x}). \quad (6)$$

Преобразуем:

$$\sum_{i=1}^n \frac{\partial F(\mathbf{a})}{\partial x_i} (x_i - a_i) = \nabla F(\mathbf{a}) \cdot (\mathbf{x} - \mathbf{a}) = -\|\nabla F(\mathbf{a})\| \|\mathbf{x} - \mathbf{a}\| \cos \alpha, \quad (7)$$

где α — угол между $\nabla F(\mathbf{a})$ и $(\mathbf{a} - \mathbf{x})$. Тогда (6) превращается в

$$0 = F(\mathbf{a}) - \|\nabla F(\mathbf{a})\| \|\mathbf{x} - \mathbf{a}\| \cos \alpha + r_1(\mathbf{x}). \quad (8)$$

Отсюда

$$\|\mathbf{x} - \mathbf{a}\| = \frac{F(\mathbf{a}) + r_1(\mathbf{x})}{\|\nabla F(\mathbf{a})\| \cos \alpha}. \quad (9)$$

Причем

$$\lim_{\mathbf{a} \rightarrow \mathbf{x}} \frac{F(\mathbf{a}) + r_1(\mathbf{x})}{\|\nabla F(\mathbf{a})\| \cos \alpha} \rightarrow \left| \frac{F(\mathbf{x})}{\|\nabla F(\mathbf{x})\|} \right|, \quad (10)$$

так как при $\mathbf{a} \rightarrow \mathbf{x}$

- $r_1(\mathbf{x}) \rightarrow 0$ по свойствам остаточного члена;
- $\cos \alpha \rightarrow \text{sign}(F(\mathbf{x}))$, поскольку $\nabla F(\mathbf{a}) \rightarrow \nabla F(\mathbf{x})$ и $F(\mathbf{x}) \parallel (\mathbf{a} - \mathbf{x})$, при этом направление $(\mathbf{a} - \mathbf{x})$ зависит от положения \mathbf{a} относительно поверхности и, соответственно, знака $F(\mathbf{x})$.

Таким образом, для \mathbf{a} вблизи поверхности можем полагать

$$\rho(\mathbf{a}, F) \approx \rho_{T1}(\mathbf{a}, F) = \left| \frac{F(\mathbf{x})}{\|\nabla F(\mathbf{x})\|} \right|. \quad (11)$$

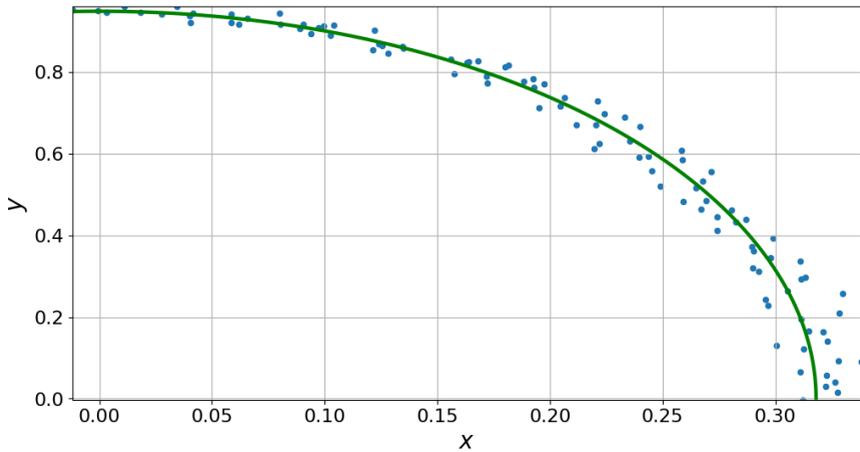


Рис. 1. Порождающий эллипс и точки с нормальным шумом, полученные из него

3. Демонстрация

Продемонстрируем различие подходов с ρ_F и ρ_{T1} на примере аппроксимации кривой второго порядка, она же квадрика. Для этого возьмем эллипс и сгенерируем множество его точек, лежащих в первой координатной четверти. После чего применим к ним нормальный шум со стандартным отклонением σ . Пример такого эллипса и множества точек можно видеть на рисунке 1.

Далее мы попробуем применить указанные подходы для нахождения аппроксимирующей кривой при различных уровнях зашумленности точек. Аппроксимирующие кривые ищутся как решение задачи (1) при соответствующих приближениях расстояния. Решение задачи производится численно, с использованием пакета `scipy` языка Python 3 [4]. Результат вы можете видеть на рисунке 2.

Как можно видеть, при нулевом шуме подходы эквивалентны. Однако, уже при $\sigma = 0.01$, подход, основанный на ρ_F , находит плохую поверхность с точки зрения исходной задачи. Это происходит из-за того, что точки, хоть и расположены далеко от найденной поверхности F^* , однако попадают в область близости F^* к 0, как показано на рисунке (3). На этом рисунке изображен контурный график квадрики, найденной с помощью ρ_F в эксперименте для $\sigma = 0.01$.

Тем не менее, при достаточно большом шуме, подход, основанный на ρ_{T1} , также перестает работать, что видно на правом графике рисунка 2.

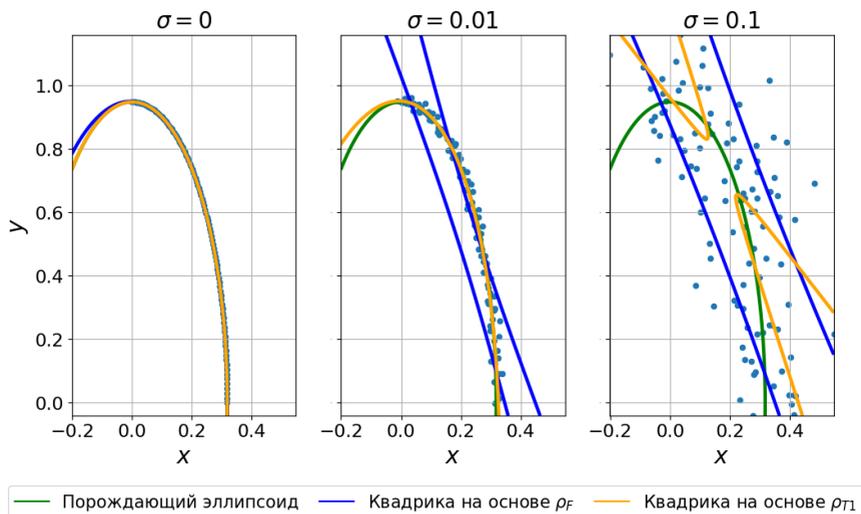


Рис. 2. Аппроксимация множества точек с разным шумом подходами с ρ_F и ρ_{T1} , а также порождающий эллипсоид

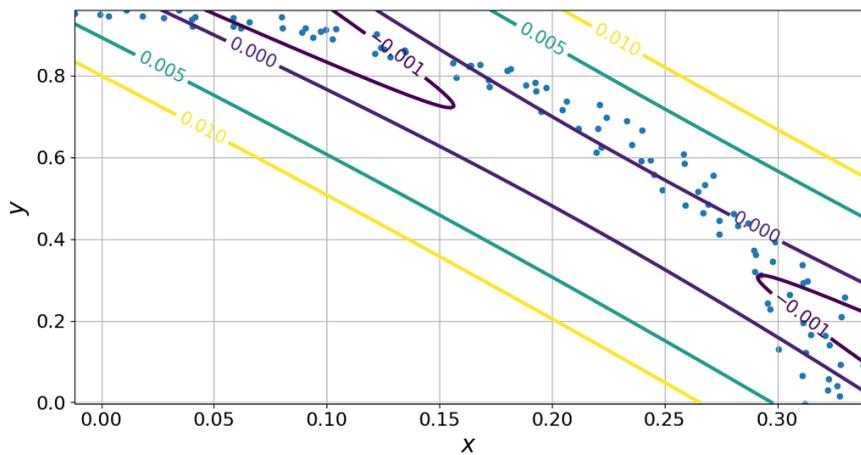


Рис. 3. Контурный график, демонстрирующий область близости к 0, в которую попадают все точки

Заключение

Таким образом, задача приближения множества точек поверхностью может быть решена с использованием приближённого расстояния, полученного из ряда Тейлора с остаточным членом первой степени. Этот метод демонстрирует более высокую устойчивость к шумам по сравнению с подходом, основанным на приближении к нулю неявной функции, определяющей поверхность.

При этом сохраняются низкие требования к поверхностям — подходят любые поверхности, задаваемые неявной дифференцируемой функцией, в отличие от «регрессионного» подхода.

В дальнейшем можно попытаться выразить приближённое расстояние из ряда Тейлора с остаточным членом второй и более высокой степени. Вероятно, использование ряда Тейлора с более высокой степенью остаточного члена повысит точность приближения расстояния, что позволит игнорировать более высокую степень шума.

СПИСОК ЛИТЕРАТУРЫ

1. *Кирилин А. М.* Об одном методе аппроксимации височной зоны черепа поверхностью второго порядка // Информационные технологии и математическое моделирование (ИТММ-2023) : Материалы XXII Международной конференции имени А.Ф. Терпугова, Томск, 2023. С. 285–290.
2. *Демиденко Е. З.* Линейная и нелинейная регрессия. Москва: Финансы и статистика, 1981. 302 с.
3. *Hörmander L.* The Analysis of Linear Partial Differential Operators I. Classics in Mathematics. Berlin: Springer, 2003. 440 p.
4. <https://docs.scipy.org/doc/scipy/reference/optimize.html> — Optimization and root finding : SciPy documentation. 2024.

Кирилин Артём Максимович — магистрант 2-го курса ИМиФИ СФУ. E-mail: kirilinartem@yandex.ru

Куликов Владимир Русланович — к.ф.-м.н., доцент ИМиФИ СФУ. E-mail: VRKulikov@sfu-kras.ru

О МОДИФИКАЦИЯХ АЛГОРИТМА ГЕНЕРАЦИИ СЕТЕВОГО ГРАФИКА ПРОЕКТОВ ДЛЯ ЗАДАЧИ RCPSP

А. М. Булавчук, Д. В. Семенова, А. А. Солдатенко

Сибирский федеральный университет, г. Красноярск, Россия

В работе предлагаются две модификации разработанного ранее алгоритма, предназначенного для генерации модельных инвестиционных проектов для задачи календарного планирования с ограниченными ресурсами и максимизацией чистой приведённой стоимости. Данная работа расширяет изначальный алгоритм на два случая: с фиксированной вероятностью появления ребра к вершинам с предыдущего слоя и “затухающей” вероятностью появления ребра к вершинам из всех предыдущих слоев. Предложенные модификации алгоритма позволяют гибко настраивать характеристики проекта. Полученные проекты могут использоваться для тестирования оптимизационных алгоритмов решения задачи календарного планирования инвестиционных проектов с ограниченными ресурсами в денежной форме.

Ключевые слова: *задача календарного планирования, инвестиционный проект, сетевой график проекта.*

Введение

Фокус наших исследований сосредоточен на задаче календарного планирования инвестиционных проектов с ограниченными ресурсами в денежной форме (RCPSP). Критерием оптимального расписания начала для каждой из работ проекта выступает максимум чистой приведённой стоимости, при котором выполняются ограничения на достаточность средств и учитываются технологические взаимосвязи между работами. Данная задача является NP -трудной в сильном смысле [1]. В [2] доказано, что каждое из расписаний проекта может быть представлено в виде произведения матрицы специального вида, рассчитанной на основе матрицы частичного порядка проекта, и вектора из идемпотентного полумодуля. В [2, 3] для решения рассматриваемой задачи были предложены эвристические алгоритмы GASPIA и SASPIA, основанные на методах идемпотентной алгебры. Тестирование алгоритмов про-

Работа поддержана Красноярским математическим центром, финансируемым Минобрнауки РФ (Соглашение 075-02-2024-1429).

дились на сгенерированных случайным образом потоках платежей и проектах, представленных в открытых базах данных [5, 6].

На этапе тестирования возникла проблема, связанная с использованием готовых экземпляров задач из открытых библиотек, поскольку все используемые в них генераторы предназначены для задач ресурсно-календарного планирования проекта с ограниченными ресурсами по критерию минимизации длины расписания выполнения всех работ проекта [7, 8, 10, 11, 9]. В [4] предложен собственный алгоритм генерации сети проектов для экземпляров задач календарного планирования инвестиционных проектов с ограниченными ресурсами. В данной работе рассматриваются две его модификации.

1. Алгоритм для генерации сети проектов типа AoN

В [4] был предложен алгоритм для генерации сети проектов типа AoN (activity-on-the-node). Псевдокод алгоритма представлен ниже.

Алгоритм 1. Алгоритм для генерации сети проектов типа AoN

Вход: n — число промежуточных работ, k — число этапов генерации

Выход: Итоговая сеть из $n + 2$ проектов (к промежуточным работам добавляются вершины источник и сток), максимально возможная глубина сети не превосходит $k + 2$, $n > k$.

- 1: Инициализировать ориентированный граф G с одной вершиной;
 - 2: $generated = 0$, $layer = 1$;
 - 3: **пока** $generated < n$:
 - 4: $genNode \sim \mathcal{U}(1, \lfloor (n - generated)/(k + 1 - layer) \rfloor)$;
 - 5: **если** $layer == k$, **то**
 - 6: $genNode = n - generated$;
 - 7: $p \sim \mathcal{U}(0, 1)$;
 - 8: **для** $i = 0$ до $genNode$:
 - 9: Добавить в граф G вершину u с номером $(generated + 2)$;
 - 10: $generated+ = 1$;
 - 11: **для** всех вершин v , сгенерированных на предыдущем этапе:
 - 12: Добавить дугу (v, u) в граф G с вероятностью p ;
 - 13: **если** в вершину u не входит ни одна дуга, **то**
 - 14: добавить дугу $(1, u)$;
 - 15: $layer+ = 1$;
 - 16: Добавить в граф G вершину с номером $(n + 2)$;
 - 17: Вернуть граф G .
-

В алгоритме 1 запись $x \sim \mathcal{U}(a, b)$ означает вызов функции, которая возвращает случайное число равномерно распределенное на отрезке $[a, b]$. Переменная *generated* отслеживает количество уже сгенерированных вершин в графе, *layer* — текущий этап алгоритма, а p — вероятность появления дуги с вершинами с предыдущего этапа. Операция $\lfloor \cdot \rfloor$ возвращает ближайшее целое число меньше или равное заданному. Алгоритм строит ориентированный граф G , содержащий $n + 2$ вершины, где вершина с номером 1 является источником и первой работой, а вершина с номером $n + 2$ является стоком и окончательной работой. Полученный граф G всегда удовлетворяет условию частичного порядка и не содержит избыточных дуг. Избыточной дугой называется дуга (u, v) такая, что также существует маршрут из u в v длины большей или равной 2. Отметим, что на последнем этапе алгоритма будут сгенерированы все недостающие вершины (шаг 5).

2. Модификации алгоритма

Данная работа расширяет алгоритм 1 на два случая: с постоянной вероятностью появления ребра к вершинам с предыдущего слоя и “затухающей” вероятностью появления ребра к вершинам из всех предыдущих слоев.

Первая модификация фиксирует вероятность появления ребра между вершинами текущего слоя и предыдущего. В этой модификации из алгоритма 1 исключается шаг 7 и вероятность p подается на вход алгоритма и далее не меняется. Пример работы алгоритма представлен на рисунках 1 и 2.

Вторая модификация дает возможность появлению ребер не только между вершинами из соседних слоев, но и между вершиной с текущего слоя и вершинами со всех предыдущих слоев. Вероятность появления ребра определяется следующим образом:

$$p(v, u) = p_{const} \cdot \frac{1}{\tau}, u \in L_i, v \in L_j, i \neq j, \quad (1)$$

где p_{const} — вероятность появления ребра между вершинами соседних слоев L_i и L_{i-1} , а τ — это расстояние между текущим слоем L_i и проверяемым L_j , которое может быть вычислено как разность их номеров, т. е. $\tau = (i - j)$. В данной модификации вероятность на шаге 7 пересчитывается для каждого проверяемого слоя отдельно по формуле (1), а цикл на шаге 11 теперь запускается для всех предыдущих сгенерированных слоев. Как и в прошлом алгоритме, вероятность появления ребра между соседними слоями p_{const} подается на вход алгоритма. На рисунках 3 и 4 представлены пример работы алгоритма.

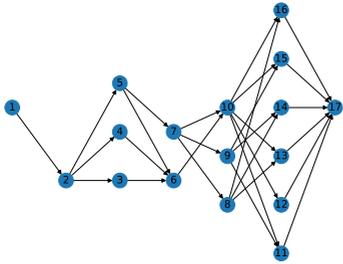


Рис. 1. Результат генерации с параметрами $n = 15$, $k = 5$, $p_{const} = 0,75$

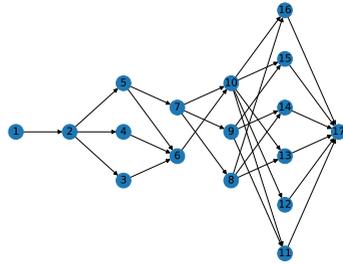


Рис. 2. Результат генерации с параметрами $n = 15$, $k = 5$, $p_{const} = 0,75$ после выполнения сортировки

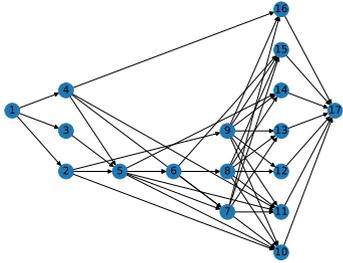


Рис. 3. Результат генерации с параметрами $n = 15$, $k = 5$, $p_{const} = 0,75$ с «затухающей» вероятностью

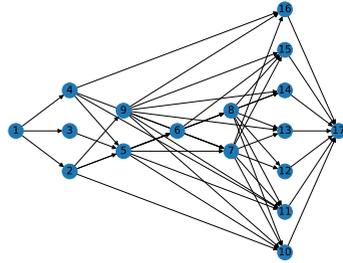


Рис. 4. Результат генерации с параметрами $n = 15$, $k = 5$, $p_{const} = 0,75$ с «затухающей» вероятностью после выполнения сортировки

Представим результаты нескольких сгенерированных сетей с популярными оценками [9] в таблице 1. В таблице использованы следующие обозначения: алгоритм A_1 – алгоритм с постоянной вероятностью появления ребра между вершинами соседних слоев, алгоритм A_2 – алгоритм с «затухающей» вероятностью появления ребра к вершинам из всех предыдущих слоев, k – количество слоев, CNC – коэффициент сложности сети, CNC' – коэффициент Кайманна, A_{red} – число избыточных ребер.

Заключение

Целью нашего исследования является разработка собственного генератора экземпляров задач календарного планирования инвестиционных проектов с ограниченными ресурсами и критерием максимизации

Таблица 1

Результаты генерации сетей с числом вершин $n = 15$

p_{const}	k	$ E $	CNC	CNC'	T -плотность	A_{red}
Алгоритм A_1						
0,25	3	28	1,647	46,118	13	0
0,5	3	29	1,705	49,471	13	0
0,75	3	48	2,824	135,529	21	0
0,25	7	27	1,588	42,882	12	0
0,5	7	27	1,588	42,882	10	0
0,75	7	26	1,529	39,765	10	0
Алгоритм A_2						
0,25	3	29	1,706	49,471	13	0
0,5	3	36	2,118	76,235	18	3
0,75	3	51	3,000	153,000	36	2
0,25	7	24	1,412	33,882	9	0
0,5	7	30	1,765	52,941	15	0
0,75	7	47	2,765	129,941	27	16

чистой приведённой стоимости. Данный генератор состоит из четырёх модулей.

- 1) Генератор сетевого графика проекта.
- 2) Генератор длительности проектных работ. В текущей версии для генерации указывается максимальная длительность в целом числе периодов.
- 3) Генератор потоков платежей. Характеристики потоков моделируются помощью иерархических копул с различными маргинальными распределениями. Кроме того, могут быть сгенерированы и нечёткие параметры потоков.
- 4) Генератор бюджета проекта.

В статье сделан акцент на первый из модулей. Предложены две модификации алгоритма генерации сети проектов типа AoN. Представлены результаты экспериментов по генерации сетей, а также вычислены популярные оценки сложности. Перспективным направлением исследований является определение теоретических оценок числа генерируемых в сети ребер и получаемых при этом значений оценок.

СПИСОК ЛИТЕРАТУРЫ

1. Казаковцева Е. А., Сервал В. В. Сложность задачи календарного планирования с кредитованием // Дискретн. анализ и исслед. опер. 2015. Т. 22. № 4. С. 35–49.

2. Булавчук А. М., Семенова Д. В. Применение методов идемпотентной алгебры в генетическом алгоритме для решения задачи календарного планирования / А. М. Булавчук, Д. В. // Прикладная дискретная математика. 2022. № 58. С. 112–124.
3. *Bulavchuk A. M., Semenova D., V.* Two Heuristic Algorithms for RCPSP with NPV Criterion // Journal of Siberian Federal Universit. Mathematics and Physics. 2023. Vol. 16, No. 5. P. 639–650.
4. Булавчук А. М., Семенова Д. В., Солдатенко А. А. Об одном генераторе проектов для задачи RCPSP // ИТ. НАУКА. КРЕАТИВ. Т. 5. Системы управления, информационные технологии и математическое моделирование : материалы I Междунар. форума (Омск, 14-16 мая 2024 г.) : в 5 Ч. / науч. ред. П. С. Ложников, отв. ред. И. Г. Ольгина. Омск : Издательство ОмГТУ, 2024. 70 с.
5. Project Scheduling Problem Library – PSPLIB <https://www.omdb.wi.tum.de/psplib/main.html>
6. *Kolisch R., Sprecher A.* PSPLIB – a Project Scheduling Problem Library // Eur. J. Oper. Res, 1997. Vol. 96. P. 205–216.
7. *Demeulemeester E., Dodin B., Herroelen W.* A random activity network generator // Operations Research. 1993. Vol. 41(5). P. 972–980.
8. *Agrawal M. K., Elmaghraby S. E., Herroelen W. S.* DAGEN: a generator of test sets for project activity nets // European Journal of Operational Research. 1996. Vol. 90(2). P. 376–382.
9. *Browning T. R., Yassine A. A.* A random generator of resource-constrained multi-project network problems // J Sched. 2010. Vol. 13. P. 143–161.
10. *Demeulemeester, E., Vanhoucke, M., Herroelen, W.* RanGen: A random network generator for activity-on-the-node networks // Journal of Scheduling. 2005. Vol. 6(1). P. 17–38.
11. *Vanhoucke M., Coelho J., Debels D., Maenhout B., Tavares L. V.* An evaluation of the adequacy of project network generators with systematically sampled networks // European Journal of Operational Research. 2008. Vol. 187(2). P. 511–524.

Булавчук Александр Михайлович — старший преподаватель, СФУ. E-mail: Bulavchuk@gmail.com

Семенова Дарья Владиславовна — кандидат физико-математических наук, доцент, доцент, СФУ. E-mail: DVSeменова@sfu-kras.ru

Солдатенко Александр Александрович — кандидат физико-математических наук, доцент, СФУ. E-mail: ASoldatenko@sfu-kras.ru

О ПРОСТРАНСТВЕННО-НЕОДНОРОДНОМ РАСПРОСТРАНЕНИИ БОЛЕЗНИ

Д. А. Подолин, А. Э. Рассадин

Национальный исследовательский университет «Высшая школа экономики», г. Нижний Новгород, Россия

Статья посвящена изучению распространения инфекционных болезней в рамках двухкомпарментной («здоровые» и «больные») модели эпидемии, альтернативной моделям типа «реакция-диффузия», а именно, в работе анализ пространственного распределения заболевших особей производится за счёт включения в модель адвективных членов для учета пространственной миграции носителей заболевания. В частности, в модели учитывается пространственная неоднородность как скорости миграции, так и скорости распространения инфекции. В статье дана общая теория решения задачи Коши на прямой для рассматриваемой модели, а также приводится пример явного решения, построенного в рамках этой теории. Если в качестве особей популяции брать людей, то предложенную модель можно использовать для повышения эффективности карантинных мероприятий на границах национальных государств в периоды пандемий. Но, вообще говоря, рассматриваемая в работе модель эпидемии применима не только к людям, но и к популяциям животных — млекопитающих, птиц, рыб и т. д. **Ключевые слова:** *метод характеристик, уравнение переноса, эйлера и лагранжева переменные.*

Введение

Анализ распространения инфекционных болезней через границы национальных государств очень важен для грамотной организации карантинных мероприятий при возникновении пандемий [1].

Обычно в математической эпидемиологии пространственное распределение заболевших исследуют с помощью моделей эпидемии типа системы «реакция-диффузия» (см. [2] и ссылки там). Однако для таких систем нахождение аналитических решений, как правило, невозможно, а получение достаточной информации об их поведении требует проведения большого объёма численных расчётов. Поэтому вопрос об альтернативном описании миграции больных популяций весьма актуален.

В работах [3, 4] для учёта пространственной миграции носителей заболевания в модель был включён однородный по пространственной

переменной адвективный член. В данной работе этот подход распространён на случай миграции заражённой популяции по пересечённой местности.

Далее статья организована следующим образом: в разделе 1 дана математическая формулировка модели. В разделе 2 изложена теория нахождения общего решения уравнений модели. В разделе 3 приведен пример их явного решения. В Заключении суммированы полученные результаты и намечены перспективы дальнейших исследований.

1. Постановка задачи

Рассмотрим следующую распределённую динамическую систему:

$$\frac{\partial I}{\partial t} + V(x) \frac{\partial I}{\partial x} = \beta(x) S I, \quad \frac{\partial S}{\partial t} + V(x) \frac{\partial S}{\partial x} = -\beta(x) S I. \quad (1)$$

В системе квазилинейных дифференциальных уравнений в частных производных первого порядка (1) неизвестные функции имеют смысл линейных плотностей здоровых $S(x, t)$ и больных $I(x, t)$ особей соответственно, а скорость миграции V зависит от пространственной координаты x : $V = V(x)$. Скорость распространения инфекции β также считается зависящей от неё: $\beta = \beta(x)$.

К системе (1) надо добавить начальные условия:

$$I(x, 0) = I_0(x) \geq 0, \quad S(x, 0) = S_0(x) \geq 0, \quad x \in \mathbb{R}. \quad (2)$$

Задача Коши (1)-(2) и является математической формулировкой модели миграции больной популяции.

2. Общее решение уравнений модели

Рассмотрим полную линейную плотность популяции $N(x, t)$:

$$N(x, t) = I(x, t) + S(x, t). \quad (3)$$

Вычислив с помощью уравнений (1) частную производную по времени от функции (3), получим, что она удовлетворяет следующему линейному уравнению в частных производных первого порядка:

$$\frac{\partial N}{\partial t} + V(x) \frac{\partial N}{\partial x} = 0. \quad (4)$$

Начальное условие к уравнению переноса (4) легко находится из формул (2) и (3):

$$N(x, 0) \equiv N_0(x) = I_0(x) + S_0(x). \quad (5)$$

Задача Коши (4)-(5) решается методом характеристик [5], а именно, уравнения характеристик имеют вид:

$$\frac{dx}{dt} = V(x), \quad x|_{t=0} = y \quad (6)$$

и

$$\frac{dN}{dt} = 0, \quad N|_{t=0} = N_0(y), \quad (7)$$

где $y \in \mathbb{R}$ — лагранжева координата [6].

Очевидно, что решение задачи Коши (6) есть:

$$t = \int_y^x \frac{d\xi}{V(\xi)}. \quad (8)$$

Выражение (8) можно преобразовать либо к виду зависимости эйлеровой координаты x от лагранжевой: $x = X(y, t)$, либо наоборот, лагранжевой координаты y от эйлеровой: $y = Y(x, t)$, причём для $\forall t \geq 0$ между этими функциями справедливо следующее соотношение [6]:

$$Y(X(y, t)) = y. \quad (9)$$

После того, как функция $y = Y(x, t)$ найдена, из формул (7) легко получается, что:

$$N(x, t) = N_0(Y(x, t)). \quad (10)$$

Далее, поскольку функция (10) теперь полностью известна, то исключение линейной плотности здоровых особей с помощью соотношения (3) приводит к следующей задаче Коши на прямой для линейной плотности больных особей:

$$\frac{\partial I}{\partial t} + V(x) \frac{\partial I}{\partial x} = \beta(x) I (N_0(Y(x, t)) - I), \quad I|_{t=0} = I_0(x). \quad (11)$$

По задаче Коши (11) снова можно выписать уравнения её характеристик [5].

Первое из этих уравнений есть:

$$\frac{dI}{dt} = \beta(x) I (N_0(Y(x, t)) - I), \quad I|_{t=0} = I_0(y), \quad (12)$$

а второе совпадает с уравнением (6).

В силу этого совпадения с помощью тождества (10) задачу Коши (12) можно переписать в более простом виде:

$$\frac{dI}{dt} = \beta(X(y, t)) I (N_0(y) - I), \quad I|_{t=0} = I_0(y). \quad (13)$$

Введём новое время:

$$T(y, t) = \int_0^t \beta(X(y, \theta)) d\theta, \quad (14)$$

тогда в этом времени задача Коши (13) станет совсем простой:

$$\frac{dI}{dT} = I(N_0(y) - I), \quad I|_{T=0} = I_0(y). \quad (15)$$

Точное решение задачи Коши (15) равно:

$$I = \frac{I_0(y) N_0(y)}{I_0(y) + S_0(y) \exp(-T N_0(y))}. \quad (16)$$

Переходя в выражении (16) от лагранжевых координат к эйлеровым, получим точное решение задачи Коши (11):

$$I(x, t) = \frac{I_0(Y(x, t)) N_0(Y(x, t))}{I_0(Y(x, t)) + S_0(Y(x, t)) \exp(-\tau(x, t) N_0(Y(x, t)))}, \quad (17)$$

где $\tau(x, t) = T(Y(x, t), t)$ — время (14) в эйлеровых координатах.

Вторая неизвестная функция $S(x, t)$ вычисляется с помощью формул (3) и (10):

$$S(x, t) = \frac{S_0(Y(x, t)) N_0(Y(x, t)) \exp(-\tau(x, t) N_0(Y(x, t)))}{I_0(Y(x, t)) + S_0(Y(x, t)) \exp(-\tau(x, t) N_0(Y(x, t)))}, \quad (18)$$

Функции (17) и (18) полностью определяют общее решение задачи Коши (1) – (2).

3. Пример точного решения уравнений модели

Выберем в системе уравнений (1) следующие переменные коэффициенты $V(x)$ и $\beta(x)$:

$$V(x) = \frac{V_m}{\operatorname{ch}(x/L)}, \quad (19)$$

и

$$\beta(x) = \beta_1 + \frac{\beta_2}{\operatorname{ch}^2(x/L)}, \quad (20)$$

где V_m, β_1, β_2 и L — положительные параметры.

Точное решение задачи Коши (6) с правой частью (19) можно записать в виде:

$$\operatorname{sh} \frac{x}{L} = \operatorname{sh} \frac{y}{L} + \frac{V_m t}{L}. \quad (21)$$

Подставляя это выражение в формулу (20) и интегрируя по времени t , получим время (14) для данного случая:

$$T(y, t) = \beta_1 t + \frac{\beta_2 L}{V_m} \left[\operatorname{arctg} \left(\operatorname{sh} \frac{y}{L} + \frac{V_m t}{L} \right) - \operatorname{arctg} \operatorname{sh} \frac{y}{L} \right]. \quad (22)$$

С помощью (21) функцию (22) несложно переписать в эйлеровой координате x :

$$\tau(x, t) = \beta_1 t + \frac{\beta_2 L}{V_m} \left[\operatorname{arctg} \operatorname{sh} \frac{x}{L} - \operatorname{arctg} \left(\operatorname{sh} \frac{x}{L} - \frac{V_m t}{L} \right) \right]. \quad (23)$$

Наконец, выразив из соотношения (21) лагранжеву координату y , получим, что:

$$Y(x, t) = L \operatorname{sh}^{-1} \left(\operatorname{sh} \frac{x}{L} - \frac{V_m t}{L} \right). \quad (24)$$

Для того, чтобы через функции (22) и (23) полностью определить точное решение (17) – (18) задачи Коши (1) – (2) в явном виде, нужно ещё задать начальные условия.

Выберем начальное условие (5) в виде гауссоиды:

$$N_0(x) = \frac{n}{l\sqrt{\pi}} \exp \left[-\frac{(x + x_m)^2}{l^2} \right], \quad (25)$$

где n — полное число особей в популяции, l — характерная ширина её начальной линейной плотности, а положительная величина x_m определяет положение максимума функции (25) слева от начала координат.

Далее, определим начальные условия (2) через выражение (25) следующим образом:

$$I_0(x) = \alpha N_0(x), \quad S_0(x) = (1 - \alpha) N_0(x), \quad (26)$$

где параметр α удовлетворяет условию: $0 < \alpha < 1$.

Заключение

В работе построено общее решение двухкомпарментной модели эпидемии, учитывающей как миграцию особей заражённой популяции по пересечённой местности, так и пространственную неоднородность скорости распространения заболевания. Для начальных условий (25)-(26) предъявлено конкретное решение модели, выражающееся в явном виде через элементарные функции.

Перспективой развития данной работы является введение в уравнения системы (1) диффузионных членов с малыми коэффициентами

диффузии. Такая нелинейная система уравнений уже не имеет аналитического решения в общем виде, тем не менее, найденные в данной работе решения задачи Коши (1) – (2) могут служить нулевым приближением для построения асимптотического решения этой системы в рамках теории сингулярных возмущений [7].

СПИСОК ЛИТЕРАТУРЫ

1. *Беляков В. Д., Яфаев Р. Х.* Эпидемиология. М.: Медицина, 1989. 416 с.
2. *Brauer F., Driessche P., Wu J.* Mathematical epidemiology. Berlin, Heidelberg: Springer-Verlag, 2008. 414 p.
3. *Аристов В. В., Строганов А. В., Ястребов А. Д.* Применение модели кинетического типа для изучения пространственного распространения COVID-19 // Компьютерные исследования и моделирование. 2021. Т. 38. Вып. 3. С. 611–627.
4. *Бугров В. О., Рассадин А. Э.* Модель распространения пандемии с двумя устойчивыми состояниями // «Математическое моделирование, численные методы и комплексы программ»: X Международная научная молодежная школа-семинар имени Е.В. Воскресенского (Саранск, 14-18 июля 2022 г.), Саранск: СВМО, 2022. С. 40–48.
5. *Лерман Л. М.* Лекции по обыкновенным дифференциальным уравнениям. М.-Ижевск: НИЦ «Регулярная и хаотическая динамика»; Институт компьютерных исследований, 2016. 280 с.
6. *Гурбатов С. Н., Руденко О. В., Саичев А. И.* Волны и структуры в нелинейных средах без дисперсии. Приложения к нелинейной акустике. М.: ФИЗМАТЛИТ, 2008. 496 с.
7. *Васильева А. Б., Бутузов В. Ф.* Асимптотические методы в теории сингулярных возмущений. М.: Высшая школа, 1990. 208 с.

Подолин Дмитрий Алексеевич — студент, кафедра фундаментальной математики. E-mail: d-podolin@mail.ru

Рассадин Александр Эдуардович — аспирант, кафедра фундаментальной математики. E-mail: brat_ras@inbox.ru

ОБЗОР МЕТОДОВ ГЕНЕРАЦИИ ЗНАКОВЫХ ГРАФОВ

Э. И. Ибрагимова

Сибирский федеральный университет, г. Красноярск, Россия

В данной статье представлен обзор методов генерации знаковых графов с заданными свойствами: модель предпочтительного присоединения, модели копирования ребер и клик, модели биномиального и степенного распределения степеней, адаптация метода Ваксмана, генерация полных и k -сбалансированных графов. Рассмотренные алгоритмы реализованы в виде комплекса программ. Полученные в результате вычислительных экспериментов знаковые графы использовались для тестирования эвристических алгоритмов решения NP-трудной задачи корреляционной кластеризации. **Ключевые слова:** *генерация знаковых графов, модель предпочтительного присоединения, модель копирования.*

Введение

Знаковым называется граф, каждое ребро которого имеет знак: «+» или «-». Данные графы имеют множество приложений, среди которых: анализ социальных сетей, анализ поведения избирателей, телекоммуникационные технологии [11]. В настоящее время активно развиваются алгоритмы и методы обработки знаковых графов, для тестирования которых требуются знаковые графы с заданными свойствами. В связи с этим актуальна разработка методов генерации и получения знаковых графов с заданными характеристиками.

В статье [1] рассматриваются такие методы получения знаковых графов из реальных данных, как непосредственное задание знакового графа, индуцирование из беззнакового графов [3, 9] и построение на основе корреляционной матрицы. В данной статье представлен краткий обзор методов генерации знаковых сетей с заданными свойствами.

1. Алгоритмы генерации знаковых графов

Можно выделить два основных типа алгоритмов генерации знаковых графов. Первый — это двухэтапная генерация, при которой сначала

Алгоритмы генерации знаковых графов

Алгоритм	Стратегия
По методу Ваксмена [8]	Беззнаковый граф генерируется по методу Ваксмена [12]. Знаки ребер формируются по схеме Бернулли с вероятностью положительного знака ребра p .
Полные графы по схеме Бернулли [8]	Для каждой пары вершин генерируется знак ребра, например, по схеме Бернулли с вероятностью положительного знака ребра p .
Генерация k -сбалансированного графа	Сначала строятся k графов с положительными ребрами, а затем соединяются между собой отрицательными ребрами.
Предпочтительное присоединение [10]	Адаптация алгоритма Барабаша-Альберта [4]
Модель копирования ребер [10]	При добавлении новой вершины к графу, она соединяется с уже существующими с помощью копирования существующих ребер.
Модель копирования клик [10]	Выбирается некоторая клика в графе. Новая вершина соединяется с вершинами клики с помощью копирования ребер некоторой вершины из выбранной клики.
Модель биномиального распределения степеней [5]	Генерируется беззнаковый граф с биномиальным распределением степеней вершин. Затем вершины делятся на две группы. Если вершины ребра находятся в одной группе, то ребро положительно, в противном случае — отрицательно.
Модель степенного распределения степеней [5]	Сначала генерируется беззнаковый граф со степенным распределением степеней вершин. Затем применяется модель копирования [7] и происходит процесс «переподключения» ребер.

генерируется беззнаковый граф, а затем знаки ребер. Для второго типа алгоритмов характерна следующая процедура построения. К некоторому изначально знаковому графу добавляется новая вершина и соединяется определенным способом с некоторыми уже существующими в графе вершинами. В таблице 1 приведены рассматриваемые алгоритмы генерации знаковый графов.

В работе [8] для генерации беззнакового графа предлагается использовать метод Ваксмана с параметрами 0,15 и 0,4. Знаки ребер для каждого графа формировались по схеме Бернулли с параметром p , который обозначает вероятность положительного знака ребра. Основная идея метода Ваксмана [12] заключается в случайном распределении вершин по прямоугольной координатной сетке и определением наличия ребра на основе евклидова расстояния между вершинами. Вероятность ребра определяется по формуле:

$$P(\{u, v\}) = \beta \exp \frac{-d(u, v)}{L\alpha}, \quad (1)$$

где $d(u, v)$ — расстояние между вершинами u, v , L — максимальное расстояние между двумя вершинами. Большие значения β соответствуют большей плотности ребер, а малые значения α увеличивают плотность коротких ребер по сравнению с длинными, при этом $\alpha, \beta \in (0, 1]$.

Для генерации полных знаковых графов в в [8] используется следующий способ: для каждой пары вершин генерируется знак ребра по схеме Бернулли с параметром p обозначающим вероятность положительного знака ребра.

Для проверки алгоритмов требуются графы с заранее заданными свойствами. Например, k -сбалансированный знаковый граф можно получить построив k графов с положительными ребрами, а затем соединив их между собой отрицательными ребрами. Также можно внести заданное число положительных и отрицательных ошибок.

В работе [10] авторы предлагают адаптацию известной модели Барабаши-Альберта для генерации ориентированного знакового графа. Алгоритм принимает на вход ориентированный знаковый граф $G_0^k = (V_0, E_0)$ размера $|V_0| = 2k + 1$. Предполагается, что в G_0^k ровно k положительных и k отрицательных ориентированных ребер. В момент времени $t + 1$ строится граф G_{t+1}^k добавлением новой вершины v_{t+1} . Новая вершина v_{t+1} соединяется с k существующими в G_t^k вершинами в качестве их положительного исходящего соседа с вероятностью

$$P[v_{t+1} \text{ is positive - out - neighbor of } v] = \frac{2d_{G_t^k}^{+out}(v)}{\sum_v d_{G_t^k}^+(v)}; \quad (2)$$

где $v \in V_t$. Вершина v_{t+1} соединяется с k существующими вершинами в качестве их отрицательного исходящего соседа с вероятностью

$$P[v_{t+1} \text{ is negative - out - neighbor of } v] = \frac{2d_{G_t^k}^{-out}(v)}{\sum_v d_{G_t^k}^-(v)}; \quad (3)$$

где $v \in V_t$. Кроме того, новая вершина v_{t+1} соединяется с k существующими вершинами в качестве их положительного заходящего соседа и с k существующими вершинами в качестве их отрицательного заходящего соседа с вероятностями

$$P[v_{t+1} \text{ is negative - out - neighbor of } v] = \frac{2d_{G_t^k}^{+in}(v)}{\sum_v d_{G_t^k}^+(v)}; \quad (4)$$

$$P[v_{t+1} \text{ is negative - out - neighbor of } v] = \frac{2d_{G_t^k}^{-in}(v)}{\sum_v d_{G_t^k}^-(v)}; \quad (5)$$

В работе [10] также предлагается модель копирования ребер. Алгоритм принимает на вход ориентированный знаковый граф G_0^k с множеством вершин размера $|V_0| = 2k + 1$. Начальный граф G_0^k содержит ровно k положительных и k отрицательных ориентированных ребер. В момент времени $t + 1$, строится G_{t+1}^k добавлением новой вершины v_{t+1} в граф. Вершина v_{t+1} соединяется с k существующими вершинами копированием k ранее не копировавшихся ребер из G_t^k . Процедура копирования выполняется следующим образом.

- 1) Случайным образом выбирается ранее не копировавшееся ребро e из G_t^k случайным образом. Пусть $e[src]$, $e[trg]$ and $e[sign]$ — начальная, конечная вершины и знак ребра e соответственно.
- 2) Добавляется новое знаковое ориентированное ребро
 - из v_{t+1} в $e[trg]$ с вероятностью $(1 - \alpha)$, копированием знака ребра e и начальной вершины $e[src]$,
 - или, из $e[src]$ в v_{t+1} с вероятностью α , копированием знака ребра e и конечной вершины $e[trg]$.
- 3) Описанная выше процедура повторяется k раз.

Согласно модели копирования ребер, в каждый момент времени в сеть добавляется новая вершина и копируется существующее ребро, чтобы соединиться с вершиной (исходной или целевой) выбранного ребра. В качестве обобщения авторы предлагают рассматривать ребро как k -клик, где $k = 2$, а вершина — это $(k - 1)$ -клика. Тогда общая форма модели копирования ребер представляется следующим образом. На вход подается знаковая ориентированная клика G_0^k размера $|V_t| = k + 1$. В момент времени $t + 1$ строится G_{t+1}^k добавлением новой вершины v_{t+1} в граф. Вершина v_{t+1} соединяется с $k - 1$ существующими вершинами и создается новая k -клика следующим образом:

- 1) Случайным образом выбирается k -клика из G_t^k .

- 2) Случайным образом выбирается вершина v из выбранной k -клик, что порождает $(k - 1)$ -клик, которой вершина v не принадлежит.
- 3) Вершина v_{t+1} соединается с вершинами из $(k - 1)$ -клик копированием знаковых ориентированных ребер, находящихся между вершиной v и вершинами $(k - 1)$ -клик.

В 2015 году Ciotti et al. в работе [5] предложили модель биномиального распределения степеней. Авторы предлагают модель генерации случайных сетей с N вершинами, которые удовлетворяют следующим требованиям.

- 1) Степени вершин однородно распределены.
- 2) Каждая вершина может быть членом одной из двух взаимоисключающих групп.
- 3) Не существует корреляций степеней до присвоения знаков ребрам.
- 4) Знаки ассоциируются с ребрами таким образом, что результирующая знаковая сеть является 2-сбалансированной.

Для получения таких сетей авторы предлагают следующий алгоритм:

- 1) Любая пара вершин соединена ребром с равномерной вероятностью p .
- 2) Если вершины делятся на две группы, то каждая вершина приписывается к одной из них с вероятностью m и к другой с вероятностью $1 - m$.
- 3) Ребра между вершинами внутри одной группы имеют положительный знак, а ребра между вершинами из разных групп — отрицательный.

Другими словами, на первом этапе модели генерируется беззнаковая сеть, в которой степени (беззнаковые) вершин имеют биномиальное распределение. На втором этапе определяется знак каждого ребра моделируемой сети путем деления вершин на две группы. Если вершины ребра находятся в одной группе, то это ребро положительное, в противном случае — отрицательное.

Другим способом генерации знаковых графов, предлагаемым в работе [5], является модель распределения степеней вершин по степенному закону. На первом этапе строится беззнаковая сеть, характеризующаяся степенным распределением. Затем к исходной связанной сети из n вершин применяется модель копирования. На каждом шаге в сеть добавляется новая вершина, а вторая вершина выбирается случайно: с вероятностью p новая вершина создаст связь с одним из соседей выбранной вершины, а с вероятностью $1 - p$ создаст связь с вершиной, выбранной случайно. После этого происходит процесс «переподключения», который моделируется следующим образом.

- 1) Два ребра выбираются случайным образом.
- 2) Четыре вершины, инцидентные этим ребрам, сортируются в порядке возрастания степени.
- 3) Если первые две вершины и последние две вершины не соединены, ребра переприсоединяются соответствующим образом.
- 4) В противном случае эти два ребра удаляются, и выбирается новая пара ребер.

В [5] доказано, что предложенные модели биномиального и степенного распределения степеней генерируют беззнаковые, неориентированные графы, характеризующиеся распределением степеней по степенному закону. Результатом работы моделей биномиального и степенного распределения степеней вершин графа является 2-сбалансированный знаковый граф, то есть граф, все циклы которого положительны [6].

2. Описание комплекса программ

Основным результатом работы является комплекс программ для работы со знаковыми графами на языке Python. Комплекс состоит из следующих четырех модулей.

- 1) Модуль построения знаковых графов на основе реальных данных. Методы и подходы, реализованные в этом модуле, описаны в [1].
- 2) Модуль генерации знаковых графов. В данном модуле реализованы все алгоритмы из таблицы 1.
- 3) Модуль корреляционной кластеризации знаковых графов [8, 2]. В данном модуле реализованы алгоритмы решения NP -трудной оптимизационной задачи корреляционной кластеризации для неориентированных и невзвешенных знаковых графов без кратных рёбер и петель, где функционал ошибки представляет собой линейную комбинацию межкластерной и внутрикластерной ошибок [8, 2].
- 4) Модуль сравнения и визуализации результатов работы алгоритмов.

Заключение

В результате исследования были изучены и реализованы различные модели генерации знаковых графов. Разработка модуля генерации знаковых графов с заданными свойствами позволяет существенно расширить библиотеку для тестирования эвристических алгоритмов решения задачи корреляционной кластеризации.

СПИСОК ЛИТЕРАТУРЫ

1. Ибрагимова Э. И., Семенова Д. В., Солдатенко А. А. Обзор методов получения знаковых графов из реальных данных // ИТ. НАУКА. КРЕАТИВ.

- Т. 5. Системы управления, информационные технологии и математическое моделирование: материалы I Междунар. форума (Омск, 14–16 мая 2024 г.): в 5 Ч. / науч. ред. П. С. Ложников, отв. ред. И. Г. Ольгина. Омск: Издательство ОмГТУ, 2024.
2. *Солдатенко А. А., Семенова Д. В., Ибрагимова Э. И.* Подход к анализу и построению алгоритмов решения одной задачи кластеризации на знаковых графах // Прикладная дискретная математика. 2024. № 64. С. 112–127.
 3. *Aniyan A., Naduvath S.* Induced signed graphs of some classes of graphs // Proceedings of the Jangjeon Mathematical Society. 2020. Vol. 23. No. 2. P. 283–291.
 4. *Barabasi A.-L., Albert R.* Emergence of scaling in random networks // Science. Vol. 286(5439). 1999. P. 509–512.
 5. *Ciotti V., Bianconi G., Capocci A., Colaiori F., Panzarasa P.* Degree correlations in signed social networks // Physica A: Statistical Mechanics and its Applications. 2015. P. 25–39
 6. *Harary F.* On the notion of balance of a signed graph // Michigan Mathematical Journal **2**. 1953. P. 143–146.
 7. *Kumar R., Raghavan P., Rajagopalan S., Sivakumar D., Tomkins A., Upfal E.* Stochastic models for the Web graph // Proceedings 41st Annual Symposium on Foundations of Computer Science. 2000. P. 57–65.
 8. *Soldatenko A., Semenova D., Ibragimova E.* On heuristic algorithm with greedy strategy for the Correlation Clustering problem solution // The 26th International Conference on Distributed Computer and Communication Networks: Control, Computation, Communications. Moscow. Lecture Notes in Computer Science. 2023. Vol. 14123. P. 462–477.
 9. *Sudheer N., George A. C., Aniyan A., Naduvath S.* Some new results on colour-induced signed graphs // Acta Univ. Sapientiae Informatica. 2022. Vol. 14. No. 2. P. 338–353.
 10. *Wahid D. F.* Random Models and Heuristic Algorithms for Correlation Clustering Problems on Signed Social Networks // A thesis submitted in partial fulfillment of the requirements for the degree of master of science in the college of graduate studies (Computer Science). The University of British Columbia(Okanagan). April 2017.
 11. *Wahid D. F., Hassini E.* A Literature Review on Correlation Clustering: Cross-disciplinary Taxonomy with Bibliometric Analysis // Oper. Res. Forum. 2022. No. 3. Vol. 47.
 12. *Warman B. M.* Routing of multipoint connections // In IEEE Journal on Selected Areas in Communications. 1988. Vol. 6(9). P. 1617–1622.

МОДЕЛИРОВАНИЕ ОТРАЖЕННОЙ СОЛНЕЧНОЙ РАДИАЦИИ С УЧЕТОМ ОТРАЖЕНИЯ ОТ ПОДСТИЛАЮЩЕЙ ПОВЕРХНОСТИ

И. Ю. Гендрина

*Национальный исследовательский Томский государственный университет,
г. Томск, Россия*

В работе описан алгоритм и представлены некоторые результаты моделирования поля отраженной солнечной радиации на верхней границе безоблачной атмосферы. Была проведена статистическая обработка данных численных экспериментов методом Монте-Карло для различных условий и построены регрессионные зависимости интенсивности отраженной радиации от различных параметров оптико-геометрической модели атмосферы. **Ключевые слова:** *метод Монте-Карло, статистическая обработка, солнечная радиация, подстилающая поверхность.*

Введение

Статистические методы исследования эффективно используются для обработки реальных данных и данных численных экспериментов для выявления закономерностей и построения прогнозов. Одним из самых распространенных методов является корреляционно-регрессионный анализ (КРА), который позволяет выявлять закономерности и формировать адекватные параметрические модели. В наших работах [2,3] КРА был использован для построения параметрической модели и исследования закономерностей поведения различных характеристик систем видения. В данной работе мы рассматриваем более сложную модель системы видения, учитывающую отражательные свойства подстилающей поверхности, что привело к необходимости разработки нового алгоритма. Представлены некоторые результаты численного моделирования солнечной радиации на верхней границе безоблачной атмосферы для различных законов отражения подстилающей поверхности. Проведена предварительная статистическая обработка с использованием КРА.

1. Оптико-геометрическая модель

Опишем схему экспериментов и оптические характеристики атмосферы. На рисунке 1 представлены геометрические параметры горизонтально однородной плоскопараллельной модели атмосферы. На верхнюю границу атмосферы падает плоскопараллельный единичный поток солнечного излучения в направлении, задаваемом зенитным и азимутальным углами $\omega_s = (\Theta_s, \varphi_s)$. Приемник, регистрирующий угловое распределение отраженной солнечной радиации, находится в точке с координатами $D(0, 0, H)$ и производит наблюдение в направлении $\omega_d = (\Theta_d, \varphi_d)$. Оптическая модель атмосферы была задана с помощью

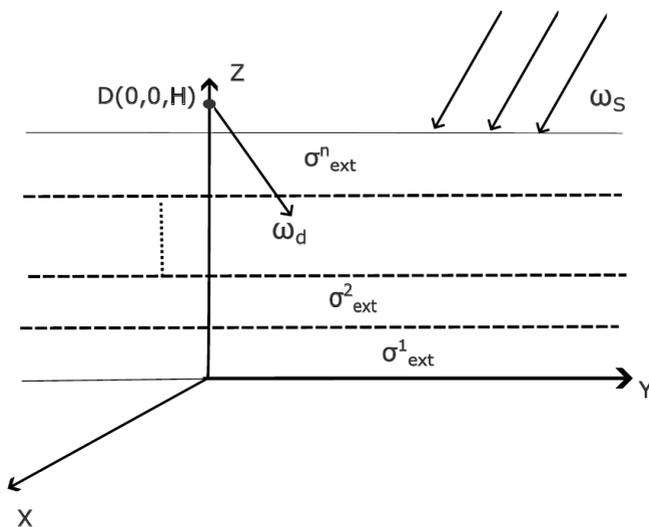


Рис. 1. Геометрическая схема наблюдения

следующих характеристик:

- 1) коэффициенты аэрозольного рассеяния $\sigma_{sc}(\mathbf{r})$ и ослабления $\sigma_{ext}(\mathbf{r})$; альbedo однократного рассеяния $\omega_0(\mathbf{r})$;
- 2) аэрозольная индикатриса рассеяния $g(\mathbf{r}, \omega, \omega')$;
- 3) вертикальные профили температуры и давления для расчета коэффициентов молекулярного рассеяния.

Подстилающая поверхность характеризуется законом отражения и альbedo подстилающей поверхности σ_{refl} .

2. Методы расчета

Теоретической базой для проведения численных экспериментов является стационарное уравнение переноса излучения [1]:

$$1(\boldsymbol{\omega}, \text{grad}I(\mathbf{r}, \boldsymbol{\omega})) = -\sigma_{ext}(\lambda, \mathbf{r}) + \sigma_{sc}(\lambda, \mathbf{r}) \int_{\Omega} I(\mathbf{r}, \boldsymbol{\omega}') g(\mathbf{r}, \boldsymbol{\omega}, \boldsymbol{\omega}') d\boldsymbol{\omega}' + \Phi_0(\mathbf{r}, \boldsymbol{\omega}). \quad (1)$$

Здесь: $X = (\mathbf{r}, \boldsymbol{\omega})$ — точка пространства $X = R \times \Omega$ координат $\mathbf{r} \in R$ и направлений $\boldsymbol{\omega} \in \Omega$. $\Phi_0(\mathbf{r}, \boldsymbol{\omega})$ — плотность распределения внутренних источников. $I(\mathbf{r}, \boldsymbol{\omega})$ — интенсивность (яркость) излучения в точке $(\mathbf{r}, \boldsymbol{\omega})$. Формально интенсивность отраженной солнечной радиации в точке $\mathbf{r}_d = (0, 0, H)$ в направлении $\boldsymbol{\omega}_d = (\Theta_d, \varphi_d)$ может быть представлена в виде:

$$I(\mathbf{r}_d, \boldsymbol{\omega}_d) = \int_R \int_{\Omega} I(\mathbf{r}, \boldsymbol{\omega}) \cdot \delta(\mathbf{r} - \mathbf{r}_d) \cdot \delta(\boldsymbol{\omega} - \boldsymbol{\omega}_d) d\mathbf{r} d\boldsymbol{\omega}. \quad (2)$$

Величина $I(\mathbf{r}, \boldsymbol{\omega})$ в (2) является решением уравнения (1) с функцией источников следующего вида:

$$\Phi_0(\mathbf{r}, \boldsymbol{\omega}) = \sigma_{ext}(\mathbf{r}) \cdot e^{-\tau(\mathbf{r})} \cdot \frac{g(\mathbf{r}, \boldsymbol{\omega}_s, \boldsymbol{\omega})}{2\pi}. \quad (3)$$

На основании (1)-(3) метод локальной оценки на сопряженных траекториях для расчета интенсивности в плоскопараллельной атмосфере [2] заключается в следующем: моделировать траекторию движения фотона из точки \mathbf{r}_d в направлении $\boldsymbol{\omega}_d$ и в каждой точке столкновения суммировать величины, характеризующие вероятность «попадания» излучения вследствие рассеяния из текущей точки столкновения $X_n = (\mathbf{r}, \boldsymbol{\omega})$ «на направление» $-\boldsymbol{\omega}_s$:

$$\psi(X_n) = e^{-\tau_s(\mathbf{r})} \cdot \frac{g(\mathbf{r}, -\boldsymbol{\omega}_s, \boldsymbol{\omega})}{2\pi}. \quad (4)$$

Здесь $\tau_s(\mathbf{r})$ - оптическая длина пути от текущей точки \mathbf{r} до верхней границы атмосферы в направлении $-\boldsymbol{\omega}_s$. Таким образом:

$$I(\mathbf{r}_d, \boldsymbol{\omega}_d) = M\xi = M \sum_{n=0}^N Q_n \cdot \frac{e^{-\tau_s(\mathbf{r})} \cdot g(\mathbf{r}, -\boldsymbol{\omega}_s, \boldsymbol{\omega})}{2\pi}. \quad (5)$$

Здесь N — количество столкновений в траектории фотона, Q_n — «вес фотона».

Для реализации отражения от подстилающей поверхности необходимо внести изменения в локальную оценку и в схему прямого моделирования на сопряженных траекториях.

Когда зафиксировано столкновение фотона с поверхностью Земли, плотность вероятности «попадания» в направлении на Солнце должна быть вычислена по формуле:

$$\psi(X_n) = e^{-\tau_s(r)} \cdot \frac{p(\mu_s)}{2\pi}. \quad (6)$$

Здесь $p(\mu)$ - плотность вероятности отражения подстилающей поверхностью падающего солнечного излучения.

Рассмотрены три варианта отражения подстилающей поверхностью:

- а) ламбертовское $p(\mu) = 2\mu$;
- б) изотропное $p(\mu) = 1$;
- в) зеркальное $p(\mu) = \delta(\mu)$.

В случае ламбертовской и изотропной поверхности направление отражения фотона не зависит от направления его падения на поверхность. Поэтому в данных случаях, учитывая то, что моделирование производится по сопряженным траекториям, μ_s - косинус острого угла между направлением попадания на Землю и осью OZ . В случае зеркального закона μ_s - косинус острого угла между направлением попадания на Солнце $-\omega_s$ и осью OZ .

Изменения в схеме моделирования заключаются в следующем.

- 1) Учитываем отражение от поверхности в соответствии с альбедо подстилающей поверхности σ_{refl} . Для этого разыгрываем случайную величину α , равномерно распределенную на $(0, 1)$. Если $\alpha > \sigma_{refl}$, то фиксируем обрыв текущей траектории и начинаем построение новой траектории. В противном случае моделируем продолжение траектории. Учет отражения может быть произведен также с помощью модификации «без отражения» по аналогии с приемом «без поглощения». В этом случае вместо розыгрыша отражения вес частицы умножаем на σ_{refl} .
- 2) Определяем новое направление движения фотона в соответствии с плотностью вероятности $p(\mu)$ с использованием метода обратной функции. Для этого разыгрываем случайную величину α , равномерно распределенную на $(0, 1)$. Косинус зенитного угла μ_α нового случайного направления вычисляем так:
 - а) ламбертовское $\mu_\alpha = \sqrt{\alpha}$;
 - б) изотропное $\mu_\alpha = \alpha$;
 - в) зеркальное $\mu_\alpha = \mu_s$, где μ_s - косинус острого угла между направле-

нием попадания на Солнце $-\omega_s$ и осью OZ .

При этом азимутальный угол нового направления для ламбертовского и изотропного законов разыгрывается равномерно в интервале $(0, 2\pi)$.

Для зеркального закона новое направление «разворачивается» на 180° по сравнению с направлением падения на поверхность Земли.

Для контроля качества результатов численных экспериментов с учетом (4)-(6) одновременно были определены величины:

$$D\xi = M\xi^2 - (M\xi)^2 \text{ и } \nu = \frac{\sqrt{D\xi}}{M\xi} \cdot 100.$$

Для формирования параметрической модели интенсивности отраженной солнечной радиации на основе метода наименьших квадратов были построены различного вида регрессии в среде RStudio.

3. Условия и результаты моделирования

Численные эксперименты были проведены для следующих условий наблюдения [2]:

Высота H , км: 100.

Длина волны, мкм: 0,44; 0,50; 0,67; 0,87; 1,06; 1,25; 2,00.

Зенитные углы Солнца, град: 0, 30, 60.

Азимутальные углы Солнца, град: 0, 45, 90, 135, 180.

Зенитные углы наблюдения, град: 0, 30, 45, 60.

Азимутальные углы наблюдения, град: 0, 90, 180.

Тип аэрозоля — континентальный, лето средних широт.

Выбор длин волн был обусловлен диапазонами окон прозрачности атмосферы, так что молекулярным поглощением в атмосфере можно было пренебречь. Длина волны (ДВ) характеризует такие оптические параметры как общая оптическая толщина (ОТ), средний косинус индикатрисы рассеяния (СК), альbedo однократного рассеяния (АОР). В таблице 1 приведены эти характеристики в зависимости от длины волны. Альbedo подстилающей поверхности: 0.80–0.95, 0.30–0.65, 0.1–0.15. Эти интервалы определяют свежесвыпавший снег, тающий снег и хвойный лес соответственно.[3]

В таблицах 2 и 3 представлены некоторые результаты моделирования. Из приведенных данных следует, в частности, что если Солнце находится над головой наблюдателя, то регистрируемая яркость на зависит от закона отражения для рассмотренных в данной работе. Если Солнце снижается над уровнем горизонта, то наблюдается увеличение яркости для ламбертовского закона отражения по сравнению с другими.

Таблица 1

Оптические параметры в зависимости от длины волны

ДВ, мкм	ООТ	СК	АОР
0,44	0,4740	0,6447	0,8996
0,50	0,3742	0,6406	0,8980
0,67	0,2736	0,6333	0,8829
0,87	0,2450	0,6339	0,8397
1,06	0,2366	0,6319	0,8043
1,25	0,2328	0,6379	0,7742
2,00	0,2303	0,7181	0,7831

Таблица 2

Яркость безоблачного неба, $\lambda = 0,5$ мкм, $\Theta_d = 0^\circ$, $\Theta_s = 0^\circ$

σ_{refl}	ламберт.	изотр.
0	0,00544	0,00544
0,35	0,01046	0,01034
0,106	0,00692	0,00686

Таблица 3

Яркость безоблачного неба, $\lambda = 0,5$ мкм, $\Theta_d = 0^\circ$, $\Theta_s = 30^\circ$

σ_{refl}	ламберт.	изотр.	зерк.
0	0,00411	0,00411	0,00411
0,35	0,05947	0,03658	0,00931
0,106	0,02064	0,01349	0,00565

Заклучение

Таким образом, в работе описан способ расчета яркости отраженной солнечной радиации с помощью статистического моделирования. Алгоритм метода Монте-Карло — локальная оценка на сопряженных траекториях — был модифицирован для случая учета отражения от подстилающей поверхности.

СПИСОК ЛИТЕРАТУРЫ

1. *Под ред. Марчука Г. И.* Метод Монте-Карло в атмосферной оптике. Новосибирск: Наука, 1976. 283 с.
2. *Irina Gendrina, Maria Alekseenko* The Monte Carlo Method for Determining the Vision System Characteristics // Applied Methods of Statistical Analysis. Nonparametric Methods in Cybernetics and System Analysis - AMSA'2017, 18-22 september 2017 Krasnoyarsk : proceedings of the international workshop. Novosibirsk: Novosibirsk State Technical University, 2017. P. 303–310.
3. *Гендрина И. Ю.* Моделирование поля отраженной солнечной радиации для безоблачной атмосферы // Новые информационные технологии в исследовании сложных структур. материалы Четырнадцатой международной конференции, Томск: Издательский Дом Томского государственного университета, 2022. С. 21–22.
4. *Павлов А. В.* Мониторинг криолитозоны Новосибирск: Акад. Изд-во «Гео», 2008. 224 с.

Гендрина Ирина Юрьевна — к.ф.-м.н., доцент, доцент кафедры прикладной математики ИПМКН ТГУ. E-mail: igendrina@bkl.ru

АНАЛИЗ КАРДИОКОМПЛЕКСОВ ГРУПП БОЛЬНЫХ С ВНЕЗАПНОЙ СЕРДЕЧНОЙ СМЕРТЬЮ И ЛЮДЕЙ С НЕЛЕТАЛЬНОЙ АРИТМИЕЙ ПО ЗАПИСЯМ ХОЛТЕРОВСКОГО МОНИТОРИРОВАНИЯ

Н. А. Лукьянова^{1,2}, А. И. Кульгаева¹, Н. В. Мишина³

¹ Красноярский государственный медицинский университет имени профессора В. Ф. Войно-Ясенецкого МЗ РФ, г. Красноярск, Россия

² Сибирский федеральный университет, г. Красноярск, Россия

³ Федеральный Сибирский научно-клинический центр Федерального медико-биологического агентства России, г. Красноярск, Россия

Одним из наиболее грозных и непредсказуемых исходов, вносящих значительный вклад в смертность от болезни системы кровообращения, является внезапная сердечная смерть, а ЭКГ продолжает оставаться основным методом диагностики патологии сердца в амбулаторно-поликлинической практике. Объектом исследования являются аннотированные электрокардиографические записи пациентов из открытого ресурса PhysioNet, а предметом исследования — морфологические и временные параметры кардиокомплексов. В качестве единицы измерения в работе выступили единичные кардиокомплексы. В качестве целевых для работы были отобраны синусовые, наджелудочковые и желудочковые этиологии кардиокомплексов. Классификация проводилась внутри группы патологии: классы разделялись по принципу тяжести исхода, а не в классическом понимании «здоровый-больной». Построено 3 классификационных модели, используя метод машинного обучения Random Forest. Доля верных классификаций среди всех моделей не опускалась ниже 97,8%. Лучшая классификация — синусовые комплексы, худшая — желудочковые. **Ключевые слова:** *внезапная сердечная смерть, холтеровское мониторирование, Random Forest.*

Введение

Внезапная сердечная смерть (ВСС) как один из исходов, составляющих внушительную долю случаев смерти населения разных возрастных, гендерных и расовых групп от кардиогенных причин, и практиче-

ски неотвратимое следствие внегоспитальной остановки сердца остается достаточно крупной и значимой проблемой мирового здравоохранения [1–4]. Поиск оптимальных решений по предупреждению и снижению частоты ВСС сохраняет свою актуальность и востребованность на протяжении всего времени существования медицинских служб.

Данная работа является продолжением исследований [5, 6]. Для выявления предсказательной значимости ЭКГ-записи в прогнозировании ВСС был проведен комплексный анализ записей мониторингования. Гипотезой для работы послужило предположение о том, что фиксируемые на ЭКГ-записи кардиокомплексы у людей с нефатальными аритмическими эпизодами и у людей со следом развившейся ВСС имеют электрофизиологические различия, способные послужить классификационным фактором для разделения этих двух разных по тяжести исхода групп. Исследовательский интерес заключался в изучении возможности классификации внутри группы патологии: классы разделялись по принципу тяжести исхода, а не в классическом понимании «здоровый-больной».

1. Материалы и методы исследования

В основу работы легли базы данных холтеровского мониторингования (ХМ) открытого репозитория PhysioNet:

1. Sudden Cardiac Death Holter Database [7] — коллекция продолжительных, от 4 до 25 часов, записей ЭКГ тридцати трех пациентов, зафиксировавших момент наступления внезапной сердечной смерти.

2. MIT-BIH Arrhythmia Database [8] — набор получасовых выдержек из 47 записей ХМ-ЭКГ, содержащих эпизоды нелетальных аритмий различного генеза.

В качестве единицы измерения в работе выступили единичные кардиокомплексы. Рабочие наборы данных по каждой из баз содержат извлеченные программным способом [9] из сигнала ЭКГ в двух отведениях холтера (II и V5) амплитудные и временные характеристики сердечных комплексов, а также два качественных идентификатора номера записи (record) и типа сердечного сокращения (type). В качестве параметра «морфологическая характеристика комплекса QRS» (QRS morph 0–4) выступало измеренное в мВ отклонение от изолинии пяти равноудаленных друг от друга точек, расставленных на линии самописца от начала до конца желудочкового комплекса QRS.

Список взятых для проведения классификации признаков представлен в таблице 1. Каждая из двух рабочих баз данных содержит 5 типов сердечных сокращений: нормальные, или синусовые (N), наджелудочковые (SVEB), желудочковые (VEB), выскальзывающие (F) и идиопати-

ческие (Q). В качестве целевых для работы были отобраны синусовые, наджелудочковые и желудочковые этиологии кардиокомплексов. В базе по ВСС для данных типов комплексов были введены обозначения классов N1, VEB1 и SVEB1, в базе с записями аритмий — N2, VEB2 и SVEB2.

Таблица 1

Перечень исследуемых признаков

Тип признаков	Группа признаков
Интервалы, мс	RR Interval, PQ Interval, QT Interval, ST Interval, QRS Interval
Амплитуды зубцов, мВ комплекса	P peak, T peak, R peak, S peak, Q peak
Морфологические признаки комплекса QRS, мВ	QRS morph feature 0, QRS morph feature 1, QRS morph feature 2, QRS morph feature 3, QRS morph feature 4

Обработка и анализ данных были реализованы при помощи языка программирования Python 3.10.12 в среде Google Colaboratory.

2. Результаты и обсуждения

Набор данных ЭКГ-комплексов по всем записям базы [7] состоит из 502 808 единичных кардиоэпизодов. В первую очередь, из базы [7] для исследования были отобраны записи ЭКГ, зафиксированные у лиц с преимущественно синусовым ритмом и ритмом фибрилляции предсердий, что составило в общем 247991 ЭКГ-комплексов. В базе записей ЭКГ по аритмиям [8] общее количество единичных кардиокомплексов составило 100 689. Далее, в базах данных по ВСС и по аритмиям поочередно выделяли группы сердечных комплексов различной этиологии: синусовой, наджелудочковой и желудочковой. После чего на базе алгоритма Random Forest строились классификационные модели.

Для различных по объему наборов данных проводилась балансировка классов: для уменьшения числа измерений применялся метод случайного отбора (sample), для увеличения — метод повторной выборки (resample). Обучение моделей осуществлялось в нескольких итерациях для возможности оценки средних показателей метрик качества классификации — accuracy, precision и recall. Ход решения проблемы дисбаланса классов представлен в таблице 2.

Таблица 2

Решение проблемы дисбаланса классов для Random Forest моделей

Модели по типам комплексов	Исходный объем комплексов	Метод балансировки	Объем после балансировки
RF Модель (N1-N2)	N1=140 963 N2=90 083	sample –	N1=90 083 N2=90 083
RF Модель (SVEB1-SVEB2)	SVEB1=1 542 SVEB2=2 779	resample sample	SVEB1=1 800 SVEB2=1 800
RF Модель (VEB1-VEB2)	VEB1=8 553 VEB2=7 009	sample –	VEB1=7 009 VEB2=7 009

Рассматривались модели по типам комплексов, в каждой из которых при помощи встроенной функции весов *feature_importances_* алгоритма RandomForestClassifier оценивался вклад в классификацию каждого из используемых признаков ЭКГ-комплексов. Поскольку практически во всех моделях, кроме модели Random Forest (VEB1-VEB2), порядка 50% природы алгоритма приходилось на первую тройку параметров, было решено выделить по три признака, наиболее значимых с точки зрения классификации кардиокомплексов по летальности исхода (табл. 3).

Таблица 3

Сводная таблица наиболее значимых признаков моделей классификации

Модели по типам комплексов	Наименование значимых признаков моделей (отведение холтера)	Вклад в построение модели
RF Модель (N1-N2)	R peak (II) pre-RR (II) S peak (II)	0,3063 0,1209 0,0630
RF Модель (SVEB1-SVEB2)	R peak (II) ST interval (V5) T peak (II)	0,3936 0,1392 0,0731
RF Модель (VEB1-VEB2)	pre-RR (II) S peak (II) T peak (II)	0,1810 0,1326 0,0693

На основе наборов данных по ВСС и аритмиям было построено 3 бинарных классификационных моделей по алгоритму Random Forest в пяти итерациях каждая. В результате построения моделей было выделено 6 наиболее значимых с точки зрения классификации электрофизиологических параметров кардиокомплексов: pre-RR, T peak, ST interval, R peak, S peak, QT interval. Выявленные параметры имеют свою практическую нишу в диагностике и уже активно используются в рутинной врачебной практике для оценки состояния сердца и выявления разнообразных патологий.

В таблице 4 представлены усредненные оценки качества моделей классификации, полученные в результате пятикратной кросс-валидации.

Таблица 4

Средние метрики качества Random Forest моделей классификации ЭКГ-комплексов на тестовой выборке

типы комплексов	Accuracy (%)	Precision (%)	Recall (%)
N1-N2	99,74	99,79	99,67
SVEB1-SVEB2	99,69	99,82	99,56
VEB1-VEB2	97,80	98,81	96,78

Таким образом, доля правильных классификаций (accuracy) по всем моделям не опускалась ниже отметки в 97,8%. Лучше всего поддавались классификации комплексы синусовой этиологии, хуже — желудочковой. Во всех трех моделях показатель точности (precision) оказался выше показателя полноты (recall), что свидетельствует о том, что обученные классификационные модели более склонны к гипердиагностике, нежели к невыявлению риска развития ВСС.

Заключение

В ходе работы была подтверждена исследовательская гипотеза о том, что между электрофизиологическими параметрами кардиокомплексов людей с исходом в виде остановки сердца и людей с нелетальными аритмическими эпизодами существуют различия. Эти различия доступны для фиксации на ЭКГ-записи и впоследствии применимы как классификационный фактор. Такой результат проведенной работы свидетельствует о необходимости продолжения исследований на тему применения ЭКГ-маркеров как предикторов и индикаторов развития сердечной остановки.

СПИСОК ЛИТЕРАТУРЫ

1. *Теплов В. М. [и др.]*. Центр по лечению внезапной сердечной смерти: результаты пяти лет работы // Вестник анестезиологии и реаниматологии. 2023. Т. 20. № 1. С. 17–23. DOI: 10.24884/2078-5658-2023-20-1-17-23
2. Национальные российские рекомендации по применению методики холтеровского мониторирования в клинической практике // Российский кардиологический журнал. 2014. № 2(106). С. 6–71.
3. *Фролов А. В.* Информационное обеспечение первичной профилактики внезапной сердечной смерти // Кардиология в Беларуси. 2022. Т. 14. № 7(3). С. 145–155. DOI: 10.34883/PI.2022.14.2.001
4. *Никифоров В. С., Метсо К. В.* Электрокардиографические предикторы внезапной сердечной смерти // Consilium Medicum. 2018. Т. 5. № 20. С. 29–33. DOI: 10.26442/2075-1753_2018.5.29-33
5. *Gusarova A. A. [et al.]*. Analysis of normal and pathological heart rate variability based on electrocardiogram data // 2022 IEEE 16th International Conference on Application of Information and Communication Technologies (AICT). Washington DC, DC, USA, IEEE, 2023. P. 1–4. DOI: 10.1109/AICT55583.2022.10013602
6. *Lukyanova N. A. [et al.]*. Classification of normal and pathological heart rate variability based on daily ECG monitoring using machine learning methods // 2023 IEEE 17th International Conference on Application of Information and Communication Technologies (AICT). Baku, Azerbaijan, IEEE, 2023. P. 1–5. DOI: 10.1109/AICT59525.2023.10313172
7. <https://physionet.org/content/sddb/1.0.0/> — Sudden Cardiac Death Holter Database // PhysioNet.
8. <https://physionet.org/content/mitdb/1.0.0/> — MIT-BIH Arrhythmia Database // PhysioNet.
9. *Sakib S., Fouda M. M., Fadlullah Z. M.* Harnessing Artificial Intelligence for Secure ECG Analytics at the Edge for Cardiac Arrhythmia Classification // Secure Edge Computing. 2021. С. 137–153. DOI: 10.1201/9781003028635-11

Лукьянова Наталья Александровна — к.ф.-м.н., доцент кафедры медицинской кибернетики и информатики, КрасГМУ, доцент кафедры высшей математики 2, СФУ. E-mail: nalukyanovakrsk@gmail.com

Мишина Наталья Владимировна — врач высшей категории отделения функциональной диагностики ФГБУ ФСНКЦ ФМБА России. E-mail: mishinatalya@gmail.com

Кульгаева Анастасия Игоревна — студент 6 курса специальности 30.05.03 Медицинская кибернетика, КрасГМУ. E-mail: tunyakul@mail.ru

МОДЕЛИРОВАНИЕ РАЗВИТИЯ ПРОФЕССИОНАЛЬНЫХ ЗАБОЛЕВАНИЙ ОТ ФИЗИЧЕСКИХ ФАКТОРОВ

М. П. Дьякович

Ангарский государственный технический университет, г. Ангарск, Россия

В статье рассматриваются результаты моделирования развития профессиональных заболеваний от воздействия шума и вибрации с использованием различных методов системного анализа.

Ключевые слова: профессиональные заболевания, физические факторы, модели.

Введение

В РФ за период с 2014 г. по 2023 г. отмечено снижение удельного веса рабочих мест, не соответствующих гигиеническим нормативам по таким физическим факторам как шум и вибрация соответственно в 1,3 и в 1,9 раз. Снизилось на 47,7% и количество впервые установленных профессиональных заболеваний (ПЗ). При этом, темп прироста к 2014 г. заболеваний и их последствий, связанных с воздействием производственных физических факторов (ПФФ) составил 2,8%. В группе ПЗ, связанных с воздействием ПФФ, вибрационная болезнь (ВБ) занимала второе место (45,3%) после нейросенсорной тугоухости (НСТ) (53,8%) [1]. В развитых странах частота случаев ВБ и НСТ также намного выше, чем других ПЗ [2,3], несмотря на меры по снижению уровней вредных факторов, активному применению средств индивидуальной защиты. Назрела необходимость разработки моделей развития ПЗ от ПФФ, позволяющих оценить их риск и динамику патогенеза, что даст возможность предотвращать инвалидизацию работников за счет ранней диагностики, своевременной профилактики и адекватного лечения.

1. Методы и средства

Для моделирования использовался созданный на основе многолетних наблюдений в ФГБНУ ВСИМЭИ информационный массив персонафицированных клинко-гигиенических показателей лиц, работающих в

Автор благодарит сотрудников лаборатории эколого-гигиенических исследований (руководитель-д.м.н. Панков В.А.) и клиники (главный врач – д.м.н., профессор Катаманова Е.В.) ФГБНУ ВСИМЭИ

контакте с вибрацией и шумом. Среди них были и практически здоровые на момент обследования работники, и те, кому был выставлен диагноз ПЗ от ПФФ. Исследовались регрессионные, информационно-энтропийные модели, модели с использованием метода нормированных интенсивных показателей НИП) с расчетом относительного риска (ОР), модели с использованием машинного обучения. Программное обеспечение синтеза структуры нейронных сетей, глубокого обучения, проверки валидности и адекватности модели, вычислительных экспериментов с моделью было разработано к.т.н. М.В. Кривовым.

2. Результаты и их обсуждение

Построение обобщенных моделей, отображающих взаимосвязи параметров функционирования организма работающих в контакте с вибрацией, которая может обусловить их неблагоприятное изменение и развитие патологии на определенном отрезке времени мы начали с регрессионных моделей. Для модельных исследований были выбраны семейства регрессионных уравнений вида:

$$f(x) = Y = b_{0i} + \frac{b_i}{x} \pm \varepsilon,$$

где x – индивидуальная стажевая доза (ИСД) вибрации, включающая в себя уровень и время воздействия в течение смены и за весь трудовой стаж;

ε – ошибка предсказания;

b_{0i} – свободный член уравнения;

b_i – коэффициент регрессии Y по X ;

$i = \overline{1, 40}$ – номер зависимой переменной из 40 рассматриваемых в процессе моделирования.

Рассчитывали ИСД с учетом эквивалентного уровня виброскорости за смену и стажа работы с виброинструментом.

$$\text{ИСД} = L_{\text{экв}} + 10 \times \lg\left(\frac{T}{T_0}\right),$$

где ИСД – стажевая доза вибрации, дБ;

$L_{\text{экв}}$ – эквивалентный уровень виброскорости за рабочую смену, дБ;

T – стаж работы с виброинструментом, лет;

T_0 – год [4].

Было показано, что специфические для ВВ показатели функционирования сердечно-сосудистой системы, болевой, вибрационной

и тактильной чувствительности, а также неспецифические показатели функционирования нейрогуморальной и иммунной систем подчиняются дозо-эффектной зависимости. Нарушения функционирования гипоталамо-гипофизарно-надпочечниковой системы отмечались и у больных ВБ, и у практически здоровых работников, что можно расценивать как ответ на воздействие вибрации. При этом с ростом ИСД у больных ВБ эти изменения становились более выраженным, чем у практически здоровых лиц. Регрессионные модельные исследования дали представление об очередности изменений в органах и системах с ростом ИСД, что не было очевидным в ходе традиционного анализа усредненных показателей. У работников с установленным диагнозом ВБ модельный прогноз развития ПЗ был адекватен реальной ситуации, в которой люди, имеющие ПЗ от ПФФ, продолжали работать, увеличивая ИСД, провоцируя прогрессирование имеющейся патологии вплоть до инвалидизации.

Далее величина вклада ИСД в развитие ВБ определялась разностью величин безусловной и условной энтропий. С ростом ИСД, наибольший вклад в формирование ВБ вносили патологически измененные параметры основной активности головного мозга и тиреоидной системы (82,6%). Результаты этих модельных экспериментов показали необходимость одновременного учета неспецифических показателей общего и кофакторного патогенеза и специфических для рассматриваемой патологии клинические и функциональные показатели с целью повышения адекватности, точности и чувствительности модели развития ПЗ от ПФФ.

С помощью модели НИП было установлено, что риск ПЗ достаточно сильно связан с профессиональной принадлежностью ($OR=61,8$) и с наличием сочетанного воздействия шума и вибрации на рабочем месте ($OR=23,8$), уровни которых часто не соответствуют гигиеническим нормативам и превышают предельно допустимые уровни из-за несовершенства технологии и/или оборудования. Степень влияния на возникновение ПЗ от ПФФ стажа работы в контакте с конкретным неблагоприятным фактором, возраста работника, отрасли промышленности оказались значительно ниже.

В ходе модельных экспериментов было установлено, что риски ПЗ у лиц до 40 лет, работающих в профессиях машиниста экскаватора и сборщика-клепальщика, имеющих контакт с шумом и вибрацией 1–4 года, составили 0,269 и 0,226, что можно отнести к неопределённому прогнозу. С ростом стажа и возраста у работников этих профессий величины риска ПЗ от ПФФ увеличиваются на 24,4 и 29,1% соответственно. Максимум риска достигается при стаже 30 и более, при

возрасте 50–59 лет (0,334 и 0,292, что было отнесено к неблагоприятному прогнозу). То, что в модель НИП не были включены такие факторы-предикторы как клиничко-функциональные, биохимические и социально-психологические показатели работников, явилось ее ограничением.

Далее нами использовалась многоклассовая кластеризация по степеням тяжести ПЗ от ПФФ на основе рекуррентных искусственных нейронных сетей с глубоким обучением. Эмпирическое сравнение качества обучения и полученной точности прогнозирования обусловило выбор архитектуры сетей. Обучение модели развития ВБ осуществлялось на наборе данных из 200 клинических наблюдений при упаковке обучающего набора из 12 наблюдений. Достигнутая точность прогноза отнесения работника к контингенту практически здоровых и больных ВБ была высока и составила 95,7%. В таблице 1 представлена конфигурация классификатора искусственной нейронной сети.

В нейросетевых моделях развития ПЗ от ПФФ нами использовались не только специфические, но и неспецифические показатели общего и кофакторного патогенеза. Показатели были сгруппированы в нейропсихологические, нейрофизиологические и социально-психологические комплексы, которые характеризовали состояние центральной нервной, периферической нервной, нейросенсорной систем и гормональный статус работника.

Таблица 1

Конфигурация классификатора искусственной нейронной сети

Параметры	Слои			
	Первый	Второй	Третий	Четвёртый
Количество нейронов	10	500	100	2
Функция активации	-	ReLU	ReLU	Softmax
Число параметров	-	5500	55802	202

Так как в моделях развития НСТ использовался комплекс качественных экспертных оценок, выполненных профпатологами, неврологами, нейрофизиологами, клиническими психологами, входной слой нейронной сети формировался на основе нечеткой нейронной продукционной сети Ванга-Менделя [5]. Структура синтезируемой для классификации искусственной нейронной сети состояла из 5-6 слоев и содержала слой прореживания данных. Количество тренируемых пара-

метров варьировало от 6600 до 7800. Для отдельных комплексов параметров функционирования нервной системы точность прогноза полученных моделей, обученных на реальных данных, составила 70–80%. Такой результат можно считать удовлетворительным.

Практическая реализация ядра моделирования была выполнена на базе фреймворка ML (Machine Learning) компании Microsoft.

3. Заключение

Регрессионные модельные исследования позволили определить очеречность неблагоприятных изменений в органах и системах при развитии ПЗ от ПФФ, доказали наличие дезадаптационных процессов в организме работников, контактирующих с вибрацией, как у лиц, имеющих уже диагноз ПЗ на момент обследования, так и у практически здоровых лиц. Были выявлены наиболее чувствительные входные переменные, которые могут быть получены в ходе скрининга состояния здоровья лиц, работающих в контакте с ПФФ. Апробация энтропийной модели показала возможность анализа параметров функциональных систем организма, участвующих в формировании ВБ в различных комбинациях, что позволит в дальнейшем использовать ее для объективизации прогноза развития ВБ и рационализации диагностического процесса.

Использование НИП позволила дать интегрированную оценку риска ПЗ как по отдельным факторам, так и по их комплексу, определить группы риска. Модель, основанная на методе НИП, позволила определить группы риска, прогнозировать вероятность возникновения ПЗ при определённых сочетаниях факторов, выделить наиболее значимые факторы для дальнейшего включения их в нейросетевые модели.

Точность прогноза полученных нейросетевых моделей НСТ меньше, чем нейросетевых моделей ВБ. Это может быть связано с недостаточностью данных для обучения, так как некоторые параметры исследования у конкретных пациентов отсутствовали по объективным причинам. В обучающей и контрольной выборках наилучшим образом были представлены параметры функционирования центральной и периферической нервной систем.

Успешность использования моделирования для оценки развития ПЗ от ПФФ определило дальнейшую направленность работ. Необходимо расширить спектр параметров, характеризующих различные функциональные системы организма и социально-психологического состояние работника для участия в моделировании, что позволит дополнительно обучить нейросети для прогнозов с более высокой точностью. Использование математического моделирования и интеллектуального анализа

данных позволит в будущем разработать систему поддержки принятия решения врача в сложной клинической задаче дифференциальной диагностики ПЗ от ПФФ, основанной на принципах персонализированной медицины. Обученные модели можно также использовать в качестве компонентов экспертной советующей системы при оценке патогенетически обоснованных методов реабилитации лиц, пострадавших от воздействия ПФФ, что также будет являться нашей дальнейшей задачей.

СПИСОК ЛИТЕРАТУРЫ

1. О состоянии санитарно-эпидемиологического благополучия населения в Российской Федерации в 2023 году: Государственный доклад. Город: Москва: Федеральная служба по надзору в сфере защиты прав потребителей и благополучия человека, 2024. 364 с.
2. *Heaver C., Goonetilleke K., Ferguson H., Shiralkar S.* Hand-arm vibration syndrome: a common occupational hazard in industrialized countries // *J. Hand SurgEur.* 2011. Vol. 3. № 5. P. 354–363.
3. *Chen K. H., Su S. B., Chen K. T.* An overview of occupational noise-induced hearing loss among workers: epidemiology, pathogenesis, and preventive measures // *Environ Health Prev Med.* 2020. Vol. 31. № 25(1). P. 65.
4. *Денисов Э. И., Овакимов В. Г.* Гигиеническая оценка производственных шумов и вибраций по экспозиции и дозе // *Гигиена труда и профессиональные заболевания.* 1985. № 2. С. 19–22.
5. *Wang L. X., Mendel J. M.* Generating Fuzzy Rules by Learning from Examples // *IEEE Transactions on Systems, Man, and Cybernetics.* 1992. Vol. 22. № 6. P. 1414–1427.

Дьякович Марина Пинхасовна — д.б.н., профессор кафедры экономики, маркетинга и психологии управления, ведущий научный сотрудник лаборатории эколого-гигиенических проблем. E-mail: *marik914@rambler.ru*

РАЗНОСТНАЯ СХЕМА ПОВЫШЕННОГО ПОРЯДКА ТОЧНОСТИ ДЛЯ МОДИФИЦИРОВАННОГО УРАВНЕНИЯ ВЛАГОПЕРЕНОСА ДРОБНОГО ПОРЯДКА

М. Х. Бештоков

*Институт прикладной математики и автоматизации КБНЦ РАН,
г. Нальчик, Россия*

Исследована первая краевая задача для модифицированного уравнения влагопереноса с двумя операторами дробного дифференцирования Герасимова-Капуто разных порядков α, β . Построена разностная схема повышенного порядка точности на равномерной сетке. Методом энергетических неравенств для решения разностной задачи получены априорные оценки при различных значениях α, β . Из полученных оценок следуют единственность и устойчивость решения по правой части и начальным данным, а также сходимости решения разностной задачи к решению исходной дифференциальной задачи со скоростью равной порядку аппроксимации. **Ключевые слова:** *первая краевая задача, априорная оценка, модифицированное уравнение влагопереноса, дифференциальное уравнение дробного порядка, дробная производная Герасимова-Капуто.*

Введение

Важной задачей вычислительной математики является построение и исследование разностных схем высокого порядка точности, аппроксимирующих уравнения математической физики. В последнее время для построения указанных разностных схем используются компактно-разностные схемы, т. е. схемы повышенных порядков аппроксимации и/или точности, записывающиеся на стандартных для данного уравнения шаблонах.

Построению и исследованию схем повышенного порядка точности посвящены работы авторов [1]-[3].

Исследованию различных краевых задач для модифицированного уравнения влагопереноса посвящены работы автора [4]-[5].

1. Постановка задачи

В замкнутом прямоугольнике $\bar{Q}_T = \{(x, t) : 0 \leq x \leq l, 0 \leq t \leq T\}$ рассмотрим задачу

$$\partial_{0t}^\alpha u = k(t) \frac{\partial^2 u}{\partial x^2} + A \partial_{0t}^\beta \frac{\partial^2 u}{\partial x^2} - q(t)u(x, t) + f(x, t), \quad 0 < x < l, \quad 0 < t \leq T, \quad (1)$$

$$u(0, t) = u(l, t) = 0, \quad 0 \leq t \leq T, \quad (2)$$

$$u(x, 0) = u_0(x), \quad 0 \leq x \leq l, \quad (3)$$

где

$$0 < c_0 \leq k(t) \leq c_1, \quad A = \text{const} > 0, \quad |q(t)| \leq c_2, \\ u(x, t) \in C^{6,3}(\bar{Q}_T), \quad k(t), q(t) \in C^1[0, T], \quad f(x, t) \in C^{6,1}(\bar{Q}_T), \quad (4)$$

$\partial_{0t}^\gamma u = \frac{1}{\Gamma(1-\gamma)} \int_0^t \frac{u_\tau(x, \tau)}{(t-\tau)^\gamma} d\tau$, – дробная производная в смысле Капуто порядка γ , $0 < \alpha < 1$, $c_i = \text{const} > 0, i = 0, 1, 2$, $\partial_{0t}^\gamma u = D_{0t}^\gamma u - \frac{u(0)}{\Gamma(1-\alpha)t^\alpha}$, $D_{0t}^\gamma u = \frac{1}{\Gamma(1-\gamma)} \frac{d}{dt} \int_0^t \frac{ud\tau}{(t-\tau)^\gamma}$ – дробная производная в смысле Римана-Лиувилля порядка γ .

По ходу изложения будем также использовать положительные постоянные числа M_i , $i = 1, 2, \dots$, зависящие только от входных данных исходной задачи.

2. Устойчивость и сходимость разностной схемы

Для решения задачи (1)-(3) применим метод конечных разностей. Для этого на равномерной сетке $\bar{\omega}_{h\tau} = \bar{\omega}_h \times \bar{\omega}_\tau$, где $\bar{\omega}_h = \{x_i = ih, i = 0, \bar{N}\}$, $h = l/N\}$, $\bar{\omega}_\tau = \{t_j = j\tau, j = 0, 1, \dots, j_0, \tau = T/j_0\}$ дифференциальной задаче (1)-(3) поставим в соответствие разностную схему порядка аппроксимации $O(h^4 + \tau^2)$ при $\alpha = \beta$ и $O(h^4 + \tau^{2-\max\{\alpha, \beta\}})$ при $\alpha \neq \beta$ [4], [7], [8]:

$$\Delta_{0t_{j+\sigma}}^\alpha \mathcal{H}_h y = a y_{\bar{x}x}^{(\sigma)} + A \Delta_{0t_{j+\sigma}}^\beta y_{\bar{x}x} - d \mathcal{H}_h y^{(\sigma)} + \mathcal{H}_h \varphi, \quad (5)$$

$$y_0^{(\sigma)} = y_N^{(\sigma)} = 0, \quad t \in \bar{\omega}_\tau, \quad (6)$$

$$y(x_i, 0) = u_0(x_i), \quad x \in \bar{\omega}_h, \quad (7)$$

где $\Delta_{0t_{j+\sigma}}^\gamma y = \frac{\tau^{1-\gamma}}{\Gamma(2-\gamma)} \sum_{s=0}^j c_{j-s}^{(\gamma, \sigma)} y_t^s$ – дискретный аналог дробной производной Герасимова-Капуто порядка γ , $0 < \alpha < 1$, обеспечивающий порядок точности $O(\tau^{3-\gamma})$, $a^j = k(t_{j+\sigma})$, $d^j = q(t_{j+\sigma})$, $\varphi_i^j = f(x_i, t_{j+\sigma})$.

$$\mathcal{H}_h y_i^j = \frac{1}{12} \left(y_{i+1}^j + 10y_i^j + y_{i-1}^j \right) = y_i^j + \frac{h^2}{12} y_{\bar{x}x, i}^j, \quad i = 1, \dots, N-1,$$

$$a_0^{(\gamma, \sigma)} = \sigma^{1-\gamma}, \quad a_l^{(\gamma, \sigma)} = (l + \sigma)^{1-\gamma} - (l - 1 + \sigma)^{1-\gamma}, \quad l \geq 1,$$

$$b_l^{(\gamma, \sigma)} = \frac{1}{2-\gamma} \left[(l+\sigma)^{2-\gamma} - (l-1+\sigma)^{2-\gamma} \right] - \frac{1}{2} \left[(l+\sigma)^{1-\gamma} + (l-1+\sigma)^{1-\gamma} \right], \quad l \geq 1,$$

при $j = 0$, $c_0^{(\gamma, \sigma)} = a_0^{(\gamma, \sigma)}$;

при $j > 0$, $c_s^{(\gamma, \sigma)} = \begin{cases} a_0^{(\gamma, \sigma)} + b_1^{(\gamma, \sigma)}, & s = 0, \\ a_s^{(\gamma, \sigma)} + b_{s+1}^{(\gamma, \sigma)} - b_s^{(\gamma, \sigma)}, & 1 \leq s \leq j-1, \\ a_j^{(\gamma, \sigma)} - b_j^{(\gamma, \sigma)}, & s = j, \end{cases}$

$$c_s^{(\gamma, \sigma)} > \frac{1-\gamma}{2} (s + \sigma)^{-\gamma} > 0, \quad \sigma = 1 - \frac{\gamma}{2}, \quad y^{(\sigma)} = \sigma y^{j+1} + (1 - \sigma) y^j.$$

Теорема 1. Пусть выполнены условия ограниченности и гладкости (4), а также выполнены условия сопряжения граничных и начальных условий (2), (3) с уравнением (1), тогда существует такое τ_0 , что если $\tau \leq \tau_0$, то для решения разностной задачи (5)-(7) справедливы априорные оценки:

1. В случае, когда $\alpha > \beta$:

$$\|y^{j+1}\|_0^2 \leq M_1 \left(\|y^0\|_0^2 + \max_{0 \leq j' \leq j} \|\varphi\|_0^2 \right). \quad (8)$$

2. В случае, когда $\alpha = \beta$:

$$\|y^{j+1}\|_{W_2^1(0,l)}^2 \leq M_2 \left(\|y^0\|_{W_2^1(0,l)}^2 + \max_{0 \leq j' \leq j} \|\varphi\|_0^2 \right), \quad (9)$$

где $\|y^{j+1}\|_{W_2^1(0,l)}^2 = \|y^{j+1}\|_0^2 + \|y_{\bar{x}}^{j+1}\|_0^2$.

3. В случае, когда $\alpha < \beta$:

$$\|y_{\bar{x}}^{j+1}\|_0^2 \leq M_3 \left(\|y_{\bar{x}}^0\|_0^2 + \max_{0 \leq j' \leq j} \|\varphi\|_0^2 \right), \quad (10)$$

где $M_1, M_2, M_3 = \text{const} > 0$, не зависящие от h и τ .

Доказательство. Априорные оценки найдем методом энергетических неравенств. Для этого введем скалярные произведения и норму:

$$(u, v) = \sum_{i=1}^{N-1} u_i v_i h, \quad (u, u) = (1, u^2) = \|u(\cdot, t)\|_0^2 = \|u\|_0^2,$$

$$(u, v) = \sum_{i=1}^N u_i v_i h, \quad (1, u^2) = \|u\|_0^2.$$

Получим некоторые вспомогательные неравенства с учетом леммы 3 [8, с. 109]:

$$\begin{aligned} (y, \mathcal{H}_h y) &= \left(y, y + \frac{h^2}{12} y_{\bar{x}\bar{x}} \right) = (y, y) - \left(1, \frac{h^2}{12} (y_{\bar{x}})^2 \right) \geq \left(1, y^2 - \frac{1}{3} y^2 \right) = \\ &= \left(1, \frac{2}{3} y^2 \right) = \frac{2}{3} \|y\|_0^2. \end{aligned} \quad (11)$$

$$\begin{aligned} \left(1, \left(y + \frac{h^2}{12} y_{\bar{x}\bar{x}} \right)^2 \right) &= \left(1, y^2 + \frac{2h^2}{12} y y_{\bar{x}\bar{x}} + \frac{h^4}{144} y_{\bar{x}\bar{x}}^2 \right) = \\ &= \left(1, y^2 - \frac{2h^2}{12} (y_{\bar{x}})^2 + \frac{h^4}{144} y_{\bar{x}\bar{x}}^2 \right). \end{aligned}$$

Из этого следует

$$(1, y^2) - \left(1, \frac{2h^2}{12} (y_{\bar{x}})^2 \right) \leq \left(1, \left(y + \frac{h^2}{12} y_{\bar{x}\bar{x}} \right)^2 \right) \leq (1, y^2) + \left(1, \frac{h^4}{144} y_{\bar{x}\bar{x}}^2 \right).$$

Тогда из последнего получим неравенство

$$(1, y^2) - \frac{2}{3} (1, y^2) \leq \left(1, \left(y + \frac{h^2}{12} y_{\bar{x}\bar{x}} \right)^2 \right) \leq (1, y^2) + \frac{1}{9} (1, y^2).$$

Итак,

$$\frac{1}{3} \|y\|_0^2 \leq \|\mathcal{H}_h y\|_0^2 \leq \frac{10}{9} \|y\|_0^2. \quad (12)$$

Найдём теперь априорную оценку, для этого воспользуемся методом энергетических неравенств. Умножим тогда уравнение (5) скалярно на $\mathcal{H}_h y^{(\sigma)}$:

$$\begin{aligned} \left(\Delta_{0t_{j+\sigma}}^\alpha \mathcal{H}_h y, \mathcal{H}_h y^{(\sigma)} \right) &= \left(a y_{\bar{x}\bar{x}}^{(\sigma)}, \mathcal{H}_h y^{(\sigma)} \right) + \left(A \Delta_{0t_{j+\sigma}}^\beta y_{\bar{x}\bar{x}}, \mathcal{H}_h y^{(\sigma)} \right) - \\ &- \left(d \mathcal{H}_h y^{(\sigma)}, \mathcal{H}_h y^{(\sigma)} \right) + \left(\mathcal{H}_h \varphi, \mathcal{H}_h y^{(\sigma)} \right). \end{aligned} \quad (13)$$

Преобразуем суммы, входящие в тождество (6), с учетом (6), (4), (5), леммы 3 [8, с. 109] и леммы 1 [8]

$$\left(\Delta_{0t_{j+\sigma}}^\alpha \mathcal{H}_h y, \mathcal{H}_h y^{(\sigma)} \right) \geq \frac{1}{2} \left(1, \Delta_{0t_{j+\sigma}}^\alpha (\mathcal{H}_h y)^2 \right) \geq \frac{1}{2} \Delta_{0t_{j+\sigma}}^\alpha \|\mathcal{H}_h y\|_0^2; \quad (14)$$

$$\left(a y_{\bar{x}\bar{x}}^{(\sigma)}, \mathcal{H}_h y^{(\sigma)} \right) = \left(a y_{\bar{x}\bar{x}}^{(\sigma)}, y^{(\sigma)} + \frac{h^2}{12} y_{\bar{x}\bar{x}}^{(\sigma)} \right) = - \left(a, \left(y_{\bar{x}}^{(\sigma)} \right)^2 \right) +$$

$$+ \left(\frac{ah^2}{12}, \left(y_{\bar{x}\bar{x}}^{(\sigma)} \right)^2 \right) \leq -c_0 \|y_{\bar{x}}^{(\sigma)}\|_0^2 + \frac{c_0}{3} \|y_{\bar{x}}^{(\sigma)}\|_0^2 = -\frac{2c_0}{3} \|y_{\bar{x}}^{(\sigma)}\|_0^2; \quad (15)$$

$$\begin{aligned} & \left(A \Delta_{0t_{j+\sigma}}^\beta y_{\bar{x}\bar{x}}, \mathcal{H}_h y^{(\sigma)} \right) \leq - \left(\frac{A}{2}, \Delta_{0t_{j+\sigma}}^\beta (y_{\bar{x}})^2 \right) + \\ & + \left(\frac{Ah^2}{24}, \Delta_{0t_{j+\sigma}}^\beta (y_{\bar{x}\bar{x}})^2 \right) \leq -\frac{A}{2} \Delta_{0t_{j+\sigma}}^\beta \|y_{\bar{x}}\|_0^2 + \frac{Ah^2}{24} \Delta_{0t_{j+\sigma}}^\beta \|y_{\bar{x}\bar{x}}\|_0^2 \leq \\ & \leq -\frac{A}{2} \Delta_{0t_{j+\sigma}}^\beta \|y_{\bar{x}}\|_0^2 + \frac{A}{6} \Delta_{0t_{j+\sigma}}^\beta \|y_{\bar{x}}\|_0^2 = -\frac{A}{3} \Delta_{0t_{j+\sigma}}^\beta \|y_{\bar{x}}\|_0^2; \end{aligned} \quad (16)$$

$$- \left(d \mathcal{H}_h y^{(\sigma)}, \mathcal{H}_h y^{(\sigma)} \right) \leq \left(d, \left(\mathcal{H}_h y^{(\sigma)} \right)^2 \right) \leq c_2 \| \mathcal{H}_h y^{(\sigma)} \|_0^2; \quad (17)$$

$$\left(\mathcal{H}_h \varphi, \mathcal{H}_h y^{(\sigma)} \right) \leq \frac{1}{2} \| \mathcal{H}_h y^{(\sigma)} \|_0^2 + \frac{1}{2} \| \mathcal{H}_h \varphi \|_0^2. \quad (18)$$

Принимая во внимание преобразования (14)-(18), из (6) находим

$$\Delta_{0t_{j+\sigma}}^\alpha \| \mathcal{H}_h y \|_0^2 + \| y_{\bar{x}}^{(\sigma)} \|_0^2 + \Delta_{0t_{j+\sigma}}^\beta \| y_{\bar{x}} \|_0^2 \leq M_1 \| \mathcal{H}_h y^{(\sigma)} \|_0^2 + M_2 \| \mathcal{H}_h \varphi \|_0^2. \quad (19)$$

Перепишем (19) в другой форме

$$\begin{aligned} & \Delta_{0t_{j+\sigma}}^\alpha \| \mathcal{H}_h y \|_0^2 + \| y_{\bar{x}}^{(\sigma)} \|_0^2 + \Delta_{0t_{j+\sigma}}^\beta \| y_{\bar{x}} \|_0^2 \leq \\ & \leq M_3^\sigma \| \mathcal{H}_h y^{j+1} \|_0^2 + M_4^\sigma \| \mathcal{H}_h y^j \|_0^2 + M_2 \| \mathcal{H}_h \varphi \|_0^2. \end{aligned} \quad (20)$$

Случай 1. Пусть $\alpha > \beta$, тогда на основании леммы 7 [6] из (20) получаем $\| \mathcal{H}_h y^{j+1} \|_0^2 \leq M_5 \left(\| \mathcal{H}_h y^0 \|_0^2 + \max_{0 \leq j' \leq j} \| \mathcal{H}_h \varphi^{j'} \|_0^2 \right)$, из последнего с учетом (5) находим (8).

Случай 2. Пусть $\alpha = \beta$, тогда на основании леммы 7 [6] и (5) из (20) получаем (9).

Случай 3. Пусть $\alpha < \beta$, тогда с учетом леммы и (5) из (20) получаем $\Delta_{0t_{j+\sigma}}^\beta \| y_{\bar{x}} \|_0^2 \leq M_6^\sigma \| y_{\bar{x}}^{j+1} \|_0^2 + M_7^\sigma \| y_{\bar{x}}^j \|_0^2 + M_8 \| \varphi \|_0^2$, из чего на основании леммы 7 [6] получаем (3). ■

Из полученных оценок (8)-(3) следуют единственность и устойчивость решения схемы (5)-(7) по начальным данным и правой части, а также сходимость решения (5)-(7) к решению исходной задачи (1)-(3) так, что существует такое τ_0 , что при $\tau \leq \tau_0$ справедливы оценки:

- 1) $\| y^{j+1} - u^{j+1} \|_0^2 \leq M_9 (h^4 + \tau^{2-\max\{\alpha, \beta\}})$ в случае, когда $\alpha > \beta$;
 - 2) $\| y^{j+1} - u^{j+1} \|_0^2 + \| y_{\bar{x}}^{j+1} - u_{\bar{x}}^{j+1} \|_0^2 \leq M_{10} (h^4 + \tau^2)$ в случае, когда $\alpha = \beta$;
 - 3) $\| y_{\bar{x}}^{j+1} - u_{\bar{x}}^{j+1} \|_0^2 \leq M_{11} (h^4 + \tau^{2-\max\{\alpha, \beta\}})$ в случае, когда $\alpha < \beta$,
- где $M_9, M_{10}, M_{11} = \text{const} > 0$ зависят только от входных данных задачи (1)-(3) и не зависят от параметров сетки h и τ .

Заключение

Настоящая работа посвящена изучению первой краевой задачи для модифицированного уравнения влагопереноса с двумя операторами дробного дифференцирования Герасимова-Капуто разных порядков α, β . Для приближенного решения поставленной задачи на равномерной сетке построена разностная схема высокого порядка точности. Методом энергетических неравенств для решения разностной задачи получены априорные оценки при различных значениях α, β . Из полученных оценок следуют единственность и устойчивость решения по правой части и начальным данным. В предположении существования точного решения в классе достаточно гладких функций, а также в силу линейности рассматриваемой задачи полученные оценки позволяют утверждать сходимость приближенного решения к точному решению со скоростью $O(h^4 + \tau^2)$ при $\alpha = \beta$ и $O(h^4 + \tau^{2-\alpha})$ при $\alpha \neq \beta$, где α, β — порядки дробных производных Герасимова-Капуто.

СПИСОК ЛИТЕРАТУРЫ

1. Микеладзе Ш. Е. О численном интегрировании уравнений эллиптического и параболического типов // Изв. АН СССР. Сер. матем. 1941. Т. 5. № 1. С. 57–74.
2. *Lele S. K.* Compact finite difference schemes with spectral-like resolution // Journal of Computational Physics. 1992. Vol. 103. № 1. pp. 16–42.
3. *Gao G. H., Sun Z. Z.* A compact finite difference scheme for the fractional sub-diffusion equations // J Comput Phys. 2011. Vol. 230. № 3. pp. 586–595.
4. *Liao H. L., Sun Z. Z.* Maximum norm error bounds of ADI and compact ADI methods for solving parabolic equations // Numer. Methods Partial Diff. Eqns. 2010. Vol. 26. № 1. pp. 37–60.
5. Бештоков М. Х. О численном решении нелокальной краевой задачи для вырождающегося псевдопараболического уравнения // Дифференц. уравнен. 2016. Т. 52. № 10. С. 1393–1406.
6. Бештоков М. Х. Численное исследование начально-краевых задач для уравнения соболевского типа с дробной по времени производной // Журн. выч. матем. и матем. физ. 2019. Т. 59. № 2. С. 185–202.
7. Самарский А. А. Теория разностных схем. М.: Наука, 1983. 616 с.
8. *Alikhanov A. A.* A new difference scheme for the time fractional diffusion equation // Journal of Computational Physics. 2015. Vol. 280. pp. 424–438.

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ ИЗГИБА ИЗОТРОПНЫХ НАНОПЛАСТИН В РАМКАХ НЕЛОКАЛЬНОЙ ТЕОРИИ ГРАДИЕНТА ДЕФОРМАЦИИ

О. В. Гермидер, В. Н. Попов

*Северный (Арктический) федеральный университет имени М. В.
Ломоносова, г. Архангельск, Россия*

В рамках нелокальной теории градиента деформации предложен новый подход к построению решения задачи моделирования изгиба изотропной нанопластины при действии нормально распределенной по поверхности этой пластины нагрузки. Решение неоднородного дифференциального уравнения в частных производных для изотропной нанопластины получено в виде частичной суммы двойного ряда по многочленам Чебышёва первого рода. В качестве точек коллокаций использованы корни этих многочленов. **Ключевые слова:** *нелокальная теория градиента деформации, моделирование изгиба нанопластин, метод коллокации, многочлены Чебышёва первого рода.*

Введение

Микро- и наноструктурные компоненты, такие как пластины и облочки, находят широкое применение в микро- и нанoeлектромеханических системах в качестве датчиков вибрации, детекторов атомной пыли, резонаторов [1]. Классические модели, основанные на гипотезах Кирхгофа, неприменимы для описания состояния нанопластин, поскольку в них не учитывается внутренняя структура материала пластин. В связи с этим в [2] и [3] согласно нелокальной теории градиента деформации [4] рассмотрены модели, которые могут учитывать микроструктурные эффекты. В [5] с использованием вариационного принципа приведен вывод граничных условий для краевой задачи моделирования изгиба нанопластины, которые обобщают классические граничные условия [6]. В [7] представлены результаты моделирования статического изгиба нанопластин с различными комбинациями граничных условий простого опирания, защемления и свободного края с использованием преобразования Фурье.

В представленной работе предлагается новый подход к построению решения задачи моделирования изгиба изотропной нанопластины при действии нормально распределенной по поверхности этой пластины нагрузки с использованием многочленов Чебышёва первого рода. Решение неоднородного дифференциального уравнения в частных производных шестого порядка для изотропной нанопластины [3] записывается в виде частичной суммы двойного ряда по многочленам Чебышёва первого рода. В качестве точек коллокаций использованы корни этих многочленов. Исследование основано на положениях теории нелокального градиента деформации второго порядка с вариационно согласованными граничными условиями на сторонах нанопластины.

1. Постановка задачи. Общие положения

Рассмотрим задачу моделирования изгиба изотропной прямоугольной нанопластины ($0 \leq x \leq d_1$, $0 \leq y \leq d_2$, $-h/2 \leq z \leq h/2$), которая находится под действием поперечной нагрузки интенсивностью $q(x, y)$. Здесь h – толщина нанопластины. Примем срединную плоскость недеформированной нанопластины за плоскость xy ; ось z направим в сторону ненагруженной внешней плоскости. Объемными силами пренебрегаем. В рамках модели отрицательного градиента деформации второго порядка при плоском распределении соотношения между компонентами напряжения и деформации в декартовых координатах x и y имеют следующий вид [3]:

$$\begin{aligned}\sigma_x &= \frac{E}{1-\nu^2}(\varepsilon_x + \nu\varepsilon_y) - l^2 \frac{E}{1-\nu^2} \Delta(\varepsilon_x + \nu\varepsilon_y), \\ \sigma_y &= \frac{E}{1-\nu^2}(\varepsilon_y + \nu\varepsilon_x) - l^2 \frac{E}{1-\nu^2} \Delta(\varepsilon_y + \nu\varepsilon_x), \\ \tau_{xy} &= \frac{E}{2(1+\nu)}\gamma_{xy} - l^2 \frac{E}{2(1+\nu)} \Delta\gamma_{xy},\end{aligned}$$

где σ_x и σ_y – компоненты нормального напряжения, E и ν – модуль Юнга и коэффициент Пуассона, соответственно, τ_{xy} – компонента касательного напряжения, γ_{xy} – компонента деформации сдвига, Δ – оператор Лапласа, l – нелокальный параметр, определяющий микроструктурные эффекты.

Изгибающие M_x и M_y и крутящий M_{xy} моменты определяем через функцию изгиба $\omega(x, y)$ [3]:

$$M_x = -D \left(\frac{\partial^2 \omega}{\partial x^2} + \nu \frac{\partial^2 \omega}{\partial y^2} \right) + l^2 D \left(\frac{\partial^4 \omega}{\partial x^4} + \nu \frac{\partial^4 \omega}{\partial y^4} + (1 + \nu) \frac{\partial^4 \omega}{\partial x^2 \partial y^2} \right),$$

$$M_y = -D \left(\frac{\partial^2 \omega}{\partial y^2} + \nu \frac{\partial^2 \omega}{\partial x^2} \right) + l^2 D \left(\frac{\partial^4 \omega}{\partial y^4} + \nu \frac{\partial^4 \omega}{\partial x^4} + (1 + \nu) \frac{\partial^4 \omega}{\partial x^2 \partial y^2} \right),$$

$$M_{xy} = D(1 - \nu) \frac{\partial^2 \omega}{\partial x \partial y} - l^2 D \left(\frac{\partial^4 \omega}{\partial x^3 \partial y} + \frac{\partial^4 \omega}{\partial y^3 \partial x} \right),$$

где D – цилиндрическая жесткость пластины [6]

$$D = \frac{Eh^3}{12(1 - \nu^2)}.$$

В этом случае дифференциальное уравнение в частных производных для определения изгиба изотропной нанопластины имеет вид [3]:

$$D \left(\frac{\partial^4 \omega}{\partial x^4} + 2 \frac{\partial^4 \omega}{\partial x^2 \partial y^2} + \frac{\partial^4 \omega}{\partial y^4} \right) - l^2 D \left(\frac{\partial^6 \omega}{\partial x^6} + 3 \frac{\partial^6 \omega}{\partial x^4 \partial y^2} + 3 \frac{\partial^6 \omega}{\partial y^4 \partial x^2} + \frac{\partial^6 \omega}{\partial y^6} \right) = q. \quad (1)$$

где второе слагаемое в скобках у $l^2 D$ обусловлено нелокальным эффектом. Рассматриваемое уравнение равновесия (1) является дифференциальным уравнением в частных производных шестого порядка, поэтому для нахождения его решения необходимо задания трех граничных условий на каждой стороне пластины. В связи с этим в дополнение к классическим граничным условиям для корректно поставленной краевой задачи моделирования изгиба нанопластины в рамках нелокальной теории градиента второго порядка должны быть удовлетворены и неклассические граничные условия. В частности, для защемленной по всему периметру нанопластины классические граничные условия имеют вид [6]:

$$\omega = 0, \quad \frac{\partial \omega}{\partial x} = 0, \quad x = 0, d_1,$$

$$\omega = 0, \quad \frac{\partial \omega}{\partial y} = 0, \quad y = 0, d_2,$$

неклассические условия [7]:

$$l^2 D \left(\frac{\partial^3 \omega}{\partial x^3} + \nu \frac{\partial^3 \omega}{\partial x \partial y^2} \right) = 0, \quad x = 0, d_1,$$

$$l^2 D \left(\frac{\partial^3 \omega}{\partial y^3} + \nu \frac{\partial^3 \omega}{\partial x^2 \partial y} \right) = 0, \quad y = 0, d_2.$$

Для шарнирного закрепления на стороне нанопластины $x = 0$ имеем согласно [7]:

$$\omega = 0, \quad \frac{\partial^2 \omega}{\partial x^2} = 0, \quad x = 0,$$

$$- D \left(\frac{\partial^2 \omega}{\partial x^2} + \nu \frac{\partial^2 \omega}{\partial y^2} \right) + \\ + D l^2 \left(\frac{\partial^4 \omega}{\partial x^4} + \nu \frac{\partial^4 \omega}{\partial y^4} + (3 - \nu) \frac{\partial^4 \omega}{\partial x^2 \partial y^2} \right) = 0, \quad x = 0.$$

2. Построение решения краевой задачи

Для нахождения решения краевой задачи уравнение (1) и граничные условия приводим к безразмерному виду посредством введения $x^* = x/d_1$, $y^* = y/d_1$, $l^* = l/d_1$, $q^* = q/q_0$ и $\omega^* = \omega D / (q_0 d_1^4)$. Раскладываем безразмерную функцию изгиба ω^* по многочленам Чебышёва первого рода $\{T_{j_i}(x_i) = \cos(j_i \arccos x_i), (j_i = 0, 1, \dots, n_i; i = 1, 2)\}$, где переменные x_1 и x_2 определяются через x и y следующим образом:

$$x_1 = \frac{2}{d_1} x - 1, \quad x_2 = \frac{2}{d_2} y - 1.$$

Функцию ω^* находим методом коллокации с использованием корней многочленов Чебышёва первого рода $T_{n_i+1}(x_i)$ в качестве точек коллокации [8]

$$x_{i,k_i} = \cos \left(\frac{\pi(2n_i - 2k_i + 1)}{2(n_i + 1)} \right), \quad k_i = 0, 1, \dots, n_i; \quad i = 1, 2.$$

Заключение

В работе предложен новый подход для решения задачи математического моделирования изгиба изотропной прямоугольной нанопластины при действии непрерывно распределенной нагрузки с использованием многочленов Чебышёва. В рамках нелокальной модели градиента деформации получены распределения изгибающих моментов. Для верификации полученных результатов проведен ряд вычислительных экспериментов.

СПИСОК ЛИТЕРАТУРЫ

1. *Babu B., Patel B. P.* A new computationally efficient finite element formulation for nanoplates using second-order strain gradient Kirchhoff's plate theory // *Composites Part B*. 2019. Vol. 168. P. 302–311.
2. *Lazopoulos K. A.* On the gradient strain elasticity theory of plates // *European Journal of Mechanics A/Solids*. 2004. Vol. 23. P. 843–852.
3. *Papargyri-Beskou S., Beskos D. E.* Static, stability and dynamic analysis of gradient elastic flexural Kirchhoff plates // *Archive of Applied Mechanics*. 2008. Vol. 78. P. 625–635.
4. *Mindlin R. D.* Second gradient of strain and surface-tension in linear elasticity // *International Journal of Solids and Structures*. 1965. Vol. 1. № 4. P. 417–438.
5. *Papargyri-Beskou S., Giannakopoulos A. E., Beskos D. E.* Variational analysis of gradient elastic flexural plates under static loading // *International Journal of Solids and Structures*. 2010. Vol. 20. P. 2755–2766.
6. *Timoshenko S., Woinowsky-Krieger S.* Theory of plates and shells. New York: McGraw-Hill Book Comp., 1959. 580 p.
7. *Babu B., Patel B. P.* Analytical solution for strain gradient elastic Kirchhoff rectangular plates under transverse static loading // *European Journal of Mechanics / A Solids*. 2019. Vol. 73. P. 101–111.
8. *Mason J., Handscomb D.* Chebyshev polynomials. Boca Raton: CRC Press, 2003. 360 p.

Гермидер Оксана Владимировна — кандидат физико-математических наук, доцент кафедры инженерных конструкций, архитектуры и графики Северного (Арктического) федерального университета имени М.В. Ломоносова. E-mail: o.germider@narfu.ru

Попов Василий Николаевич — доктор физико-математических наук, профессор, профессор кафедры высшей и прикладной математики Северного (Арктического) федерального университета имени М.В. Ломоносова. E-mail: v.popov@narfu.ru

МОДЕЛИРОВАНИЕ ИНТЕНСИВНОСТЕЙ ТРАНСПОРТНЫХ ПОТОКОВ ДЛЯ МОДУЛЯ ВАЛИДАЦИИ КОМПЛЕКСА MARLIN24

Т. И. Тисленко

Сибирский федеральный университет, г. Красноярск, Россия

В статье представлены результаты разработки модуля валидации для программного комплекса MARLIN24, предназначенного для адаптивного управления светофорными объектами. Структура комплекса включает имитационный модуль «разумный водитель» (Intelligent Driver Model, IDM), который использует данные об интенсивности движения через детектируемые участки дорожной сети. Это и послужило стимулом для создания модели, учитывающей зависимость входных интенсивностей движения. В ходе разработки модуля валидации в работе [8] был предложен метод, который позволял моделировать количество машин, проехавших через поперечное сечение участка дороги за фиксированный промежуток времени. Данный подход обладал существенным недостатком: приходилось учитывать связь шага симуляции имитационной модели и длину интервала наблюдения при построении интересующей нас статистики. Преимущество представленного в текущей работе подхода заключается в рассмотрении временных промежутков между появлением автомобилей. В работе представлены результаты численных экспериментов по моделированию многомерных распределений интенсивностей движения на основе маргинальных эмпирических распределений, полученных с оптических детекторов города Красноярска в период с 2019 по 2020 год.

Ключевые слова: *копулы Маршалла-Олкина, методы Монте-Карло, интенсивности транспортных потоков.*

Введение

Разработка и внедрение программных и математических инструментов для моделирования транспортных потоков и управления светофорными объектами для наиболее нагруженных участков дорожной сети учитывает общесоциальные целевые ориентиры: подвижность населения, снижение аварийности, рисков и угроз безопасности транспорта.

Работа поддержана Красноярским математическим центром, финансируемым Минобрнауки РФ (Соглашение 075-02-2024-1429).

В работах [2, 1, 3] был реализован комплекс программных средств MARLIN24 для адаптивного управления светофорными объектами, который решает задачу разгрузки транспортных узлов.

Для проведения численных экспериментов сравнения показателей эффективности управления MARLIN24 [1] был разработан модуль IDM (Intelligent Driver Model, IDM) [4]. В статье предлагается способ задания входных данных модели, использующий зависимые интенсивности движения, для его валидации.

1. Моделирование движения транспортных средств

Система уравнений, описывающих текущую скорость i -го автомобиля и расстояние до $(i - 1)$ -го автомобиля в модели IDM имеет вид:

$$\begin{cases} \frac{dv_i}{dt} = a_i \left(1 - \left(\frac{v_i}{v_{0,i}} \right)^\delta \right) - a_i \left(\frac{s^*(v_i, \Delta v_i)}{s_i} \right)^2, & (1) \\ s^*(v_i, \Delta v_i) = s_{0,i} + v_i T_i + \frac{v_i \Delta v_i}{\sqrt{2a_i b_i}}, & (2) \end{cases}$$

где $s_{0,i} = 4m$ — минимальная желательная дистанция между i и $i - 1$, $v_{0,i} = 14 \frac{M}{C}$ — максимальная желательная скорость i , $\delta = 1$ — компонента ускорения, $T_i = 1c$ — время реакции i -го водителя, $a_i = 1,44 \frac{M}{C^2}$ — максимальное ускорение i , $b_i = 8,61 \frac{M}{C^2}$ — комфортное ускорение при торможении водителя i , s^* — возможное расстояние между i и $i - 1$.

2. Моделирование интенсивностей транспортных потоков

Выше была приведена модель движения для каждого транспортного средства по отдельности. Опишем движение транспортных потоков.

Будем рассматривать транспортный поток как ординарный случайный поток событий. Известны модели, описывающие движение транспортных потоков, на языке теории массового обслуживания [5, 6, 7]. Как правило, используются предположения о том, что законы распределения случайных величин, описывающих время между появлением транспорта, представимы в виде экспоненциальных распределений.

Рассмотрим линейный участок дороги с двумя типами автомобилей, которые задаются точками на этом участке. Автомобили первого типа движутся с постоянной скоростью $v_{0,1}$, а автомобили второго типа — с постоянной скоростью $v_{0,2}$, где $v_{0,1} < v_{0,2}$. Появление транспортных средств первого типа в зоне, оборудованной оптическими датчиками, и, соответственно, время между обнаружениями транспортных средств является случайной величиной, имеющей заданное распределе-

ние. Пусть медленные автомобили (второго типа) расположены в точках $x_0 < x_1 < \dots < x_k < x_{k+1} < \dots$. Эти автомобили перемещаются без столкновений и «взаимодействуют» с остальными транспортными средствами при совпадении их координат. Более быстрые автомобили обгоняют медленные в порядке очереди, скорость при обгоне изменяется согласно (1) и равна $v_{0,1}$. Обгоны происходят через временные промежутки, которые распределены по экспоненциальному закону.

В данной работе под интенсивностью транспортного потока будем понимать число автомобилей, проезжающих через поперечное сечение участка дорожной сети в единицу времени, а под среднесуточной интенсивностью — усредненное количество машин по рабочим дням в течение года. Ранее в работе [8] при описании интенсивности транспортного потока использовалась статистка, описывающая количество машин. В текущей работе используется величина временного интервала между проездом двух автомобилей через сечение участка дорожной сети.

Рассмотрим один из возможных подходов введения в модель зависимых случайных величин. В основе подхода лежит использование двухпараметрической копулы Маршалла-Олкина [10] для совместного распределения временных интервалов появления автомобилей для двух полос.

Пусть случайная величина X с функцией распределения $F(x)$ и случайная величина Y с функцией распределения $G(y)$ описывают временной интервал между проездом двух автомобилей через сечение детектируемого участка на полосах 1 и 2 соответственно. По теореме Склера [9] совместную функцию распределения можно представить копулой C

$$H_{XY}(x, y) = C(F(x), G(y)), \quad \forall x, y \in \mathbb{R}. \quad (3)$$

Далее будем использовать двухпараметрическую копулу Маршалла-Олкина [10, 11] с коэффициентами $0 \leq \alpha, \beta \leq 1$

$$C_{\alpha, \beta}(u, v) = uv \min(u^{-\alpha}, v^{-\beta}). \quad (4)$$

Моделирование зависимых распределений среднесуточных интенсивностей состоит из следующих этапов: предобработки данных, имитационного моделирования временных интервалов и вычисления числа полученных интервалов, входящих в 1-секундное разбиение.

Этап предобработки включает оценивание плотности распределений случайных величин X и Y , которые описывают временные интервалы между появлением транспортных средств на полосах 1 и 2 соответственно. На первом шаге строятся ядерные оценки плотности с использованием ядра Епанечникова [12]. Затем делается упрощающее предположение, что каждая из рассматриваемых случайных величин может быть

представлена в виде смеси экспоненциального и нормального распределений. Параметры смесей определяются с помощью EM-алгоритма [13].

Результаты этапа предобработки на 10-й итерации EM алгоритма и вид маргинальных плотностей распределений случайных величин X и Y приведены на рисунке 1а) и рисунке 1б) соответственно.

Для этапа имитационного моделирования временных интервалов была разработана модификация метода дискретной суперпозиции Монте-Карло [14] для генерации случайной величины (X, Y) :

шаг 1: моделируем случайные величины $w_1, w_2 \sim U_{[0; 1]}$;

шаг 2: если выполняется $\sum_{m=1}^M p_m < w_1 < \sum_{m=1}^{M+1} p_m$, то зафиксируем функцию распределения M -й компоненты смеси как $F = F_M$;

шаг 3: аналогично этапу 2, для w_2 определяем компоненту смеси G ;

шаг 4: восстанавливаем пару случайных величин (U, V) , совместное распределение которых является копулой Маршалла-Олкина [11];

шаг 5: восстанавливаем U -квантиль X_U распределения F ;

шаг 6: восстанавливаем V -квантиль Y_V распределения G .

Гистограмма выборки, полученной моделированием копулой Маршалла-Олкина с параметрами $\alpha = 0.9, \beta = 0.2$ представлена на рисунке 2.

3. Заключение

Текущая работа продолжает исследование задачи моделирования интенсивности движения транспортных средств [8] и расширяет данный подход, позволяя более точно моделировать различные зависимости как между числом моделируемых транспортных средств, так и временными интервалами между их появлениями, что позволяет более гибко моделировать интенсивности транспортных потоков на полосах.

Для демонстрации зависимости моделируемых данных в таблице 1 приведены формулы для копул из семейства Маршалла-Олкина [11] для

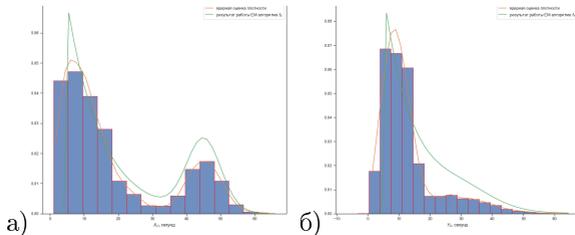


Рис. 1. Результаты этапа предобработки: оценка плотностей маргинальных распределений: а) случайной величины X ; б) случайной величины Y

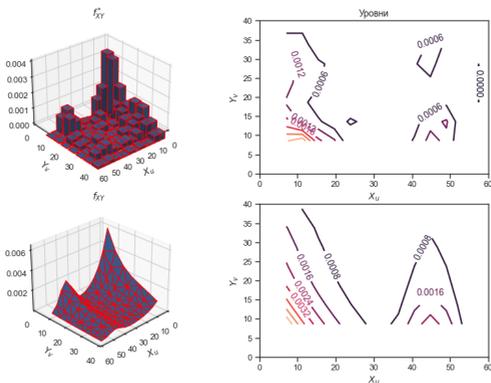


Рис. 2. Сравнение эмпирической f_{XY}^* и теоретической f_{XY} функции плотности распределения (X, Y) при моделировании значений (U, V) для копулы Маршалла-Олкина с параметрами $\alpha = 0.9, \beta = 0.25$

коэффициентов корреляции Спирмена ρ_S , Кендалла τ_K , коэффициента верхней хвостовой зависимости λ_U . Для проверки сгенерированных копулой значений был использован критерий согласия Колмогорова-Смирнова. Было получено расстояние Колмогорова $D_n = 0.59$. С помощью метода генерации бутстреп выборок [15] был превышен уровень значимости 0.1 и получено $pvalue = 0.57$.

Таблица 1

Полученные статистики для копул Маршалла-Олкина

Теоретические значения	Расчетные значения
$\rho_S = \frac{3\alpha\beta}{2\alpha - \alpha\beta + 2\beta} = 0.27$	$\rho_S = 0.215$
$\tau_K = \frac{\alpha\beta}{\alpha - \alpha\beta + \beta} = 0.2$	$\tau_K = 0.15$
$\lambda_U = \min(\alpha, \beta) = 0.2$	$\lambda_U = 0.14$

СПИСОК ЛИТЕРАТУРЫ

1. Тисленко Т. И., Семенова Д. В., Сергеева Н. А. Оптимизация планов координаций для светофорных объектов участка дорожной сети // Системы управления, информационные технологии и математическое моделирование: материалы IV Всероссийской научно-практической конференции с международным участием (Омск, 19 мая 2022 г.): в 2 т., Омск: Изд-во ОмГТУ, 2022. С. 255–261.
2. Tislenko T. I., Semenova D. V., Sergeeva N. A., Goldenok E. E., Kononova N. V. Multiagent Reinforcement Learning for Integrated Network: Applying to a Part of the Road Network of Krasnoyarsk City // 2022

- IEEE 16th International Conference on Application of Information and Communication Technologies (AICT), Washington DC, 2022. P. 1–5.
3. *Tislenko T. I., Semenova D. V., Soldatenko A. A., Sergeeva N. A., Goldenok E. E., Kononova N. V.* Modeling and Comparison of Different Management Approaches on the Intersections Network // In: Vishnevskiy V. M., Samouylov K. E., Kozyrev D. V. (eds) Distributed Computer and Communication Networks. DCCN 2023. Communications in Computer and Information Science, vol 2129. Springer, Cham, 2024. P. 1–15.
 4. *Treiber M., Hennecke A., Helbing D.* Congested traffic states in empirical observations and microscopic simulations // *Transportation Physics Reviews E*, 2000, Vol. 62, P. 1805–1824.
 5. *Замятин А. А., Малышев В. А.* Транспортные автомобильные потоки – введение в вероятностный подход // Труды МФТИ. 2010. Т. 4. № 4.
 6. *Задорожный В. Н.* Имитационное и статистическое моделирование: учебное пособие. Омск: Изд-во ОмГТУ, 2013. 136 с.
 7. *Замятин А. А., Малышев В. А.* Транспортные автомобильные потоки - введение в вероятностный подход // ТРУДЫ МФТИ. Труды Московского физико-технического института (национального исследовательского университета), 2010, Т. 2, № 4(8), С. 58–74.
 8. *Тисленко Т. И.* Моделирование интенсивностей транспортных потоков при помощи копул Маршала-Олкина // ИТ. НАУКА. КРЕАТИВ. Т. 5. Системы управления, информационные технологии и математическое моделирование : материалы I Междунар. форума (Омск, 14–16 мая 2024 г.) : в 5 Ч. / науч. ред. П. С. Ложников, отв. ред. И. Г. Ольгина. Омск : Издательство ОмГТУ, 2024. 70 с.
 9. *Nelsen Roger B.* An Introduction to Copulas. Springer, 2006, 270 с.
 10. *Marshall A.* Copulas, Marginals, and Joint Distributions // *Lecture Notes-Monograph Series. Distributions with Fixed Marginals and Related Topics*, 1996, no. 28, P. 213–222.
 11. *Quesada-Molina J.J., Rodríguez-Lallena J.A.* Bivariate copulas with quadratic sections // *Journal of Nonparametric Statistics*, 1995, pp. 323–337.
 12. *Епанечников В. А.* Непараметрическая оценка многомерной плотности вероятности // *Теория вероятностей и ее применения*. 1969. Т. 14. С. 156–161.
 13. *Bishop C. M.* Pattern Recognition and Machine Learning. Springer, 2007.
 14. *Войтшик А. В.* Основы метода Монте-Карло: Учебное пособие, ред. В. Н. Задорожный. Новосибирск: НГУ, 2010. 108 с.
 15. *Davison A. C., Hinkley D. V.* Bootstrap methods and their application. Cambridge University Press, 1997.

Прикладной
вероятностный и
статистический анализ

ON FIXED POINT OF A MAPPING RELATED TO ISING MODEL ON \mathbb{Q}_p^2

Z. Abdukaxorova

Namangan State University, Namangan, Uzbekistan

In this paper, we study a mapping related to the Ising model on a Cayley tree of order two on \mathbb{Q}_p^2 . We find fixed points of this mapping and analyze the characters of these fixed points. **Keywords:** *p-adic numbers, fixed point, repelling, attracting, saddle, Ising model.*

Introduction

Starting in the 20th century, mathematicians have been developing theories based on number systems other than the real numbers.

One of the new areas of math research is focused on p -adic dynamical systems and their chaotic behavior, which are different from the usual dynamical systems we study. These new systems are based on a type of number called p -adic numbers. Scientists Herman and Yoccoz have been exploring this topic since the 1980s, and they've found that p -adic dynamical systems behave in very different ways compared to their real-number counterparts [1]. In addition to the usual patterns we see in real-number systems, p -adic systems can have unique patterns called "fuzzy cycles."

The study of p -adic dynamical systems has been useful in various fields like biology, economics, and physics. In biology, one example is a model created by Khrennikov which uses p -adic numbers to describe how populations grow. It shows that even though different groups within a population might grow very quickly, the overall population growth can still be limited. This model is based on the special properties of p -adic numbers.

Our paper is devoted to extending the characterization of fixed points of nonlinear mappings related to the Ising model on the Cayley tree of order two in \mathbb{Q}_p^2 . In the actual setting, p -adic Ising model has broad theoretical and practical applications in many areas of applied and theoretical sciences [4], [5].

p -adic number

Let \mathbb{Q} be the field of rational numbers. For a fixed prime number p , every rational number $x \neq 0$ can be represented in the form $x = p^r \frac{n}{m}$, where, $r, n \in \mathbb{Z}$, $m \in \mathbb{N}$ and n and m are relatively prime with p : $(p, n) =$

1, $(p, m) = 1$. The p -adic norm of x is given by

$$|x|_p = \begin{cases} p^{-r}, & x \neq 0, \\ 0, & x = 0. \end{cases}$$

This norm is non-Archimedean and satisfies the strong triangle inequality

$$|x + y|_p \leq \max\{|x|_p, |y|_p\}, \quad \forall x, y \in \mathbb{Q}.$$

Moreover, this norm has the following properties:

- 1) if $|x|_p \neq |y|_p$, then $|x \pm y|_p = \max\{|x|_p, |y|_p\}$;
- 2) if $|x|_p = |y|_p$, then $|x \pm y|_p \leq |x|_p$.

The completion of \mathbb{Q} with respect to the p -adic norm defines the p -adic field \mathbb{Q}_p . The elements of the p -adic field \mathbb{Q}_p are called *p -adic numbers*.

Any p -adic number $x \neq 0$ can be uniquely represented in the canonical form

$$x = p^{\gamma(x)}(x_0 + x_1p + x_2p^2 + \dots)$$

where $\gamma(x) \in \mathbb{Z}$ and the integers x_j satisfy that $x_0 \neq 0$, $x_j \in \{0, 1, \dots, p-1\}$, $j \in \mathbb{N}$.

In this case $|x|_p = p^{-\gamma(x)}$.

For $a \in \mathbb{Q}_p$ and $r > 0$ we denote

$$B(a, r) = \{x \in \mathbb{Q}_p : |x - a|_p < r\}.$$

p -adic exponential is defined by

$$\exp_p(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!},$$

which converges for $x \in B(0, p^{-1/(p-1)})$. Denote

$$\mathcal{E}_p = \left\{ x \in \mathbb{Q}_p : |x - 1|_p < p^{-1/(p-1)} \right\} = B(1, p^{-1/(p-1)})$$

(see [2]).

The dynamics on \mathbb{Q}_p^2

Consider a non-linear mapping

$$F : x = (x_1, x_2, \dots, x_m) \in \mathbb{Q}_p^m \rightarrow x' = F(x) = (x'_1, x'_2, \dots, x'_m) \in \mathbb{Q}_p^m$$

given by $x'_k = F_k(x_1, x_2, \dots, x_m)$, $k = 1, 2, \dots, m$, where $F_i : \mathbb{Q}_p^m \rightarrow \mathbb{Q}_p^m$, $i = 1, 2, \dots, m$ is a continuously differentiable single-valued function.

Definition 1. If $F(x^*) = x^*$ then $x^* \in \mathbb{Q}_p^m$ is called a *fixed point* of the mapping F .

The eigenvalues of the Jacobian matrix J_F determine the local behavior of the non-linear dynamical system (see [3]).

Definition 2. A fixed point x^* of a mapping F is called hyperbolic point of its Jacobian J_F at x^* has no eigenvalues on the unit circle.

Definition 3. The hyperbolic point is called

- attracting point if all the eigenvalues of the Jacobi matrix $J_F(x^*)$ are less than 1 in absolute value;
- repelling point if all the eigenvalues of the Jacobi matrix $J_F(x^*)$ are greater than 1 in absolute value;
- a saddle point otherwise.

Problem

The problem of finding H_A -weakly periodic p -adic generalized Gibbs measure for the p -adic Ising model on the Cayley tree brings to consider the following mapping:

$$W_{\theta,\eta} : \mathbf{h} = (\widehat{h_0}, \widehat{h_1}) \in \mathbb{Q}_p^2 \rightarrow \mathbf{h}' = \left(\eta^3 \frac{\theta \widehat{h_0} + 1}{\widehat{h_0} + \theta} \cdot \frac{\theta \widehat{h_1} + 1}{\widehat{h_1} + \theta}; \eta^3 \left(\frac{\theta \widehat{h_0} + 1}{\widehat{h_0} + \theta} \right)^2 \right) \in \mathbb{Q}_p^2.$$

We study the dynamics of $W_{\theta,\eta}$. Firstly, we find fixed points of $W_{\theta,\eta}$

$$\mathbf{h}^{(1)} = (\widehat{h_0^{(1)}}, \widehat{h_1^{(1)}}), \mathbf{h}^{(2)} = (\widehat{h_0^{(2)}}, \widehat{h_1^{(2)}}), \mathbf{h}^{(3)} = \left(-\frac{\theta + \eta^3}{\theta \eta^3 + 1}, \frac{1}{\eta^3} \right)$$

where $\widehat{h_0^{(1)}} \equiv 1 \pmod{p}$, $\widehat{h_0^{(2)}} \equiv -1 \pmod{p}$, and $\eta, \theta \in \mathcal{E}_p$.

Main result

Stability of these fixed points are as follows:

Theorem 1. For the fixed points of $W_{\theta,\eta}$ the following statements hold:

- 1) $\mathbf{h}^{(1)}$ is attracting fixed point;
- 2) $\mathbf{h}^{(2)}$ is repelling fixed point;
- 3) The fixed point $\mathbf{h}^{(3)}$ is saddle.

Conclusion

In [7], we investigate the characteristics of fixed points for a mapping associated with the p -adic Ising model without an external field on the Cayley tree of order two in \mathbb{Q}_p^2 . In [6], the authors examine certain properties of the fixed points of a mapping related to the p -adic Ising model in an external field on the Cayley tree in \mathbb{Q}_p . In this paper, we study a mapping connected to the Ising model on a Cayley tree of order two in \mathbb{Q}_p^2 . We determine fixed points of this mapping and analyze their properties.

REFERENCES

1. *Herman M.R. and Yoccoz J.C.* Generalization of some theorem of small divisors to non-Archimedean fields. Berlin: Springer, 1983. P.408-447.
2. *Vladimirov V.S., Volovich I.V., Zelenov, E.I.* p -adic Analysis and Mathematical Physics Singapore: World Scientific, 1994.
3. *Rozikov U.A.* An introduction to mathematical billiards Singapore: World Scientific, 2019.
4. *Baxter R.J.* Exactly solved models in statistical mechanics. London: Academic Press, Inc., 1982.
5. *Eggarter T.P.* Cayley trees, the Ising problem, and the thermodynamic limit. Phys. Rev. B, 1974. Vol. 9. P. 2989–2992.
6. *Mukhamedov M.M., Rahmatullaev M.M., Tukhtabaev A.M. and Mamadjonov M.* The p -adic Ising model in an external field on a Cayley tree: periodic Gibbs measures. Theoretical and Mathematical Physics, 2023. Vol. 216(2). P. 1238–1253.
7. *Rahmatullaev M.M. and Abdukaxorova Z.* H_A -weakly periodic p -adic generalized Gibbs measures for the p -adic Ising model on the Cayley tree of order two. p -Adic Numbers, Ultrametric Analysis and Applications, 2024. Vol. 16. No. 3. P. 233-263

Abdukaxorova Zulxumor Tuxtasinovna — student of PhD, teacher, Namangan state university. E-mail: zulxumorabdukaxorova@gmail.com

METHOD OF FACTORIZATION OF SIMULTANEOUS EQUATIONS FOR UNBIASED ESTIMATION OF THEIR COEFFICIENTS

A. Minitaeva

Bauman Moscow State Technical University, Moscow, Russian Federation

The paper deals with equations with differently directed explanatory effects of simultaneous variables which are interdependent or endogenous variables. The goal is to research and develop estimation methods that can circumvent the problem of correlated variables and errors. A method is proposed and the efficiency of it is compared to that of both the usual method of least squares and two-, three-step methods of least squares. A violation of the basic assumption of least-squares methods about the independence of predetermined variables and errors leads to the fact that least-squares estimation methods in this case lose their best properties: unbiasedness, efficiency, and consistency. The paper proves the theorem about the unbiasedness of estimates obtained using the method of factorization of simultaneous equations. **Keywords:** *exogenous variables, endogenous variables, simultaneous equations, method of factorization of simultaneous equations, two-step least squares method, unbiased estimation.*

Introduction

In contrast to regression equations, which exhibit a unidirectional causal relationship from independent variables to dependent variables, systems analysis involves equations where explanatory effects of variables are bidirectional. This approach to structural modeling and system analysis is effective not only in regression studies of sociometric and econometric systems but also in exploring physicochemical dependencies. It accounts for the multifactorial mutual influence of physical laws and chemical properties, where the variables can act as both regressors and predictors [1].

Such interdependent variables are known as endogenous variables. A distinctive feature of simultaneous equation systems is that an endogenous variable, which serves as a regressor in one equation, can also appear as a regressor in another equation. Consequently, endogenous variables become correlated with the random disturbances in the equations [2,3,6]. The num-

ber of equations in such a system must match the number of endogenous variables.

1. Materials and methods

Here, we present the general form of the system of simultaneous equations:

$$\begin{cases} y_{t1} = \beta_{21}y_{t2} + \beta_{31}y_{t3} + \dots + \beta_{k1}y_{tk} + y_{11}x_{t1} + y_{21}x_{t2} + \dots + y_{n1}x_{tn} + u_{t1}, \\ y_{t2} = \beta_{12}y_{t1} + \beta_{32}y_{t3} + \dots + \beta_{k2}y_{tk} + y_{12}x_{t1} + y_{22}x_{t2} + \dots + y_{n2}x_{tn} + u_{t2}, \\ \dots \\ y_{tk} = \beta_{1k}y_{t1} + \beta_{k2}y_{t2} + \dots + \beta_{k-1,k}y_{t,k-1} + y_{1k}x_{t1} + y_{2k}x_{t2} + \dots + \\ + y_{n2}x_{tn} + u_{tk}. \end{cases} \quad (1)$$

The regression coefficients are selected to minimize the sum of squared errors. However, when the dependent variables correlate with the error, the objective function of the least squares method becomes dependent on the variable values themselves. Consequently, the least squares estimates of regression coefficients are conditional, valid only for the given set of variable values. This issue persists in multi-step methods [17-19], as they replace simultaneous equations with predetermined ones, which they are not. Such a model cannot effectively analyze an economic system because it requires sensitivity analysis of the resulting indicators to all possible values of the explained variables [5].

The random deviation, an element on the right-hand side of the regression equation, depends on all its components: the regression coefficients, and the values of the dependent and "independent" variables. The relationship between the dependent variable and the random deviation is determined by the equation's form. According to one of the main postulates of the least squares method, the independent variables on the right-hand side are assumed to be predetermined, deterministic, and independent of randomness [4,5]. Such a model is unsuitable for studying economic systems [11-16], as this analysis requires examining the sensitivity of the resulting indicators to all possible values of the explained variables [6].

2. Unbiased estimation of the coefficients

2.1. The Method of Factorization of Simultaneous Equations

Let us describe our own method for obtaining estimates of the parameters of the model of simultaneous equations. It is based on the de-

composition of the simultaneous variables into stochastic factors, which can be called factorization, and the method itself is the method of factorization of simultaneous equations (MFSE) [3,6]. The representation of the variables of equation (3), for which the covariances (4) and (5) are valid, in the form of some combination of instrumental factors is performed on the basis of stochastic factor analysis using the model:

$$(X_i \quad X_i = F \cdot A_{Y_i}^T \quad A_{X_i}^T + \sum_i), \tag{2}$$

where F – is the matrix of values of the stochastic factors; $A = \begin{pmatrix} A_{Y_i} \\ A_{X_i} \end{pmatrix}$ – is the matrix of factor loadings (its blocks correspond to explanatory endogenous and exogenous variables); \sum_i – is the matrix of random deviations of the factor model for the variables Y_i and X_i ; $E \{ \sum_i \} = 0$; $E \{ \sum_i^T \sum_i \} = D_i^2$ (D – is diagonal matrix); the stochastic factors satisfy the conditions:

$$\begin{cases} E \{ F^T u_i \} = 0; \\ E \{ F^T \sum_i \} = 0. \end{cases} \tag{3}$$

In this case the equation is transformed to the form:

$$\begin{aligned} y_i &= AA_{Y_i}^T : A_{X_i}^T + \sum_i \begin{pmatrix} \beta_i \\ \gamma_i \end{pmatrix} + u_i = \\ &= FA_{Y_i}^T : A_{X_i}^T \begin{pmatrix} \beta_i \\ \gamma_i \end{pmatrix} + \sum_i \begin{pmatrix} \beta_i \\ \gamma_i \end{pmatrix} + u_i = \\ &= FA_{Y_i}^T : A_{X_i}^T \begin{pmatrix} \beta_i \\ \gamma_i \end{pmatrix} + w_i, \end{aligned} \tag{4}$$

where $w_i = \sum_i \begin{pmatrix} \beta_i \\ \gamma_i \end{pmatrix}$ – is the vector of stochastic deviations of the i -th equation.

Ultimately, the following equation

$$y_i = AA_{Y_i}^T : FA_{X_i}^T \begin{pmatrix} \beta_i \\ \gamma_i \end{pmatrix} + w_i \tag{5}$$

is an ordinary regression equation in which the regressors, due to conditions (3), are not correlated with random deviations. Therefore, the vector of coefficients $\begin{pmatrix} \beta_i \\ \gamma_i \end{pmatrix}$ can be estimated by the usual method of least squares.

Thus, the estimation of the coefficients of the equation by factorization of simultaneous variables has the form:

$$\begin{pmatrix} \hat{\beta}_i \\ \hat{\gamma}_i \end{pmatrix} = \begin{pmatrix} A_{Y_i} A_{Y_i}^T & A_{Y_i} A_{X_i}^T \\ A_{X_i} A_{Y_i}^T & A_{X_i} A_{X_i}^T \end{pmatrix}^{-1} \begin{pmatrix} A_{Y_i} F^T y_i \\ A_{X_i} F^T y_i \end{pmatrix}. \quad (6)$$

In the equation (6) the orthogonality of the stochastic factors [7] is taken into account.

2.2. Proof of the MFSE unbiasedness

Theorem 1. The estimate (6) of the coefficients of the simultaneous equation obtained using the method of factorization of simultaneous equations is unbiased.

Proof. Let us find the mathematical expectation of the estimate (6):

$$\begin{aligned} E \left\{ \begin{pmatrix} \hat{\beta}_i \\ \hat{\gamma}_i \end{pmatrix} \right\} &= E \left\{ \begin{pmatrix} A_{Y_i} A_{Y_i}^T & A_{Y_i} A_{X_i}^T \\ A_{X_i} A_{Y_i}^T & A_{X_i} A_{X_i}^T \end{pmatrix}^{-1} \begin{pmatrix} A_{Y_i} F^T \\ A_{X_i} F^T \end{pmatrix} y_i \right\} = \\ &= E \left\{ \begin{pmatrix} A_{Y_i} A_{Y_i}^T & A_{Y_i} A_{X_i}^T \\ A_{X_i} A_{Y_i}^T & A_{X_i} A_{X_i}^T \end{pmatrix}^{-1} \begin{pmatrix} A_{Y_i} F^T \\ A_{X_i} F^T \end{pmatrix} \left(x_i; X_i \left(\frac{\hat{\beta}_i}{\hat{\gamma}_i} \right) + w_i \right) \right\} = \\ &= E \left\{ \begin{pmatrix} A_{Y_i} A_{Y_i}^T & A_{Y_i} A_{X_i}^T \\ A_{X_i} A_{Y_i}^T & A_{X_i} A_{X_i}^T \end{pmatrix}^{-1} \begin{pmatrix} A_{Y_i} F^T \\ A_{X_i} F^T \end{pmatrix} \right. \\ &\quad \left. \left(A A_{Y_i}^T; F A_{Y_i}^T \left(\frac{\hat{\beta}_i}{\hat{\gamma}_i} \right) + w_i \right) \right\} = \\ &= \begin{pmatrix} \beta_i \\ \gamma_i \end{pmatrix} + E \left\{ \begin{pmatrix} A_{Y_i} A_{Y_i}^T & A_{Y_i} A_{X_i}^T \\ A_{X_i} A_{Y_i}^T & A_{X_i} A_{X_i}^T \end{pmatrix}^{-1} \begin{pmatrix} A_{Y_i} F^T w_i \\ A_{X_i} F^T w_i \end{pmatrix} \right\} = \\ &= \begin{pmatrix} \beta_i \\ \gamma_i \end{pmatrix} + \begin{pmatrix} A_{Y_i} A_{Y_i}^T & A_{Y_i} A_{X_i}^T \\ A_{X_i} A_{Y_i}^T & A_{X_i} A_{X_i}^T \end{pmatrix}^{-1} \begin{pmatrix} A_{Y_i} F^T E \{w_i\} \\ A_{X_i} F^T E \{w_i\} \end{pmatrix} = \begin{pmatrix} \beta_i \\ \gamma_i \end{pmatrix}. \end{aligned}$$

Thus, for the case of simultaneity of exogenous variables according to the unbiasedness criterion, the estimate obtained by the method of factorization of simultaneous equations is better than the estimate of the two-step method of least squares. ■

Conclusion

Similarly, we can demonstrate that the proposed method is more efficient than the three-step least squares method, which is the second most commonly used method [18]. Therefore, when the variables of the system are interdependent, it is appropriate to use the method of factorization of

simultaneous equations. Research into the nature of simultaneity in econometric equations has shown that both endogenous and exogenous variables can be interrelated. Existing methods for estimating parameters in simultaneous equations systems overlook this fact, leading to biased estimates [8,9]. The method of factorization of simultaneous equations (MFSE) has been developed to provide unbiased parameter estimates and is more effective than the most widely used estimation methods. The MFSE also enhances the multi-model approach for predicting nonlinear nonstationary processes in optimal control problems [10].

REFERENCES

1. *Kanatnikov A., Tkacheva O.* State Observer for a Four-Dimensional System with Vector Output // *Differential Equations*. 2021. Vol. 57. № 11. P. 1460–1467.
2. *Upton G., Cook I.* *Dictionary of Statistics*. (3 ed.) Oxford: Oxford University Press, 2014. 496 p.
3. *Minitaeva A.* Coefficient estimation of simultaneous equations in the case of stochasticity of exogenous variables // *Journal of Physics Conference Series*. 2023. Vol. 2543. DOI 10.1088/1742-6596/2543/1/012005
4. *Zimin V., Krylov A., Kuvyrkin G., Shakhverdov A.* Mathematical Modeling of Deployment Dynamics of Large Transformable Space Structures // *Advanced Structured Materials*. 2023. Vol. 176. P. 259–277.
5. *Chatfield C.* *The Analysis of Time Series: An Introduction*. (6 ed.) New York: Chapman and Hall/CRC, 2003. 352 p.
6. *Minitaeva A.* Unbiased Estimation Method for Coefficients of Simultaneous Equations with Stochasticity of Exogenous Variables: The paper establishes the unbiasedness and efficiency of estimates obtained through proposed Method of Factorization of Simultaneous Equations. // In *2024 6th International Conference on Big-data Service and Intelligent Computation (BDSIC 2024)*, May 29–31, 2024, Hong Kong, Hong Kong. ACM, New York, NY, USA, 5 pages. 2024. DOI 10.1145/3686540.3686554
7. *Kotz S., Read C., Balakrishnan N., Vidakovic B.* *Encyclopedia of Statistical Sciences*. (2 ed.) New York: John Wiley & Sons, 2005.
8. *Suyatinov S., Buldakova T., Vishnevskaya Yu.* Identification of Situations Based on Synergetic Model // *2021 3rd International Conference on Control Systems, Mathematical Modeling, Automation and Energy Efficiency (SUMMA)*, Lipetsk: IEEE, 2021. DOI 10.1109/SUMMA53307.2021.9632207
9. *Syuzev V., Smirnova E., Proletarsky A.* Algorithms of multidimensional random process simulation // *Computer Optics*. 2021. Vol. 45. № 4. P. 627–637. DOI 10.18287/2412-6179-CO-770
10. *Minitaeva A.* Multimodel Approach to Forecasting Nonlinear Nonstationary Processes in Optimal Control Problems // *2024 6th International Youth Con-*

- ference on Radio Electronics, Electrical and Power Engineering (REEPE), Moscow: IEEE, 2024. DOI 10.1109/REEPE60449.2024.10479796
11. *Salam R., Basuki P.* Simultaneous Equation: The Case of Inflation & Rupiah Exchange Rate in Indonesia 2001-2022 // International journal of multidisciplinary research and analysis. 2023. Vol. 6. № 12. P. 5531–5538. DOI 10.47191/ijmra/v6-i12-12
 12. *Warrener C., Silva F., Guimarães L.* Financial system and economic development: a study for emerging countries through a system of simultaneous equations // Internext. 2023. Vol. 18. № 3. DOI 10.18568/internext.v18i3.741
 13. *Yilanci V., Çütücü İ., Araci S.* The causality relationship between trade and environment in G7 countries: evidence from dynamic symmetric and asymmetric bootstrapped causality tests // Mathematics. 2022. Vol. 10. № 15. DOI 10.3390/math10152553
 14. *Liu Y., Jin M., Zhao S., Qi G.* How do volatile and non-volatile energy factors affect energy OFDI? evidence from simultaneous equation model // Front. Energy Res. 2024. Vol. 11. № 1302374. P. 627–637. DOI 10.3389/fenrg.2023.1302374
 15. *Kheireddine H., Lacombe I., Jarboui A.* The moderating effect of environmental performance on the relationship between sustainability assurance quality and firm value: a simultaneous equations approach // Benchmarking An International Journal. 2024. DOI 10.1108/BIJ-06-2022-0389
 16. *Garba M., Akanni S., Banjoko A., Oladele T.* The nexus between foreign aid and foreign direct investment in Nigeria: simultaneous equations approach // Malaysian Journal of Computing. 2023. Vol. 8. № 2. P. 1620–1638. DOI 10.24191/mjoc.v8i2.16635
 17. *Egger P., Prucha I.* Refined GMM estimators for simultaneous equations models with network interactions // Empirical Economics. 2023. Vol. 64. № 6. P. 2535–2542. DOI 10.1007/s00181-023-02408-8
 18. *Hadjiantoni S., Kontoghiorghe E.* A recursive three-stage least squares method for large-scale systems of simultaneous equations // Linear Algebra and its Applications. 2018. Vol. 536. P. 210–227. DOI 10.1016/j.laa.2017.08.019
 19. *Hernández-Sanjaimé R., González M., López-Espín J.* Multilevel simultaneous equation model: A novel specification and estimation approach // Journal of Computational and Applied Mathematics. 2020. Vol. 366. № 112378. DOI 10.1016/j.cam.2019.112378

A POWER-LAW ADJUSTABLE COEFFICIENT DYNAMIC PRICING MODEL CONSIDERING SHORTAGES

A. Kitaeva¹, Y. Cao¹

¹*Tomsk State University, Tomsk, Russian Federation*

We consider a single product with a fixed lifetime and price sensitive intensity of a Poisson customers flow. A model with a power-law adjustable coefficient of a dynamic retail price control through the intensity of the demand is considered accommodating scenario of shortages. We conduct the diffusion approximation of the demand process and a linear approximation of the intensity-of-price dependence to find the probabilistic characteristics of the selling process, and the expected revenue. To validate the accuracy and effectiveness of the proposed model, simulations and numerical analysis are performed. **Keywords:** *dynamic pricing, price sensitive demand, Poisson demand, diffusion approximation, shortages.*

1. Introduction and problem statement

In [1], we studied a zero-ending inventory price control model with a weight function depending on a power-law coefficient without considering cost of shortages. Here, we expand our analysis: if the lot is sold out before the end of the selling period T , resulting in shortages, we consider a case of lost sales.

The demand is modeled as a compound Poisson process with intensity $\lambda(c)$, where $c = c(t)$ is a dynamic retail price per unit, the orders are independent identically distributed continuous random variables with the first and second moments a_1 and a_2 respectively.

We are going to consider the following model of expiration date-based price control

$$a_1 \lambda(c(t)) = \frac{Q(t)}{T(1-t/T)^\gamma},$$

where coefficient $\gamma \neq 1$.

Consistent with our earlier papers, we utilize the diffusion approximation for analyzing the stock level process. It follows that the stock level process satisfies the following equation:

$$dQ(t) = -\frac{Q(t)}{T(1-t/T)^\gamma} dt + \sqrt{\frac{a_2}{a_1} \frac{Q(t)}{T(1-t/T)^\gamma}} dw(t). \quad (1)$$

where $w(\cdot)$ is the Wiener process.

2. Probabilistic characters of the stock level process

Expectation and variance of process $Q(\cdot)$

$$E \{Q(t)\} = \bar{Q}(t) = Q_0\alpha(t/T),$$

$$Var \{Q(t)\} = V_Q(t) = \frac{a_2 Q_0}{a_1} \alpha(t/T) (1 - \alpha(t/T)),$$

where $\alpha(t/T) = \exp \left\{ \frac{(1 - t/T)^{1-\gamma} - 1}{1 - \gamma} \right\}$.

Thus, the lot is completely sold during its lifetime almost surely for $\gamma > 1$, but shortages are possible.

Consider the Laplace transform of the probability density function (PDF) of $Q(\cdot)$ $\Phi(p, t) = E \{ \exp(-pQ(t)) \}$.

By applying Ito's formula, we get from (1)

$$\begin{aligned} d(e^{-pQ}) &= \left(\frac{Qpe^{-pQ}}{T(1-t/T)^\gamma} + \frac{a_2 Qp^2 e^{-pQ}}{2a_1 T(1-t/T)^\gamma} \right) dt - \\ &\quad - pe^{-pQ} \sqrt{\frac{a_2 Q}{a_1 T(1-t/T)^\gamma}} dw(t). \end{aligned} \quad (2)$$

After averaging (2) with respect to $Q = Q(\cdot)$, and solving by the method of characteristic, we get

$$\Phi(p, t) = \varphi \left(\frac{p}{p + \beta} \alpha(t/T) \right),$$

where $\varphi(\cdot)$ is an unknown function and parameter $\beta = 2a_1/a_2$.

Finally, we get

$$\Phi(p, t) = \exp \left(- \frac{Q_0 \beta p \alpha(t/T)}{p + \beta - p \alpha(t/T)} \right).$$

Using inverse Laplace transform, we obtain PDF of $Q(\cdot)$

$$\begin{aligned} f(q, t) &= \exp \left(- \frac{\beta \alpha Q_0}{1 - \alpha} \right) \left\{ \delta(q) + \exp \left(- \frac{\beta Q_0}{1 - \alpha} \right) \times \right. \\ &\quad \left. \times \sqrt{\frac{\beta^2 \alpha Q_0}{(1 - \alpha)^2 q}} I_1 \left(2 \sqrt{\frac{\beta^2 \alpha Q_0 q}{(1 - \alpha)^2}} \right) \right\}, \end{aligned} \quad (3)$$

where $\alpha = \alpha(t/T)$, $I_1(\cdot)$ is the modified Bessel function of the first kind and first order, $\delta(\cdot)$ is the Dirac delta function.

From (3), it follows that the cumulative distribution function of the length of time τ it takes to sell lot Q_0

$$F_\tau(t) = \exp\left(-\frac{\beta Q_0 \alpha(t/T)}{1 - \alpha(t/T)}\right). \tag{4}$$

Figure 1 shows the plot of $F_\tau(t)$ for $\frac{Q_0}{a_1} = 100, \frac{a_2}{a_1^2} = \frac{4}{3}, \gamma = 1.5, 2, 3$.

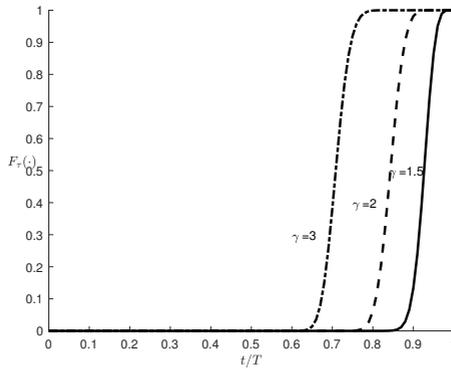


Figure 1. $F_\tau(\cdot)$ dependence on t/T for $\gamma = 1.5, 2, 3$

Thus, for $\gamma > 1$, the shortages can happen in practice. Let us denote $T_1 = Tx_0$, where x_0 is the smallest value that satisfies inequality $\alpha(x_0) < a_1/Q_0$; $P(\tau \leq T_1) = 1$ can be high in practice for some purchases' distributions. We compare time points of shortages occurrence in the cases of $\alpha(x_0) < a_1/Q_0 = 0.01$, $\alpha(x_0) < a_1/Q_0 = 0.001$, and $\alpha(x_0) \approx 0$ in MATLAB package. The results are given in Table 1.

Table 1

Values of x_0 for $\gamma = 1.5, 2, 3, 4, 5$

γ	1.5	2	3	4	5
$x_0 \alpha(x_0) < 0.01$	0.6922	0.7226	0.6276	0.5478	0.4868
$x_0 \alpha(x_0) < 0.001$	0.8854	0.8307	0.7091	0.6165	0.5464
$x_0 \alpha(x_0) \approx 0$	0.9711	0.9078	0.7814	0.6783	0.5966

The expectation of the selling duration

$$E\{\tau\} = \int_0^{T_1} (1 - F_\tau(t))dt = T_1 \left(1 - e^{\beta Q_0} \int_0^{1-\alpha(x_0)} \frac{\exp(-\beta Q_0/x) ((1-\gamma) \ln(1-x) + 1)^{\frac{\gamma}{1-\gamma}}}{1-x} dx \right).$$

Thus, $E\{\tau\} = T_1 \left(1 + O\left(\frac{1}{\beta Q_0}\right) \right)$ for $\beta Q_0 \gg 1$.

3. The expected revenue and its optimization

Let us consider linear approximation of the intensity-of-price dependence

$$\lambda(c) = \lambda_0 - \lambda_1 \frac{c(t) - c_0}{c_0},$$

where c_0 is a stationary (basic) price corresponding stationary intensity λ_0 and parameter $\lambda_1 > 0$ characterizes the sensitivity of $\lambda(\cdot)$ to the relative price's deviations from the stationary price. Linear dependence of the customers' flow intensity on the price is common in the literature, for example, in [2]; the demand rate is supposed to be a linear function of the price.

It follows

$$c(t) = c_0 \left(1 + \frac{\lambda_0}{\lambda_1} - \frac{Q(t)}{a_1 \lambda_1 T (1 - t/T)^\gamma} \right).$$

We assume that the demand intensity during the shortages is the mean intensity corresponding to the stationary price c_0 and the cost of shortages

$$\bar{S}_l = -a_1 c_0 \lambda_0 (T - T_1).$$

Thus, we can write the expected revenue over the whole cycle considering shortages as follow

$$\begin{aligned} \bar{S}_{\gamma>1} &= \int_0^{T_1} E\{c(t)\lambda(t)\} dt + \bar{S}_l = \frac{c_0 Q_0}{\lambda_1} \left\{ (\lambda_0 + \lambda_1) - \frac{1}{a_1 T} \left(Q_0 - \frac{a_2}{a_1} \right) \times \right. \\ &\quad \left. \times \int_{1-x_0}^1 \left(\alpha^2(1-z) + \frac{a_2}{a_1 Q_0 - a_2} \alpha(1-z) \right) \frac{dz}{z^{2\gamma}} \right\} - a_1 c_0 \lambda_0 (T - T_1). \end{aligned} \tag{5}$$

4. Numerical results

Weighted expected revenue $\frac{\bar{S}_{\gamma>1}}{a_1c_0}$ is a concave with respect to γ . It depends on four dimensionless values $\frac{Q_0}{a_1}, \lambda_0T, \lambda_1T, \frac{a_2}{a_1^2}$. Extensive numerical results are carried out for different sets of the inventory system parameters.

On Figure 2 the results of simulation of weighted revenue $\frac{\bar{S}_{\gamma>1}}{a_1c_0}$ with respect to γ are presented. For generating a nonhomogeneous Poisson process, the thinning algorithm is applied. The number of iterations for simulation of revenue is 1000 and we take mean values for each γ ; $\frac{Q_0}{a_1} = 100, \lambda_0T = 400, \lambda_1T = 100, \frac{a_2}{a_1^2} = \frac{4}{3}$. On Figure 2 black curve represents theoretical result (5) and red points are the simulation results.

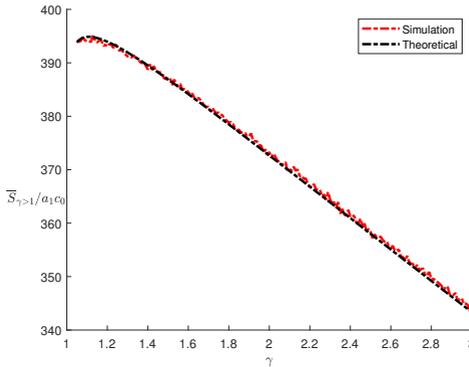


Figure 2. Simulation of weighted revenue $\frac{\bar{S}_{\gamma>1}}{a_1c_0}$

Figure 3 depicts weighted revenue $\frac{\bar{S}_{\gamma>1}}{a_1c_0}$ dependence on $\frac{Q_0}{a_1}$ for different values of γ .

5. Conclusion

The proposed more complicated dynamic price control model not only addresses the issue of shortages, but also allows us to optimize the sales process using the introduced power-law adjustable coefficient. Moreover, the coefficient can help us to solve the problem of fitting the demand rate in

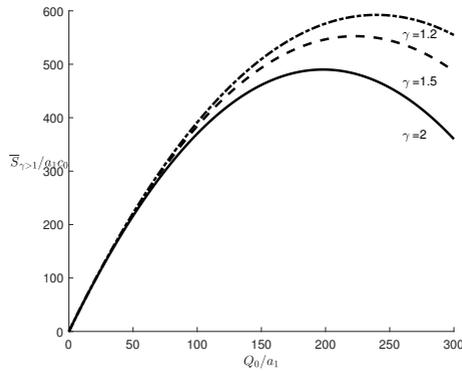


Figure 3. $\frac{\bar{S}_{\gamma>1}}{a_1c_0}$ dependence on $\frac{Q_0}{a_1}$ for $\gamma = 1.2, 1.5, 2$

real-life situations, which is the most challenging problem in implementing dynamic pricing in practice, see [3].

In this paper, we only considered the situation of shortages, it would be interesting to solve the revenue optimization problem considering leftovers with positive and negative values for $\gamma < 1$.

REFERENCES

1. *Kitaeva A., Stepanova N., Zhukovskiy O.* Profit Optimization with a Power-Law Adjustable Coefficient for Zero Ending Inventories Dynamic Pricing Model, Stochastic Demand, and Fixed Lifetime Product // IFAC-PapersOnLine. 2022. Vol. 55. № 10. P. 1793–1797.
2. *Li S., Zhang J., Tang W.* Joint dynamic pricing and inventory control policy for a stochastic inventory system with perishable products // International Journal of Production Research. 2015. Vol. 53. № 10. P. 2937–2950.
3. *Dolgui A., Proth J.* Pricing strategies and models // Annual Reviews in Control. 2010. Vol. 34. № 1. P. 101–110.

Kitaeva Anna — Applied Mathematics, Professor, Institute of Applied Mathematics and Computer Science. E-mail: *kit1157@yandex.ru*

Cao Yu — Applied Mathematics, PhD, Institute of Applied Mathematics and Computer Science. E-mail: *ch.cy@stud.tsu.ru*

LOCAL LIMIT THEOREMS FOR THE NONCRITICAL GALTON-WATSON BRANCHING SYSTEMS

A. A. Imomov^{1,2}, M. S. Murtazaev², E. E. Tukhtaev¹

¹*Karshi State University, Karshi, Uzbekistan*

²*V.I.Romanovskiy Institute of Mathematics, Tashkent, Uzbekistan*

Consider the noncritical Galton-Watson stochastic branching system. We are interested in the asymptotic expansion of local probabilities of the system. Our discussion relies heavily on our recent results, which explicitly calculate the famous constant from the theory of subcritical systems, announced in 1938 by A.Kolmogorov. We establish the rate of convergence to the Kolmogorov constant. This, in turn, contributes to establishing a local limit theorem with an estimate of the remainder term for noncritical Galton-Watson branching systems. **Keywords:** *Galton-Watson branching system, Markov chain, extinction time, Kolmogorov constant, invariant distribution, limit theorems, convergence rate.*

1. Introduction

Let $Z(n)$ be a population size at the moment $n \in \mathbb{N}_0$ in the Galton-Watson branching (GWB) system with branching rates $\{p_k, k \in \mathbb{N}_0\}$, where $\mathbb{N}_0 = \{0\} \cup \mathbb{N}$ and $\mathbb{N} = \{1, 2, \dots\}$. Each individual in the system lives a unit length life time and then gives $k \in \mathbb{N}_0$ descendants with probability p_k . This is a reducible, homogeneous-discrete-time Markov chain with a state space consisting of two classes: $\mathcal{S}_0 = \{0\} \cup \mathcal{S}$, where $\{0\}$ is absorbing state, and $\mathcal{S} \subset \mathbb{N}$ is the class of possible essential communicating states. We assume throughout this paper that $p_0 > 0$ and $p_1 > 0$ which implies the Schröder case, i.e. $p_0 + p_1 > 0$. Define the offspring generating function (GF)

$$f(s) = \sum_{k \in \mathcal{S}_0} p_k s^k \quad \text{for } s \in [0, 1).$$

Let q be an extinction probability of the system initiated by the single founder-individual and $m := \sum_{k \in \mathcal{S}} k p_k < \infty$. Then $m = f'(1-)$ is the average number of direct descendants of one individual in one-step generation. The chain $\{Z(n)\}$ is called subcritical, critical or supercritical if $m < 1$, $m = 1$ and $m > 1$ respectively. The extinction probability $q < 1$ if $m > 1$ and $q = 1$ if $m \leq 1$. Moreover, it is smallest nonnegative root of the fixed-point equation $f(s) = \sum_{k \in \mathcal{S}_0} p_k s^k = s$ on the domain of $\{s : s \in [0, 1]\}$.

We are interested in the noncritical case.

Let $P_i \{*\} := P \{* \mid Z(0) = i\}$ and consider n -step transition probabilities $P_{ij}(n) = P_i \{Z(n) = j\}$. Using the branching property and the Kolmogorov-Chapman equation, we observe that the probability GF

$$E_i s^{Z(n)} := \sum_{j \in S_0} P_{ij}(n) s^j = [f_n(s)]^i,$$

for $s \in [0, 1)$, where the GF $f_n(s)$ is the GF of probabilities $p_k(n) := P_{1k}(n)$, i.e.

$$f_n(s) = \sum_{k \in S_0} p_k(n) s^k.$$

It is known that $f_n(s)$ is an n -fold iteration of $f(s)$; see [2, Ch.I].

In what follows, we will write E and P instead of E_1 and P_1 respectively.

Let $\mathcal{H} := \min \{n \in \mathbb{N} : Z(n) = 0\}$ be an extinction time of the system initiated by a single founder. Put $R_n(s) := q - f_n(s)$. Then we see that

$$Q(n) := R_n(0) = P \{n < \mathcal{H} < \infty\}$$

is a bridled-survival probability of the system. For the subcritical case $P \{\mathcal{H} < \infty\} = 1$ and, hence $Q(n) = P \{Z(n) > 0\}$. In this case, [9] proved that if $f''(1-) < \infty$, then $Q(n)$ and $EZ(n) = m^n$ vary (decrease) asymptotically equivalently, i.e., the following asymptotic representation holds:

$$Q(n) = \mathcal{K} m^n (1 + o(1)) \quad \text{as } n \rightarrow \infty, \tag{1}$$

where \mathcal{K} is an equivalence ratio called the Kolmogorov constant, but it does not have an explicit form here. The same situation is obstinate in the continuous-time case. In this regard, Zolotarev [14] regretted the absence of an explicit expression for this constant. This drawback undoubtedly, makes it difficult to complete many limit theorems in which the Kolmogorov constant is to be attended. In [2] was restored Kolmogorov's assertion under a much weaker condition $\sum_{k \in S} p_k k \ln k < \infty$. But even so, the equivalence ratio remained explicitly not calculated.

2. Main Results

Recently, in our work [6], the constant in (1) was explicitly calculated and generalized to the noncritical case under the condition

$$\left[m \neq 1 \quad \text{and} \quad f''(1-) < \infty \quad \text{for } m < 1 \right], \tag{K}$$

which we will further call the Kolmogorov condition. In our other work [4], we estimated the rate of convergence to the extended Kolmogorov constant and found the asymptotic expansion for $R_n(s)$ for all $s \in U_q[0, 1)$, where

$$U_q[0, 1) = \left\{ [0, q) \cup (q, 1) \right\}$$

is a unit interval with a punctured point q . Below we recall this result.

Lemma 1. Let Kolmogorov condition [K] holds. Then

$$\sup \left\{ \left| \frac{\beta^n}{R_n(s)} - \frac{1}{\mathcal{A}_q(s)} - B_2\beta^n \right| : s \in U_q[0, 1) \right\} = o(\beta^n)$$

as $n \rightarrow \infty$, where

$$\frac{1}{\mathcal{A}_q(s)} = \frac{1}{q - s} + \gamma_q,$$

at that $B_2 = f''(q)/(2\beta)$, $\beta = f'(q)$ and

$$\gamma_q = \frac{f''(q)}{2\beta(1 - \beta)}.$$

In this report, using Lemma 1 and results from above-mentioned papers, we state and prove the following local limit theorem.

Theorem 1. Let Kolmogorov condition [K] holds. Then

$$p_j(n) = \pi_j \frac{1}{q^2} \mathcal{K}_q^2 \beta^n \cdot \left(1 + 2\gamma_q \mathcal{K}_q \beta^n (1 + o(1)) \right) \quad \text{as } n \rightarrow \infty$$

for all $j \in \mathcal{S}$, where the GF $\mathcal{P}(s) := \sum_{j \in \mathcal{S}} \pi_j s^j$ is form of

$$\mathcal{P}(s) = \frac{q^2}{\mathcal{K}_q^2} \left(\mathcal{K}_q - \mathcal{A}_q(s) \right),$$

therein $\mathcal{K}_q = \mathcal{A}_q(0)$ is Kolmogorov constant.

3. Conclusion

The paper considers the discrete-time noncritical Galton-Watson branching system. Our research is continuous of our recent results that exactly calculate the famous Kolmogorov constant. Our main goal is to establish the convergence rate to the Kolmogorov constant, which allows us to estimate the remainder term for noncritical Galton-Watson branching systems.

REFERENCES

1. *Athreya K. B. and Ney P. E.* Branching processes. Springer, New York, 1972.
2. *Nagaev A. V and Badalbaev I. S.*. A refinement of certain theorems on branching random process. *Litovskiy Mat. Sbornik*, 1967. Vol. 7(1). P. 129–136. (Russian)
3. *Imomov A. A. and Murtazaev M.* Renewed Limit Theorems for Non-critical Galton-Watson Branching Systems. *J Theor. Probab.*, 2024, <https://doi.org/10.1007/s10959-024-01330-8>.
4. *Imomov A. A. and Murtazaev M.*. On the Kolmogorov constant explicit form in the theory of Discrete-time Stochastic Branching Systems. *J. Appl. Probab.* 2024, <https://doi.org/10.1017/jpr.2023.85>.
5. *Kolmogorov A. N.*. K resheniyu odnoy biologicheskoy zadachi. *Reports of SRI Math. and Mech. at Tomsk Univ.* 1938. Vol. 2. P. 7–12. (Russian)
6. *Zolotarev V. M.* (1957). More exact statements of several theorems in the theory of branching processes. *Theory Prob. Appl.* Vol. 2. P. 245–253.

Azam A. Imomov — Doctor of Sciences, Head of Department, Mathematics and Computer Sciences faculty. E-mail: imomov-azam@mail.ru

Misliddin Murtazaev — post-doctoral student, V.I.Romanovskiy Institute of Mathematics. E-mail: misliddin1991@mail.ru

Erkin Tukhtaev — post-doctoral student, Karshi State University. E-mail: tukhtaev-erkin@mail.ru

COMPARATIVE ANALYSIS OF UNKNOWN PARAMETER ESTIMATION OF THE EXPONENTIAL DISTRIBUTION WITH RIGHT-CENSORED DATA IN INCOMPLETE STATISTICAL MODELS

N. Nurmukhamedova, M. Berdimuradov

*National University of Uzbekistan named after Mirzo Ulugbek, Tashkent,
Uzbekistan*

In this paper, we analyze parameter estimation for Exponential and Gamma distributions under Type I right-censoring using Maximum Likelihood Estimation. We apply both the Nelder-Mead simplex and Expectation-Maximization algorithms to estimate the unknown parameters for both distributions. Numerical simulations are conducted using Python and MATLAB, demonstrating the effectiveness of these algorithms. **Keywords:** *exponential distribution, gamma distribution, Nelder-Mead simplex algorithm, expectation-maximization, maximum likelihood estimation, right-censoring.*

Introduction

When estimating unknown parameters using the Maximum Likelihood Estimation method, numerical optimization techniques often replace analytical methods due to the complexity of calculating integrals and derivatives of the cumulative distribution function. In this paper, we focus on parameter estimation for both the Exponential and Gamma distributions under Type I right-censoring, a common challenge in fields like reliability and life data analysis. Specifically, we apply the Maximum Likelihood Estimation approach, combined with both the Nelder-Mead simplex and Expectation-Maximization algorithms for estimating the unknown parameters of both distributions. Numerical simulations are conducted using Python and MATLAB to evaluate the performance of these algorithms.

Our work builds on existing studies, such as those by Luís Andrade Ferreira and José Luís Silva, who have demonstrated the robustness of the Expectation-Maximization algorithm in handling censored data. We compare the estimated results with their initial guesses, examining the effectiveness of these algorithms for censored and fully observed datasets. Our findings are presented in both tabular and graphical formats.

1. Type I right-censoring

Let X random variable (r.v.) with distribution function (d.f.) $F(X, \theta)$, where θ is unknown parameter $\theta \in \Theta \subseteq \mathbb{R}^1$. We observed the sample $C^n = \{(Z_i, \delta_i), 1 \leq i \leq n\}$, where $Z_i = \min(X_i, T_i)$ and $\delta_i = I(X_i \leq T)$. Interpretations for T_i are considered: T_i - **constant** and T_i - **r.v.**.

Let X have a probability density function $f(X; \theta)$, then the joint probability density function of the r.v. $Z_i = \min(X_i, T_i)$ and $\delta_i = I(X_i \leq T_i)$ can be written as follows:

$$L(Z_i; \theta) = \begin{cases} f(Z_i; \theta), & \text{if } \delta_i = 1, \\ 1 - F(Z_i; \theta), & \text{if } \delta_i = 0. \end{cases}$$

Accordingly, the likelihood function is:

$$L(Z^n; \theta) = \prod_{i=1}^n \left[f(Z_i; \theta)^{\delta_i} \cdot (1 - F(Z_i; \theta))^{1-\delta_i} \right]. \quad (1)$$

Where $Z^n = ((Z_1, \delta_1), \dots, (Z_n, \delta_n))$. Loglikelihood function is

$$\log L(Z^n; \theta) = \sum_{i=1}^n [\delta_i \log f(Z_i; \theta) + (1 - \delta_i) \log(1 - F(Z_i; \theta))]. \quad (2)$$

If the likelihood function is differentiable with respect to θ , the MLE $\hat{\theta}$ satisfies the equation:

$$\left. \frac{\partial \log L(Z^n; \theta)}{\partial \theta} \right|_{\theta=\hat{\theta}} = 0. \quad (3)$$

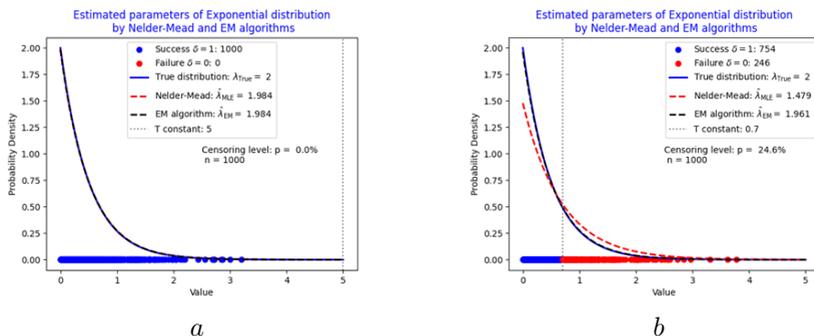
In this work, various estimation methods have been considered. In particular, equation (3) is solved using numerical methods such as the EM algorithm and the Nelder-Mead simplex algorithm. The obtained estimates are compared, and graphs are drawn (see Table 1 and 2, and Figures A and B below). These estimates are based on the Exponential distribution, and since the Exponential distribution is a special case of the Gamma distribution, the unknown parameters of the Gamma distribution have also been estimated using these algorithms based on their initial values, as shown in Table 2 and Figures B.

Table 1
 Estimation of an unknown parameter of the Exponential distribution using two algorithms

Exponential Distribution, Initial Guess $\lambda = 2$; T_i - constant							
n Sample data	T_i	$p\%$	MLE by EM algorithm	Difference $\Delta = \lambda - \hat{\lambda}_{EM} $	MLE by Nelder-Mead simplex algorithm	Difference $\Delta = \lambda - \hat{\lambda}_{NM} $	
100	5	0	1.962	0.038	1.962	0.038	
100	0.7	25	1.931	0.069	1.521	0.479	
100	0.35	50	1.828	0.172	1.103	0.897	
500	5	0	1.964	0.036	1.964	0.036	
500	0.7	25	1.944	0.056	1.533	0.467	
500	0.35	50	1.839	0.161	1.105	0.895	
1000	5	0	1.988	0.012	1.988	0.012	
1000	0.7	25	1.961	0.039	1.536	0.464	
1000	0.35	50	1.902	0.098	1.113	0.887	

*Note: The sample data is simulated randomly from Exponential distribution. Therefore, if the code is executed, it may display different results compared to those presented in Table 1.

Figures A



*Note: Figures a and b show the estimated unknown parameter of the Exponential distribution by the Nelder-Mead and EM algorithms with **complete data** and **censored data**, respectively. In both graphs, the results of the algorithms are represented by the fitted red dashed and black dashed lines, and are compared to the true blue smooth line.

Table 2
 Estimation of unknown parameters of the Gamma distribution using two algorithms

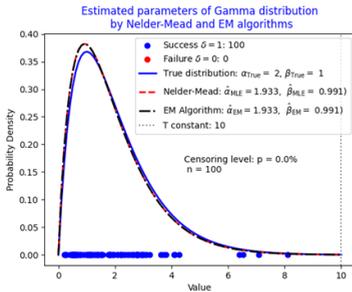
Gamma Distribution, Initial Guess $\alpha = 2, \beta = 1; T_i$ - constant

n Sample data	T_i	$p\%$	MLE by EM algorithm	Difference $\Delta_1 = \alpha - \hat{\alpha}_{EM} $ $\Delta_2 = \beta - \hat{\beta}_{EM} $	MLE by Nelder-Mead simplex algorithm	Difference $\Delta_1 = \alpha - \hat{\alpha}_{NM} $ $\Delta_2 = \beta - \hat{\beta}_{NM} $
100	5	0	$\alpha = 1.933, \beta = 0.991$	0.067, 0.009	$\alpha = 1.933, \beta = 0.991$	0.067, 0.009
100	0.7	25	$\alpha = 2.157, \beta = 0.895$	0.157, 0.105	$\alpha = 1.499, \beta = 1.634$	0.501, 0.634
100	0.35	50	$\alpha = 2.163, \beta = 0.882$	0.163, 0.118	$\alpha = 1.135, \beta = 2.645$	0.865, 1.645
500	5	0	$\alpha = 2.061, \beta = 0.994$	0.061, 0.006	$\alpha = 2.061, \beta = 0.994$	0.061, 0.006
500	0.7	25	$\alpha = 2.112, \beta = 0.904$	0.112, 0.096	$\alpha = 1.595, \beta = 1.618$	0.405, 0.618
500	0.35	50	$\alpha = 2.128, \beta = 0.899$	0.128, 0.101	$\alpha = 1.142, \beta = 2.324$	0.858, 1.324
1000	5	0	$\alpha = 2.019, \beta = 1.005$	0.019, 0.005	$\alpha = 2.019, \beta = 1.005$	0.019, 0.005
1000	0.7	25	$\alpha = 2.119, \beta = 0.913$	0.119, 0.087	$\alpha = 1.623, \beta = 1.527$	0.377, 0.527
1000	0.35	50	$\alpha = 2.142, \beta = 0.918$	0.142, 0.082	$\alpha = 1.201, \beta = 2.197$	0.799, 1.197

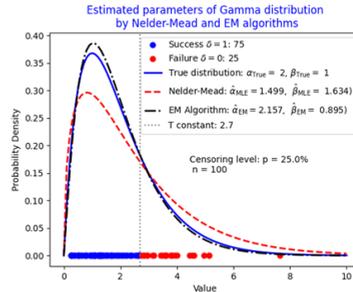
*Note: The sample data is simulated randomly from Gamma distribution. Therefore, if the code is executed, it may display different results compared to those presented in

Table 2.

Figures B



c



d

*Note: Figures c and d show the estimated unknown parameters of the Gamma distribution by the Nelder-Mead and EM algorithms with **complete data** and **censored data**, respectively. In both graphs, the results of the algorithms are represented by the fitted red dashed and black dashed lines, and are compared to the true blue smooth line.

Now we consider T_i as a r.v. denote it as $T_i = Y_i$.

Let the d.f. $G(x)$ and probability density function $g(x)$ of the r.v. Y not depend on the unknown parameter. For instance, $N(0, 1)$, $\text{Exp}(1)$, and $R[0, 1]$.

Consider the sample $C^n = \{(Z_i, \delta_i), 1 \leq i \leq n\}$, where $Z_i = \min(X_i, Y_i)$ and $\delta_i = I(X_i \leq Y_i)$; Accordingly, the likelihood function is:

$$\begin{aligned} L(Z^{(n)}; \theta) &= \prod_{i=1}^n \left\{ (f(X_i; \theta)(1 - G(Y_i))^{\delta_i} \cdot ((1 - F(X_i; \theta)) \cdot g(Y_i))^{(1-\delta_i)}) \right\} \\ &= \prod_{i=1}^n \left\{ f(X_i; \theta)^{\delta_i} (1 - F(X_i; \theta))^{(1-\delta_i)} \right\} \cdot \prod_{i=1}^n \left\{ (1 - G(Y_i))^{\delta_i} g(Y_i)^{(1-\delta_i)} \right\}. \quad (4) \end{aligned}$$

Loglikelihood function is:

$$\begin{aligned} \log L(Z^n; \theta) &= \sum_{i=1}^n (\delta_i \log f(X_i; \theta) + (1 - \delta_i) \log(1 - F(X_i; \theta))) \\ &\quad + \sum_{i=1}^n (\delta_i \log(1 - G(Y_i)) + (1 - \delta_i) \log g(Y_i)). \quad (5) \end{aligned}$$

Since d.f. of r.v. Y does not depend on the unknown parameter θ , the maximum likelihood estimation equations are the same as in the case when T is constant.

In this study, we explore different estimation methods, with a focus on solving equation (3) using numerical techniques. The resulting estimates are compared, and graphs are made about their application across different distribution families (see Table 3, 4 and Figures C, D below).

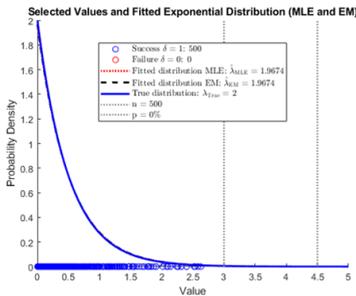
Table 3

Estimation of an unknown parameter of the Exponential distribution using two algorithms

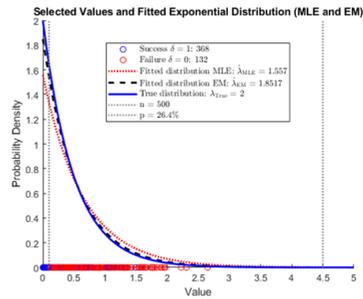
Exponential Distribution, Initial Guess $\lambda = 2$; $Y_i \sim \text{Exp}(1), 1 \leq i \leq n$.					
n Sample data	$p\%$	MLE by EM algorithm	Difference $\Delta = \lambda - \hat{\lambda}_{EM} $	MLE by Nelder-Mead simplex algorithm	Difference $\Delta = \lambda - \hat{\lambda}_{NMF} $
100	0	1.9205	0.0795	1.9205	0.0795
100	25	1.8231	0.1769	1.5416	0.4584
100	50	1.7943	0.2057	1.1056	0.8944
500	0	1.9674	0.0326	1.9674	0.0326
500	25	1.8517	0.1483	1.5671	0.4329
500	50	1.8062	0.1938	1.1223	0.8777
1000	0	1.9994	0.0006	1.9994	0.0006
1000	25	1.8618	0.1382	1.5831	0.4169
1000	50	1.8173	0.1827	1.2134	0.7866

**Note: The sample data is simulated randomly from Exponential distribution. Therefore, if the code is executed, it may display different results compared to those presented in Table 3.*

Figures C



e



f

**Note: Figures e and f show the estimated unknown parameter of the Exponential distribution by the Nelder-Mead and EM algorithms with complete data and censoring data, respectively. In both graphs, the results of the algorithms are represented by the fitted red dashed and black dashed lines, and are compared to the true blue smooth line.*

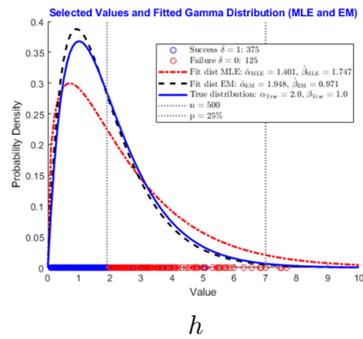
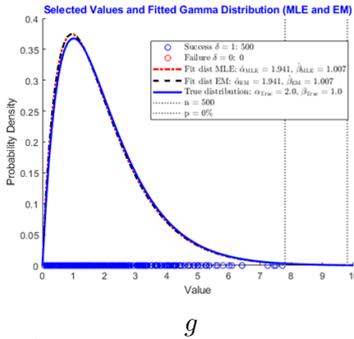
Table 4

Estimation of unknown parameters of the Gamma distribution using two algorithms

Gamma Distribution, Initial Guess $\alpha = 2, \beta = 1$;						
$Y_i \sim \text{Exp}(1), 1 \leq i \leq n$						
n	$p\%$	MLE by EM algorithm	Difference $\Delta_1 = \alpha - \hat{\alpha}_{EM} $ $\Delta_2 = \beta - \hat{\beta}_{EM} $	MLE by Nelder-Mead simplex algorithm	Difference $\Delta_1 = \alpha - \hat{\alpha}_{NM} $ $\Delta_2 = \beta - \hat{\beta}_{NM} $	
100	0	$\alpha = 1.912, \beta = 0.923$	0.088, 0.077	$\alpha = 1.912, \beta = 0.923$	0.088, 0.077	
100	25	$\alpha = 1.869, \beta = 0.817$	0.131, 0.183	$\alpha = 1.424, \beta = 1.697$	0.576, 0.697	
100	50	$\alpha = 1.724, \beta = 0.809$	0.276, 0.191	$\alpha = 1.026, \beta = 2.847$	0.974, 1.847	
500	0	$\alpha = 1.941, \beta = 1.007$	0.059, 0.007	$\alpha = 1.941, \beta = 1.007$	0.059, 0.007	
500	25	$\alpha = 1.948, \beta = 0.971$	0.052, 0.029	$\alpha = 1.401, \beta = 1.747$	0.599, 0.747	
500	50	$\alpha = 1.889, \beta = 0.902$	0.111, 0.098	$\alpha = 1.104, \beta = 2.753$	0.896, 1.753	
1000	0	$\alpha = 1.959, \beta = 0.995$	0.041, 0.005	$\alpha = 1.959, \beta = 0.995$	0.041, 0.005	
1000	25	$\alpha = 1.892, \beta = 0.901$	0.108, 0.099	$\alpha = 1.432, \beta = 1.962$	0.568, 0.962	
1000	50	$\alpha = 1.824, \beta = 0.848$	0.176, 0.152	$\alpha = 1.116, \beta = 2.514$	0.884, 1.514	

**Note: The sample data is simulated randomly from Exponential and Gamma distributions. Therefore, if the code is executed, it may display different results compared to those presented in Table 4.*

Figures D



Note: Figures g and h show the estimated unknown parameters of the Gamma distribution by the Nelder-Mead and EM algorithms with **complete data and **censoring data**, respectively. In both graphs, the results of the algorithms are represented by the fitted red dashed and black dashed lines, and are compared to the true blue smooth line.*

2. Conclusion

In conclusion, the estimation results for both Exponential and Gamma distributions reveal that the Expectation-Maximization algorithm achieves higher estimation accuracy in censored states, while the Nelder-Mead and

Expectation-Maximization algorithms produce identical results under uncensored full observed states. This outcome is because, in uncensored data, both algorithms essentially solve the same optimization problem where all data points are fully observed, leading to convergence at the same parameter estimate. In the case of censored data, the Expectation-Maximization algorithm iterative handles the incomplete data by maximizing the expected likelihood, which provides a more robust estimate compared to Nelder-Mead, which directly optimizes the likelihood function and might be less effective when handling missing or censored data points. To illustrate these findings, we present a comparison of the results in both a table and a graph obtained from the simulations. The table summarizes the estimated parameters under different conditions, and the graph visually compares the true distribution against the estimated distributions from both algorithms.

REFERENCES

1. *Berdimuradov M.B.* Estimation of unknown parameter of gamma distribution in incomplete models of statistics // ACTA NUUz. 2024. Vol. 1. № 1. P. 43–52.
2. *Ferreira LA, Silva JL.* Parameter estimation for Weibull distribution with right censored data using EM algorithm // Eksploatacja i Niezawodnosc – Maintenance and Reliability. 2017. Vol. 19. № 2. P. 310–315.
3. *Nelder J., Mead R.* A simplex method for function minimization // The Computer Journal. 1965. Vol. 7. № 4. P. 308–313.
4. *Srikanth K. Iyer, Debasis Kundu, S. Rao Jammalamadaka* Analysis of Middle Censored Data with Exponential Lifetime Distributions // Journal of Statistical Planning and Inference. 2008. Vol. 138. № 11 P. 3550–3560.

Nargiza Nurmukhamedova — Doctor of Philosophy in Physical and Mathematical Sciences, Associate Professor, Econometrics and Economic Modeling. E-mail: *rasulova.nargiza@mail.ru*

Mirkamol Berdimuradov — PhD student, Department of Applied mathematics and Intellectual technologies. E-mail: *mirkamol1403@gmail.com*

ON A LIMITING STRUCTURE OF DISCRETE-TIME STOCHASTIC BRANCHING SYSTEMS WITH THE EVENTUALLY AWAITING DEATH MOMENT

A. A. Imomov^{1,2}, Y. Ibrohimova¹

¹*Karshi State University, Karshi, Uzbekistan*

²*V.I.Romanovskiy Institute of Mathematics, Tashkent, Uzbekistan*

We examine the population growth model called Galton-Watson Branching system. We are interested in the limits on the probability of population distribution along the trajectories awaiting its death moment. **Keywords:** *Branching system, Markov chain, generating function, transition probabilities, extinction time, invariant distribution*

1. Introduction

One of an important and descriptive mathematical model describing a population growth is stochastic branching system. The origin of mathematical models of the theory of branching systems is due to the prospect of interest of an estimation of survival probability of family trees of individuals. The simplest branching model was initiated by the English statisticians Henry Watson and Francis Galton in second half of 19th century. Studying statistics on the reproductive rates of English peers, they have offered mathematical model which describes population family growth. Their model is now called the Galton-Watson Branching (GWB) system. By present time numerous branching schemes have been developed depending on a context of considering problems, which are necessary to be solved; see [1], [2], [3], [5], [6], [7] and [13].

Consider the set of natural numbers \mathbb{N} and $\mathbb{N}_0 = \{0\} \cup \mathbb{N}$. Let $Z(n)$ be a population size at the moment n in the GWB system. An evolution of the individuals' population occurs according to the branching rates law $\{p_k, k \in \mathbb{N}_0\}$. It means, that each individual lives a unit length-life time and then produces $k \in \mathbb{N}_0$ descendants with probability p_k . Newborn individuals undergo to evolve under the same scheme. In the paper we consider the system $\{Z(n)\}$ in the Schröder case, i.e. $p_0 > 0$, $p_1 > 0$ and $p_0 + p_1 < 1$. The stochastic system $\{Z(n), n \in \mathbb{N}_0\}$ is a reducible, homogeneous-discrete-time Markov chain with a state space consisting of two classes: $\mathcal{S}_0 = \{0\} \cup \mathcal{S}$, where $\{0\}$ is absorbing state, and $\mathcal{S} \subset \mathbb{N}$ is the class of possible essential

communicating states. Let

$$P_{ij}(n) := \mathbb{P}\{Z(n+k) = j \mid Z(k) = i\} \quad \text{for any } k \in \mathbb{N}_0$$

be the n -step transition probabilities of the chain. Putting $\mathfrak{p}_k(n) := P_{1k}(n)$, and using the Kolmogorov-Chapman equation, we observe that the probability generating function (GF)

$$\sum_{k \in \mathcal{S}_0} P_{ij}(n) s^k = [f(n; s)]^i, \quad (1)$$

where the GF

$$f(n; s) = \sum_{k \in \mathcal{S}_0} \mathfrak{p}_k(n) s^k,$$

is n -fold iteration of the offspring GF

$$f(s) := \sum_{k \in \mathcal{S}_0} p_k s^k.$$

Throughout the paper symbols $\mathbb{P}_i\{\cdot\}$ and $\mathbb{E}_i[\cdot]$ always stand for the probability and mean functions provided that the system is initiated by i individual-founder, i.e. $\mathbb{P}\{Z(0) = i\} = 1$. Therewith we will write $\mathbb{P}\{\cdot\}$ and $\mathbb{E}[\cdot]$ instead of $\mathbb{P}_1\{\cdot\}$ and $\mathbb{E}_1[\cdot]$ respectively.

We assume that $m := \sum_{k \in \mathcal{S}} k p_k < \infty$. Then $m = f'(1-)$ is the mean per-capita offspring number, which regulates the classification of \mathcal{S} . In fact, it follows from (1), that $\mathbb{E}[Z(n)] = m^n$, i.e., the mathematical expectation of $Z(n)$ asymptotically behaves differently depending on the value of the parameter m . So, the state space of the chain $\{Z(n)\}$ is classified as subcritical, critical and supercritical if $m < 1$, $m = 1$ and $m > 1$ respectively. Evidently that $f(n; 0) = \mathfrak{p}_0(n)$ is a vanishing probability of the system at the moment n initiated by a single founder. This probability approaches monotonously to a finite limit q as $n \rightarrow \infty$, which is called the extinction probability of the system, i.e. $\lim_{n \rightarrow \infty} \mathfrak{p}_0(n) = q$. Secondly, this probability is a smallest nonnegative root of the fixed-point equation $f(s) = s$ on the domain of $\{s : s \in [0, 1]\}$, therewith $q = 1$ if $m \leq 1$ and $q < 1$ if $m > 1$. Moreover $f(n; s) \rightarrow q$ as $n \rightarrow \infty$ uniformly in $s \in [0, r]$ for any $r < 1$; see [2, Ch.I, §§1–5]. Let the variable \mathcal{H} be the time of extinction of the direct line of descendants of one individual. Then it can be written as $\mathcal{H} = \min\{n \in \mathbb{N} : Z(n) = 0\}$ provided that the system is initiated by single individual-founder. The event $\{n < \mathcal{H} < k\}$ means that the system survives at moment n , but at the same time expects its death in the next k steps. The classical extinction theorem asserts that $\mathbb{P}_i\{\mathcal{H} < \infty\} = q^i$.

In this report, we observe an asymptotic properties of trajectories of non-critical, $m \neq 1$, GWB system on the event $\{n < \mathcal{H} < k\}$. We use the following conditioned probability measure:

$$\mathbb{P}_i^{\mathcal{H}(n,k)}\{\cdot\} := \mathbb{P}_i\{\cdot | n < \mathcal{H} < k\} \quad \text{and} \quad \mathbb{P}_i^{\mathcal{H}(n)}\{\cdot\} := \mathbb{P}_i\{\cdot | n < \mathcal{H} < \infty\}$$

for any $n, k \in \mathbb{N}$. The first function denotes the probability distribution of the system awaiting its death in the last k steps. The second one is the probability distribution of a system that will survive at the current moment but will eventually die.

2. Invariant measure

In this step we are interested in an invariant properties of the state space \mathcal{S} of GWB system. Recall that the set of nonnegative numbers $\{\mu_k, k \in \mathbb{N}\}$ is called an invariant measure with respect to transition probabilities $P_{ij}(\cdot)$, if it satisfied to the functional equation $\mu_j = \sum_i \mu_i P_{ij}(1)$. If $\sum_j \mu_j < \infty$, then it is called an invariant distribution. The monotone ratio limit property, proved in [2, Ch.I, §7], asserts that

$$\frac{p_j(n)}{p_1(n)} \uparrow \pi_j < \infty \quad \text{as} \quad n \rightarrow \infty, \tag{2}$$

and an appropriate GF $\mathcal{P}(s) := \sum_{j \in \mathcal{S}} \pi_j s^j$ converges for $s \in [0, 1)$ and it satisfies to the following Abel functional equation

$$\mathcal{P}(f(n; s)) = \beta^n \mathcal{P}(s) + \frac{1 - \beta^n}{1 - \beta} \mathcal{P}(p_0) \tag{3}$$

for all $n \in \mathbb{N}$, where $\beta = f'(q)$. The equation (3) expresses the invariance property of the numbers $\{\pi_j, j \in \mathcal{S}\}$ defined in (2) with respect to the transition probabilities $P_{ij}(\cdot)$.

In what follows we consider the non-critical case, i.e. $m \neq 1$.

Remark 1. In [2, Ch.I, §7, Th 2] it was stated that the series $\sum_{j \in \mathcal{S}} \pi_j$ converges if $m < 1$ and diverges if $m > 1$. However, the statement about the divergence of $\sum_{j \in \mathcal{S}} \pi_j$ in the case of $m > 1$ is incorrect. Below we will prove that $\sum_{j \in \mathcal{S}} \pi_j$ converges without any assumption the case of $m > 1$.

Theorem 1. If $m > 1$ then the series $\sum_{j \in \mathcal{S}} \pi_j$ converges.

Proof. We follow discussions of the proof in [2, Ch.I, §7, Th 2] in which erroneous conclusions led to incorrect statements. First, putting $s = 1$ in (3), we obtain that

$$\sum_{j \in \mathcal{S}} \pi_j = \frac{\mathcal{P}(p_0)}{1 - \beta}. \tag{4}$$

At the same time, equation (3) entails

$$\beta^n \pi_j = \sum_{k \in \mathcal{S}} \pi_k P_{kj}(n)$$

for all $n \in \mathbb{N}$ and $j \in \mathcal{S}$. Then, since $\pi_1 = 1$, it follows from (3) that

$$\begin{aligned} \beta^n = \beta^n \pi_1 &= \sum_{k \in \mathcal{S}} \pi_k P_{k1}(n) = \sum_{k \in \mathcal{S}} \pi_k k p_0^{k-1}(n) p_1(n) \\ &\geq \frac{p_1(n)}{p_0(n)} \sum_{k \in \mathcal{S}} \pi_k p_0^k(n) = \frac{p_1(n)}{p_0(n)} \mathcal{P}(p_0(n)). \end{aligned}$$

Therefore

$$\mathcal{P}(p_0(n)) \leq \frac{p_0(n)}{p_1(n)} \beta^n \tag{5}$$

for all $n \in \mathbb{N}$. In particular, $p_j(1) = p_j$, and therefore, combining (4) and (5) we obtain that

$$\sum_{j \in \mathcal{S}} \pi_j \leq \frac{p_0}{p_1} \frac{\beta}{1 - \beta} < \infty.$$

The theorem is proved. ■

Remark 2. Recently, in [4] it was proved that if $f''(1-) < \infty$ for $m < 1$, then the GF $\mathcal{P}(s)$ converges for all $s \in [0, 1) \setminus \{q\}$ and

$$\mathcal{P}(s) = \frac{q^2}{\mathcal{K}_q^2} (\mathcal{K}_q - \mathcal{A}_q(s)), \tag{6}$$

where \mathcal{K}_q is the extended Kolmogorov constant and

$$\frac{1}{\mathcal{A}_q(s)} = \frac{1}{q - s} + \frac{f''(q)}{2\beta(1 - \beta)}.$$

Relation (6) implies that $\sum_{j \in \mathcal{S}} \pi_j < \infty$.

Another invariant distribution can be constructed follows. If $p_1 > 0$ and $q > 0$, then for all $j \in \mathcal{S}$ and $k \in \mathbb{N}$

$$\nu_j(k) := \lim_{n \rightarrow \infty} \mathbb{P}^{\mathcal{H}(n+k)} \{Z(n) = j\} = \frac{q^j - f^j(k; 0)}{\beta^k \sum_{k \in \mathcal{S}} \pi_k q^k} \pi_j \tag{7}$$

is a probability function. For $\nu_j := \nu_j(0)$, the GF $\mathcal{V}(s) := \sum_{j \in \mathcal{S}} \nu_j s^j$ has the form of $\mathcal{V}(s) = \mathcal{P}(qs)/\mathcal{P}(q)$; see [2, Ch.I, §8,§14].

Theorem 2. The set of numbers $\{\nu_j(k), j \in \mathcal{S}\}$ forms an invariant distribution with respect to the probabilities $P_{ij}(\cdot)$, and for an appropriate GF $\mathcal{V}_k(s) := \sum_{j \in \mathcal{S}} \nu_j(k) s^j$ the following relation holds:

$$\beta^k \mathcal{V}_k(s) = \mathcal{V}(s) - \mathcal{V}\left(\frac{\mathbb{P}_0(k)}{q} s\right), \tag{8}$$

where $\mathcal{V}(s) = \sum_{j \in \mathcal{S}} \nu_j s^j$. If $f''(1-) < \infty$ for $m < 1$, then

$$\mathcal{V}(s) = \frac{s}{1 + (1 - s)q\gamma_q} \tag{9}$$

and $\gamma_q = f''(q)/2\beta(1 - \beta)$.

Proof. Consider

$$\mathcal{V}_k(n; s) := \sum_{j \in \mathcal{S}} \mathbb{P}^{\mathcal{H}(n+k)}\{Z(n) = j\} s^j$$

and denoting $f_q(s) := f(qs)/q$, we rewrite

$$\begin{aligned} \mathcal{V}_k(n; s) &= \frac{f(n + k, qs) - f(n + k, 0)}{q - f(n + k, 0)} \\ &= 1 - \mathcal{G}(n; f_q(k, s)) \frac{1 - f_q(n, 0)}{1 - f_q(n + k, 0)}, \end{aligned}$$

where $\mathcal{G}(n; x) = [1 - f_q(n, x)]/[1 - f_q(n, 0)]$ and $f_q(n, s)$ is the n -fold iteration of $f_q(s)$. We note that $f'_q(1) = \beta < 1$, i.e. $f_q(s)$ is the GF of subcritical GW system. From the arguments in [2, Ch.I, §8, Th 1] it follows that there exists a limit function $\mathcal{G}(x) := \lim_{n \rightarrow \infty} \mathcal{G}(n; x)$ such that $\mathcal{G}(f(x)) = \beta \mathcal{G}(x)$, iteration of this produces

$$\mathcal{G}(f_q(k, s)) = \beta^k \mathcal{G}(x). \tag{10}$$

On the other hand, it is easy to check that

$$\frac{1 - f_q(n, 0)}{1 - f_q(n + k, 0)} \longrightarrow \frac{1}{\beta^k} \quad \text{as } n \rightarrow \infty.$$

Then

$$\mathcal{V}_k(s) = \lim_{n \rightarrow \infty} \mathcal{V}_k(n; s) = 1 - \mathcal{G}(f_q(k, s)) \frac{1}{\beta^k}. \tag{11}$$

A combination of (10) and (11) implies $\mathcal{V}_k(s) = 1 - \mathcal{G}(s)$ and hence $\mathcal{V}_k(f_q(s)) = 1 - \beta\mathcal{G}(s)$. This leads to the Schröder functional equation

$$1 - \mathcal{V}_k(f_q(s)) = \beta \cdot (1 - \mathcal{V}_k(s)), \tag{12}$$

which expresses the invariance property of the limiting-measure $\{\nu_j(k), j \in \mathcal{S}\}$ with respect to the probabilities $P_{i_j}(\cdot)$. Setting now $s = 1$ in (12), we immediately obtain that $\sum_{j \in \mathcal{S}} \nu_j(k) = 1$.

To prove (8), we rewrite (7) as follows:

$$\nu_j(k) = \frac{\nu_j}{\beta^k} \left[1 - \frac{\mathbf{p}_0^j(k)}{q^j} \right]. \tag{13}$$

where $\nu_j = \pi_j q^j / \sum_{k \in \mathcal{S}} \pi_k q^k$. Multiplying both sides of equality (13) by s^j and summing over $j \in \mathcal{S}$, we obtain equation (8), where $\mathcal{V}(s) = \sum_{j \in \mathcal{S}} \nu_j s^j$. The form of (9) is in [4]. ■

Since the event $\{n < \mathcal{H} < n + k\}$ is complementary to $\{n + k < \mathcal{H} < \infty\}$, the following result is a logical continuation of the construction of the invariant distribution $\{\nu_j(k), j \in \mathcal{S}\}$ mentioned above.

Theorem 3. There exists $\sigma_j(k) := \lim_{n \rightarrow \infty} \mathcal{P}^{\mathcal{H}(n, n+k)} \{Z_n = j\}$ and

$$\sigma_j(k) = \frac{\pi_j \mathbf{p}_0^j(k)}{\sum_{i \in \mathcal{S}} \pi_i \mathbf{p}_0^i(k)}.$$

An appropriate GF $\mathcal{C}_k(s) := \sum_{j \in \mathcal{S}} \sigma_j(k) s^j$ has the form of

$$\mathcal{C}_k(s) = \frac{\mathcal{P}(\mathbf{p}_0(k)s)}{\mathcal{P}(\mathbf{p}_0(k))}$$

for any $k \in \mathbb{N}$. Moreover, $\mathcal{C}_k(s) \rightarrow \mathcal{V}(s)$ as $k \rightarrow \infty$.

3. Conclusion

The paper considers discrete-time Galton-Watson branching systems. We study the asymptotic behavior of the transition probabilities of these systems. We observe the properties of the population size distribution in the system on positive trajectories. One of the interesting results of this paper is Theorem 1, which corrects an error missed in the classical monograph by K.Athreya and P.Ney [2, Ch.I, §7, Th 2]. Recall that in this monograph the statement of the above-mentioned theorem was confirmed incorrectly.

REFERENCES

1. *Asmussen S. and Hering H.* Branching processes. Birkhäuser, Boston, 1983.
2. *Athreya K. B. and Ney P. E.* Branching processes. Springer, New York, 1972.
3. *Harris T. E.* The theory of branching processes. Springer-Verlag, Berlin, 1963.
4. *Imomov A. A. and Murtazaev M.* Renewed Limit Theorems for Non-critical Galton-Watson Branching Systems. J Theor. Probab., 2024, <https://doi.org/10.1007/s10959-024-01330-8>.
5. *Jagers P.* Branching Processes with Biological applications, JW & Sons, Pitman Press, UK. 1975.
6. *Karpenko A. V. and Nagaev S. V.* Limit theorems for the total number of descendents for the Galton-Watson branching process. Theory Probab. Appl., 1994. Vol. 38. P. 433–455.
7. *Kennedy D. P.* The Galton-Watson process conditioned on the total progeny. Jour. Appl. Prob., 1975, Vol. 12. P. 800–806.
8. *Sevastyanov B. A.* Branching processes. Nauka, Moscow. (Russian)

Azam A. Imomov — Doctor of Sciences, Head of Department, Mathematics and Computer Sciences faculty. E-mail: imomov_azam@mail.ru

Yorqinoy Ibrohimova — post-doctoral student, Mathematics and Computer Sciences faculty. E-mail: yorqinoyibrohimova2@gmail.com

p -ADIC QUASI GIBBS MEASURES FOR THE FOUR STATES p -ADIC SOS MODEL ON THE CAYLEY TREE OF ORDER TWO

M. Rahmatullaev^{1,2}, O. Axmedov³

¹ *V.I.Romanovskiy Institute of Mathematics, Tashkent, Uzbekistan*

² *Namangan State University, Namangan, Uzbekistan*

³ *Ferghana State University, Ferghana, Uzbekistan*

In the present paper, we consider a for the four states p -adic SOS model on a Cayley tree of order two. The existence of translation-invariant and $G_2^{(2)}$ -periodic p -adic quasi Gibbs measures of this model is investigated. Moreover, the boundedness of such kinds of measures is established, which yields the occurrence of a phase transition. **Keywords:** *p -adic numbers, p -adic SOS model, quasi Gibbs measure, phase transition.*

Introduction

The SOS (solid-on-solid) model can be considered as a generalization of the Ising model [1] or a less symmetric variant of the Potts model [2]. In [3] the SOS model on the cubic lattice \mathbb{Z}^d was studied and it was shown that the phase transition depends on the number of spin values. In [4], the SOS model on the Cayley tree was studied. In [5] the SOS model on the lattice \mathbb{Z}^d with an external field was studied and conditions for the occurrence of phase transitions were found.

In this paper, we studied p -adic SOS models with four states on a Cayley tree of order two. If $p \equiv 1 \pmod{4}$, then for the p -adic SOS model with four states on a Cayley tree of order two, there exist translation-invariant (TI) and G_2^2 -periodic p -adic quasi-Gibbs measures. Because the phase transition happens when there are at least two unique measures, one of which is bounded and the other is unbounded, we investigate the boundedness of the obtained measures.

1. p -adic number

Let \mathbb{Q} be the field of rational numbers. For a fixed prime number p , every rational number $x \neq 0$ can be represented in the form $x = p^r \frac{n}{m}$, where, $r, n \in \mathbb{Z}$, $m \in \mathbb{N}$ and n and m are relatively prime with p : $(p, n) =$

$1, (p, m) = 1$. The *p*-adic norm of x is given by

$$|x|_p = \begin{cases} p^{-r}, & x \neq 0, \\ 0, & x = 0. \end{cases}$$

This norm is non-Archimedean and satisfies the strong triangle inequality

$$|x + y|_p \leq \max\{|x|_p, |y|_p\}, \quad \forall x, y \in \mathbb{Q}.$$

The completion of \mathbb{Q} with respect to the *p*-adic norm defines the *p*-adic field \mathbb{Q}_p . The elements of the *p*-adic field \mathbb{Q}_p are called *p*-adic numbers.

Any *p*-adic number $x \neq 0$ can be uniquely represented in the canonical form

$$x = p^{\gamma(x)}(x_0 + x_1p + x_2p^2 + \dots)$$

where $\gamma(x) \in \mathbb{Z}$ and the integers x_j satisfy that $x_0 \neq 0, x_j \in \{0, 1, \dots, p-1\}, j \in \mathbb{N}$.

In this case $|x|_p = p^{-\gamma(x)}$.

For $a \in \mathbb{Q}_p$ and $r > 0$ we denote

$$B(a, r) = \{x \in \mathbb{Q}_p : |x - a|_p < r\}.$$

p-adic exponential is defined by

$$\exp_p(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!},$$

which converges for $x \in B(0, p^{-1/(p-1)})$. Denote

$$\mathcal{E}_p = \left\{ x \in \mathbb{Q}_p : |x - 1|_p < p^{-1/(p-1)} \right\} = B(1, p^{-1/(p-1)})$$

(see [8]).

2. Cayley tree

The Cayley tree Γ^k (See [6],[7]) of order $k \geq 1$ is an infinite tree, i.e., a graph without cycles, from each vertex of which exactly $k + 1$ edges issue. Let $\Gamma^k = (V, L, i)$, where V is the set of vertices of Γ^k, L is the set of edges of Γ^k and i is the incidence function associating each edge $l \in L$ with its endpoints $x, y \in V$. If $i(l) = \{x, y\}$, then x and y are called *nearest neighboring vertices*, and we write $l = \langle x, y \rangle$.

It is known (see [8]) that there exists a one-to-one correspondence between the set V of vertices of the Cayley tree of order $k \geq 1$ and the group G_k of the free products of $k + 1$ cyclic groups $\{e, a_i\}, i = 1, \dots, k + 1$ of the second order (i.e. $a_i^2 = e, a_i^{-1} = a_i$) with generators a_1, a_2, \dots, a_{k+1} .

We denote by $S(x)$ the set of direct successors of $x \in G_k$.

3. The model

Let \mathbb{Q}_p be the field of p -adic numbers and Φ be a finite set. A configuration σ on a set $A \subset V$ is defined as a function assigning each $x \in A$ a value $\sigma(x) \in \Phi$. Let Ω_A denote the set of all configurations on A and let Ω denote the set Ω_V . We consider the p -adic SOS model on the Cayley tree, where spins take values from the set $\Phi := \{0, 1, 2, \dots, m\}$, $m \geq 2$.

The Hamiltonian of the p -adic SOS model on Ω_{V_n} is given by (see [9]):

$$H_n(\sigma) = J \sum_{\langle x,y \rangle \in L_n} |\sigma(x) - \sigma(y)|_\infty, \quad \sigma \in \Omega_{V_n}, \tag{1}$$

where $J \in B(0, p^{-1/(p-1)})$ is the coupling constant, L_n is the set of edges in V_n , and $|\cdot|_\infty$ denotes the usual absolute value.

4. The compatibility condition

Let $z : x \mapsto z_x = (z_{0,x}, z_{1,x}, z_{2,x}, \dots, z_{m,x}) \in \mathbb{Q}_p^{m+1}$, and the p -adic quasi Gibbs measure $\mu : \Omega \rightarrow \mathbb{Q}_p$ is defined as:

$$\mu_{\tilde{z}}^{(n)}(\sigma) = Z_{n,\tilde{z}}^{-1} \exp_p\{H_n(\sigma)\} \prod_{x \in W_n} \tilde{z}_{\sigma(x),x}, \tag{2}$$

where $x \in V \setminus \{x^0\}$, $\sigma \in \Omega_{V_n}$, $n \in \mathbb{N}$, and $Z_{n,\tilde{z}}$ is the corresponding partition function, i.e.,

$$Z_{n,\tilde{z}} = \sum_{\omega \in \Omega_{V_n}} \exp_p\{H_n(\omega)\} \prod_{x \in W_n} \tilde{z}_{\omega(x),x}. \tag{3}$$

Note that if $z_{\sigma(x),x} \in \mathcal{E}_p$ for any $x \in V$, then the quasi Gibbs measure coincides with the p -adic Gibbs measure.

Let for p -adic distributions (2) the compatibility condition holds for all $n \geq 1$ and $\sigma_{n-1} \in \Phi^{V_{n-1}}$:

$$\sum_{\omega_n \in \Omega_{W_n}} \mu_{\tilde{z}}^{(n)}(\sigma_{n-1} \vee \omega_n) = \mu_{\tilde{z}}^{(n-1)}(\sigma_{n-1}), \tag{4}$$

then from the non-Archimedean p -adic analogue of Kolmogorov’s theorem in Ω , there exists a unique limiting measure $\mu_{\tilde{z}}$ satisfying condition (4), where $\sigma_{n-1} \vee \omega_n \in \Omega_{V_n}$ is the union of configurations and is defined as follows:

$$(\sigma_{n-1} \vee \omega^{(n)})(x) = \begin{cases} \sigma_{n-1}(x), & \text{if } x \in V_{n-1}, \\ \omega^{(n)}(x), & \text{if } x \in W_n. \end{cases}$$

The following theorem expresses the requirement on function z that the distributions $\mu^{(n)}(\sigma_n)$ hold the compatibility conditions.

Theorem 1. *The *p*-adic probability distributions $\mu_z^{(n)}(\sigma_n)$, which are defined by (2) are compatible if and only if for any $x \in V \setminus \{x^0\}$ the following system of equations holds (see [9]):*

$$z_{i,x} = \prod_{y \in S(x)} \frac{\sum_{j=0}^{m-1} \theta^{|i-j|_\infty} z_{j,y} + \theta^{m-i}}{\sum_{j=0}^{m-1} \theta^{m-j} z_{j,y} + 1}, \quad i = 0, 1, \dots, m - 1, \quad (5)$$

here $\theta = \exp_p(J)$ and $z_{i,x} = \tilde{z}_{i,x} / \tilde{z}_{m,x}$, $i = 0, 1, \dots, m - 1$.

In this case, by the *p*-adic analogue of the Kolmogorov theorem there exists a unique *p*-adic quasi Gibbs measure μ_z on the set Ω such that $\mu_z(\{\sigma|_{V_n} \equiv \sigma_n\}) = \mu_z^{(n)}(\sigma_n)$ for all n and $\sigma_n \in \Omega_{V_n}$ (see [2]).

The set of all *p*-adic quasi Gibbs measures associated with functions $\mathbf{z} = \mathbf{z}_x, x \in V$ is denoted by $\mathcal{QG}(H)$.

Definition 1. A *phase transition* is said to occur if there exist at least two different *p*-adic quasi Gibbs measures $\mu, \nu \in \mathcal{QG}(H)$ such that μ is bounded and ν is unbounded. In other words, two different functions s and h may be found defined on \mathbb{N} , for which corresponding measures μ_s and μ_h exist; one of these measures is bounded, while the other is unbounded (See [10]).

Namely, for any collection of functions satisfying the functional equation (5) there exists a unique quasi Gibbs measure, the correspondence being one-to-one.

Let $G_k/G_k^* = \{H_1, \dots, H_r\}$ be a factor group, where G_k^* is a normal subgroup of index $r \geq 1$.

Definition 2. Let G_k^* be a subgroup of G_k . A collection of vectors $z = \{z_x \in R^m : x \in G_k\}$ is said to be G_k^* -periodic, if $z_{yx} = z_x$ for all $x \in G_k$ and $y \in G_k^*$. A G_k -periodic collections is said to be *translation-invariant*.

Definition 3. A Gibbs measure is called G_k^* -periodic (*translation-invariant*), if it corresponds to G_k^* -periodic (*translation-invariant*) collection h .

5. The main results

Let $G_k^{(2)}$ be the subgroup in G_k consisting of all words of even length. Clearly, $G_k^{(2)}$ is a subgroup of index 2.

Theorem 2. If $p \equiv 1 \pmod{4}$, then for the four states p -adic SOS model on the Cayley tree of order two, there are at least three translation-invariant quasi Gibbs measures.

Theorem 3. If $p \equiv 1 \pmod{4}$, then for the four states p -adic SOS model on the Cayley tree of order two, there are at least two $G_2^{(2)}$ -periodic quasi Gibbs measures.

Theorem 4. If $p \equiv 1 \pmod{4}$, a phase transition occurs for the four states p -adic SOS model on a Cayley tree of order two.

6. Conclusion

In conclusion, if $p \equiv 1 \pmod{4}$, then the four-state p -adic SOS model on the Cayley tree of order two has translation-invariant and $G_2^{(2)}$ -periodic quasi-Gibbs measures. The existence of a phase transition has been determined.

In [4], the existence of a phase transition was determined for $m = 2$, $k = 2$, and $p = 3$. In our case, we have identified the conditions under which a phase transition occurs for $m = 3$, $k = 2$ and $p \equiv 1 \pmod{4}$.

All the results obtained in Section 5 are new results.

REFERENCES

1. *Hans-Otto Georgii* Gibbs measures and phase transitions // Walter de Gruyter. 2011. Vol. 9.
2. *Potts R., B.* Some generalized order-disorder transformations // Mathematical Proceedings of the Cambridge Philosophical Society. 1952. Vol.48 №1. P. 106–109.
3. *Mazel A.E. and Suhov Yu.M.* Random surfaces with two-sided constraints: An application of the theory of dominant ground states // Journal of Statistical Physics. 1991. Vol. 64. P. 111–134.
4. *Dinaburg E. I., Mazel A. E.* Layering transition in SOS model with external magnetic field // Journal of Statistical Physics. 1994. Vol. 74. P. 533–563.
5. *Rozikov U., Suhov Y.* Gibbs measures for SOS model on a Cayley tree // Infinite Dimensional Analysis, Quantum Probability and Related Topics. 2006. №9. P. 471–488.
6. *Rozikov U.* Gibbs measures on Cayley trees. Singapore: World scientific, 2013. 398 p.
7. *Ganikhodzhaev, N.* Group representation and automorphisms of the Cayley tree // Dokl. Akad. Nauk Resp. Uzbekistan. 1994. №4. P. 3–5.
8. *Vladimirov V.S., Volovich I.V. and Zelenov E. V.* p -Adic Analysis and Mathematical Physics Singapore: World scientific, 1994.

9. *Khakimov O.* On *p*-adic Solid-On-Solid model on a Cayley tree // Theoretical and Mathematical Physics. 2017. 193(3) P.547–562.
10. *Mukhamedov F.* On dynamical systems and phase transitions for $q + 1$ -state *p*-adic Potts model on the Cayley tree // Mathematical Physics, Analysis and Geometry. 2013. Vol. 16. P.49–87.

Rahmatullaev Muzaffar — DSc, associate professor, Head of the Naman-gan branch of the Institute of Mathematics, mathematics. E-mail: *mrahmatul-laev@rambler.ru*

Akhmedov Olimkhon — student of PhD, teacher, Ferghana state university, mathematics.. E-mail: *olimxonaxmedov5@gmail.com*

ON APPLICATION OF KARAMATA SLOWLY VARYING FUNCTIONS IN THE THEORY OF NONCRITICAL MARKOV BRANCHING SYSTEMS

A. A. Imomov^{1,2}, Z. A. Nazarov²

¹*Karshi State University, Karshi, Uzbekistan*

²*V.I.Romanovskiy Institute of Mathematics, Tashkent, Uzbekistan*

The paper discusses the continuous-time Markov Branching System. We deal only with the noncritical case. The primary task of this paper is to extend to the continuous case our recent result, which explicitly calculates the famous constant in the theory of subcritical Galton-Watson branching systems, announced by Kolmogorov in 1938. We demonstrate the rate of convergence to the Kolmogorov constant for the continuous-time Markov Branching System. This result contributes to determining the approximation rate in several classical limit theorems of the theory of Markov Branching Systems.

Keywords: *Markov Branching System, Markov chain, extinction time, Kolmogorov constant, slowly varying functions.*

1. Background and purpose

The branching stochastic models describe an evolution of the population size of the reproductive individuals system. These models most clearly illustrate numerous stochastic phenomena occurring both in nature and in human activity. The simple Galton-Watson model, originally evolved as a family survival model in the second half of the 19th century, today has numerous generalizations and modifications; see [1], [2], [3] [8].

The advanced current standing of the modern theory of stochastic branching systems has achieved largely due to its compatibility with the theory of slowly varying functions. The conception of slow variation or, more general, regular variation, was first initiated by the famous Serbian mathematician Jovan Karamata in the early 30s of the XX century; see, for instance, [4], [5] and [11]. The integration of these two theories began with the publication of the excellent work of Zolotarev [14].

Recall that real-valued, positive and measurable function $\ell(x)$ is said to be slowly varying (SV) at infinity, in the sense of Karamata, if $\ell(\lambda x)/\ell(x) \rightarrow 1$ as $x \rightarrow \infty$ for each $\lambda > 0$. The SV property can be defined at any finite point by shifting the origin of the function to this point. In what follows, we use the symbol \mathcal{SV}_α to denote the class of Karamata SV-functions at point

α . The Representation Theorem [11] asserts that each function $\ell(\cdot) \in \mathcal{SV}_\infty$ may be written in the form

$$\ell(x) = c(x) \exp \left(\int_a^x \frac{\varepsilon(u)}{u} du \right) \tag{1}$$

for some $a > 0$, where $c(x)$ has a finite positive limit, $c(x) \rightarrow c > 0$, the function $\varepsilon(x)$ is a continuous called the SV index, such that $\varepsilon(x) \rightarrow 0$ as $x \rightarrow \infty$. In a special case, $\ell(\cdot)$ is said to be normalised if $c(x) \equiv \text{const}$. Any SV-function is associated with an appropriate regular varying function. A function $R(x)$ is called regularly varying at infinity with index δ , if it can be expressed as $R(x) = x^\delta L(x)$ for some $\ell(\cdot) \in \mathcal{SV}_\infty$. Then $R(\lambda x)/R(x) \rightarrow \lambda^\delta$ as $x \rightarrow \infty$ for each $\lambda > 0$; see also, [4] and [11] for more information.

In this paper we are interested in the application of SV-functions to enhance and generalize some classical limit results in the theory of stochastic continuous-time Markov Branching system. The states of the system are determined by the number of individuals and its growth occurs on a continuous-time axis $\mathcal{T} := [0, +\infty)$ according to the following random mechanism. The individuals undergo transformations in accordance with the branching rates $\{a_k, k \in \mathbb{N}_0\}$, where $\mathbb{N}_0 = \{0\} \cup \mathbb{N}$ and \mathbb{N} is a set of natural numbers.

Denoting by $Z(t)$ the population size at time $t \in \mathcal{T}$, we have a *homogeneous-continuous-time Markov Branching (MB) system* with branching rates $\{a_k, k \in \mathbb{N}_0\}$. The appropriate q-matrix $Q = \{q_{ij}\}$ of the system $\{Z(t), t \in \mathcal{T}\}$ is given as follows:

$$q_{ij} = \begin{cases} ia_{j-i+1} & \text{if } j \geq i \geq 0, \\ ia_0 & \text{if } j = i - 1, \\ 0 & \text{otherwise.} \end{cases} \tag{2}$$

This in essence regulates the further evolution of the system; see [10].

Consider transition probabilities of the Markov chain $\{Z(t), t \in \mathcal{T}\}$

$$P_{ij}(t) := \mathbb{P}\left\{Z(\tau + t) = j \mid Z(\tau) = i\right\} \quad \text{for any } \tau \in \mathcal{T}.$$

They are i -fold convolution of the distribution $\mathbf{p}_j(t) := P_{1j}(t)$. Define generating functions (GFs)

$$f(s) = \sum_{j \in \mathbb{N}_0} a_j s^j \quad \text{and} \quad F(t; s) := \sum_{j \in \mathbb{N}_0} \mathbf{p}_j(t) s^j$$

for $s \in [0, 1)$. An important fact is that the GFs $F(t; s)$ and $f(s)$ are related by the Kolmogorov backward equation $\partial F(t; s)/\partial t = f(F(t; s))$ which is solvable under the boundary condition $F(0; s) = s$; see [13].

If $m := \sum_{j \in \mathcal{S}} j a_j$ converges, then $m = f'(1-)$ is the average intensity of the branching rate in the system, which determines an asymptotic classification of the system trajectories. The MB system is divided into the subcritical, critical and supercritical types, depending on $m < 0$, $m = 0$ and $m > 0$ respectively. Letting q be the extinction probability of the system initiated by the one founder-individual, it is the smallest root of the equation $f(s) = 0$ for $s \in [0, 1]$. This probability also satisfies the fixed-point equation $F(t; s) = s$ for all $t \in \mathcal{T}$. We note, that always $q > 0$ being that $a_0 > 0$ by default. In what follows we consider the noncritical system, i.e. $m \neq 0$.

Let $\mathbb{P}\{Z(t) > 0\}$ be the survival probability at time t of the system initiated by the single founder-individual. In the subcritical situation, one of the classical problems is to study the asymptote of this probability. This problem was first processed by Sevastyanov [12] in 1951. He proved that if $f''(1-) < \infty$, then the following asymptotic formula is valid:

$$\mathbb{P}\{Z(t) > 0\} \sim \mathcal{K}e^{mt} \quad \text{as } t \rightarrow \infty, \quad (3)$$

where \mathcal{K} is some constant depending on a form of $f(s)$. In the discrete-time case, the limiting coefficient \mathcal{K} is the famous Kolmogorov constant in the theory of subcritical Galton-Watson branching systems, which has been announced by Kolmogorov [9] in 1938. In the recent paper [6], the constant in Galton-Watson case was calculated explicitly depending on the second moment of the offspring law. By analogy with the Galton-Watson case, we refer to \mathcal{K} in (3) as the Kolmogorov constant.

In this report we are purposed in an explicit-calculation analysis of the constant \mathcal{K} in the asymptotic formula (3) and its extending to the supercritical case, depending on the moments of the branching rates of the system under consideration.

2. Basic assumption and main results

Return to the MB system $\{Z(t), t \in \mathcal{T}\}$. Let $U_q[0, 1) := \{[0, q) \cup (q, 1)\}$ be a unit interval with a punctured point q . Taking $R(t; s) := q - F(t; s)$, we note that in the subcritical case $q = 1$ and $Q(t) := R(t; 0) = \mathbb{P}\{Z(t) > 0\}$ is the survival probability at time t of the system initiated by the one founder-individual. In this case, bypassing the second moment condition

$f''(1-) < \infty$, Zolotarev [14] proved that the existence of the integral

$$I(m) = - \int_0^1 \frac{f(u) + m(1-u)}{(1-u)f(u)} du \tag{4}$$

is necessary and sufficient for validity of the asymptotic representation of type (3) and the constant \mathcal{K} is related to $I(m)$ by the equality $I(m) = \ln \mathcal{K}$. For convenience, we wrote the integral in (4) in a slightly different form. As shown in [13], that the convergence of integral $I(m)$ is equivalent to that

$$\sum_{k \in \mathcal{S}} a_k k \ln k < \infty. \tag{5}$$

Zolotarev’s result motivates us to think about extending the task he has processed, to the noncritical situation. Define

$$\mathcal{A}_q(t; s) := \frac{R(t; s)}{\beta^t}$$

for $s \in U_q[0, 1)$, where $\beta = \exp\{f'(q)\}$. Let $\mathcal{H} := \min\{t : Z(t) = 0\}$ be an extinction time of the genealogical tree of one founder-individual. The classical extinction theorem implies that $\mathbb{P}\{t < \mathcal{H} < \infty | Z(t) = j\} = q^j$ for any $t \in \mathcal{T}$. Taking this into account and using the total probability formula, we obtain $\mathbb{P}\{t < \mathcal{H} < \infty\} = q - F(t; 0)$. So, $Q(t) = R(t; 0) = \mathbb{P}\{t < \mathcal{H} < \infty\}$ is the bridled-survival probability of the genealogical tree of one founder-individual in the noncritical situation. Let

$$\mathcal{A}_q(t) := \mathcal{A}_q(t; 0) = \frac{Q(t)}{\beta^t}.$$

In [7] it was proved that

$$\ln \frac{\mathcal{A}_q(t; s)}{q - s} \longrightarrow \int_s^q \frac{f(u) - f'(q)(u - q)}{f(u)(u - q)} du. \tag{6}$$

In this case, the convergence of the integral in (6) is also equivalent to Zolotarev’s condition (5). To verify this, we define the transformed MB system $\{Z_q(t)\}$, generated by the Harris-Sevastyanov transformation $f_q(s) = f(qs)/q$, provided that $q \neq 0$ for the supercritical case. It is easy to see that $\{Z_q(t)\}$ is a subcritical system initiated by one founder-individual, the branching rates of which is $\varphi_k = a_k q^{k-1}$ and $f_q(s) = \sum_{k \in \mathcal{S}} \varphi_k s^k$. Therefore the average intensity of the branching rate is $f'_q(1-) = f'(q) < 0$. Next,

letting $F_q(t; s) := \mathbb{E}s^{Z_q(t)}$, it immediately follows from the Kolmogorov backward equation that $F_q(t; s) = F(t; qs)/q$. Then $\mathbb{E}Z_q(t) = \beta^t$. Now, using standard transformations, we write the integral in (6) in the form

$$\int_1^s \frac{f_q(u) + f'_q(1)(1-u)}{(1-u)f_q(u)} du. \tag{7}$$

for $s \in [0, 1)$. Since $q \leq 1$ it follows from (5) that $\sum_{k \in \mathcal{S}} \varphi_k k \ln k < \infty$. Then, in accordance with Zolotarev’s conclusion, integral (7) converges. So, Zolotarev’s condition (5) is sufficient for converging the integral on the right-hand side of (6). Thus, since $R(t; qs) = qR_q(t; s)$ it follows that $Q(t) = q\mathbb{P}\{Z_q(t) > 0\}$ and hence, according to formula (3), we have

$$\frac{1}{\mathcal{A}_q(t)} = \frac{\mathbb{E}Z_q(t)}{\mathbb{P}\{t < \mathcal{H} < \infty\}} = \frac{1}{q} \mathbb{E}[Z_q(t) \mid Z_q(t) > 0] \rightarrow \frac{1}{q\mathcal{K}} =: \frac{1}{\mathcal{K}_q}$$

as $t \rightarrow \infty$. So \mathcal{K}_q can be interpreted as the coefficient of asymptotic equivalence between the expectation $\mathbb{E}Z_q(t) = \beta^t$ of the population size and the survival probability $\mathbb{P}\{Z_q(t) > 0\}$ of the transformed subcritical MB system generated by the Harris-Sevastyanov transformation $f_q(s)$. By analogy with the discrete-time case we refer to \mathcal{K}_q as the extended Kolmogorov constant; see [6]. Further, since $f_q(s)$ is a power series, in order for Zolotarev’s condition (5) to be satisfied, it is necessary and sufficient that the integrand in (7) for some $\nu > 0$ has the form

$$\frac{f_q(s) + f'_q(1)(1-s)}{(1-s)f_q(s)} = \mathcal{O}(1-s)^{\nu-1} \quad \text{as } s \uparrow 1 \tag{8}$$

This necessity proposes considering the case when the Harris-Sevastyanov transformation $f_q(s)$ to be admitted the following representation:

$$f_q(s) = |f'_q(1)|(1-s) + (1-s)^{1+\nu} \mathcal{L}\left(\frac{1}{1-s}\right), \tag{9}$$

where $\mathcal{L}(\cdot) \in \mathcal{SV}_\infty$ and we choose that $\nu \in (0, 1)$. This implies that $f''(1) = \infty$ if $\nu < 1$. The following moment assumption is related to condition (9):

$$m \neq 0 \quad \text{and} \quad \sum_{k \in \mathbb{N}} k^{1+\nu} a_k < \infty \quad \text{for } \nu \in (0, 1). \tag{A}$$

We state our main result in the following theorem.

Theorem 1. Let the condition (9) with [A] is satisfied. Then

$$\mathcal{A}_q(t) = \frac{q}{1 + q\Delta(t)}, \tag{10}$$

where

$$\Delta(t) = \frac{1}{|\ln \beta| \nu q} L_{\Delta}(t), \quad (11)$$

and $L_{\Delta}(\cdot) \in \mathcal{SV}_{\infty}$ such that $L_{\Delta}(t)/\mathcal{L}(\beta^{-t}) \rightarrow 1$ as $t \rightarrow \infty$.

3. Conclusion

The paper considers the continuous-time Markov Branching System. In noncritical case, we state and prove an asymptotical expansion of bridled-survival probabilities of the genealogical tree of one founder-individual. The principal novelty of our result is explained by the fact that it contributes to determine the approximation rate of convergence in several classical limit theorems of the theory of noncritical Markov Branching Systems. In the future, we intend to find an explicit form of the famous Kolmogorov constant depending on the finite second-order branching rates moment.

REFERENCES

1. *Asmussen S. and Hering H.* Branching processes. Birkhäuser, Boston, 1983.
2. *Athreya K. B. and Ney P. E.* Branching processes. Springer, New York, 1972.
3. *Harris T. E.* The theory of branching processes. Springer-Verlag, Berlin, 1963.
4. *Bingham N. H., Goldie C. M. and Teugels J. L.* Regular variation, Cambridge, 1987.
5. *Imomov A. A. and Meyliev A.Kh.* On the Application of Slowly Varying Functions with Remainder in the Theory of Markov Branching Processes with Mean One and Infinite Variance. Ukrainian Math. Journal. 2022. Vol.73(8). P. 1225–1237.
6. *Imomov A. A. and Murtazaev M.* On the Kolmogorov constant explicit form in the theory of Discrete-time Stochastic Branching Systems. J. Appl. Probab. 2024.
7. *Imomov A. A.* On conditioned limit structure of the Markov branching process without finite second moment. Malaysian J. Math. Sci. 2017. Vol.11(3). P. 393–422.
8. *Jagers P.* Branching Processes with Biological applications, JW & Sons, Pitman Press, UK. 1975.
9. *Kolmogorov A. N.* K resheniyu odnoy biologicheskoy zadachi. Reports of SRI Math. and Mech. at Tomsk Univ. 1938. Vol.2. P. 7–12. (Russian)
10. *Li J., Cheng L. and Li L.* Long time behaviour for Markovian branching-immigration systems. Discrete Event Dynamic Systems. 2021. Vol.31. P. 37–57.
11. *Seneta E.* Regularly varying functions. Springer-Verlag, Berlin; Translated Russian, Nauka, Moscow, 1985.

12. *Sevastyanov B. A.* The theory of Branching stochastic process. *Uspekhi Mat. Nauk.* 1951. Vol.6(46). P. 47–99.
13. *Sevastyanov B. A.* Branching processes. Nauka, Moscow. (Russian)
14. *Zolotarev V. M.* More exact statements of several theorems in the theory of branching processes. *Theory Prob. Appl.* 1957. Vol. 2. P. 245–253.

Azam A. Imomov — Doctor of Sciences, Head of Department, Mathematics and Computer Sciences faculty. E-mail: *imomov_azam@mail.ru*

Zuhridin A. Nazarov — post-doctoral student, Stochastic analysis department. E-mail: *zuhrov13@gmail.com*

MAXIMUM A POSTERIORI PREDECESSORS IN STATE OBSERVATION MODELS

B. Rudić¹, M. Pichler-Scheder¹, D. Efrosinin²

¹*Linz Center of Mechatronics GmbH, Linz, Austria*

²*JKU, Institute of Stochastics, Linz, Austria*

Maximum A Posteriori (MAP) inference in state observation models is typically viewed as either decoding marginal MAP state estimates or the joint MAP state sequence estimate. Instead, this paper addresses a novel yet fundamental MAP inference method denoted as predecessor decoding. This method decodes the most probable predecessors of a chosen initial state using only the marginal distributions from a forward filtering pass. It has recently been shown that MAP predecessor results exhibit joint MAP characteristics and achieve superior error statistics over marginal MAP estimates. This paper elaborates on the relationship and, more specifically, the differences between marginal MAP, joint MAP, and MAP predecessors. **Keywords:** *state observation models, MAP inference, recursive Bayesian inference.*

1. Introduction

Dynamic systems are frequently abstracted into state observation models, which are utilized in various fields such as signal processing, navigation, telecommunications, finance, and more [1]. Inference within state observation models is a well-researched area, encompassing both marginal and Maximum A Posteriori (MAP) inference, applicable to models with either discrete or continuous state spaces [2].

For instance, in discrete-state Hidden Markov Models (HMM), marginal inference is performed by the forward-backward algorithm, while the Viterbi algorithm addresses the joint MAP problem. However, discrete-state as well as particle-based approaches, such as [3, 4], suffer from the curse of dimensionality. In continuous-state models, efficient marginal inference techniques have been discovered presuming Gaussian distributions [5, 1]. In such models the marginal means coincide with the joint MAP assignment [6]. Gaussian assumptions can lead to relatively poor results, e.g., in dynamic systems with nonlinear or non-Gaussian characteristics.

This work has been supported by the COMET-K2 Center of the Linz Center of Mechatronics (LCM) funded by the Austrian federal government and the federal state of Upper Austria.

For example, consider an application designated to track the positions, linear and angular velocities and accelerations, orientations, and possible other features of a dynamically moving object indirectly observed through measurements affected by multimodal noise. Instead of discretizing or oversimplifying such systems to fit into a HMM, particle filter or Gaussian model, more accurate approximations can be achieved using Gaussian mixtures. For such models, a MAP predecessor decoder was recently introduced in [7], which results exhibit joint MAP characteristics, i.e., coherent state estimates whose sequence is consistent with the presumed transition model. Given that Gaussian mixtures can be used to approximate density functions to an arbitrary degree, an appropriate decoder holds considerable potential for a range of applications. In this paper, the relationship between MAP predecessors and the classic MAP inference methods in state observation models is discussed more generally without assuming specific distribution models.

2. Marginal Inference in State Observation Models

For general state observation models the joint distribution of hidden states $\mathbf{x}_{0:T}$ and given observations $\mathbf{o}_{1:T}$ factorizes into

$$p(\mathbf{x}_{0:T}, \mathbf{o}_{1:T}) = p(\mathbf{x}_0) \prod_{t=1}^T p(\mathbf{x}_t | \mathbf{x}_{t-1}) p(\mathbf{o}_t | \mathbf{x}_t). \quad (1)$$

Marginal inference in such models can be distinguished in predicting, filtering and smoothing, which refer to computing the marginal distributions

$$p(\mathbf{x}_t | \mathbf{o}_{1:s}) \equiv \begin{cases} \text{predicting} & \iff s < t \\ \text{filtering} & \iff s = t \\ \text{smoothing} & \iff s > t \end{cases} \quad (2)$$

The marginals can be inferred by marginalization of the joint posterior, or recursively based on the factors in (1) thereby exploiting independencies among the random variables of a state observation model. In most cases, only the latter option is feasible. Note that there are two fundamentally different approaches when inferring the smoothing marginals, based on a recursive or *fixed-lag* factorization or on a non-recursive *two-filter* factorization of $p(\mathbf{x}_t | \mathbf{o}_{1:T})$. Both apply a form of forward/backward-principle, i.e., fixed-lag smoothing implements forward-filtering/backward-smoothing while a two-filter smoother performs forward-filtering/backward-filtering. The formulas for the marginals referenced in (2) can for instance be found in [8], which also includes their derivations.

3. Marginal MAP and Joint MAP Inference

MAP inference is traditionally performed either by maximizing the marginals or the joint posterior

$$\overleftarrow{\mathbf{x}}_t := \arg \max_{\mathbf{x}_t} p(\mathbf{x}_t | \mathbf{o}_{1:T}) \quad \forall t, \tag{3}$$

$$\widehat{\mathbf{x}}_{0:T} := \arg \max_{\mathbf{x}_{0:T}} p(\mathbf{x}_{0:T} | \mathbf{o}_{1:T}). \tag{4}$$

Figure 1 and 2 illustrate exemplary distributions and their marginal and joint MAP assignments within the model graph. Computing $\overleftarrow{\mathbf{x}}_t$ and $\widehat{\mathbf{x}}_{0:T}$ is also denoted posterior decoding and sequence decoding, respectively. While posterior decoding can be reduced to marginal inference and mode-finding, decoding the joint assignment $\widehat{\mathbf{x}}_{0:T}$ presents a greater challenge in state observation models. To our knowledge, analytic solutions are limited to special cases, such as discrete-state and Gaussian models [4, 6].

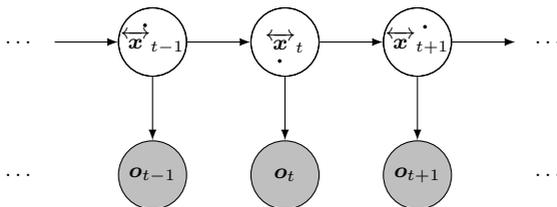


Figure 1. Model graph with incoherent marginal MAP assignments (dots).

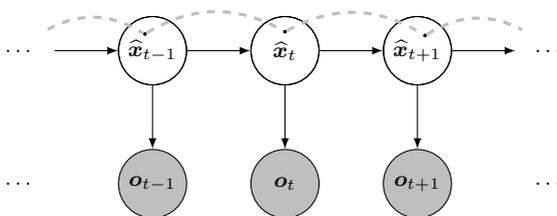


Figure 2. Model graph with coherent joint MAP assignments (dashed).

4. MAP Predecessors

Contrary to the traditional inference methods from section 3, this section is concerned with MAP predecessors.

Definition 1. The MAP predecessor of a given \mathbf{x}_{t+1}^* w.r.t. $\mathbf{o}_{1:t}$ is

$$\vec{\mathbf{x}}_t := \arg \max_{\mathbf{x}_t} p(\mathbf{x}_t | \mathbf{x}_{t+1}^*, \mathbf{o}_{1:t}). \tag{5}$$

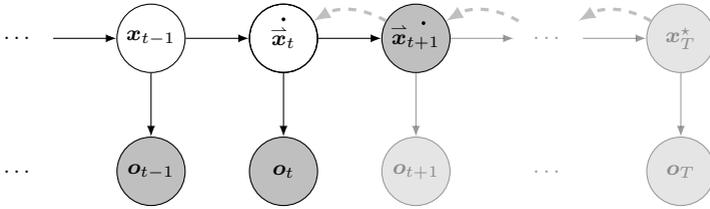


Figure 3. Predecessor decoding of $\vec{\mathbf{x}}_t$ given $\mathbf{o}_{1:t}$ and $\vec{\mathbf{x}}_{t+1}$.

We are interested in inferring the MAP predecessors iteratively, i.e., the given state for which the MAP predecessor is to be inferred is $\mathbf{x}_{t+1}^* := \vec{\mathbf{x}}_{t+1}$. In this manner, and as illustrated in Figure 3, we recursively obtain a MAP predecessor sequences $\vec{\mathbf{x}}_{0:T}$, where only the initial \mathbf{x}_T^* is specified explicitly.

Lemma 1. The MAP predecessor can be inferred recursively via

$$\vec{\mathbf{x}}_t = \arg \max_{\mathbf{x}_t} p(\mathbf{x}_t | \mathbf{o}_{1:t}) p(\vec{\mathbf{x}}_{t+1} | \mathbf{x}_t). \quad (6)$$

Proof. By Bayes' theorem the distributions in (5) and (6) are proportional, differing only by scalar factor $p(\vec{\mathbf{x}}_{t+1} | \mathbf{o}_{1:t})$, i.e., $p(\mathbf{x}_t | \vec{\mathbf{x}}_{t+1}, \mathbf{o}_{1:t}) p(\vec{\mathbf{x}}_{t+1} | \mathbf{o}_{1:t}) = p(\mathbf{x}_t | \mathbf{o}_{1:t}) p(\vec{\mathbf{x}}_{t+1} | \mathbf{x}_t)$. ■

By induction on (6) a predecessor sequence $\vec{\mathbf{x}}_{0:T}$ is obtained based on the filtering marginals, the most likely transitions, and where the initialization $\vec{\mathbf{x}}_T$ can be chosen freely. In an application, the initialization would typically be a point of interest where one aims to identify a probable sequence of preceding states with regard to a given sequence of observations.

Note that $\vec{\mathbf{x}}_t$ equivalently maximizes $p(\mathbf{x}_t, \vec{\mathbf{x}}_{t+1} | \mathbf{o}_{1:t})$ and by conditional independence (or *Bayes-Ball*) also $p(\mathbf{x}_t, \vec{\mathbf{x}}_{t+1:T} | \mathbf{o}_{1:T})$. With regard to the joint posterior $p(\mathbf{x}_{0:T} | \mathbf{o}_{1:T})$, note that

$$\overleftarrow{\mathbf{x}}_t = \arg \max_{\mathbf{x}_t} \int p(\mathbf{x}_{0:t-1}, \mathbf{x}_t, \mathbf{x}_{t+1:T} | \mathbf{o}_{1:T}) d\mathbf{x}_{-t}, \quad (7)$$

$$\vec{\mathbf{x}}_t = \arg \max_{\mathbf{x}_t} \int p(\mathbf{x}_{0:t-1}, \mathbf{x}_t, \vec{\mathbf{x}}_{t+1:T} | \mathbf{o}_{1:T}) d\mathbf{x}_{0:t-1}, \quad (8)$$

$$\hat{\mathbf{x}}_t = \arg \max_{\mathbf{x}_t} p(\hat{\mathbf{x}}_{0:t-1}, \mathbf{x}_t, \hat{\mathbf{x}}_{t+1:T} | \mathbf{o}_{1:T}). \quad (9)$$

The problem with marginal MAP (7) is that $\overleftarrow{\mathbf{x}}_{0:T}$ can contain implausible or zero transitions, while MAP predecessor decoding (8) yields coherent sequences $\vec{\mathbf{x}}_{0:T}$ as demonstrated in [7]. The joint MAP assignment $\hat{\mathbf{x}}_{0:T}$ in (9) is also coherent but hard to compute for general models [4, 2].

5. Conclusion

We elaborate on a novel MAP inference method, denoted predecessor decoding, distinguishing it from conventional inference methods used in discrete-time state observation models. Similar to marginal MAP, predecessor decoding is performed by maximizing marginal distributions, yet the sequence of predecessors is coherent, a characteristic akin to joint MAP. This is achieved by inferring the predecessors recursively where each marginal is conditioned on the previous result. Beyond an initialization, which can be chosen freely, MAP predecessor decoding only requires the distributions encountered in forward filtering, thus its applicability is inherent wherever a forward filter is already in use.

REFERENCES

1. *Särkkä S.* Bayesian Filtering and Smoothing. Cambridge: Univ. Press, 2013.
2. *Koller D., Friedman N.* Probabilistic Graphical Models: Principles and Techniques. Cambridge: MIT Press, 2009. 1208 p.
3. *Gustafsson F.* Particle Filter Theory and Practice with Positioning Applications // IEEE Aerosp. Electron. Syst. Mag. 2010. Vol. 25. № 7. P. 53–82.
4. *Godsill S., Doucet A., West M.* Maximum a Posteriori Sequence Estimation Using Monte Carlo Particle Filters // Ann. Inst. Statist. Math. 2001. Vol. 53. № 1. P. 82–96
5. *Rauch H., Tung F., Striebel C.* Maximum likelihood estimates of linear dynamic systems // AIAA Journal. 1965. Vol. 3. № 8. P. 1445–1450.
6. *Ainsleigh P., Kehtarnavaz N., Streit R.* Hidden Gauss-Markov Models for Signal Classification // IEEE Trans. Signal Process. 2002. Vol. 50. № 6.
7. *Rudić B., Pichler-Scheder M., Efrrosinin D.* Valid Decoding in Gaussian Mixture Models // 2024 IEEE 3rd Conference on Information Technology and Data Science (CITDS), Debrecen, Hungary: accepted, 2024. P. 1–6.
8. *Rudić B., Sturm V., Efrrosinin D.* On the Analytic Solution to Recursive Smoothing in Mixture Models // 2024 58th Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, CA, USA: submitted, 2024. P. 1–5.

Branislav Rudić — Dipl.-Ing., R&D Engineer, Sensors and Communication, Linz Center of Mechatronics. E-mail: branislav.rudic@lcm.at

Markus Pichler-Scheder — Dr. Dipl.-Ing., Team Leader, Sensors and Communication, Linz Center of Mechatronics. E-mail: markus.pichler-scheder@lcm.at

Dmitry Efrrosinin — Assoz. Univ.-Prof. Dr., Deputy Head, Institute of Stochastics, Johannes Kepler University Linz. E-mail: dmitry.efrosinin@jku.at

MESHFREE LEAST SQUARES MODELLING APPLIED TO THE TIME-DEPENDENT SCHRÖDINGER EQUATION

M. A. Sánchez-Benavides¹, F. R. Saucedo-Zendejo²

¹*Facultad de Ciencias Físico Matemáticas, Universidad Autónoma de Coahuila.
Unidad Saltillo, 25280, Saltillo, Coahuila, México*

²*Centro de Investigación en Matemáticas Aplicadas, Universidad Autónoma de
Coahuila. Unidad Saltillo, 25280, Saltillo, Coahuila, México*

In this report a Generalized Finite Difference Method (GFDM) based on the least squares method is presented and discussed in order to solve the time-dependent Schrödinger equation. In this truly meshfree formulation, the approximation of the field variables depends on both the governing equations and the local problem discretization. It is applied for the modelling of three benchmark examples, whose numerical results suggest that the proposed mesh-free formulation is feasible and promising for solving problems governed by the time-dependent Schrödinger equation. **Keywords:** *Schrödinger equation, meshfree method, GFDM.*

Introduction

There have been several methods that were developed or have evolved in the last few years with the objective of solving linear and non-linear Partial Differential Equations (PDEs). The main problem of solving these types of equations is to solve the spatial part, specially when the spatial domain is evolving through time [1]. By adding the spatial solution to a time discretization, one can obtain a full solution to the problem.

Some of the methods used to solve this kind of equations include Haar Wavelets [2], Crank-Nicolson finite difference [3], Fourier Pseudo-Spectral method [4], Semi-implicit finite difference method [6], and the one presented to this article, the Least Squares Method [5], among others.

1. Problem Statement

The problem to be solved is the Linear Schrödinger Equation (LSE). To solve it, the meshfree Least Squares Method, or simply referred to as Least Squares Method (LSM), is employed to discretize the second order partial derivatives in the space domain, and to discretize the time domain, a Runge-Kutta of fourth order scheme is considered [7]. Schrödinger's model can be expressed as follows

$$i \frac{\partial \psi(\mathbf{x}, t)}{\partial t} = -\frac{1}{2} \Delta \psi + \hat{V}(\mathbf{x}) \psi(\mathbf{x}, t), \tag{1}$$

where $i = \sqrt{-1}$, Δ is the laplacian operator, $\hat{V}(\mathbf{x})$ is the potential operator, $\psi(\mathbf{x}, t)$ is the wave function. By means of simplicity in the present article the one-dimensional case is considered and the problem is solved in cartesian coordinates. The boundary condition is of Dirichlet type, which can be expressed as

$$\psi(x, t) = f(x) | x = a \vee x = b \tag{2}$$

To be precise the function $f(x)$ used was:

$$\psi(x, t) = 0 | x = a \vee x = b. \tag{3}$$

One important conservation law that the equation implies is the conservation of charge which can be written as:

$$Q(t) = \int_a^b |\psi(x, t)|^2 dx = 1 \quad \forall \quad t > 0. \tag{4}$$

Also, if the total energy is a constant over time, this means:

$$E = \int_a^b \left(-\frac{1}{2} \frac{\partial^2 \psi}{\partial x^2} + \hat{V}(x) \psi \right) dx \quad \forall \quad t, \tag{5}$$

$$\frac{\partial E}{\partial t} = 0. \tag{6}$$

For the numerical solution, it is assumed that the function starts at $t = 0$, and the initial condition is:

$$\psi(x, t = 0) = \psi^0(x). \tag{7}$$

The superscript indicates the time layer $m | m = 0, 1, 2, \dots, \tau$, the time will be split by τ , so that the time step will be given by $\Delta t = T/\tau$, where T is the time where the simulation ends. ψ_0 is a solution to the LSE, the challenge then becomes and see how it evolves through time. The domain will be conformed by the set of points $x_i | i = 1, 2, 3, \dots, N$, they will be separated a distance Δx from their adjacents and each point will have associated a discretization of the wave function $\psi_i^m = \psi^m(x_i)$.

2. Numerical Formulation

As it was indicated above, for the solution of the LSE, the LSM is employed to discretize the second order partial derivatives in the space domain, and to discretize the time domain, a Runge-Kutta of fourth order is considered. Therefore, along this section, the LSM is introduced. As said before the challenge resides in computing the second order partial derivative in the LSE, since the potential function can be calculated directly. Supposing the domain has been discretized so that the position vectors are given by $\mathbf{x}_i | i = 1, 2, 3, \dots, N$ and the adjacent points to a point x_i are $\mathbf{x}_j | j = 1, 2, 3, \dots, n$, where these adjacent points have a weight gaussian function given by:

$$\omega(\mathbf{x}_i, \mathbf{x}_j, h, \alpha) = \begin{cases} \exp(-\alpha \frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{h^2}) & \frac{\|\mathbf{x}_i - \mathbf{x}_j\|}{h} \leq 1 \\ 0 & \text{else.} \end{cases} \tag{8}$$

Now by doing a Taylor expansion around \mathbf{x}_i until the second term of the expansion and evaluating this expansion on one of the adjacent points, the expression would be

$$f(\mathbf{x}_j) = f(\mathbf{x}_i) + \sum_{k=1}^3 f_k(\mathbf{x}_i)(x_{kj} - x_{ki}) + \frac{1}{2} \sum_{k,l=1}^3 f_{kl}(\mathbf{x}_i)(x_{kj} - x_{ki})(x_{lj} - x_{li}). \tag{9}$$

where $\mathbf{x}_{1i}, \mathbf{x}_{2i}$ and \mathbf{x}_{3i} are the x, y and z components of the vector \mathbf{x}_i , and the unknowns f, f_k and f_{kl} , f is the function evaluated at point \mathbf{x}_i , f_k are the derivatives of first order and f_{kl} are the derivatives of second order. Viewing the set of equations on its matrix form, the following equation is found:

$$M\mathbf{a} \approx \mathbf{b}, \tag{10}$$

where:

$$M = \begin{pmatrix} 1 & dx_1 & dy_1 & dz_1 & \frac{1}{2}dx_1^2 & dx_1dy_1 & dx_1dz_1 & \frac{1}{2}dy_1^2 & dy_1dz_1 & \frac{1}{2}dz_1^2 \\ 1 & dx_2 & dy_2 & dz_2 & \frac{1}{2}dx_2^2 & dx_2dy_2 & dx_2dz_2 & \frac{1}{2}dy_2^2 & dy_2dz_2 & \frac{1}{2}dz_2^2 \\ \vdots & \vdots \\ 1 & dx_n & dy_n & dz_n & \frac{1}{2}dx_n^2 & dx_ndy_n & dx_ndz_n & \frac{1}{2}dy_n^2 & dy_ndz_n & \frac{1}{2}dz_n^2 \end{pmatrix}$$

$$\mathbf{a} = (f \quad f_1 \quad f_2 \quad f_3 \quad f_{11} \quad f_{12} \quad f_{13} \quad f_{22} \quad f_{23} \quad f_{33})^T \tag{11}$$

$$\mathbf{b} = (f(\mathbf{x}_1) \quad f(\mathbf{x}_2) \quad \dots \quad f(\mathbf{x}_n))^T. \tag{12}$$

The error then would be given by $e = M\mathbf{a} - \mathbf{b}$. By minimizing it with a weighted least squares method through the quadratic form $J = \sum_{i=1}^n \omega_i e_i^2$, the vector \mathbf{a} is obtained as:

$$\mathbf{a} = (M^T W M)^{-1} (M^T W) \mathbf{b}. \tag{13}$$

where W is given by

$$W = \begin{bmatrix} \omega(\mathbf{x}, \mathbf{x}_1) & 0 & \dots & 0 \\ 0 & \omega(\mathbf{x}, \mathbf{x}_2) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \omega(\mathbf{x}, \mathbf{x}_n) \end{bmatrix}. \tag{14}$$

The fifth element of vector \mathbf{a} is the second order derivative that has to be computed and is the only element that is going to be needed in the Runge-Kutta scheme. The method should be applied until a desired error is achieved.

3. Numerical Results

In this section the numerical model presented above is assessed considering the benchmark example of a particle in a box, which is given when the potential in the LSE takes the form

$$\hat{V}(x) = 0.$$

For this case the domain will be of length $L = 1$ and $x \in [0, 1]$, the time step was chosen to be $\Delta t = 1 \times 10^{-3}$ and the spacing between points to be $\Delta x = 0.01$. The analytical solution for the case $n = 1$ is given by:

$$\psi_1(x, t) = \sqrt{2} \sin(\pi x) e^{-\frac{it}{2}}.$$

A comparison of the real part of the numerical and the analytical solutions, and their evolutions over time are shown in Figure 1.

This figure shows that both solutions match very well. Similar results could be observed if the imaginary part of the solutions are compared. Moreover, if the case $n = 2$ is considered, a similar behaviour is observed. Therefore, it indicates that the numerical results computed with the presented numerical approach converge to the exact solutions and do not diverge through time for the particle in the box example.

To further validate these results, two additional numerical examples were considered. They correspond to the Quantum Harmonic Oscillator, and the

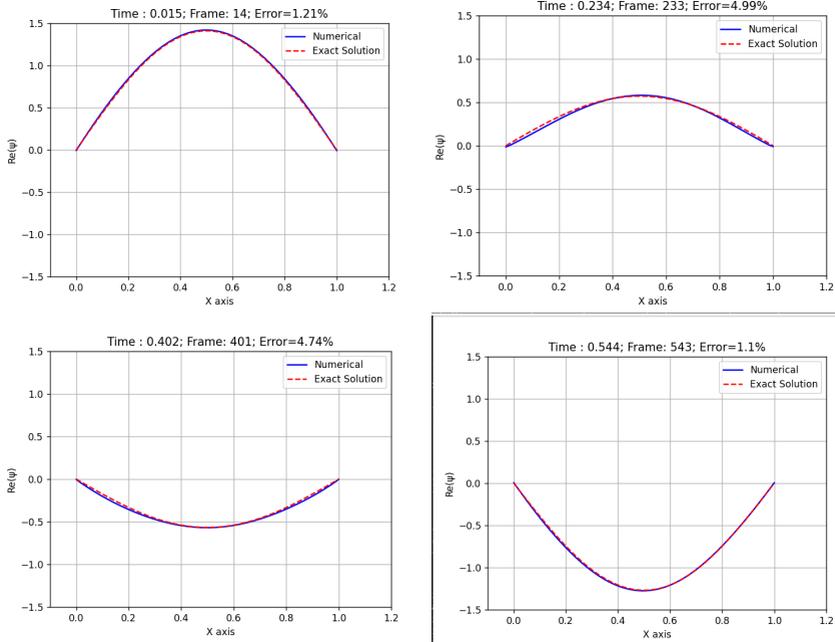


Figure 1. Comparison of the numerical and analytical solutions for $n = 1$.

inverse squared type of potential. For the second example, the domain will be of length $L = 6$ and $x \in [-3, 3]$, the time step was chosen to be $\Delta t = 1 \times 10^{-3}$ and the spacing between points to be $\Delta x = 0.06$, whilst for the third, the domain will be of length $L = 2$, the domain will be $x \in [-1, 1]$ divided in 10 points, the time step was chosen to be $\Delta t = 0.001$. As for the example of the particle in the box, the numerical solutions match very well with their analytical counterparts.

4. Conclusion

The meshfree numerical formulation discussed and presented in this report for the first time has demonstrated to be feasible for the numerical modelling of time-dependent Schrödinger Equation governed problems. Throughout the test examples it could be observed that the method is capable of reproducing with good accuracy the shape of the analytical solution over time. It has a good behaviour to capture the nature of the problem being modelled. Therefore, the range of application of meshfree methods can be considered extended with the results of this work.

REFERENCES

1. *Nguyen V. P., and Rabczuk T., and Bordas S. and Duflot M.* Meshless methods: a review and computer implementation aspects // Mathematics and computers in simulation 2000. Vol. 79. P. 763–813.
2. *Ahsan M., Ahmad I., Ahmad M., Hussian I.* A numerical Haar wavelet-finite difference hybrid method for linear and non-linear Schrödinger equation // Mathematics and Computers in Simulation 2019. Vol. 165. P. 13–25.
3. *Crank J. and Nicolson P.* A practical method for numerical evaluation of solutions of partial differential equations of the heat-conduction type // Cambridge University Press, 2008.
4. *Gong Y., Wang Q., Wang Y., Cai J.* A conservative Fourier pseudo-spectral method for the nonlinear Schrödinger equation // Journal of Computational Physics 2017. Vol. 328. P. 354–370.
5. *Saucedo-Zendejo F. R., Resendiz-Flores E. O.* A new approach for the numerical simulation of free surface incompressible flows using a meshfree method // Computer Methods in Applied Mechanics and Engineering 2017. Vol. 324. P. 619–639.
6. *Tiwari S. and Kuhnert J.* Finite pointset method based on the projection method for simulations of the incompressible Navier-Stokes equations // Meshfree methods for partial differential equations, 2003. Springer Berlin Heidelberg. P. 354–370.
7. *Quarteroni A., Sacco R., Saleri F.* Numerical mathematics // Springer, 2010.

Sánchez-Benavides, Marco A. — Bachelor in Physics Engineering, MSc Student, Facultad de Ciencias Físico Matemáticas, Universidad Autónoma de Coahuila, Mexico. E-mail: *marco.aurelio@live.com.mx*

Saucedo-Zendejo, Félix R. — Doctor of Science, Dr., Full Time Professor, Centro de Investigación en Matemáticas Aplicadas, Universidad Autónoma de Coahuila, Mexico. E-mail: *fesaucedoz@uadec.edu.mx*

PERIODIC GIBBS MEASURES FOR THE POTTS-SOS MODEL WITH AN EXTERNAL FIELD

M. Rahmatullaev^{1,2}, M. Rasulova^{1,2}

¹*V.I.Romanovskiy Institute of Mathematics, Tashkent, Uzbekistan*

²*Namangan State University, Namangan, Uzbekistan*

We consider Potts-SOS model with an external field, with spin values $0, 1, 2, \dots, m$, on the Cayley tree of order $k \geq 1$. We study the structure of all periodic Gibbs measures with respect to arbitrary normal divisors. We prove that periodic Gibbs measures with respect to an arbitrary normal divisor are either translation-invariant or have a period equal to two. In addition we find a condition for the normal divisor that ensures that all periodic Gibbs measures are translation-invariant. **Keywords:** *Cayley tree, Potts-SOS model with an external field, Gibbs measure, translation-invariant Gibbs measure, periodic Gibbs measure.*

Introduction

In recent years, worldwide interest in phase transition problems for statistical mechanics models on the Cayley tree has been growing. In particular, works [1], [2] are devoted to these issues.

[2] covers the issues discussed on the Potts model, In [4], [5], [9], phase transition issues for the SOS model were solved, Gibbs measures for the Potts-SOS model are studied in [6]-[8].

In this work, Potts-SOS model with an external field is considered, this model is a generalization of the Potts and SOS model with an external field added.

1. Cayley tree

The Cayley tree Γ^k (See [1]-[3]) of order $k \geq 1$ is an infinite tree, i.e., a graph without cycles, from each vertex of which exactly $k + 1$ edges issue. Let $\Gamma^k = (V, L, i)$, where V is the set of vertices of Γ^k , L is the set of edges of Γ^k and i is the incidence function associating each edge $l \in L$ with its endpoints $x, y \in V$. If $i(l) = \{x, y\}$, then x and y are called *nearest neighboring vertices*, and we write $l = \langle x, y \rangle$.

It is known (see [3]) that there exists a one-to-one correspondence between the set V of vertices of the Cayley tree of order $k \geq 1$ and the group G_k of the free products of $k + 1$ cyclic groups $\{e, a_i\}$, $i = 1, \dots, k + 1$ of the second order (i.e. $a_i^2 = e$, $a_i^{-1} = a_i$) with generators a_1, a_2, \dots, a_{k+1} .

We denote by $S(x)$ the set of direct successors of $x \in G_k$.

2. The model

We consider model, where the spin takes values in the set $\Phi = \{0, 1, 2, \dots, m\}$, $m \geq 1$.

For $A \subseteq V$ a spin configuration σ_A on A is defined as a function $x \in A \mapsto \sigma_A(x) \in \Phi$; the set of all configurations coincides with $\Omega_A = \Phi^A$.

The Hamiltonian of the Potts-SOS model with an external field has the form

$$H(\sigma) = -J_p \sum_{\langle x,y \rangle \in L} \delta_{\sigma(x)\sigma(y)} - J \sum_{\langle x,y \rangle \in L} |\sigma(x) - \sigma(y)| + \alpha \sum_{x \in V} \sigma(x) \quad (1)$$

where $J_p, J \in \mathbb{R}$ are nonzero coupling constants, $\alpha \in \mathbb{R}$ is an external field, $\sigma \in \Phi^V$.

Remark 1. Note that model (1) coincides with the Potts-SOS model, if $\alpha = 0$. For the Potts-SOS model, translation-invariant Gibbs measures were studied in [6],[7], periodic Gibbs measures were studied in [8].

3. The compatibility condition

Let $h : x \mapsto h_x = (h_{0,x}, h_{1,x}, \dots, h_{m,x}) \in \mathbb{R}^{m+1}$ be a real vector-valued function of $x \in V$.

We consider the probability distributions $\mu^{(n)}$ on Φ^{V_n} for given $n = 1, 2, \dots$ defined by

$$\mu^{(n)}(\sigma_n) = Z_n^{-1} \exp(-\beta H(\sigma_n) + \sum_{x \in W_n} h_{\sigma(x),x}), \quad (2)$$

where $\sigma_n \in \Phi^{V_n}$ and the partition function Z_n can be expressed as follows:

$$Z_n(\tilde{\sigma}_n) = \sum_{\tilde{\sigma}_n \in \Phi^{V_n}} \exp(-\beta H(\tilde{\sigma}_n) + \sum_{x \in W_n} h_{\tilde{\sigma}(x),x}). \quad (3)$$

We say that the probability distributions $\mu^{(n)}$ are compatible if $\forall n \geq 2$ and $\sigma_{n-1} \in \Phi^{V_{n-1}}$:

$$\sum_{\omega_n \in \Phi^{W_n}} \mu^{(n)}(\sigma_{n-1} \vee \omega_n) = \mu^{(n-1)}(\sigma_{n-1}), \quad (4)$$

here $(\sigma_{n-1} \vee \omega_n) \in \Phi^{V_n}$ is the concatenation of σ_{n-1} and ω_n .

If for probability distributions $\mu^{(n)}$ on Φ^{V_n} the equality (4) holds, then then by the Kolmogorov extension theorem there is a unique measure μ on Φ^V such that, $\forall n$ and $\sigma_n \in \Phi^{V_n}$,

$$\mu(\{\sigma \mid \sigma_{V_n} = \sigma_n\}) = \mu^{(n)}(\sigma_n).$$

The measure μ is called splitting Gibbs measure (SGM) corresponding to Hamiltonian H and function $h : x \mapsto h_x, x \neq x^0$.

4. The main results

The following theorem expresses the requirement on function h that the distributions $\mu^{(n)}(\sigma_n)$ hold the compatibility conditions.

Theorem 1. The probability distributions (2) satisfy the compatibility condition iff

$$h_x^* = \sum_{y \in S(x)} F(h_y^*, m, \rho, \theta, \gamma) \quad (5)$$

is satisfied for any $x \in V \setminus \{x^0\}$, where

$$\rho = \exp(J_p \beta), \quad \theta = \exp(J \beta), \quad \gamma = \exp(-\alpha \beta)$$

and also $\beta = 1/T$ is the inverse temperature. Here h_x^* represents the vector

$$(h_{0,x} - h_{m,x}, h_{1,x} - h_{m,x}, \dots, h_{m-1,x} - h_{m,x})$$

and the vector function $F(\cdot, m, \rho, \theta, \gamma) : \mathbb{R}^m \rightarrow \mathbb{R}^m$ is defined as follows $F(h, m, \rho, \theta, \gamma) = (F_0(h, m, \rho, \theta, \gamma), F_1(h, m, \rho, \theta, \gamma), \dots, F_{m-1}(h, m, \rho, \theta, \gamma))$, where

$$F_i(h, m, \rho, \theta, \gamma) = \ln \frac{\sum_{j=0}^{m-1} \rho^{\delta_{ij}} \theta^{|i-j|} \gamma^j \exp(h_j) + \rho^{\delta_{mi}} \theta^{m-i} \gamma^m}{\sum_{j=0}^{m-1} \rho^{\delta_{mj}} \theta^{m-j} \gamma^j \exp(h_j) + \rho \gamma^m},$$

$$h = (h_0, h_1, \dots, h_{m-1}), i = 0, 1, 2, \dots, m-1.$$

Namely, for any collection of functions satisfying the functional equation (5) there exists a unique splitting Gibbs measure, the correspondence being one-to-one.

Let $G_k/G_k^* = \{H_1, \dots, H_r\}$ be a factor group, where G_k^* is a normal subgroup of index $r \geq 1$.

Definition 1. Let G_k^* be a subgroup of G_k . A collection of vectors $h = \{h_x \in \mathbb{R}^m : x \in G_k\}$ is said to be G_k^* -periodic, if $h_{yx} = h_x$ for all $x \in G_k$ and $y \in G_k^*$. A G_k -periodic collections is said to be *translation-invariant*.

Definition 2. A Gibbs measure is called G_k^* -periodic (*translation-invariant*), if it corresponds to G_k^* -periodic (*translation-invariant*) collection h .

Let $G_k^{(2)}$ be the subgroup in G_k consisting of all words of even length. Clearly, $G_k^{(2)}$ is a subgroup of index 2.

Theorem 2. Let G_k^* be a finite-index normal subgroup in G_k . Then each G_k^* -periodic GM for the Potts-SOS model with an external field is either translation-invariant or $G_k^{(2)}$ -periodic.

Let G_k^* be a normal subgroup of finite index in G_k . What condition on G_k^* will guarantee that each G_k^* -periodic GM is translation-invariant? We put $I(G_k^*) = G_k^* \cap \{a_1, a_2, \dots, a_{k+1}\}$, where $a_i, i = 1, 2, \dots, k+1$ are generators of G_k .

Theorem 3. If $I(G_k^*) \neq \emptyset$ then each G_k^* -periodic GM for Potts-SOS model with an external field is translation-invariant.

5. Conclusion

In conclusion, it is proved that all periodic Gibbs measures with respect to the normal divisor with an arbitrary finite index for the Potts-SOS model with an external field on the Cayley tree are translation-invariant or $G_k^{(2)}$ -periodic. These scientific results summarize similar results obtained for the Potts and SOS models.

REFERENCES

1. *Rozikov U.* Gibbs measures on Cayley trees. Singapore: World scientific, 2013. 398 p.
2. *Rozikov U.* Gibbs Measures in Biology and Physics: The Potts Model. Singapore: World scientific, 2023. 314 p.
3. *Ganikhodzhaev, N.* Group representation and automorphisms of the Cayley tree // Dokl. Akad. Nauk Resp. Uzbekistan. 1994. № 4. P. 3–5.
4. *Rozikov U., Suhov Y.* Gibbs measures for SOS model on a Cayley tree // Infinite Dimensional Analysis, Quantum Probability and Related Topics. 2006. № 9. P. 471–488.
5. *Kulske C., Rozikov U.* Extremality of translation-invariant phases for three-state SOS-model on the binary tree // Journal of Statistical Physics. 2015. № 160. P. 659–680.
6. *Saygılı H.* Gibbs measures for the Potts-SOS model with three states of spin values // Asian Journal of Current Research. 2017. Vol. 1. № 3. P. 114–121.
7. *Rahmatullaev M., Rasulova M.* Extremality of translation-invariant Gibbs measures for the Potts-SOS model on the Cayley tree // Journal of Statistical Mechanics: Theory and Experiment. 2021. № 7. 073201. P. 1–18.
8. *Rasulova M.* Periodic Gibbs measures for the three-state Potts-SOS model on a Cayley tree // Uzbek Mathematical Journal. 2022. Vol. 66. № 2. P. 150–155.

9. *Karshiboev O.* Periodic Gibbs measures for the three-state SOS model on a Cayley tree with a translation-invariant external field // Theoretical and Mathematical Physics. 2022. Vol. 212. № 3. P. 1276–1283.

Rahmatullaev Muzaffar — DSc, associate professor, Head of the Naman-gan branch of the Institute of Mathematics, mathematics. E-mail: *mrahmatullaev@rambler.ru*

Rasulova Muhayyo — PhD, lead teacher, senior researcher, mathematics. E-mail: *m_rasulova_a@rambler.ru*

FISHER INFORMATION UNDER RANDOM CENSORING ON THE RIGHT

A. Abdushukurov¹, N. Nurmukhamedova²

¹*Moscow State University named after M. V. Lomonosov, Tashkent Branch, Tashkent, Uzbekistan*

²*National University of Uzbekistan named after Mirzo Ulugbek, Tashkent, Uzbekistan*

In this paper, we consider Fisher information under the conditions of random censoring. We present numerical results for calculating Fisher information for the multiple random right censoring model. Special attention is given to analyzing the impact of various parameters on the estimation of Fisher information, as well as a comparative analysis with traditional estimation methods. The results obtained can be useful for further research and application in the statistical analysis of censored data. **Keywords:** *Fisher information, distribution family, censorship, incomplete observation, numerical analysis.*

Introduction

R.A. Fisher [1] introduced the concept of Fisher information in 1925 as a means of comparing statistical data. He studied the loss of accuracy using different estimates of the unknown parameter, such as median or maximum likelihood estimates. Fisher's information is used in the Cramer-Rao inequality and provides a lower bound for the variance of the unbiased estimator, and is related to the asymptotic variance of the maximum likelihood estimates. In [2], Fisher information is calculated for both the Weibull distribution and the generalized exponential distribution under scenarios that include complete and right-censored samples, where the censoring random variable (r.v.) is assumed to be constant. And in [3], Fisher information is studied for a Type I and II censored sample, in which the censoring random variable is also a constant. In [4], the lower bounds, specifically the Cramer-Rao and Bhattacharyya lower bounds system, are studied under random censoring. This paper explains the Fisher information for different models of censoring.

1. Fisher information

Fisher information is a fundamental concept in statistical theory, quantifying the amount of information that observed data provide about an

unknown parameter. It plays a crucial role in various aspects of statistical inference, such as estimation, hypothesis testing, and the construction of confidence intervals.

Let ξ be the random variable of interest, defined on the probability space (Ω, \mathcal{A}, P) , with values in $\mathcal{X} \subseteq \mathbb{R}^1 = (-\infty, \infty)$. By $F(x)$, $x \in \mathbb{R}^1$, we denote the distribution function (d.f.) of r.v. ξ . Let us consider the parametric case when the d.f. F is defined with precision to the parameter θ : $F(x; \theta) = P_\theta(\xi < x)$, $P = \{P_\theta, \theta \in \Theta\}$ where $\theta \in \Theta$, θ -open interval in \mathbb{R}^1 .

Definition: The Fisher information $I(\theta)$ for the parameter θ in a model with probability density $f(x, \theta)$ is defined as:

$$I(\theta) = \mathbb{E} \left[\left(\frac{\partial}{\partial \theta} \log f(\xi, \theta) \right)^2 \right],$$

where the expectation is taken with respect to the distribution $f(\xi, \theta)$.

The Fisher information $I(\theta)$ can be interpreted as the variance of the score function $U(\theta) = \frac{\partial}{\partial \theta} \log f(\xi, \theta)$, which measures the sensitivity of the likelihood function to changes in the parameter θ . A higher Fisher information indicates that the data provide more information about the parameter, leading to more precise parameter estimates.

Fisher information has several important properties:

- *nonnegativity*: the Fisher information is always nonnegative, $I(\theta) \geq 0$;
- *additivity*: for independent and identically distributed (i.i.d.) random variables $\xi_1, \xi_2, \dots, \xi_n$, the Fisher information for the entire sample is the sum of the Fisher information for each individual observation. Formally, if each ξ_i has Fisher information $I(\theta)$, then the total Fisher information $I_n(\theta)$ for the sample is given by: $I_n(\theta) = nI(\theta)$.

This property reflects the fact that independent observations provide additive information about the parameter.

This relationship underscores the pivotal role of Fisher information in parameter estimation, particularly in quantifying the efficiency of estimators.

We consider Fisher's information in different models of censoring. Censoring observations is a situation where information about some values of observations is missing or limited. This is common in survival analysis, where, for example, data may be right-sidedly censored if the event of interest (e.g., death) did not occur before the end of the observation period. Fisher's information for censored data takes into account both the observed values and the censoring structure. Fisher's information measures the amount of

information that observations contain about a model parameter. Censoring reduces the amount of information available about a parameter, which reduces the value of Fisher’s information. Complete observations contain more information than censored observations.

Consider a time-to-event dataset where the data is censored on the right. Let ξ_i be the time to event for the i th individual, η_i be the censored time, $\zeta = \min(\xi_i, \eta_i)$, $\delta_i = I(\xi_i \leq \eta_i)$. Suppose that $\xi_i \sim \text{Exp}(\theta)$ (exponential distribution with parameter θ). Then the logarithmic likelihood function is:

$$\begin{aligned} \ell(\theta) &= \sum_{i=1}^n [\delta_i \log f(X_i, \theta) + (1 - \delta_i) \log(1 - F(X_i, \theta))] = \\ &= \sum_{i=1}^n [\delta_i \log(\theta e^{-\theta X_i}) + (1 - \delta_i) \log(e^{-\theta X_i})] = \sum_{i=1}^n [\delta_i \log \theta - \theta X_i], \end{aligned}$$

and Fisher Information:

$$I(\theta) = -\mathbb{E} \left[\frac{\partial^2 \ell(\theta)}{\partial \theta^2} \right] = \sum_{i=1}^n \frac{\delta_i}{\theta^2}.$$

In this model, under the assumption of complete observations, the Fisher information is $I(\theta) = \frac{n}{\theta^2}$, i.e. $\frac{n}{\theta^2} \geq \frac{\sum_{i=1}^n \delta_i}{\theta^2}$. But for some distributions this information is not explicitly computed and we have to use numerical methods. Now let us consider the Fisher information in the multiple random censoring model on the right.

2. Fisher information in the multiple random censoring model on the right

Let ξ be a non-negative r.v., denoting the duration life of the tested object is defined on the probability space (Ω, \mathcal{A}, P) with values in the measurable space $(\mathcal{X}, \mathcal{B})$, where $\mathcal{X} \subseteq R^+$ and $\mathcal{B} = \sigma(\mathcal{X})$. Let $F(x; \theta) = P_\theta(\xi < x)$ and $\xi_{1n} < \xi_{2n} < \dots < \xi_{nn}$ be the variational series obtained from r.v. $\{\xi_1, \dots, \xi_n\}$. Then the joint density of the vector $\xi^{(n)}$ is given by the formula

$$p_n(x^{(m)}; \theta) = \frac{n!}{(n - m)!} \prod_{i=1}^m f^{(i)}(x_{in}; \theta) [1 - F(x_{mn}; \theta)]^{n-m},$$

where $x^{(m)} = (x_{1n}, \dots, x_{mn})$ is the realization of the vector $\xi^{(n)}$. Let $\beta \in (0, 1]$ be such that at $n \rightarrow \infty$, $\frac{x}{n} \rightarrow \beta$ and $h(x; \theta) = f(x; \theta)/(1 - F(x; \theta))$ -intensity function. In this scheme, the Fisher information is defined as follows:

$$I_\beta(\theta) = \int_0^{\lambda\beta} \left(\frac{\partial \log h(\xi; \theta)}{\partial \theta} \right)^2 f(\xi; \theta) dx, \tag{1}$$

where λ_β is the quantile of the level β of the distribution $F(x; \theta)$.

Consider the case when the random variable ξ has a gamma distribution ($Gamma(\theta, k)$), where θ is scale parameter. For this distribution, the integral (1) cannot be solved analytically, necessitating numerical methods to compute its value. Let us show on the graph the approximate values of Fisher information at different values of β (see Figure 6)).

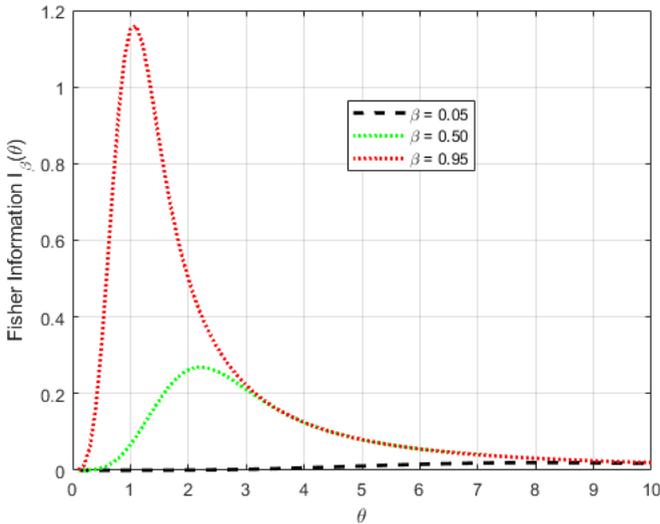


Figure 1. Fisher information in the multiple random censoring model on the right

The following table summarizes Fisher information values in the multiple random censoring model on the right. The table shows that the Fisher information value is larger when the quantile tends to 1 (see Table 1).

3. Conclusion

In the context of estimating unknown parameters from incomplete observations, the relationship between the censoring distribution and the parameter of interest plays a significant role. This relationship distinguishes two primary cases: when the distribution of censoring random variables depends on the parameter of interest, and when it does not. In statistical theory, it is recognized that censoring, especially when the censoring mechanism is parameter-dependent, can diminish the Fisher information available for parameter estimation. Fisher information quantifies the amount of information that observed data provide about a parameter, and its loss due to censoring underscores the challenges in accurately estimating parameters

Table 1

Fisher information values in the multiple random censoring model on the right

θ	$\beta = 0.05$	$\beta = 0.50$	$\beta = 0.95$
0.1000	0.0000	0.0000	0.0009
0.2000	0.0000	0.0002	0.0131
0.3000	0.0000	0.0009	0.0594
0.4000	0.0000	0.0026	0.1609
0.5000	0.0000	0.0061	0.3229
0.6000	0.0000	0.0119	0.5294
0.7000	0.0000	0.0207	0.7475
0.8000	0.0000	0.0327	0.9404
0.9000	0.0000	0.0481	1.0795
1.0000	0.0000	0.0666	1.1517
1.5000	0.0002	0.1833	0.8569
2.0000	0.0006	0.2620	0.4997
2.5000	0.0014	0.2581	0.3200
3.0000	0.0026	0.2106	0.2222
3.5000	0.0042	0.1617	0.1633
4.0000	0.0063	0.1249	0.1250
4.5000	0.0086	0.0988	0.0988
5.0000	0.0111	0.0800	0.0800
5.5000	0.0135	0.0661	0.0661
6.0000	0.0157	0.0556	0.0556
6.5000	0.0174	0.0473	0.0473
7.0000	0.0188	0.0408	0.0408
7.5000	0.0196	0.0356	0.0356
8.0000	0.0199	0.0312	0.0313
8.5000	0.0198	0.0277	0.0277
9.0000	0.0194	0.0247	0.0247
9.5000	0.0186	0.0222	0.0222
10.0000	0.0177	0.0200	0.0200

from incomplete data. Numerical methods are often employed to illustrate and quantify the impact of censoring on Fisher information. These methods demonstrate empirically that as censoring becomes more dependent on the parameter being estimated, the resulting Fisher information decreases. This reduction in Fisher information highlights the complexity introduced

by censoring mechanisms and underscores the importance of understanding and appropriately modeling censoring effects in statistical inference.

REFERENCES

1. *Fisher, R.A.* Theory of Statistical Estimation // Proceedings of the Cambridge Philosophical Society, 22 1925. P. 700–725.
2. *Rameshwar D. Gupta, Debasis Kundu.* On the comparison of Fisher information of the Weibull and GE distributions // Journal of Statistical Planning and Inference. 2006. Vol. 136. № 9. P. 3130–3144.
3. *Yanhua Wang, Shuyuan He.* Fisher information in censored data // Statistics & Probability Letters. 2005. Vol. 73. P. 199–206.
4. *Abdushukurov, I.A., Kim, L.V.* Lower Cramer-Rao and Bhattacharya bounds for randomly censored observations // J. Soviet Math. 1987. P. 2171–2185.

Abdushukurov Abdurakhim — Doctor of Physical and Mathematical Sciences, Professor, Applied Mathematics and Computer Science. E-mail: *a.abdushukurov@rambler.ru*

Nurmukhamedova Nargiza — Doctor of Philosophy in Physical and Mathematical Sciences, Associate Professor, Econometrics and Economic Modeling. E-mail: *rasulova.nargiza@mail.ru*

МОДУЛЯРНОСТЬ ГРАФОВ, ЭВОЛЮЦИОНИРУЮЩИХ МОДЕЛЬЮ КЛАСТЕРНОГО ПРИСОЕДИНЕНИЯ

Н. М. Маркович, М. С. Рыжов

*Институт Проблем Управления им. В.А. Трапезникова РАН,
117997, Москва, ул. Профсоюзная, д. 65*

Кластерное присоединение (КП) используется как модель эволюции случайных графов. В отличие от предпочтительного присоединения, КП ведет к легким хвостам распределений числа связей узлов и кластерам превышений модулярностью достаточно высокого уровня. Модулярность показывает связность узлов и позволяет разделять графы на сообщества. Экстремальный индекс аппроксимирует средний размер кластеров превышений достаточно высокого уровня. Рассматривая изменение модулярности на каждом шаге эволюции, экстремальный индекс случайных последовательностей модулярностей указывает последовательные большие связности узлов. Это отражает появление сообществ во время эволюции. Приведенное моделирование показывает, как параметры модели КП влияют на всплески модулярности. Дается сравнение эволюции КП с удалением узлов и связей в графе и без удаления. **Ключевые слова:** эволюция, кластерное присоединение, модулярность.

Введение

Пусть $G_t = (V_t, E_t)$ ненаправленный граф на шаге эволюции $t \geq 0$, где V_t и E_t множества узлов и ребер, соответственно. Эволюция графа представляется, как последовательность графов G_0, G_1, \dots , где G_0 начальный граф, с которого начинается эволюция. Наша цель исследовать кластерное присоединение (КП), как модель эволюции графа. КП предложено в [1] и развито в [7], [8] для ненаправленных графов. Присоединение нового узла может происходить с вероятностью¹

$$P_{CA}(i, t) \propto c_{i,t}^\alpha + \varepsilon, \quad \alpha, \varepsilon \geq 0, \quad (1)$$

Работа выполнена при финансовой поддержке гранта Российского научного фонда (грант №24-21-00183).

¹ $x \propto y$ означает, что существует ненулевая константа C такая, что $x = Cy$.

пропорциональной кластерному коэффициенту узла i

$$c_{i,t} = \begin{cases} 0, & k_{i,t} = 0 \text{ or } k_{i,t} = 1, \\ 2\Delta_{i,t}/(k_{i,t}(k_{i,t} - 1)), & k_{i,t} \geq 2, \end{cases}$$

$c_{i,t} \in [0, 1]$. $\Delta_{i,t}$ - число треугольников соединенных узлов, включающих узел i , а $k_{i,t}$ степень узла i (число его связей) в момент времени t . По КП новый узел может только присоединяться к существующему узлу i , включенному в некоторое число треугольников $\Delta_{i,t}$ в момент t для $\varepsilon = 0$. КП ведет к распределениям с легкими хвостами для числа связей узлов и всплескам последовательности модулярностей, которые формируются в ходе эволюции. Разбиение на сообщества, например, из максимума модулярности, важно для анализа случайных графов. Модулярность $Q(t)$ ненаправленного графа показывает, как много ребер существует внутри сообществ и между ними:

$$Q(t) = \frac{1}{2||E_t||} \sum_{ij} \left[A_{ij} - \frac{k_{i,t}k_{j,t}}{2||E_t||} \right] \mathbf{1}(i, j),$$

где A_{ij} элементы матрицы смежности A графа G_t , $||E_t|| = \frac{1}{2} \sum_{ij} A_{ij}$ число ребер в G_t , $\mathbf{1}(i, j)$ равно 1, когда узлы i и j в G_t принадлежат одному сообществу и равно нулю в противном случае, [3]. Цель - исследовать кластеры превышения уровня последовательностью модулярности, возникающие во время эволюции КП. Для этого оценивается экстремальный индекс (ЭИ), который аппроксимирует средний размер кластера, т.е. число превышений уровня, приходящихся на кластер [2]. ЭИ модулярности показывает последовательные большие значения связности узлов. Он отражает появление сообществ во время эволюции. Изучаются модели КП с удалением и без удаления узлов и связей.

Структура статьи следующая. Обзор известных результатов приводится в разделе 1. В разделе 2 исследуются последовательности модулярностей графов, моделируемых КП, с удалением узлов и связей и без. В разделе 3 приводится оценивание ЭИ модулярности. Статья завершается выводами.

1. Обзор известных результатов

Говорят, что стационарная последовательность $\{Y_n\}_{n \geq 1}$ случайных величин (сл.в.) с функцией распределения $F(x)$ и $M_n = \max_{1 \leq j \leq n} Y_j$ имеет ЭИ $\theta \in [0, 1]$, если для любого $0 < \tau < \infty$ существует последователь-

ность вещественнозначных чисел $u_n = u_n(\tau)$ такая, что

$$\lim_{n \rightarrow \infty} n(1 - F(u_n)) = \tau \quad \text{и} \quad \lim_{n \rightarrow \infty} P\{M_n \leq u_n\} = e^{-\tau\theta},$$

[5]. ЭИ играет ключевую роль в теории экстремальных величин и анализе данных, так как позволяет получить предельное распределение максимумов зависимых сл.в., т.е. $P\{M_n \leq u_n\} = F^{n\theta}\{u_n\} + o(1)$ при $n \rightarrow \infty$. ЭИ равен 1 для независимых одинаково распределённых сл.в.. Обратное неверно. Чем ближе θ к нулю, тем сильнее так называемая локальная зависимость (кластерность) в данных. $1/\theta$ аппроксимирует среднее число превышений достаточно высокого уровня u в кластере, т.е. средний размер кластера. Кластер определяется, например, как блок данных с хотя бы одним превышением u . Для оценивания θ будем использовать интервальную оценку [4], как одну из наиболее точных. Определив моменты превышения уровня $1 \leq S_1 < \dots < S_{N_u} \leq n$ для $\{Y_n\}_{n \geq 1}$, можно получить интервалы между последовательными превышениями $T_i = S_{i+1} - S_i$ при $i \in \{1, \dots, N_u - 1\}$. $N_u = \sum_{i=1}^n \mathbf{1}\{Y_i > u\}$ общее число превышений уровня u , $L \equiv L(u) = N_u - 1$. Интервальная оценка определяется, как

$$\hat{\theta}_n(u) = \begin{cases} \min(1, \hat{\theta}_n^1(u)), & \text{if } \max\{T_i : 1 \leq i \leq L\} \leq 2, \\ \min(1, \hat{\theta}_n^2(u)), & \text{if } \max\{T_i : 1 \leq i \leq L\} > 2, \end{cases} \quad (2)$$

$$\hat{\theta}_n^1(u) = \frac{2(\sum_{i=1}^L T_i)^2}{L \sum_{i=1}^L T_i^2}, \quad \hat{\theta}_n^2(u) = \frac{2(\sum_{i=1}^L (T_i - 1))^2}{L \sum_{i=1}^L (T_i - 1)(T_i - 2)}. \quad (3)$$

Уровень u может быть найден методом невязки, предложенным в [6].

2. Эволюция модулярности графа при кластерном присоединении с удалением и без удаления узлов и связей

Рассмотрим эволюцию КП, когда на каждом шаге эволюции вновь добавленный узел присоединяется к $m_0 \geq 2$ существующим узлам с помощью m_0 новых связей. Вероятность добавить связь между существующим узлом i и новым узлом $v \notin V_{t-1}$ равна

$$P_{CA}(i, t) = \frac{c_{i,t}^\alpha + \varepsilon}{\sum_{j \in V(t)} c_{j,t}^\alpha + \|V(t)\|\varepsilon}. \quad (4)$$

Предполагается, что не все $c_{j,t}$, $j \in V_t$ и ε равны нулю одновременно [8]. Заметим, что (4) совпадает с (1) при $C = 1 / \left(\sum_{j \in V(t)} c_{i,t}^\alpha + \|V(t)\|\varepsilon \right)$.

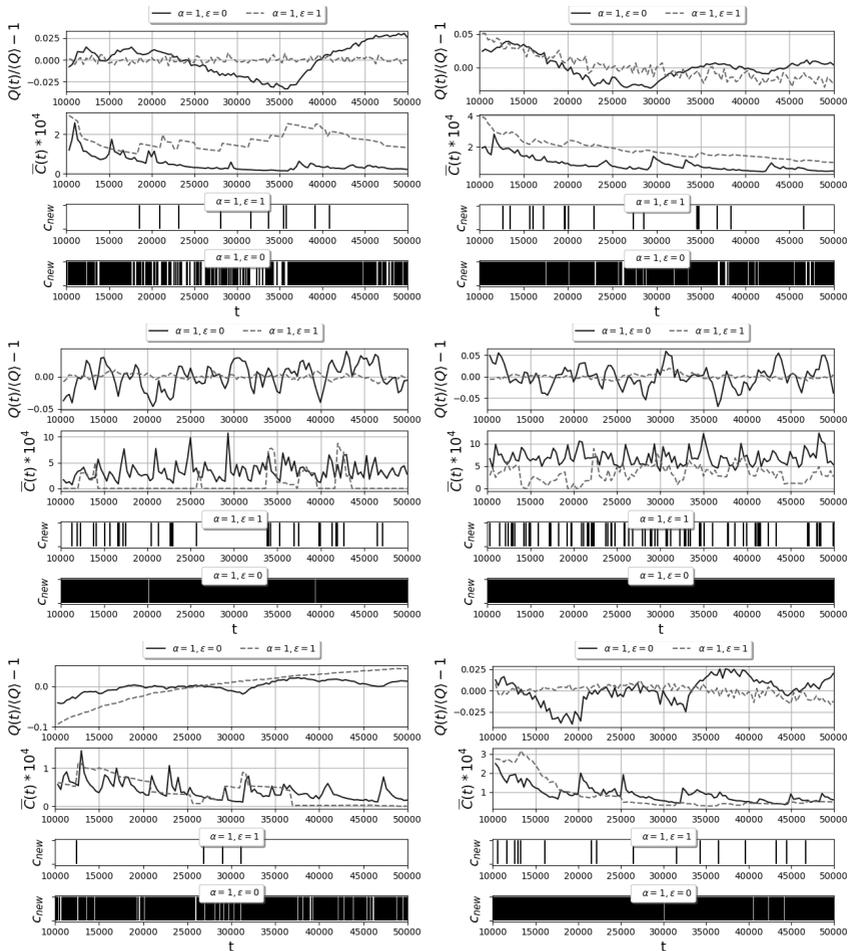


Рис. 1. Эволюция нормализованной модулярности графа (5) и среднего кластерного коэффициента (6) против шага эволюции КП, и лента штрихов, соответствующих точечно процессу шагов эволюции, когда вновь добавляемые узлы создают треугольник с существующими узлами: без удаления узлов и связей (верхние рисунки), с равномерным удалением узлов (средние рисунки) и с равномерным удалением связей (нижние рисунки), где в колонке слева $m_0 = 2$ и справа $m_0 = 3$.

В работе исследуются всплески $Q(t)$ в момент времени t при эволюции КП. Всплески наблюдаются в кластерах последовательных превышений

нормализованной модулярностью

$$\psi(t) = Q(t)/\langle Q \rangle - 1 \tag{5}$$

уровня u . $\langle Q \rangle$ обозначает среднюю модулярность за некоторый период эволюции. Кластеризация $\psi(t)$ увеличивается с ростом параметра α из (1) [1]. Кластерная структура $\psi(t)$ и среднего кластерного коэффициента

$$\bar{C}_t = \frac{1}{\|V_t\|} \sum_{i \in V_t} c_{i,t} \tag{6}$$

против числа шагов эволюции t для параметров КП $(\alpha, \varepsilon) \in \{(1, 0), (1, 1)\}$ и различных стратегий удаления узлов и связей представлена на Рис. 1, где $m_0 \in \{2, 3\}$. \bar{C}_t убывает с ростом t для эволюции без удаления узлов и связей и с удалением связей, Рис. 1. Поскольку удаление узлов приводит к графу с фиксированным числом узлов, \bar{C}_t становится большим и из-за увеличения числа изолированных узлов стабилизируется около относительно большого значения по сравнению с двумя другими стратегиями, Рис. 1. Штрихи на Рис. 1 указывают на моменты эволюции с созданием нового треугольника. На Рис. 1 показано, что случай $(\alpha, \varepsilon) = (1, 1)$ приводит к существенно меньшему числу треугольников, чем $(\alpha, \varepsilon) = (1, 0)$ для любой стратегии удаления. $\{\psi(t)\}$ может быть близко к нулю (т.е. Q близко к $\langle Q \rangle$) или может расти без видимой кластеризации, как на Рис. 1 (центр, низ) для $(\alpha, \varepsilon) = (1, 1)$. Новые узлы присоединяются с почти одинаковой вероятностью к существующим узлам, поскольку ε доминирует в (4), и новые треугольники появляются редко.

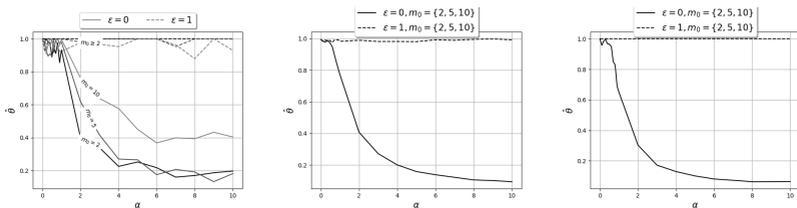


Рис. 2. Интервальные оценки ЭИ $\hat{\theta}$ для нормализованной модулярности $\psi(t) = Q(t)/\langle Q \rangle - 1$, усреднённые по 100 моделируемым графам, полученным КП без удаления узлов и связей (слева), с равномерным удалением узлов (центр) и равномерным удалением связей (справа) от параметра α в (4).

3. Эволюция экстремального индекса модулярности

Оценки ЭИ модулярности против значений параметра α в (4) для КП без удалений узлов и связей и с равномерными удалениями узлов или связей приведены на Рис. 2. При $\varepsilon = 0$ ЭИ убывает с ростом α . Убывание тем медленнее, чем больше значение m_0 . При $\varepsilon \geq 1$ ЭИ стабилизируется около значений, близких к 1. Случай $\hat{\theta} \approx 1$ может означать отсутствие кластеров превышений модулярности и то, что превышения в пределе единичны. При $\alpha \in [0, 1]$ и $\varepsilon = 0$ ЭИ тоже стабилизируется близко к 1, см. Рис. 2, поскольку значения $c_{i,t}^\alpha$ больше, чем соответствующие значения при $\alpha > 1$. Если узел или ребро удаляются на каждом шаге эволюции, то оценки ЭИ совпадают для разных m_0 , и их значения зависят только от параметра ε , см. Рис. 2 (центр, справа). Чем ε меньше, тем сильнее кластеризация (или локальная зависимость) последовательности модулярности. При $\varepsilon = 0$ значения ЭИ близки к нулю при удалении узла или ребра, и кластеризация модулярности сильнее. Локальная зависимость модулярности слабее (и средний размер кластера меньше), когда КП проводится без удаления узлов и связей. Последнее согласуется с кластеризацией модулярности на Рис. 1.

4. Заключение

С помощью моделирования исследовалась модель эволюции КП с возможным удалением узлов и ребер. Изучались всплески последовательности модулярности. Оценивался экстремальный индекс модулярности. Расширенная версия может включать исследование взаимосвязи модулярности и количества сообществ.

СПИСОК ЛИТЕРАТУРЫ

1. *Bagrow J., Brockmann D.* Natural Emergence of Clusters and Bursts in Network Evolution // *Physical Review X*. 2012. Vol. 3. № 2. P. 021016.
2. *Beirlant J., Goegebeur Y., Teugels J., Segers J.* Statistics of Extremes: Theory and Applications. Chichester, West Sussex: Wiley, 2004. 498 p.
3. *Dugué N., Perez A.* Directed Louvain : maximizing modularity in directed networks // In Research Report Université d'Orléans. hal-01231784; 2015. P. 1–15.
4. *Ferro C.A.T., Segers J.* Inference for Clusters of Extreme Values // *J. R. Statist. Soc. B*. 2003. Vol. 65. P. 545–556.
5. *Leadbetter M.R., Lingren G., Rootz én, H.* Extremes and Related Properties of Random Sequence and Processes; ch.3, Springer: New York, 1983.

6. *Markovich N.M., Rodionov I.V.* Threshold selection for extremal index estimation // *Journal of Nonparametric Statistics*. 2023. Vol. 36(3). P. 527–546.
7. *Markovich N.M., Vaičiulis M.* Investigation of triangle counts in graphs evolved by uniform clustering attachment // *arXiv: 2401.11548v1*. 2024. P. 1–16.
8. *Markovich N.M., Ryzhov M.S., Vaičiulis M.* Inferences for Random Graphs Evolved by Clustering Attachment // *arXiv: 2403.00551v1*. 2024. P. 1–25.
9. *Hagberg A.A., Schult D.A., Swart P.J.* Exploring network structure, dynamics, and function using NetworkX // *In Proceedings of the 7th Python in Science Conference (SciPy2008)*, Gäel Varoquaux, Travis Vaught, and Jarrod Millman (Eds), (Pasadena, CA USA), 2008. P. 11–15.

Маркович Наталья — доктор физико-математических наук, главный научный сотрудник, Институт Проблем Управления им. В.А. Трапезникова РАН. E-mail: *nat.markovich@gmail.com*

Рыжов Максим — научный сотрудник, Институт Проблем Управления им. В.А. Трапезникова РАН. E-mail: *ryzhov@phystech.edu*

EXPLICIT INTEGRAL REPRESENTATION OF SOME PATH-DEPENDENT BROWNIAN FUNCTIONALS

V. Jokhadze, O. Purtukhia

Ivane Javakishvili Tbilisi State University, Tbilisi, Georgia

We consider some path-dependent Brownian functionals and derive constructive formulas for the stochastic integral representation. **Keywords:** *Brownian functional, Malliavin derivative, integral representation.*

Introduction

The question of representing Brownian functionals as a stochastic Itô integral with an explicit form of the integrand is investigated. One of the pioneering works on the existence of an integral representation belongs to Clark (1970). Let a Brownian Motion $B = (B_t)$, $t \in [0, T]$, be given on a probability space $(\Omega, \mathfrak{F}, P)$, and let

$$\mathfrak{S}_t := \mathfrak{S}_t^B = \sigma\{B_u : 0 \leq u \leq t\}.$$

It should be noted that finding an explicit expression for the integrand $\varphi(t, \omega)$ in the integral representation is a very complex task. In this direction, one general result is known, called the Clark-Ocone formula (1984), according to which

$$\varphi(t, \omega) = E[D_t F | \mathfrak{S}_t],$$

where $D_t F$ is the so-called stochastic (Malliavin) derivative of the functional F under consideration.

Glonti and Purtukhia (2017) generalized the Clark-Ocone formula to the case when the functional is not stochastically smooth, but its conditional mathematical expectation is stochastically differentiable, and proposed a method for finding the integrand. Here we consider some path-dependent Brownian functionals and, building on an earlier generalization of the Clark-Ocone formula by Jaoshvili-Purtukhia (2005), derive constructive formulas for the stochastic integral representation.

1. Auxiliary results

Theorem 1.1. (see [4]). Let f and its generalized derivative ∂f be a square-integrable functions with weight function $\exp\{-\frac{x^2}{2}\}$ and $\xi \in D_{1,2}$,

then the following stochastic integral representation is valid:

$$f(\xi) = Ef(\xi) + \int_0^T E[\partial f(\xi) D_t \xi | \mathfrak{F}_t] dB_t \quad (P - a.s.)$$

In addition, we needed two technical results about ordinary integrals.

Proposition 1.1. For any real number a and non-negative integer $n \geq 0$ we have

$$\begin{aligned} & \frac{1}{\sqrt{2\pi}} \int_a^{+\infty} x^{2n} \exp\left\{-\frac{x^2}{2}\right\} dx = \\ & = \varphi(a) \sum_{k=1}^n \frac{(2n-1)!!}{(2k-1)!!} a^{2k-1} + (2n-1)!! (1 - \Phi(a)), \end{aligned}$$

where Φ is the standard normal distribution function and φ is its density function, $(2n-1)!! := 1 \cdot 3 \cdot \dots \cdot (2n-1)$, $(-1)!! := 1$.

Proof. To verify the proposition, we will use the method of mathematical induction. For $n = 1$, using the formula of integration by parts, it is obvious that

$$\begin{aligned} & \frac{1}{\sqrt{2\pi}} \int_a^{\infty} x^2 \exp\left\{-\frac{x^2}{2}\right\} dx = -\frac{1}{\sqrt{2\pi}} \int_a^{\infty} x d(\exp\left\{-\frac{x^2}{2}\right\}) = \\ & = -\frac{x}{\sqrt{2\pi}} \exp\left\{-\frac{x^2}{2}\right\} \Big|_a^{\infty} + \frac{1}{\sqrt{2\pi}} \int_a^{\infty} \exp\left\{-\frac{x^2}{2}\right\} dx = a\varphi(a) + (1 - \Phi(a)). \end{aligned}$$

Now suppose that this relation is true for n and check that it is also true for $n + 1$. Again, thanks to the integration by parts formula, it is easy to see that

$$\begin{aligned} & \frac{1}{\sqrt{2\pi}} \int_a^{\infty} x^{2(n+1)} \exp\left\{-\frac{x^2}{2}\right\} dx = -\frac{1}{\sqrt{2\pi}} \int_a^{\infty} x^{2n+1} d(\exp\left\{-\frac{x^2}{2}\right\}) = \\ & = -\frac{x^{2n+1}}{\sqrt{2\pi}} \exp\left\{-\frac{x^2}{2}\right\} \Big|_a^{\infty} + \frac{2n+1}{\sqrt{2\pi}} \int_a^{\infty} x^{2n} \exp\left\{-\frac{x^2}{2}\right\} dx = \\ & = \frac{a^{2n+1}}{\sqrt{2\pi}} \exp\left\{-\frac{a^2}{2}\right\} + \\ & + (2n+1) \left[\varphi(a) \sum_{k=1}^n \frac{(2n-1)!!}{(2k-1)!!} a^{2k-1} + (2n-1)!! (1 - \Phi(a)) \right] = \\ & = \varphi(a) \sum_{k=1}^{n+1} \frac{(2n+1)!!}{(2k-1)!!} a^{2k-1} + (2n+1)!! (1 - \Phi(a)). \end{aligned}$$

By that, the proof of proposition is complete. ■

In the same way, the validity of the following proposition is verified.

Proposition 1.2. For any real number a and natural n we have

$$\frac{1}{\sqrt{2\pi}} \int_a^{+\infty} x^{2n-1} \exp\left\{-\frac{x^2}{2}\right\} dx = \varphi(a) \sum_{k=0}^{n-1} \frac{(2n-2)!!}{(2k)!!} a^{2k},$$

where $(2n)!! := 2 \cdot 4 \cdot \dots \cdot (2n)$, $0!! := 1$.

2. Main results

For any non-negative integer n , consider the following path-dependent Brownian functional:

$$F(n) := \left(\int_0^T B_s ds \right)^{2n+1}.$$

Let us denote $F^+(n) := [F(n)]^+$ and $F^-(n) := [F(n)]^-$.

Theorem 2.1. For any non-negative integer n the following stochastic integral representation is valid:

$$F^+(n) = EF^+(n) + (2n+1) \sum_{k=0}^{2n} \int_0^T (T-t) C_{2n}^k \sigma^k \eta^{2n-k} I_k^+(\sigma, \eta) \Big|_{\eta = \int_0^t (T-s) dB_s} dB_t,$$

where

$$I_{2k-1}^+(\sigma, \eta) = \varphi\left(\frac{\eta}{\sigma}\right) \sum_{i=0}^{k-1} \frac{(2k-2)!!}{(2i)!!} \left(\frac{\eta}{\sigma}\right)^{2i},$$

$$I_{2k}^+(\sigma, \eta) = (2k-1)!! \Phi\left(\frac{\eta}{\sigma}\right) + \varphi\left(\frac{\eta}{\sigma}\right) \sum_{i=1}^k \frac{(2k-1)!!}{(2i-1)!!} \left(-\frac{\eta}{\sigma}\right)^{2i-1}.$$

Proof. According to the Theorem 1.1, we have

$$\begin{aligned} \varphi^+(t, \omega) &= E[\partial F^+(n) D_t \left(\int_0^T B_s ds \right) | \mathfrak{S}_t] = \\ &= E[I_{\{(\int_0^T B_s ds)^{2n+1} > 0\}} (2n+1) \left(\int_0^T B_s ds \right)^{2n} \cdot \int_0^T I_{[0,s]}(t) ds | \mathfrak{S}_t] = \\ &= (2n+1)(T-t) E[I_{\{\int_0^T (T-s) dB_s > 0\}} \left(\int_0^T (T-s) dB_s \right)^{2n} | \mathfrak{S}_t]. \end{aligned}$$

It's obvious that

$$\int_0^T (T-s)dB_s = \int_t^T (T-s)dB_s + \int_0^t (T-s)dB_s,$$

where

$$\int_t^T (T-s)dB_s \sim N(0, (T-t)^3/3) := N(0, \sigma^2).$$

Further, thanks to the well-known properties of conditional mathematical expectation, using the Newton binomial formula and the values of the moments of the normal distribution, it is not difficult to establish that

$$\begin{aligned} \varphi^+(t, \omega) &= (2n+1)(T-t) \times \\ &\times E[I_{\{\int_t^T (T-s)dB_s > -y\}} (\int_t^T (T-s)dB_s + y)^{2n}]|_{y=\int_0^t (T-s)dB_s} = \\ &= \frac{(2n+1)(T-t)}{\sqrt{2\pi}\sigma} \sum_{r=0}^{2n} C_{2n}^r y^{2n-r} \int_{-y}^{+\infty} x^r \exp\{-\frac{x^2}{2\sigma^2}\} dx|_{y=\int_0^t (T-s)dB_s} = \\ &= (2n+1)(T-t) \sum_{r=0}^{2n} C_{2n}^r \sigma^r y^{2n-r} I_{2n}^+(r, \sigma, y)|_{y=\int_0^t (T-s)dB_s}, \end{aligned}$$

where

$$I_{2n}^+(r, \sigma, y) := \frac{1}{\sqrt{2\pi}} \int_{-y/\sigma}^{+\infty} x^r \exp\{-\frac{x^2}{2}\} dx.$$

On the other hand, from Proposition 1.1 and 1.2, for any natural r we have

$$\begin{aligned} I_{2n}^+(2r-1, \sigma, y) &:= \frac{1}{\sqrt{2\pi}} \int_{-y/\sigma}^{+\infty} x^{2r-1} \exp\{-\frac{x^2}{2}\} dx = \\ &= \varphi(-\frac{y}{\sigma}) \sum_{k=0}^{r-1} \frac{(2r-2)!!}{(2k)!!} (-\frac{y}{\sigma})^{2k} = \varphi(\frac{y}{\sigma}) \sum_{k=0}^{r-1} \frac{(2r-2)!!}{(2k)!!} (\frac{y}{\sigma})^{2k}. \end{aligned}$$

and

$$\begin{aligned} I_{2n}^+(2r, \sigma, y) &:= \frac{1}{\sqrt{2\pi}} \int_{-y/\sigma}^{+\infty} x^{2r} \exp\{-\frac{x^2}{2}\} dx = \\ &= \varphi(-\frac{y}{\sigma}) \sum_{k=1}^r \frac{(2r-1)!!}{(2k-1)!!} (-\frac{y}{\sigma})^{2k-1} + (2r-1)!! (1 - \Phi(-y/\sigma)) = \end{aligned}$$

$$= \varphi\left(\frac{y}{\sigma}\right) \sum_{k=1}^r \frac{(2r-1)!!}{(2k-1)!!} \left(-\frac{y}{\sigma}\right)^{2k-1} + (2r-1)!! \Phi(y/\sigma). \quad \blacksquare$$

Using similar arguments, the following result is proven.

Theorem 2.2. For any non-negative integer n the following stochastic integral representation is valid:

$$F^-(n) = EF^-(n) - (2n+1) \sum_{k=0}^{2n} \int_0^T (T-t) C_{2n}^k \sigma^k \eta^{2n-k} I_k^-(\sigma, \eta) \Big|_{\eta=\int_0^t (T-s)dB_s} dB_t,$$

where

$$I_{2k-1}^-(\sigma, \eta) = -\varphi\left(\frac{\eta}{\sigma}\right) \sum_{i=0}^{k-1} \frac{(2k-2)!!}{(2i)!!} \left(\frac{\eta}{\sigma}\right)^{2i},$$

$$I_{2k}^-(\sigma, \eta) = (2k-1)!! [1 - \Phi\left(\frac{\eta}{\sigma}\right)] - \varphi\left(\frac{\eta}{\sigma}\right) \sum_{i=1}^k \frac{(2k-1)!!}{(2i-1)!!} \left(-\frac{\eta}{\sigma}\right)^{2i-1}.$$

Combining Theorems 2.1 and 2.2, we obtain the following result.

Theorem 2.3. For any non-negative integer n the following stochastic integral representation is valid:

$$|F(n)| = E|F(n)| + (2n+1) \sum_{k=0}^{2n} \int_0^T (T-t) C_{2n}^k \sigma^k \eta^{2n-k} I_k(\sigma, \eta) \Big|_{\eta=\int_0^t (T-s)dB_s} dB_t,$$

where

$$I_{2k-1}(\sigma, \eta) = 2\varphi\left(\frac{\eta}{\sigma}\right) \sum_{i=0}^{k-1} \frac{(2k-2)!!}{(2i)!!} \left(\frac{\eta}{\sigma}\right)^{2i},$$

$$I_{2k}(\sigma, \eta) = (2k-1)!! [2\Phi\left(\frac{\eta}{\sigma}\right) - 1] + 2\varphi\left(\frac{\eta}{\sigma}\right) \sum_{i=1}^k \frac{(2k-1)!!}{(2i-1)!!} \left(-\frac{\eta}{\sigma}\right)^{2i-1}.$$

Corollary 2.1. The following stochastic integral representation is true

$$\left| \int_0^T B_s ds \right| = \sqrt{\frac{2T^3}{3\pi}} + \int_0^T (T-t) [2\Phi(\sigma^{-1} \int_0^t (T-s)dB_s) - 1] dB_t.$$

Acknowledgements. This research was partially funded by the Shota Rustaveli National Science Foundation of Georgia (SRNSFG), grant # STEM-22-226.

REFERENCES

1. *Clark M.C.* The representation of functionals of Brownian motion by stochastic integrals // *Ann. Math. Stat.* 1970. Vol. 41. P. 1282–1295.
2. *Ocone D.* Malliavin calculus and stochastic integral representation formulas of diffusion processes // *Stochastics*. 1984. Vol. 12. P. 161–185.
3. *Glonti O., Purtukhia O.* On One Integral Representation of Functionals of Brownian Motion // *SIAM J. Theory of Probability and Its Applications*. 2017. Vol. 61. № 1. P. 133–139.
4. *Jaoshvili O., Purtukhia O.* Stochastic Integral Representation of Functionals of Wiener Processes // *Bulletin Georgian Acad. Sci.* 2005. Vol. 171. № 1. P. 17–20.

Jokhadze Valeriane — Master of Economics, PhD student, Faculty of Economics and Business. E-mail: *Jokhadze.valeriane@gmail.com*

Purtukhia Omar — Doctor of Sciences, Professor, Head of Department, Department of Mathematics. E-mail: *o.purtukhia@gmail.com*

MODELING REFLECTIONLESS TRANSMISSION OF THE KINK SOLITON OF KLEIN-GORDON EQUATION ON NETWORKS

Q. U. Asadov¹, K. K. Sabirov²

¹*University of management and future technology, Tashkent, Uzbekistan*

²*EMU University, Tashkent, Uzbekistan*

In this paper we consider the nonlinear Klein-Gordon equation on the metric star graph with three semi-infinite bonds. At the branched point we put two types of vertex boundary conditions: the weight continuity and the condition for derivatives of wave functions as the generalized Kirchhoff rule. We solve this equation satisfying vertex boundary conditions. We also show reflectionsless propagations of the kink soliton solution. **Keywords:** *Klein-Gordon equation, vertex boundary conditions, Kirchhoff rule, reflectionsless transmission, kink soliton, graphs.*

Introduction

Nonlinear dynamics of solitary waves, that occur in many scientific areas, can be described by nonlinear partial differential equations [15], [7]. The nonlinear Klein-Gordon equation is one of the important classes of partial differential equations and is used to model many different phenomena, including the propagation of dislocations in crystals and the behavior of elementary particles [4]. The exact and numerical soliton solutions of Klein-Gordon equation are given in the pioneer work [1], in that numerically solved with the time and coordinate steps $k = h = 0.05$ and found the discretization scheme using Taylor series. The energy and momentum conservating schemes used to integrate the nonlinear Klein-Gordon equation in [2] and invariant-conserving finite difference algorithms are given [6].

The nonlinear partial differential equations on the branched structures are attached the most interest in the last decade years. Such attention was caused by the possibility for obtaining soliton solution of nonlinear partial differential equations such as nonlinear Schrödinger, Dirac equations and its numerous applications in different branches of physics [3, 4]. The branched structures can be modeled metric graphs [14]. Metric graphs are consist of two sets: the set of points (vertices) and the set of intervals that connect vertices.

In this paper we focus on one of the exact solutions and transmission of the kink soliton of the nonlinear Klein-Gordon equation through the

vertices of the networks. The present work is organized as follows: in the section 1 the nonlinear Klein-Gordon equation on the metric star graph with derivations the vertex boundary conditions is obtained; in the section 2 the kink soliton solution of the formulated problem with the sum rule for the reflectionless transmission is analytically given; in the last section the conclusions are given.

1. Formulation of the problem

We consider a star graph with three bonds e_j , for which a coordinate x_j is assigned. Choosing the origin of coordinates at the vertex, 0, for bond e_1 we put $x_1 \in (-\infty, 0]$ and for $e_{2,3}$ we fix $x_{2,3} \in [0, +\infty)$. In what follows, we use the shorthand notation $q_j(x)$ for $q_j(x_j)$ where x is the coordinate on the bond j to which the component q_j refers. Klein-Gordon equation on the each bond e_j of the star graph is written as

$$\partial_{tt}^2 q_j - \partial_{xx}^2 q_j - q_j + b_j q_j^3 = 0. \tag{1}$$

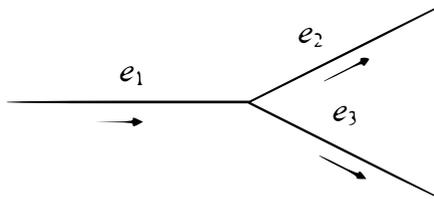


Figure 1. The metric star graph

Now we define the vertex boundary conditions at the branched point of the star graph, therefore we derive these boundary conditions from conservation laws is energy. The energy conservation law is defined as

$$E = \sum_{j=1}^3 E_j, \tag{2}$$

where

$$E_j = \int_{e_j} \left[\frac{1}{2} (\partial_t q_j)^2 + \frac{1}{2} (\partial_x q_j)^2 - \frac{1}{2} q_j^2 + \frac{b_j}{4} q_j^4 \right] dx. \tag{3}$$

From $\dot{E} = 0$ we can get the following nonlinear boundary condition as

$$\partial_x q_1 \partial_t q_1 |_{x=0} = \partial_x q_2 \partial_t q_2 |_{x=0} + \partial_x q_3 \partial_t q_3 |_{x=0}. \tag{4}$$

We need two types of boundary conditions to find a solution of (1) and to fulfil the nonlinear vertex boundary condition (4). Therefore the first type is the following weight continuity

$$\alpha_1 q_1|_{x=0} = \alpha_2 q_2|_{x=0} = \alpha_3 q_3|_{x=0}, \quad (5)$$

the second type is given derivatives of wave functions at the branched point as Kirchoff rule

$$\frac{1}{\alpha_1} \partial_x q_1|_{x=0} = \frac{1}{\alpha_2} \partial_x q_2|_{x=0} + \frac{1}{\alpha_3} \partial_x q_3|_{x=0}. \quad (6)$$

2. The kink soliton solution of Klain-Gordon equation on the star graph with three edges

The kink soliton solution of Klein-Gordon equation (1) on the each bond e_j of the metric star graph is the following

$$q_j(x, t) = \mp \frac{1}{\sqrt{b_j}} \tanh \left(\frac{x - l - vt}{\sqrt{2(1 - v^2)}} \right), \quad (7)$$

where l is the initial center of mass of soliton. Fulfilling the vertex boundary conditions (5)-(6) we can get the following constrains

$$\frac{\alpha_1}{\sqrt{b_1}} = \frac{\alpha_2}{\sqrt{b_2}} = \frac{\alpha_3}{\sqrt{b_3}}, \quad (8)$$

$$\frac{1}{\alpha_1 \sqrt{b_1}} = \frac{1}{\alpha_2 \sqrt{b_2}} + \frac{1}{\alpha_3 \sqrt{b_3}}. \quad (9)$$

Theorem 1. From (5) and (6) for the transmissionless propagation of the kink soliton (7) of Klein-Gordon equation (1) on the star graph we obtain the following sum rule for nonlinearities

$$\frac{1}{b_1} = \frac{1}{b_2} + \frac{1}{b_3}. \quad (10)$$

Proof. If we satisfy the soliton solution (7) of Eq. (1) into (5) and (6), we can obtain (8) and (9). It is clear that from (8) and (9) we can get (10). ■

Using the kink soliton solution of Klein-Gordon equation (1) with the sum rule for the nonlinearities (10) we can also show that another conservation law, i.e. the momentum is conserved from the last expression the momentum is constant.

3. Conclusion

In this paper we studied the nonlinear Klein-Gordon equation on the star graph with three semi-infinite bonds. First of all we derived the nonlinear boundary condition from the energy conservation law. Satisfying this boundary conditions at the vertex we obtained the weight continuity and the condition for derivatives of the wave function as Kirchhoff rule. We obtained the soliton solution on the metric star graph and the constrain as inverses of nonlinearities for the reflectionless transmission.

REFERENCES

1. *Ablowitz M.J. and Segur H.* Solitons and the Inverse Scattering Transform. City: Philadelphia, 1981.
2. *Ablowitz M.J. and Clarkson P.A.* Solitons, Nonlinear Evolution Equations and Inverse Scattering. City: London, Cambridge University Press 1991.
3. *Greiner W.* Relativistic Quantum Mechanics-Wave Equations // Springer-Verlag, 3rd edition, City: Berlin, Germany, 2000.
4. *Ablowitz M.J., Kruskal M.D. and Ladik J.F.* Solitary wave collisions SIAM // Appl. Math. june 1979. Vol. 36. № 3.
5. *Jimenez S. and Vazquez L.* Analysis of Four Numerical Schemes for a Nonlinear Klein-Gordon Equation // Appl. Math. june 1990. Vol. Comp.35. P. 61–94.
6. *Los Vu-Quos and Shaofan Li* Invariant-conserving finite difference algorithms for the nonlinear Klein-Gordon equation // Comp. Meth. in App. Mech. and Eng. (North-Holland, CMA 403), 1993. Vol. 107. P. 341–391.
7. *Sabirov K.K., Babajanov D.B., Matrasulov D.U. and Kevrekidis P.G.* Dynamics of dirac solitons in networks // Phys. A: Math. Theor. 2018. Vol. 51. P. 435203.
8. *Sabirov K.K., Yusupov J.R., Matyokubov Kh.Sh., Susanto H., Matrasulov D.U.* Networks with point-like nonlinearities // Nanosystems: Phys. Chem. Math. 2022. Vol. 13. № 1. P. 3035.
9. *Yusupov J.R., Sabirov K.K., Asadov Q.U., Ehrhardt M., and Matrasulov D.U.* Dirac particles in transparent quantum graphs: Tunable transport of relativistic quasiparticles in branched structures // Phys. Rev. E. 2020. Vol. 101. P. 062208.

Asadov Quvonchbek Urinovich — senior teacher, department of Fundamental science. E-mail: quvonchbekurinovichasadov@gmail.com

Sabirov Karimjan Kabiljanovich — PhD, dotcent, department of Natural science. E-mail: karimjonsabirov80@gmail.com

ON ESTIMATION OF A MIXTURE OF DISTRIBUTIONS WITH INCOMPLETE DATA

R. Muradov¹, M. Ziyoidinova²

¹*Namangan Institute of Engineering and Technology, Namangan, Uzbekistan*

²*Andijan State University, Andijan, Uzbekistan*

Estimating a mixture of distributions in statistics, especially when dealing with incomplete data, can be challenging but often follows a structured approach. The article discusses the estimation of the mixture of distribution function with incomplete data. The estimate is constructed using substitution method and the property of estimate has been studied. **Keywords:** *mixture of distributions, incomplete data, estimation, empirical distribution function.*

Introduction

Most sample surveys yield data sets that are incomplete in various ways and for various reasons. The current attitude and methodology to handle the missing information issue can be summarized as trying to avoid it in the data collection phase, trying to make up for it during initial data screening, and pretending it never existed, in the final analysis. Some people will always refuse to participate in the survey, some, although selected into the sample, will not be found and many will be missed by the sampling procedure, giving them unintentionally too little or no chance at all of being selected. Some others, although they are available for the survey, choose not to respond to certain questions, leading to a different kind of missing information. All that is far too important and potentially informative, to be considered a nuisance.

Incomplete data in statistics refers to datasets that have missing values or observations. This can occur for various reasons, such as non-response in surveys, data entry errors, or loss of data during collection. A random right censoring model is used in survival analysis to handle situations where the event of interest (like failure or death) is not observed for all subjects within the study period. In this model, some subjects may drop out or the study ends before the event occurs, leading to incomplete data. Incomplete data can occur in many forms, such as missing values in covariates, or unobserved component labels in the case of mixture models. The goal is to esti-

The authors thank Professor Abdurakhim Abdushukurov, Tashkent branch of Moscow State University, Department of Applied Mathematics and Informatics.

mate the parameters of the mixture model despite this incompleteness (see, [1,2,5]). Estimating a mixture of distributions with incomplete data typically refers to the problem of estimating parameters of a statistical model where not all data points are fully observed or are censoring (see, [3,4]). Let for a fixed l integer $\left\{ \mathbf{X}^{(j)} = \left[X_i^{(j)} = \left(X_{1i}^{(j)}, X_{2i}^{(j)} \right), i \geq 1 \right], j = 1, \dots, l \right\}$ and $\left\{ \mathbf{Y}^{(j)} = \left[Y_i^j = \left(Y_{1i}^{(j)}, Y_{2i}^{(j)} \right), i \geq 1 \right], j = 1, \dots, l \right\}$ be mutually independent sequences of independent random variables (its may be dependent also) with X_i^j and Y_i^j having continuous random distribution functions for $(s, t) \in R^2$:

$$F^{(j)}(s, t) = P \left(X_{11}^{(j)} \geq s, X_{21}^{(j)} \geq t \right), j = 1, \dots, l,$$

$$G^{(j)}(s, t) = P \left(Y_{11}^{(j)} \geq s, Y_{21}^{(j)} \geq t \right), j = 1, \dots, l.$$

1. Random right censoring model

In this model, the sequence $\mathbf{X}^{(j)}$ is censored from the right [1] by the sequence $\mathbf{Y}^{(j)}$ and l samples are observed:

$$\mathbf{V}_j^{(n)} = \left\{ \left(Z_i^{(j)}, \delta_i^{(j)} \right), i = 1, \dots, n \right\}, j = 1, \dots, l, \tag{1}$$

where $Z_i^{(j)} = \left(Z_{1i}^{(j)}, Z_{2i}^{(j)} \right)$, $Z_{ki}^{(j)} = X_{ki}^{(j)} \wedge Y_{ki}^{(j)}$, $\delta_i^{(j)} = \left(\delta_{1i}^{(j)}, \delta_{2i}^{(j)} \right)$, $\delta_{ki}^{(j)} = I \left(Z_{ki}^{(j)} = X_{ki}^{(j)} \right)$, $k = 1, 2; j = 1, \dots, l$. The task is to estimate the mixture of distributions

$$F(s, t) = \sum_{j=1}^l \theta_j F^{(j)}(s, t), (s, t) \in R^2, \tag{2}$$

where the weights $\{\theta_j, j = 1, \dots, l\}$ are unknown nuisance parameters that constitute the common probability distribution of a sequence of independent identically distributed random variables $\{v_1, v_2, \dots\}$ with the set of values $\{1, \dots, l\}$

$$\theta_j = P(v_i = j) > 0, i = 1, \dots, n, \sum_{j=1}^l \theta_j = 1. \tag{3}$$

The sequence $\{v_j\}$ is independent of $\left\{ \left(\mathbf{X}_i^{(j)}, \mathbf{Y}_i^{(j)} \right), j = 1, \dots, l \right\}$. Let $K(t) = P(v_j = t)$, $t \in R$. Then, $\theta_j = K(j + 1) - K(j)$, $j = 1, \dots, l$ and $K(l + 1) = 1$. Thus, estimation of $F(s, t)$ is possible from the joint sample

$$\mathbf{V}_*^{(n)} = \left\{ \left(Z_i^{(j)}, \delta_i^{(j)}, v_i \right), j = 1, \dots, l, i = 1, \dots, n \right\}. \tag{4}$$

Obviously, a natural estimate of θ_j is

$$\theta_{jn} = \frac{1}{n} \sum_{i=1}^n I(v_i = j) = K_n(j + 1) - K_n(j), \quad j = 1, \dots, l. \tag{5}$$

where $K_n(t)$ is the empirical distribution function corresponding to $K(t)$.

2. The main result

Let $F_{kn}^{(j)}(s, t)$, $k = 1, 2, 3$, $j = 1, \dots, l$, be the estimates of the exponential, product-limit and power types structure (see, [1]) constructed from samples (1) for the r.f. $F^{(j)}(s, t)$, $j = 1, \dots, l$ respectively. Then, with the account of (5), the corresponding estimates from sample (4) for mixture (2) are

$$F_{nk}(s, t) = \sum_{j=1}^l \theta_{jn} \cdot F_{kn}^{(j)}(s, t), \quad (s, t) \in R^2, \quad k = 1, 2, 3. \tag{6}$$

Let $Q^{(j)} = \{(s, t) \in R^2 : 0 < F^{(j)}(s, t) \cdot G^{(j)}(s, t) < 1\}$, $j = 1, \dots, l$, and $Q = \bigcap_{j=1}^l Q^{(j)} \neq \emptyset$.

Theorem 1. For $k = 1, 2, 3$ the following equality's are valid

$$\sup_{(s,t) \in Q} |F_{kn}(s, t) - F(s, t)| \stackrel{a.s.}{=} O(\lambda_n).$$

Proof. We have the chains of inequalities for $k = 1, 2, 3$:

$$\begin{aligned} \sup_{(s,t) \in Q} |F_{kn}(s, t) - F(s, t)| &\leq \sum_{j=1}^l \left[|\theta_{jn} - \theta_j| \cdot \sup_{(s,t) \in Q} (F_{kn}^{(j)}(s, t)) + \right. \\ &+ \theta_j \sup_{(s,t) \in Q} \left| F_{kn}^{(j)}(s, t) - F^{(j)}(s, t) \right| \Big] \leq \sum_{j=1}^l [|K_n(j + 1) - K(j + 1)| + \\ &+ |K_n(j) - K(j)| + \sup_{(s,t) \in Q^{(j)}} |F_{kn}^{(j)}(s, t) - F^{(j)}(s, t)|] \leq \\ &\leq 2l \cdot \sup_{-\infty < t < \infty} |K_n(t) - K(t)| + \\ &+ \sum_{j=1}^l \sup_{(s,t) \in Q^{(j)}} \left| F_{kn}^{(j)}(s, t) - F^{(j)}(s, t) \right| \stackrel{a.s.}{=} O(\lambda_n). \end{aligned}$$

where the last equality follows from the ILL. Theorem 1 is proved.

For $k = 1, 2, 3$ introduce the Gaussian processes

$$\begin{aligned} \Lambda_k(s, t) &= \sum_{j=1}^l \Lambda_k^{(j)}(s, t) = \\ &= \sum_{j=1}^l \left[\theta_j \cdot \mathbf{W}_k^{(j)}(s, t) + F^{(j)}(s, t) \cdot (\theta_j(1 - \theta_j))^{\frac{1}{2}} \xi_0 \right], \end{aligned}$$

where $\{\mathbf{W}_k^{(j)}, k = 1, 2, 3\}$ are Gaussian processes from Theorem 1, and ξ_0 is the standard normal random variable independent of them (see, [2,3]).

Theorem 2. When $n \rightarrow \infty$

$$l_{kn}(s, t) = n^{\frac{1}{2}} \cdot (F_{kn}(s, t) - F(s, t)) \xrightarrow{D} A_k(s, t), \text{ in } \mathbf{D}^2(Q).$$

3. Conclusion

In summary, this paper constructs a new statistical estimator based on incomplete data for a mixture of distributions and studies the asymptotic properties of these estimators.

REFERENCES

1. *Abdushukurov A.* Estimation of unknown distributions from incomplete observations and their properties. Berlin: LAMBERT Acad. Publ., 2011. 264 p.
2. *Little R., Rubin D.* Statistical analysis with missing data. 2nd ed. New York: Wiley; 2002. 389 p.
3. *Muradov R.* Estimation mixture of two-dimensional reliability functions on the plane by the random right censoring // XI International conference "Financial and actuarial mathematics and eventology security Krasnoyarsk: Russia, 2012. P. 262–265.
4. *Abdushukurov A., Muradov R.* Estimation of bivariate survival function from random censored data with Poisson sample size // Calcutta Statistical Association Bulletin. 2020. Vol. 72. № 2. P. 111–122.
5. *Murray G.D., Titterington D.M.* Estimation Problems with Data from a Mixture // Journal of the Royal Statistical Society. Series C (Applied Statistics) 1978. Vol. 27, No. 3, P. 325–334.

Muradov Rustamjon — DSc., docent, department of higher mathematics. E-mail: *rustamjonmuradov@gmail.com*

Ziyoitdinova Matluba — Ph.D., student, department of mathematics. E-mail: *ziyoitdinov@yahoo.com*

NUMERICAL MODELING OF FLUID FLOW BASED ON THE RAKHMATULIN MODEL TAKING INTO ACCOUNT THE FORCHHEIMER EQUATION

U. Dalabaev¹, M. Ikramova²

¹*University of World Economy and Diplomacy, Tashkent, Uzbekistan*

²*Scientific Research Institute of Irrigation and Water Problems, Tashkent,
Uzbekistan*

This paper considers a mathematical model of the movement of liquid in a region consisting of two subregions: a porous medium dried by a liquid and the second subregion - a free zone. An interpenetrating model is used to mathematically describe the behavior of a fluid. For the strength of interaction between phases, Darcy's law and Forchheimer's equation are adopted. For the numerical implementation of the resulting equation, a SIMPLE-like algorithm was used. Examples are considered in which the influence of the Forchheimer term, the Darcy number, and the influence of porosity on the flow characteristics are studied. **Keywords:** *Flow, Darcy law, Rakhmatulin model, Forchheimer equation, porous media, numerical method.*

Introduction

In nature and in many technological processes, viscous fluid flows in areas with and without porous inclusions. An important element of such research is the mathematical description of such phenomena. The complexity of such flows lies in the consideration of the flow in two subregions: a free zone and a porous subregion. In the free zone, the flow is modeled by the Stokes/Navier–Stokes equations, and in the porous subregion there are various approaches for modeling the flow in the porous region [1-2].

1. Mathematical model

Developing the approach given in [3], some problems of joint fluid motion in a complex area are studied. In this case, the flow simulation is described by a single equation. The derivation of such an equation is based on Rakhmatulin's interpenetrating flow model of two-phase media [4-5]. The equations in the dimensionless two-dimensional case are described using systems of equations:

$$\begin{aligned}
 Shf \frac{\partial u}{\partial x} + fu \frac{\partial u}{\partial x} + fv \frac{\partial u}{\partial x} &= -\frac{f}{Re} \frac{\partial p}{\partial x} + \frac{4}{3Re} \frac{\partial}{\partial x} \left(f \frac{\partial u}{\partial x} \right) + \\
 + \frac{1}{Re} \frac{\partial}{\partial y} \left(f \frac{\partial u}{\partial y} \right) - \frac{2}{3Re} \frac{\partial}{\partial x} \left(f \frac{\partial v}{\partial y} \right) + \frac{1}{Re} \frac{\partial}{\partial y} \left(f \frac{\partial v}{\partial x} \right) - Cu, & \quad (1)
 \end{aligned}$$

$$\begin{aligned}
 Shf \frac{\partial v}{\partial x} + fu \frac{\partial v}{\partial x} + fv \frac{\partial v}{\partial x} &= -\frac{f}{Re} \frac{\partial p}{\partial x} + \frac{1}{Re} \frac{\partial}{\partial x} \left(f \frac{\partial v}{\partial x} \right) + \\
 + \frac{4}{3Re} \frac{\partial}{\partial y} \left(f \frac{\partial u}{\partial y} \right) + \frac{1}{Re} \frac{\partial}{\partial x} \left(f \frac{\partial v}{\partial y} \right) - \frac{2}{3Re} \frac{\partial}{\partial y} \left(f \frac{\partial v}{\partial x} \right) - Cv, & \quad (2)
 \end{aligned}$$

$$\frac{\partial(fu)}{\partial x} + \frac{\partial(fv)}{\partial y} = 0. \quad (3)$$

Here u, v velocity components; p is pressure, f is volume concentration, Re is Reynolds number.

The coefficient C is taken in the form

$$C = \frac{f^2}{Da \cdot Re} + \frac{C_f f^3}{\sqrt{Da}} |\bar{u}|. \quad (4)$$

In this expression, the first term is related to Darcy’s law (Da is Darcy’s number), and the second is related to Forchheimer’s relation.

The numerical solution (1)-(3) was carried out by analogy with the SIMPLE method [6]. Note that in [7] flows in the combined region were studied using a single equation.

Various problems of fluid flow with partially filled porous inclusions are considered. The influence of the Darcy number, Reynolds number and porosity on the flow structure is considered.

2. The problem of fluid flow in a flat pipe with a partially filled porous medium

Let us consider the flow of liquid in a flat pipe, the lower half of which is filled with a porous medium. A uniform flow is set in the inlet part of the pipe. The no slip conditions are set on the pipe walls. At the output boundary we use a soft boundary condition. Equations (1)-(3) will be solved in the area $D : 0 \leq x \leq 5, 0 \leq y \leq 1$ Number of grid nodes: 50x30.

An uneven mesh was used with condensation at the beginning of the pipe, where significant deformation of the flow occurs. Figure 1 shows the formation of velocity across sections with flow parameters: $Re = 1, Da =$

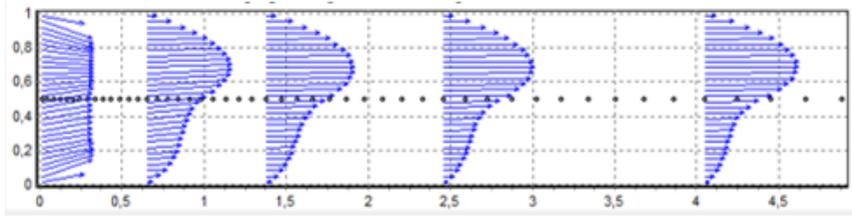
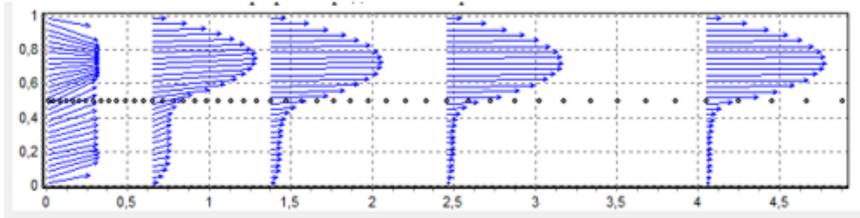
a) $Re = 1; Da = 0,0001; f = 0,9$ b) $Re = 1; Da = 0,00001; f = 0,9$

Figure 1. Formation of flow velocity.

0,0001, $f = 0,9$. The graphs shown in Figure 1 correspond to the sections: 0,017; 0,657; 1,384; 2,460; 4,053.

From Fig. 1 it is clear that the Darcy number significantly affects the formation of the velocity profile. It is clear from the graphs that the formation of a transit zone near the interphase boundary is in qualitative agreement with the results discussed in various articles.

Figure 2 shows the particle streamlines at flow parameters: $Re = 200, Da = 0,00001, f = 0,9, \psi_1 = 0,1; \psi_2 = 0,2; \psi_3 = 0,4; \psi_4 = 0,8$.

Figure 3 shows the nature of the change in transverse velocity at the interface depending on the Reynolds number. In the graph, represented as a diamond, $Re=1$ corresponds, asterisks correspond to $Re=10$, and triangles correspond to $Re = 100$. ($f = 0,7, Da = 0,001$). Figure 3 shows that an increase in the Reynolds number will lead to an increase in the transverse velocity through the interface and an increase in the length of the flow stabilization zone.

Fig.4 Transverse velocity through the interface at different Darcy numbers.

In Fig. 4, the points represented in the form of a diamond correspond to $Da = 0,00001$, the asterisks – $Da = 0,001$, and the triangles – $Da = 0,1$. Figure 4 shows that increasing the Darcy number will reverse the behavior of the transverse velocity across the interface.

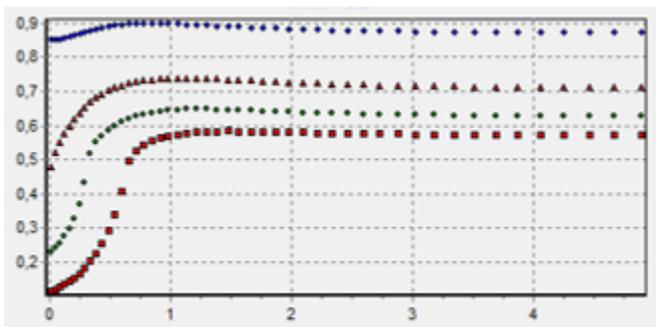


Figure 2. Streamlines.

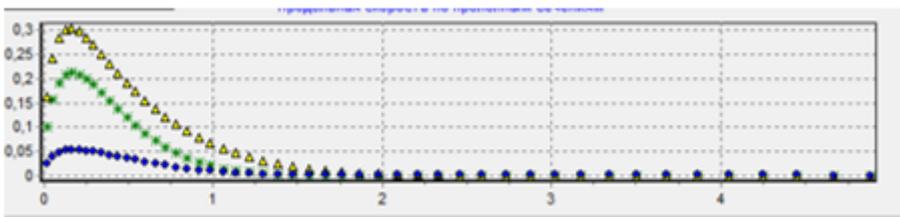


Figure 3. Transverse velocity through the interface at different Reynolds numbers.

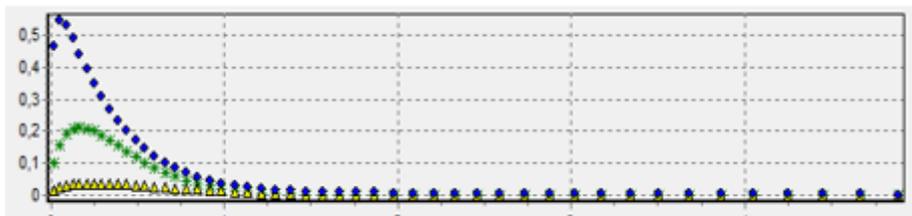


Figure 4. Transverse velocity through the interface at different Darcy numbers.

Figure 5 shows the dimensionless change in the friction distribution $\frac{1}{Re} \frac{\partial u}{\partial y}$ over a section multiplied by Re (for clarity, the value of the friction distribution was multiplied by the number 0,03). An increase in the Reynolds number will lead to an increase in the amplitude of the change in friction in each section.

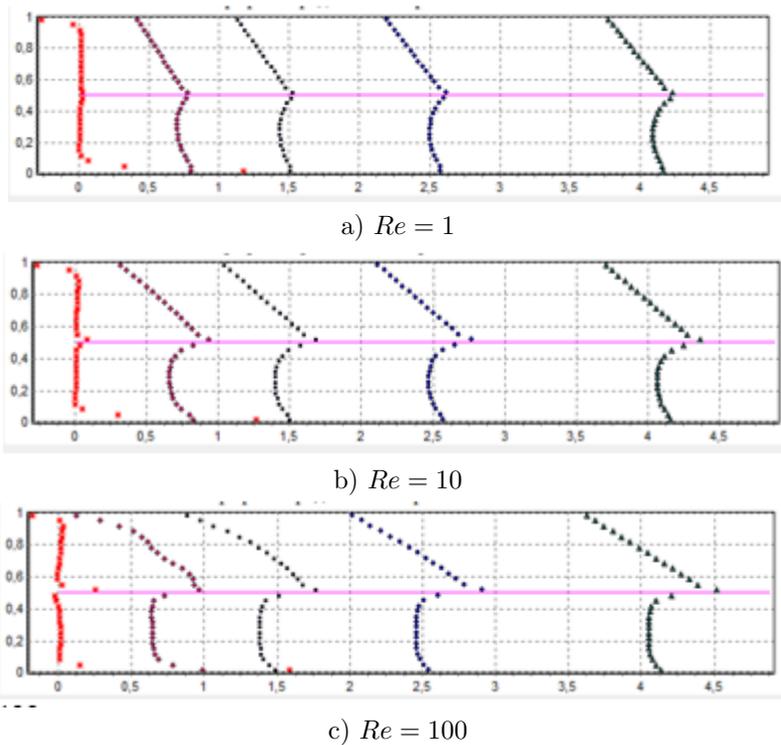


Figure 5. Change in shear stress across pipe sections 0,017; 0,657; 1,384; 2,460; 4,053.

Figure 6 compares the graphs of the transverse velocity at the interface with environmental parameters: $Re = 10$, $f = 0,9$, $Da = 0,00001$. Dot plots with an asterisk correspond to the numerical results obtained taking into account the Forchheimer force. The diamond-shaped scatter plot does not take this force into account. The Forchheimer force prevents fluid leakage through the interface.

3. Conclusion

The application of Rakhmatulin's interpenetrating model can be successfully applied to study problems related to the movement of fluid in heterogeneous areas. Taking into account the Forchheimer force in obtaining a solution in high porous media.

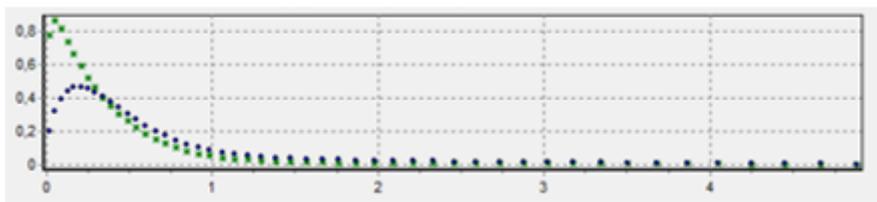


Figure 6. Transverse velocity through the interface at different Darcy numbers.

REFERENCES

1. *Donald A. Nield, Adrian Bejan.* Convection in Porous Media. Springer International Publishing AG, 2017.
2. *Fernando A. Morales, Ralph E. Showalter* A Darcy–Brinkman model of fractures in porous media *Journal of Mathematical Analysis and Applications.* 2017. Vol. 452. P. 1332–1358.
3. *Dalabaev U.* Investigation of fluid flow in a heterogeneous region based on the Rakhmatulin model *International Conference: Mechanics, Earthquake Engineering,* May 27–28, 2021 Tashkent, Uzbekistan.
4. *Faizullaev, Dzharulla F.* *Laminar Motion of Multiphase Media in Conduits.* Springer US, ISBN 978-1-4899-4832-8, 1969.
5. *Nigmatulin, R. I.* *Fundamentals of the mechanics of heterogeneous media.* Moscow, Izdatel'stvo Nauka, 336 p. In Russian, 1978.
6. *Patankar S.* *Numerical Heat Transfer and fluid Flow.* ISBN 9780891165224 Published January 1, by CRC Press, 1980.
7. *S.V. Polyakov, M.A. Trapeznikova, A.G. Churbanov, N.G. Churbanova.* Calculation of incompressible flows in the porous body-free flow system *Preprints of the Institute of Mechanical Engineering named after. M.V.Keldysh.* 2021. № 71. P. 19.

Dalabaev Umurdin — Doctor of Physical and Mathematical Sciences, Professor, University of World Economy and Diplomacy, Department System Analysis and Mathematical Modeling. E-mail: *udalabaev@mail.ru*

Ikramova Malika — Doctor of Technical Sciences, Professor, Scientific Research Institute of Irrigation and Water Problems, Head of Water Resources Complex Use Department. E-mail: *malika.ikramova56@gmail.com*

ПРИМЕНЕНИЯ ВЕТВЯЩИХСЯ ПРОЦЕССОВ ДЛЯ РЕШЕНИЯ КРАЕВЫХ ЗАДАЧ ДЛЯ ПОЛУЛИНЕЙНОГО УРАВНЕНИЯ ГЕЛЬМГОЛЬЦА

А. С. Расулов, Г. М. Раимова

Университет мировой экономики и дипломатии, г. Ташкент, Узбекистан

Рассматриваются вероятностные модели для решения краевых задач для полулинейного уравнения Гельмгольца. Получены вероятностные представления решения задач в виде математического ожидания некоторой случайной величины. В соответствии с вероятностными представлениями построены ветвящиеся процессы, заданы моделирующие формулы для каждого из ветвящихся процессов. Доказано, что построенные ветвящиеся процессы с вероятностью единица вырождаются и среднее число частиц каждого n -ого поколения меньше либо равно единице. Для рассмотренных задач построены несмещенные оценки на траекториях соответствующих случайных процессов. Полученные оценки решений имеют ограниченную дисперсию, строятся на траекториях ветвящегося процесса с ограниченным средним числом ветвлений и легко моделируется. Результаты вычислительного эксперимента показывают, что при помощи построенного алгоритма можно строить практические оценки, эффективно реализуемые на компьютере. **Ключевые слова:** *метод Монте-Карло, несмещенные оценки, ветвящийся процесс.*

1. Вероятностная модель для решения задачи Дирихле для уравнения Гельмгольца в полулинейном случае

Несмещенные и смещенные оценки решения краевых задач для уравнения Гельмгольца Г.А.Михайловым и Б.С.Елеповым в работах [2, 3], Н.А.Симоновым в [9], А.С.Сипиным в [10], Г.А.Михайлова и Р.Н.Макарова [6], Г.А.Михайлова [5], А.С.Расуловым в [8], Г.М. Раимовой в [7]. Используя результаты вышеупомянутых работ предлагаются алгоритмы метода Монте Карло для задачи Дирихле для следующего уравнения: $\Delta u - c u = -f(x, u)$ с полиномиальной нелинейностью в правой части относительно неизвестной функции $f(x, u) = \sum_{n=0}^{\infty} a_n(x)u^n(x)$. Далее при предположении существования решения [1, 4] строится несмещенная оценка на траекториях ветвящегося процесса блуждания по сферам.

Пусть D ограниченная область в R^3 с регулярной границей Γ , причем коэффициенты $f(x, u)$ удовлетворяют условию $\lim_{n \rightarrow \infty} \sqrt[n]{\bar{a}_n} = 0$, где $\sup_{x \in D} |a_n(x)| \leq \bar{a}_n$. Рассмотрим следующую задачу Дирихле:

$$-\Delta u(x) + cu(x) = f(x, u), \quad x \in D, \quad u|_{\Gamma} = \varphi. \quad (1)$$

Предполагается, что функции $\varphi(x) \in C(\bar{\Gamma})$, $a_n(x) \in C(\bar{D})$ ($n = 0, 1, \dots$) и коэффициент $c > 0$ таковы, что существует [4] единственное непрерывное решение полулинейной задачи. Пусть

$$\sum_{n=0}^{\infty} \bar{a}_n n < \infty. \quad (2)$$

При указанных предположениях, строится несмещенная оценка решения с ограниченной дисперсией в некоторой произвольной точке $x \in D$.

2. Вероятностное представление решения задачи

Используя результаты [2, 3] к уравнению $-\Delta u(x) + cu(x) = f(x, u)$, получим следующее интегральное представление:

$$u(x) = \frac{R\sqrt{c}}{4\pi R^2 \operatorname{sh}(R\sqrt{c})} \int_{S_R} u(y) dy + \int_{K_R} \frac{\operatorname{sh}[(R - |x - y|)\sqrt{c}]}{4\pi|x - y|\operatorname{sh}(R\sqrt{c})} f(y, u(y)) dy. \quad (3)$$

Здесь $R(x) = \min_{y \in \Gamma} |x - y|$ - расстояние от точки x до границы, K_R - шар радиуса R с центром в точке x , S_R - соответствующая сфера. Первый интеграл в (3) есть интеграл по поверхности сферы S_R . Так как площадь поверхности сферы S_R равна $4\pi R^2$, то после введения этого выражения под знак первого интеграла, он будет представлять равномерное распределение на сфере:

$$u(x) = \frac{R\sqrt{c}}{\operatorname{sh}(R\sqrt{c})} \int_{S_R} u(x + R\omega) d\omega + \frac{\operatorname{sh}(R\sqrt{c}) - R\sqrt{c}}{\operatorname{sh}(R\sqrt{c})} \int_{K_R} \frac{\operatorname{sh}[(R - |x - y|)\sqrt{c}]}{4\pi|x - y|(\operatorname{sh}(R\sqrt{c}) - R\sqrt{c})} \cdot \frac{f(y, u(y))}{c} dy.$$

Перепишем данное выражение следующим образом

$$u(x) = q \int_{S_R} u(y) d\omega + (1 - q) \int_{K_R} p(x, y) \frac{f(y, u(y))}{c} dy, \quad (4)$$

здесь были использованы следующие обозначения: $q = R\sqrt{c} \setminus (\text{sh}(R\sqrt{c}))$, ω - равномерное распределение на S_R , $p(x, y) = \frac{\text{sh}[(R - |x - y|)\sqrt{c}]}{4\pi|x - y|(\text{sh}(R\sqrt{c}) - R\sqrt{c})}$ - плотность перехода из x в y ($x, y \in K_R$). Случайный ветвящегося процесс, согласованный с вероятностным представлением (4), и определение процесса блуждания по сферам с ветвлением можно использовать из работ [7, 13]. Формулу (4) можно переписать следующим образом:

$$u(x) = q \int_{S_R(x)} u(y) d\omega + (1 - q) \int_{K_R(x)} p(x, y) \sum_{n=0}^{\infty} \frac{a_n(y)}{c} u^n(y) dy. \quad (5)$$

В соответствии с (5) несмещенные оценки для решения задачи (1) естественно строить на траекториях процесса блуждания по сферам с ветвлением, в котором участвуют частицы одного типа.

3. Случайный ветвящийся процесс, согласованный с вероятностным представлением

В соответствии с представлением (5) построим процесс блуждания с ветвлением в фазовом пространстве D , в котором участвуют частицы одного типа. Пусть $M = \sum_{n=1}^{\infty} n\bar{a}_n$, и α - некоторое число, $0 < \alpha < 1$. В начальный момент имеется одна частица в точке $x_0 = x$. Пусть $n > 0$ и x_n известно. За единицу времени частица с вероятностью $q = \frac{R\sqrt{c}}{\text{sh}(R\sqrt{c})}$ переходит в точку x_{n+1} , равномерно распределенную на сфере $S_R(x)$ и с вероятностью $1 - q$ переходит в точку x_{n+1} , распределенную в шаре $K_R(x)$ с плотностью $p(x, y)$. Во втором случае с вероятностью $\pi_n = \frac{\alpha}{M} \bar{a}_n$, ($n = 1, 2, \dots$) частица делится на n новых частиц либо с вероятностью $\pi_0 = 1 - \sum_{n=1}^{\infty} \pi_n$ гибнет. Новые частицы ведут себя независимо друг от друга аналогично рассмотренной частице. Процесс обрывается, если все частицы в \bar{D} погибают, либо, если все частицы попадают в Γ . Параметр α позволяет регулировать число ветвлений.

4. Построение несмещенной оценки решения

Для построения оценки метода Монте-Карло для задачи (1) реализуется процесс блуждания по сферам с ветвлением из точки $x_0 = x$, т.е. в начале имеется одна частица в точке x . За единицу времени частица с вероятностью $q = \frac{R\sqrt{c}}{\text{sh}(R\sqrt{c})}$ переходит в точку $x_1 = y_1$, которая выбирается равномерно по сфере $S_R(x)$ и с вероятностью $1 - q$ в точку

$x_1 = y_2$, которая выбирается внутри шара $K_R(x)$ соответственно плотности $p(x, y)$. Во втором случае с вероятностью π_n ($n \neq 0$) происходит ветвление, т.е. частица делится на n частиц и вес умножается на величину $\frac{Ma_n(x_1)}{c\bar{a}_n\alpha}$. В случае $n = 0$ происходит поглощение частицы в точке

x_1 и вес помножается на величину $\frac{a_0(x_1)}{c\pi_0}$. Из точки x_1 строятся n независимых траекторий блуждания. Новые частицы ведут себя независимо друг от друга аналогично рассмотренной частице. Процесс обрывается, либо, если все частицы в \bar{D} погибают, либо, если все частицы выходят на ε окрестность границы. Определим оценку рекуррентным способом. Пусть $\zeta_0(x) = u(x)$, $\zeta_k(x) = \Psi(\zeta_{k-1}(x))$, где

$$\Psi(\zeta(x)) = \begin{cases} \zeta(y_1), & \text{с вероятностью } q(x) \\ W_n(y_2), \prod_{i=1}^n \zeta^{i-1}(y_2), & \text{с вероятностью } (1 - q(x))\pi_n, \quad n \neq 0; \\ W_0(y_2), & \text{с вероятностью } (1 - q(x))\pi_0. \end{cases}$$

Здесь $\zeta^{(i)}(y)$ - независимые реализации случайной величины $\zeta(y)$. Так называемые «веса», т.е. множители, на которые при каждом шаге помножается оценка, определим следующим образом: $W_n(y) = \frac{Ma_n(y)}{ca_n\alpha}$, $W_0(y) = \frac{a_0(y)}{c\pi_0}$. Пусть \mathfrak{R}_k - σ -алгебра, порожденная последовательностями $\{\omega_i\}_{i=0}^{k-1}$, $\{\alpha_i^0\}_{i=0}^{k-1}$, $\{\alpha_i^1\}_{i=0}^{k-1}$, $\{\alpha_i^2\}_{i=0}^{k-1}$. Справедливо следующее утверждение.

Теорема 1. Последовательность $\{\zeta_k(x)\}_{k=0}^\infty$ образует мартингал относительно $\{\mathfrak{R}_k\}_{k=0}^\infty$. Если $M < c$, то $\zeta_k(x, t)$ - равномерно интегрируемый мартингал.

Доказательство теоремы опубликовано в [7]. С целью уменьшения времени получения одной реализации оценки решения задачи (1), рассмотрим несколько видоизмененный процесс с меньшим числом ветвлений и оценки на нем со сколь угодно малым смещением. Возьмем ε - достаточно малым и рассмотрим внутреннюю ε - окрестность границы: Γ_ε . Пусть N_1 - момент обрыва процесса внутри области, и N_ε - момент первого попадания всех частиц в Γ_ε . $N = \min\{N_1, N_\varepsilon\}$ - момент остановки процесса. Тогда вероятность обрыва траектории в точке x_n будет равна:

$$g(x_n) = \begin{cases} 1, & \text{если } x_n \in \Gamma_\varepsilon, \\ (1 - q(x_{n-1}))\pi_n, & \text{если } x_n \in \bar{D} \setminus \Gamma_\varepsilon. \end{cases}$$

Можно доказать, что $N < \infty$ и если $M < c$, то $\zeta_N(x)$ является несмещенной оценкой для $u(x)$ с конечной дисперсией.

5. Вычислительный эксперимент

В данном разделе приводятся результаты численных экспериментов на основе предложенных оценок решения начально-краевой задачи:

$$\begin{cases} -\Delta u(x) + cu(x) = f(x, u), & x \in D, \\ u(x) = \varphi(x), & x \in \Gamma. \end{cases}$$

Область D - либо шар радиуса R с центром в начале координат $D_1 = \{(x_1, x_2, x_3) : x_1^2 + x_2^2 + x_3^2 \leq R^2\}$ либо единичный куб $D_2 = \{(x_1, x_2, x_3) : 0 \leq x_1, x_2, x_3 \leq 1\}$; Правая часть уравнения $f(x, u)$ определяется как одно из следующих выражений: $g \exp(u) + f_0$, $g \sin(u) + f_0$, $g \cos(u) + f_0$, $(g - const)$. Результаты вычислительного эксперимента приведены ниже.

Таблица 1

Результаты вычислительного эксперимента

D	R	правая часть	x_0	$U_{\text{Точ}}$	S	err
D_1	0,9	$g \exp(u) + f_0$	(-0,4;0,2;-0,6)	0,89022	0,89204	0,00182
D_1	0,9	$g \exp(u) + f_0$	(-0,7;0,2;-0,1)	-0,5646	-0,5674	0,00276
D_1	0,9	$g \cos(u) + f_0$	(0,4;0,2;-0,6)	0,0023	0,00282	0,00052
D_1	0,9	$g \cos(u) + f_0$	(-0,4;0,2;-0,6)	0,89022	0,89553	0,00531
D_2	1	$g \exp(u) + f_0$	(0,7;0,2;0,1)	0,84147	0,83611	0,00536
D_2	1	$g \sin(u) + f_0$	(0,9;0,2;0,1)	0,65244	0,64132	0,01112

Пояснения к таблице 1. D - область, в которой рассматривается задача; R - радиус шара, если область D шар, длина грани, если D куб; x_0 - точка, в которой решается задача; $N_t = 5000$ - количество реализаций оценки; $U_{\text{Точ}}$ - точное решение; S - выборочная оценка; DS - выборочная дисперсия; err - разность между точным решением и оценкой.

Заключение

Результаты вычислительного эксперимента показывают, что при помощи построенного алгоритма можно строить оценки, эффективно реализуемые на компьютере. Параметр α используемых при расчетах вероятностей ветвлений, дает возможность контроля за средним количеством ветвлений и средним количеством частиц дерева. Напомним, что параметр выбирался из условия $M/c \leq \alpha < 1$. Для численного эксперимента использовано значение $\alpha = 0.9$. Так как в выбранных примерах известны точные решения, мы имеем возможность убедиться, что все оценки попадают в доверительный интервал.

СПИСОК ЛИТЕРАТУРЫ

1. *Дубинский Ю. А.* Квазилинейные эллиптические и параболические уравнения любого порядка // УМН, Москва, 1968, Т.23, вып.1(139). С. 45--90.
2. *Елепов Б. С., Михайлов Г. А.* Алгоритм «блуждания по сферам» для уравнения $\Delta u - cu = -g$ // ДАН СССР. Москва, 1973. Т.212, № 1. С.15--18.
3. *Елепов Б. С., Михайлов Г. А.* Использование фундаментальных решений эллиптических уравнений для построения алгоритмов метода Монте-Карло // ЖВМиМФ. Москва, 1974. Т.14, № 3. С.728-736.
4. *Курант Р.* Уравнения с частными производными. М: Мир, 1964. 789с.
5. *Михайлов Г. А.* Решение задачи Дирихле для нелинейных эллиптических уравнений методом Монте-Карло // Сиб. матем. журн. Новосибирск, 1994. Т.35, № 5. С.1085-1093.
6. *Михайлов Г. А., Макаров Р. Н.* Решение краевых задач 2-го и 3-го рода методами Монте-Карло // Сиб. мат.жур. Н.,1997. Т.38, № 3. С.603-614.
7. *Раимова Г. М.* Вероятностные модели для решения задачи Дирихле для нелинейных эллиптических уравнений // Узб. мат. журн. Ташкент, 2017. № 1, С. 114-123.
8. *Расулов А. С.* Метод Монте-Карло для решения нелинейных задач. Ташкент: Фан, 1992. 104с.
9. *Симонов Н. А.* Алгоритмы случайного блуждания по сферам для решения смешанной краевой задачи и задачи Неймана // Сиб. журн. вычисл. матем. Новосибирск, 2007. Т. 10, № 2. С.209-220.
10. *Сипин А. С.* Решение двух краевых задач Дирихле методом Монте-Карло // ЖВМиМФ. Москва, 1979. Т.19, № 2. С.388-401.
11. *Фридман А.* Уравнения с частными производными параболического типа. Москва: Мир, 1968. 428с.
12. *Харрис Т. Е.* Теория ветвящихся процессов. Москва: Мир, 1966. 355с.
13. *Ermakov S. M., Nekrutkin V. V., Sipin A. S.* Random Processes for Classical Equations of Mathematical Physics. Kluwer Acad. Publ., 1989. 282p.

Расулов Абдужабар Саттарович — доктор физ.-мат. наук, профессор, профессор кафедры «Системный анализ и математическое моделирование», Университет мировой экономики и дипломатии. E-mail: asrasulov@gmail.com

Раимова Гулнора Мирвалиевна — доктор физ.-мат. наук, профессор, зав.кафедрой «Системный анализ и математическое моделирование», Университет мировой экономики и дипломатии. E-mail: raimova27@gmail.com

ИССЛЕДОВАНИЕ АЛГОРИТМОВ ЧИСЛЕННОГО МОДЕЛИРОВАНИЯ ДВУМЕРНЫХ СЛУЧАЙНЫХ ВЕКТОРОВ С ЗАВИСИМЫМИ КОМПОНЕНТАМИ

А. В. Войтишек^{1,2}, Д. Е. Кирильчик²

¹ *Институт вычислительной математики и математической геофизики
СО РАН, г. Новосибирск, Россия*

² *Новосибирский государственный университет, г. Новосибирск, Россия*

В работе представлен детальный анализ реализации т. н. «технологии распределенного (взвешенного) параметра», предназначенной для конструирования совместных плотностей двумерных случайных векторов, для которых есть возможность построить эффективные алгоритмы численного моделирования выборочных значений. Анализ проведен с помощью авторской диалоговой системы NMTDD (Numerical Modelling of Two-Dimensional Distributions), представляющей собой двумерный аналог известной (и тоже авторской) системы NMPUD (Numerical Modelling of Probabilistic Univariate Distributions).

Ключевые слова: *численное моделирование двумерных случайных векторов, зависимые компоненты, технология распределенного (взвешенного) параметра, технология последовательных (вложенных) замен, исследовательские диалоговые компьютерные системы.*

Введение

Численное моделирование случайных величин и векторов (компьютерная реализация выборочных значений) является ключевым элементом алгоритмов метода Монте-Карло (см., например, [1]). В целом ряде приложений требуется моделировать отрезки (фиксированной или случайной длины) выборочных значений состояний цепей Маркова. Здесь происходит последовательная компьютерная реализация состояний, при этом на каждом шаге происходит моделирование двумерного вектора (ξ, η) с зависимыми компонентами. Учитывая относительно медленную сходимость метода Монте-Карло, такое моделирование должно быть экономичным.

Исследования выполнены в рамках государственного задания ИВМиГ СО РАН FWNM-2022-0002.

В разделе 1 данной работы сформулирован алгоритм моделирования двумерного вектора (ξ, η) и обозначена проблема построения «моделируемых» плотностей для компонент ξ и η . В разделе 2 предлагается решать эту проблему с помощью специальной авторской технологии распределенного (взвешенного) параметра. Здесь также приговора разработана нами технология последовательных (вложенных) замен, которая позволяет получать практически неограниченное количество элементарных плотностей (в том числе, содержащих числовой параметр). Свойства используемых технологий удается изучить с помощью авторских исследовательских компьютерных систем NMPUD (Numerical Modelling of Probabilistic Univariate Distributions) и NMTDD (Numerical Modelling of Two-Dimensional Distributions), кратко описанных в разделе 3. В разделе «Заключение» сформулирован основной вывод работы.

1. Моделирование двумерного случайного вектора

Рассмотрим двумерный случайных вектор $(\xi, \eta) \in (a, b) \times (c, d)$, распределенный на прямоугольной области $(a, b) \times (c, d)$ (здесь $-\infty \leq a < b \leq +\infty, -\infty \leq c < d \leq +\infty$) согласно совместной плотности

$$f_{(\xi, \eta)}(u, v) = f_{\eta}(v) \times f_{\xi}(u|v); \quad a < u < b, \quad c < v < d; \quad (1)$$

здесь

$$f_{\eta}(v) = \int_a^b f_{(\xi, \eta)}(u, v) du, \quad f_{\xi}(u|v) = \frac{f_{(\xi, \eta)}(u, v)}{f_{\eta}(v)}. \quad (2)$$

АЛГОРИТМ 1 (см., например, раздел 1.13 книги [1]). 1. Моделируем выборочное значение η_0 случайной величины η согласно плотности $f_{\eta}(v)$.

2. Моделируем выборочное значение ξ_0 случайной величины ξ согласно плотности $f_{\xi}(u|\eta_0)$.

В данной работе мы будем полагать, что сформулированный алгоритм 1 эффективен, если плотности $f_{\eta}(v)$ и $f_{\xi}(u|\eta_0)$ *элементарны*, т. е. для этих плотностей можно выписать экономичные формулы метода обратной функции распределения

$$\eta_0 = F_{\eta}^{-1}(\alpha_1); \quad \xi_0 = F_{\xi|\eta_0}^{-1}(\alpha_2); \quad \alpha_1, \alpha_2 \in U(0, 1) \quad (3)$$

(см. раздел 1.4 книги [1]); здесь $F_{\eta}(x) = \int_{-\infty}^x f_{\eta}(v) dv$, $F_{\xi|\eta_0}(x) = \int_{-\infty}^x f_{\xi}(u|\eta_0) du$ – это соответствующие функции распределения, а α_1, α_2 – стандартные случайные числа (выборочные значения случайной величины $\alpha \in U(0, 1)$, равномерно распределенной на интервале $(0, 1)$ –

такие значения получаются на компьютере с помощью соответствующего генератора – см., например, раздел 1.15 книги [1]).

Возникает вопрос о том, как получить набор плотностей (1), (2) с экономичными формулами (3).

2. Технологии распределенного (взвешенного) параметра и последовательных (вложенных) замен

Достаточно большое разнообразие требуемых плотностей (1), (2) можно получить с помощью следующей *технологии распределенного (взвешенного) параметра*.

ТЕХНОЛОГИЯ 1. Рассмотрим элементарную плотность распределения $f_\xi(u; \lambda)$, $u \in (a, b)$, зависящую от параметра λ , допустимые значения которого принадлежат интервалу (C, D) . Элементарность распределения означает существование простой (элементарной) формулы $\xi_0 = \psi_\xi(\alpha_0; \lambda) = F_{\xi|\eta_0=\lambda}^{-1}(\alpha_0)$, $\alpha_0 \in U(0, 1)$ для получения выборочного значения случайной величины ξ (здесь $\psi_\xi(\cdot)$ – композиция элементарных функций).

Рассмотрим также еще одну элементарную плотность $f_\eta(v)$ случайной величины η , принимающей значения в интервале $(c, d) \subseteq (C, D)$; при этом имеется соответствующая элементарная моделирующая формула $\eta_0 = \psi_\eta(\alpha_0) = F_\eta^{-1}(\alpha_0)$, $\alpha_0 \in U(0, 1)$.

Теперь поставим задачу построения эффективного (экономичного) алгоритма моделирования выборочного значения (ξ_0, η_0) двумерного случайного вектора (ξ, η) , имеющего плотность распределения

$$f_{(\xi, \eta)}(u, v) = f_\eta(v) \times f_\xi(u; v), \quad a < u < b; \quad c < v < d. \quad (4)$$

Это результат формального умножения плотностей $f_\eta(v)$ и $f_\xi(u; v)$ (здесь происходит подстановка переменной v вместо параметра λ). В представлении (1) для плотности (4) имеем $f_\xi(u|v) = f_\xi(u; v)$.

Для этого представления получаем эффективный алгоритм:

$$\eta_0 = \psi_\eta(\alpha_1), \quad \xi_0 = \psi_\xi(\alpha_2; \eta_0); \quad \alpha_1, \alpha_2 \in U(0, 1). \quad (5)$$

В свою очередь, большое разнообразие моделирующих формул метода обратной функции распределения вида (5) можно получить, применяя следующую *технологии последовательных (вложенных) замен*.

ТЕХНОЛОГИЯ 2. Пусть $f_\zeta(z; \lambda)$ – плотность случайной величины ζ , имеющей элементарное распределение в интервале (z_1, z_2) , т. е. из соотношения вида $\int_{z_1}^{\zeta_0} f_\zeta(z; \lambda) dz = \alpha_0$, $\alpha_0 \in U(0, 1)$ (здесь λ – некото-

рый параметр) для соответствующего выборочного значения ζ_0 случайной величины ζ можно получить формулу вида: $\zeta_0 = \psi_\zeta(\alpha_0; \lambda)$, где $\psi_\zeta(w; \lambda)$ – простая композиция элементарных функций.

Рассмотрим взаимно-однозначное преобразование, задаваемое монотонно возрастающей дифференцируемой функцией $\varphi(y) \uparrow$ и переводящее интервал (y_1, y_2) в интервал (z_1, z_2) ; при этом $\varphi(y_1) = z_1$, $\varphi(y_2) = z_2$. Полагая также, что саму функцию $\varphi(y)$, ее производную $\varphi'(y)$ и обратную к функции $\varphi^{-1}(y)$ можно представить в виде простых композиций элементарных функций.

Пусть случайная величина γ имеет плотность распределения

$$f_\gamma(y; \lambda) = f_\zeta[\varphi(y); \lambda] \times \varphi'(y), \quad y \in (y_1, y_2). \quad (6)$$

При сделанных выше предположениях можно утверждать, что $f_\gamma(y; \lambda)$ является плотностью элементарного распределения с моделирующей формулой $\gamma_0 = \varphi^{-1}[\psi_\zeta(\alpha_0; \lambda)]$; $\alpha_0 \in U(0, 1)$.

При реализации технологии 2 можно использовать убывающую функцию $\varphi(y) \downarrow$; при этом в формуле (6) второй множитель нужно брать по модулю: $f_\gamma(y; \lambda) = f_\zeta[\varphi(y); \lambda] \times |\varphi'(y)|$, $y \in (y_1, y_2)$.

Название *технология последовательных (вложенных) замен* для технологии 2 здесь связано с тем, что, полученную плотность (6) можно взять в качестве исходной плотности $f_\zeta(z; \lambda)$ и осуществить еще одно взаимно-однозначное преобразование типа $\varphi(y)$. С помощью таких вложенных замен можно получать неограниченное количество новых плотностей элементарных распределений.

3. Диалоговые компьютерные системы для исследования свойств технологий 1 и 2

Технология последовательных (вложенных) замен (технология 2) обладает тем существенным недостатком, что введение каждой из функций $\varphi(y)$ из формулы (6), как правило, делает соответствующую моделирующую формулу $\gamma_0 = \varphi^{-1}[\psi_\zeta(\alpha_0; \lambda)]$ существенно более трудоемкой, чем исходная формула $\zeta_0 = \psi_\zeta(\alpha_0; \lambda)$. Это показано в работе [2] на основании исследований с помощью разработанной нами системы NMPUD (Numerical Modelling of Probabilistic Univariate Distributions) [3].

Важнейшим новым элементом данной работы является то, что нам удалось создать проект [4], а затем и реализовать двумерный аналог системы NMPUD [3] – исследовательскую диалоговую компьютерную систему NMTDD (Numerical Modelling of Two-Dimensional Distributions).

На рисунке 1 представлена основная страница разработанной нами системы. Внизу страницы (в левой части) демонстрируется (или заного-

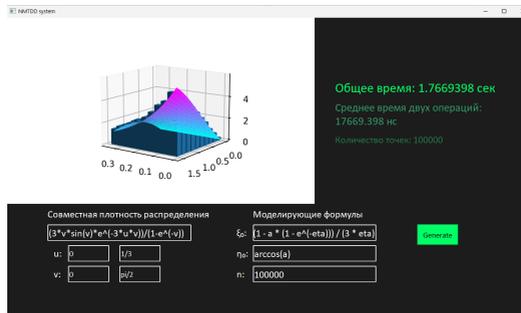


Рис. 1. Основная страница компьютерной системы NMTDD

сится) формула совместной плотности (1), полученная с помощью технологии 1, и соответствующие интервалы распределения переменных $u \in (a, b)$ и $v \in (c, d)$. Конкретнее, на рисунке 1 представлена плотность

$$f_{(\xi, \eta)}(u, v) = \frac{3v \sin v e^{-3uv}}{1 - e^{-v}} = f_{\eta}(v) \times f_{\xi}(u|v) = \sin v \times \frac{3v e^{-3uv}}{1 - e^{-v}} \quad (7)$$

для $0 < u < \frac{1}{3}$ и $0 < v < \frac{\pi}{2}$. В той же левой части страницы, но уже сверху, представлен график этой функции (гладкая поверхность).

Моделирующие формулы (алгоритмы) (3) демонстрируются (или заносятся) внизу в правой части основной страницы системы; конкретнее, они на рисунке 1 имеют вид

$$\eta_0 = \arccos \alpha_1, \quad \xi_0 = -\frac{\ln [1 - \alpha_2 (1 - e^{-\eta_0})]}{3\eta_0}; \quad \alpha_1, \alpha_2 \in U(0, 1). \quad (8)$$

С помощью многократного обращения к этим моделирующим формулам строится гистограмма – отображается темными столбиками (параллелепипедами). Правильность моделирующих формул (алгоритмов) (3) определяется визуальной близостью построенной гистограммы и графика плотности (1).

Также вычисляется и показывается с правой стороны экрана (крупными цифрами) общее время в секундах, затраченное на моделирование с помощью формул (8) заданного количества точек (ξ_0, η_0) (рисунок 1 построен для $n = 10^5$ выборочных значений случайного вектора (ξ, η)). Чуть ниже можно увидеть показатель среднего времени одного обращения к двум формулам (алгоритмам) (3) в наносекундах.

Соображения из работы [2], основанные на расчетах, проведенных с помощью компьютерной системы NMPUD, позволяют сделать вывод о

том, что затраты моделирующих формул (8) для моделирования двумерного случайного вектора (ξ, η) с совместной плотностью распределения (7) являются «умеренно большими».

Заключение

В работе представлена авторская исследовательская диалоговая компьютерная система NMTDD (Numerical Modelling of Two-Dimensional Distributions), позволяющая подробно изучать свойства технологии распределенного (взвешенного) параметра, предназначенной для конструирования совместных плотностей $f_{(\xi, \eta)}(u, v)$; $(u, v) \in (a, b) \times (c, d)$ двумерных случайных векторов (ξ, η) , для которых есть возможность построить эффективные алгоритмы численного моделирования выборочных значений (ξ_0, η_0) .

СПИСОК ЛИТЕРАТУРЫ

1. Михайлов Г. А. Некоторые вопросы теории методов Монте-Карло. Новосибирск: Наука, 1974. 143 с.
2. Рыжов И. А., Трофимов И. А. Исследование экономичности алгоритмов моделирования случайных величин с помощью компьютерной системы NMPUD // Математика: Материалы 59-й Международной науч. студ. конф. 12–23 апреля 2021 г. / Новосибир. гос. ун-т. Новосибирск: ИПЦ НГУ, 2021. С. 126.
3. Васильев Т. В., Постовалов Я. С., Черкашин Д. А. Проект компьютерной системы для выбора и исследования моделируемых вероятностных распределений // Математика: Материалы 58-й Международной науч. студ. конф. 10–13 апреля 2020 г. / Новосиб. гос. ун-т. Новосибирск: ИПЦ НГУ, 2020. С. 112.
4. Кирильчик Д. Е. Проект компьютерной системы моделирования двумерных случайных векторов // Математика: Материалы 62-й Международной науч. студ. конф. 17–23 апреля 2024 г. / Новосиб. гос. ун-т. Новосибирск: ИПЦ НГУ, 2024. С. 199.

Войтишек Антон Вацлавович — доктор физико-математических наук, профессор, главный научный сотрудник, лаборатория стохастических задач ИВМиМГ СО РАН; профессор, кафедра вычислительной математики НГУ. E-mail: vav@osmf.sccc.ru

Кирильчик Данил Евгеньевич — студент 4 курса механико-математического факультета НГУ. E-mail: d.kirilchik@g.nsu.ru

ОЦЕНИВАНИЕ НЕКОТОРЫХ ФУНКЦИОНАЛОВ ФУНКЦИЙ РАСПРЕДЕЛЕНИЯ ПО ЗАВИСИМЫМ НЕПОЛНЫМ НАБЛЮДЕНИЯМ В ТЕОРИИ МАССОВОГО ОБСЛУЖИВАНИЯ

Р. С. Мурадов¹, Н. Т. Душатов²

¹ *Наманганский инженерно-технологический институт, г. Наманган, Узбекистан*

² *Алмалыкский филиал Ташкентского государственного технического университета имени Ислама Каримова, г. Алмалык, Узбекистан*

В теории массового обслуживания (ТМО) часто возникает задача оценки функций среднего остаточного времени жизни в зависимых моделях при случайном цензурировании наблюдений справа. Это важно для анализа временных параметров системы массового обслуживания, таких как время ожидания, время обслуживания и другие характеристики. В статье рассматривается оценивание функции среднего остаточного времени жизни в зависимой модели при случайном цензурировании наблюдений справа. Оценка построена с использованием архимедова копула функций. Доказана свойство состоятельности оценки. **Ключевые слова:** *неполные-цензурированные данные, среднее остаточное время жизни, архимедовы-копула функции, функции риска, теории массового обслуживания.*

Введение

В ТМО при работе с зависимыми моделями и неполными наблюдениями оценка функции среднего остаточного времени жизни может быть весьма сложной, но решающей для понимания надежности и производительности системы. В ТМО зависимости между событиями (например, временем обслуживания, временем между прибытиями) являются общими и могут быть смоделированы с использованием копул или других структур зависимостей.

Рассмотрим последовательность $\{(X_k, Y_k), k \geq 1\}$ - независимых и одинаково распределенных (н.о.р.) пар неотрицательных случайных величин (с.в.) с общей совместной функцией распределения (ф.р.) $H(x, y) = P(X_1 \leq x, Y_1 \leq y)$, $(x, y) \in \bar{R}^{+2} = [0, \infty) \times [0, \infty)$. Пусть маргинальные ф.р. $F(x) = P(X_1 \leq x) = H(x, +\infty)$ и $G(y) = P(Y_1 \leq$

Авторы благодарят профессора Ташкентского филиала МГУ, кафедры прикладной математики и информатики Абдурахима Абдушукурова.

$y) = H(+\infty, y)$, $x, y \in \bar{R}^+$, являются непрерывными, $F(0) = G(0) = 0$. Рассмотрим случай, когда последовательность $\mathbb{X} = \{X_k, k \geq 1\}$ цензурируется справа последовательностью $\mathbb{Y} = \{Y_k, k \geq 1\}$ и в n -м шаге эксперимента наблюдается выборка $\mathbb{V}^{(n)} = \{(Z_k, \delta_k), 1 \leq k \leq n\}$, где $Z_k = \min(X_k, Y_k)$ и $\delta_k = I(Z_k = X_k)$, т.е. интересующие нас с.в. X_k наблюдаемы лишь в случае $\delta_k = 1$. Следует отметить, что последовательности \mathbb{X} и \mathbb{Y} могут быть и зависимыми. В данной статистической модели задача состоит в оценивании функционалов от неизвестной ф.р. F при мешающей неизвестной ф.р. G . Такие задачи возникают не только в ТМО, но и в области биомедицины, социологических исследований, инженерии, страхового дела и других прикладных исследований (см. [1]).

1. Зависимые модели и метод оценивания

Рассмотрим оценку, исследованную в работах [2,3] для ф.р. F . Пусть $C(u; v)$ -копула-функция, соответствующая паре (X_1, Y_1) и $C^*(u; v) = u+v-1+C(1-u, 1-v)$, $(u, v) \in [0, 1]^2$, отвечающая ей копула выживания, является сама архимедовой копулой: $C^*(u; v) = \varphi^{-1}[\varphi(u) + \varphi(v)]$ (см. [4]). Здесь φ -функция-генератор копулы $\varphi: [0, 1] \rightarrow \bar{R}^+$ - непрерывная, строго убывающая функция такая, что $\varphi(1) = 0$, $\varphi(0) = \infty$ и φ^{-1} -обратная для φ функция. Отметим, что в случае независимого цензурирования, т.е. когда последовательности \mathbb{X} и \mathbb{Y} независимы и копула-генератор $\varphi(u) = -\ln u$, $u \in (0, 1)$, $\varphi^{-1}(t) = \exp(-t)$, $t \geq 0$. Определим оценку F_n формулой $F_n(x) = 1 - S_n(x)$, где

$$S_n^X(x) = \varphi^{-1} \left[\varphi \left(\widehat{S}_n^Z(x) \right) \mu_n(x) \right] = 1 - F_n(x), \tag{1}$$

$$\mu_n(x) = \varphi(S_n^X(x)) / \varphi(\widehat{S}_n^Z(x)), \quad \varphi(S_n^X(x)) = - \int_0^x \varphi'(S_n^Z(u)) dH_n^{(1)}(u),$$

$$\varphi(\widehat{S}_n^Z(x)) = - \int_0^x n \left[\varphi(S_n^Z(u)) - \varphi\left(S_n^Z(u) - \frac{1}{n}\right) \right] dH_n^{(1)}(u),$$

$$\varphi(\widetilde{S}_n^Z(x)) = - \int_0^x \varphi'(S_n^Z(u)) dH_n(u), \quad H_n(x) = \frac{1}{n} \sum_{i=1}^n I(Z_i \leq x) = 1 - \widetilde{S}_n^Z(x),$$

$$H_n^{(1)}(x) = \frac{1}{n} \sum_{i=1}^n I(Z_i \leq x, \delta_i = 1).$$

2. Оценивание функции среднего остаточного времени жизни

Рассмотрим функционал

$$\begin{aligned} E(x) &= E(x; F) = E(X_1 - x / X_1 > x) = \\ &= (S^X(x))^{-1} \cdot \int_x^{+\infty} S^X(u) du, \quad x \in [0, T_X], \end{aligned} \quad (2)$$

где $T_X = \inf \{x \geq 0 : S^X(x) = 0\}$, $S^X = 1 - F$.

При анализе данных типа времени жизни функционал $E(\bullet; F)$ называется средним остаточным временем жизни, который существует при условии $\mu = E(0; F) = EX_1 < +\infty$. Здесь в модели зависимого случайного цензурирования справа построим следующую оценку для $E(x)$:

$$E_n(x) = E(x; F_n) = \begin{cases} (S_n^X(x))^{-1} \cdot \int_x^\infty S_n^X(u) du, & x \in [0, Z^{(n)}], \\ 0, & x \geq Z^{(n)}, \end{cases} \quad (3)$$

где $Z^{(n)} = \inf \{x \geq 0 : S_n^Z(x) = 0\} = \max \{Z_1, \dots, Z_n\}$.

Пусть $T_Z = \inf \{x \geq 0 : S^Z(x) = 0\}$, $S^Z(x) = P(Z_1 > x)$. Тогда $T_Z = \min \{T_X, T_Y\}$. Поскольку оценки (1) и (3) определены на интервале $[0, Z^{(n)}]$, а также при $n \rightarrow \infty$

$$Z^{(n)} \xrightarrow{\text{П.Н.}} T_Z, \quad (4)$$

то доопределим функцию (2), полагая

$$E(x) = \begin{cases} E(x; F), & x < T_Z, \\ 0, & x \geq T_Z. \end{cases} \quad (5)$$

Определим последовательность с.в. $\varepsilon_n(F) = \sup_{0 \leq x < \infty} \chi(F(x)) |E_n(x) - E(x)|$,

где весовая функция $\chi : [0, 1] \rightarrow \overline{R}^+$ удовлетворяет условиям:

(А) χ — измеримая функция и при каждом $\eta > 0$:

$$\sup \{\chi(u) : u \in [0, 1 - \eta]\} < \infty;$$

(В) Функция $\chi^*(u) = \chi(u)(1 - u)^{-1}$ не убывает в окрестности 1;

$$(C) \int_0^{T_X} \left\{ (S^X(x))^{-1} \cdot \int_x^{T_X} \chi(F(y)) dy \right\} dF(x) < \infty.$$

В нижеприводимых условиях на распределение H и копулы φ под $\Lambda(x)$ подразумеваются интегральные интенсивности $\Lambda^X(x) = -\log S^X(x)$ и $\Lambda^Z(x) = -\log S^Z(x)$:

(D) Функция φ строго убывает на $(0, 1]$ и является достаточно гладкой в следующем смысле: первые две производные функций $\varphi(x)$ и $\psi(x) =$

$-x\varphi'(x)$ ограничены для $x \in [\varepsilon, 1]$, где $\varepsilon > 0$ произвольное число. Более того, производное φ' ограничено и отделено от нуля на $[0, 1]$;

$$(E) \int_0^{Tz} [\psi(S^Z(x))]^2 d\Lambda(x) < \infty;$$

$$(F) \int_0^{Tz} |\psi'(S^Z(x))| d\Lambda(x) < \infty.$$

Имеет место следующее утверждение о равномерной состоятельности $E_n(x)$ с весом.

Теорема 1. Пусть $\mu = EX_1 < \infty$ и выполнены условия (A)-(F). Тогда при $n \rightarrow \infty$

$$\varepsilon_n(F) \xrightarrow{P} 0.$$

Заключение

Оценка функции среднего остаточного времени жизни в теории массового обслуживания в рамках зависимых моделей с неполными данными требует интеграции моделирования зависимостей на основе копул с методами анализа выживаемости. Такой подход позволяет более точно оценить надежность и производительность системы, принимая во внимание сложные зависимости, присущие системам массового обслуживания.

СПИСОК ЛИТЕРАТУРЫ

1. *Abdushukurov A. A.* Nonparametric estimation of distribution function based on relative-risk function // Commun.Statist.: Theory and Methods. 1998. Vol. 27. № 8. P. 1991–2012.
2. *Abdushukurov A. A.* Estimation of conditional survival function under dependent right random censored data // Austrian J. Statist. 2020. Vol. 49. P. 1–8.
3. *Абдушукуров А. А.* Обобщение степенной оценочной функции относительного риска при зависимых случайно цензурированных данных // Современная математика. Фундаментальные направления. 2022. Т. 68. № 1. С. 1–13.
4. *Nelsen R. B.* An Introduction to Copulas. New York: Springer, 2006. 272 p.

Муратов Рустамжон Собитхонович — доктор физико-математических наук, доцент. E-mail: rustamjonmuradov@gmail.com

Душатов Нурлан Тоирбекович — ассистент. E-mail: n_dushatov@rambler.ru

СРАВНЕНИЕ ОЦЕНОК ФУНКЦИИ ВЫЖИВАНИЯ С УЧЁТОМ ИХ СВОЙСТВА ИДЕНТИФИЦИРУЕМОСТИ С МОДЕЛЬЮ НЕПОЛНЫХ НАБЛЮДЕНИЙ

А. А. Абдушукуров¹, С. Б. Бозоров²

¹ *Московский государственный университет, филиал в
г. Ташкенте, Ташкент, Узбекистан*

² *Гулистанский государственный университет, г. Гулистан, Узбекистан*

Рассматривается модель случайно-цензурированных справа наблюдений. В этой модели вводятся три типа оценок функции выживания: экспоненциальная, множительная, степенная. Введенные оценки проверяются на идентифицируемость с рассматриваемой моделью. Представлены численные примеры сравнения оценок на основе естественной выборки. Построены доверительные полосы для неизвестной функции выживания. Численные результаты подтверждают то, что степенные оценки определены на всей прямой в отличие от остальных двух. Новизна работы заключается в том, что только степенные оценки обладают свойством идентифицируемости.

Ключевые слова: *случайное цензурирование, функции выживания, непараметрическое оценивание, идентифицируемость.*

Введение

Задача построения хорошей оценки для функции распределения на базе неполных данных, которая обладала бы аналогичными свойствами что и эмпирическая функция распределения, была предметом исследований многих статистиков. Существующие в статистической литературе оценки для функции распределения по неполным данным имеют более сложные структуры чем эмпирическая функция распределения. Целью работы является сравнение экспоненциальной, множительной и степенной оценок функции выживания при случайном цензурировании справа.

1. Построение оценок

Рассмотрим модель случайно-цензурированных справа наблюдений. Предположим, что X_1, X_2, \dots и Y_1, Y_2, \dots — последовательности двух независимых случайных величин (с.в.) с соответствующими функциями распределения (ф.р.) $F(t)$ и $G(t)$, $t \in R$. Основная задача статистики

заключается в оценивании ф.р. $F(t)$ при мешающей ф.р. $G(t)$. Наблюдению доступна выборка объёма n : $\mathbb{C}^{(n)} = \{(Z_i, \delta_{(i)}), 1 \leq i \leq n\}$, где $Z_{(i)} = \min(X_i, Y_i)$ и индикатор $\delta_{(i)} = I(X_i \leq Y_i) = I(Z_i = X_i)$. Таким образом, в выборке $\mathbb{C}^{(n)}$ число наблюдаемых X_i равно случайному числу $\nu(n) = \delta_{(1)} + \delta_{(2)} + \dots + \delta_{(n)}$ с биномиальным распределением $Bi(n; p)$, где $p = P(\delta_i = 1)$.

Первая оценка множительного типа (product-limit) была предложена в 1958 году Капланом-Мейером [8] и она задаётся формулой:

$$F_n^{KM}(t) = \begin{cases} 1 - \prod_{\{j: Z_{(j)} \leq t\}} \left(1 - \frac{\delta_{(j)}}{n-j+1}\right), & t \leq Z_{(n)}, \\ 1, & t > Z_{(n)}, \quad \delta_{(n)} = 1, \\ \text{неопределена,} & t > Z_{(n)}, \quad \delta_{(n)} = 0. \end{cases} \quad (1)$$

Здесь $Z_{(1)} \leq Z_{(2)} \leq \dots \leq Z_{(n)}$ — вариационный ряд по Z — подвыборке $\{Z_i, 1 \leq i \leq n\}$, а $\delta_{(1)}, \delta_{(2)}, \dots, \delta_{(n)}$ сопровождают члены $Z_{(i)}$ этого вариационного ряда. Оценка (1) имеет огромное влияние на развитие статистической теории цензурированных наблюдений. В литературе встречаются различные модификации оценки (1) и они исследуются и используются во множестве работ [2,3,8]. Для рассмотрения других типов оценок введём в рассмотрение ф.р. наблюдаемого минимума $H(t) = P(Z_i \leq t) = 1 - (1 - F(t))(1 - G(t))$ и субраспределения $H^{(0)}(t) + H^{(1)}(t) = H(t)$. Введём также кумулятивные функции интенсивности распределений $H(t)$, $G(t)$ и $F(t)$ равенствами

$$\Lambda(t) = \int_{-\infty}^t \frac{dH(u)}{1 - H(u)} = \Lambda^{(0)}(t) + \Lambda^{(1)}(t), \quad (2)$$

где

$$\Lambda^{(0)}(t) = \int_{-\infty}^t \frac{dH^{(0)}(u)}{1 - H(u)}, \quad \Lambda^{(1)}(t) = \int_{-\infty}^t \frac{dH^{(1)}(u)}{1 - H(u)}.$$

А их оценки имеют вид

$$\Lambda_n(t) = \int_{-\infty}^t \frac{dH_n(u)}{1 - H_n(u) + \frac{1}{n}} = \Lambda_n^{(0)}(t) + \Lambda_n^{(1)}(t). \quad (3)$$

Здесь

$$\Lambda_n^{(0)}(t) = \int_{-\infty}^t \frac{dH_n^{(0)}(u)}{1 - H_n(u) + \frac{1}{n}}, \quad \Lambda_n^{(1)}(t) = \int_{-\infty}^t \frac{dH_n^{(1)}(u)}{1 - H_n(u) + \frac{1}{n}},$$

где слагаемое $\frac{1}{n}$ в знаменателе этих оценок добавлено с целью предотвращения операции деления на ноль. В этих формулах

$$H_n(t) = \frac{1}{n} \sum_{i=1}^n I(Z_i \leq t) = H_n^{(0)}(t) + H_n^{(1)}(t),$$

$$H_n^{(m)}(t) = \frac{1}{n} \sum_{i=1}^n I(Z_i \leq t, \delta_i = m), m = 1, 2$$

— эмпирические оценки $H(t)$ и $H^{(m)}(t)$ соответственно. С использованием функций (2) и (3) легко убедиться в том, что ещё одной, естественной экспоненциальной является оценка Альтшулера-Бреслоу, предложенная в работах [5, 6]:

$$\begin{aligned} F_n^{AB}(t) &= 1 - \exp\left(-\Lambda_n^{(1)}(t)\right) = \\ &= 1 - \exp\left\{-\sum_{i=1}^n \frac{I(Z_i \leq t, \delta_i = 1)}{n \left[1 - H_n(Z_i) + \frac{1}{n}\right]}\right\} = \\ &= 1 - \exp\left(-\sum_{\{i: Z_{(i)} \leq t\}} \left(\frac{\delta_{(i)}}{n-i+1}\right)\right) = \\ &= 1 - \left[\prod_{\{i: Z_{(i)} \leq t\}} \exp\left(-\frac{1}{n-i+1}\right)\right]^{\delta_{(i)}}. \end{aligned} \quad (4)$$

Для сравнения с (4), используем следующую модификацию оценки (1):

$$\begin{aligned} \hat{F}_n^{KM}(t) &= \begin{cases} 1 - \prod_{\{i: Z_{(i)} \leq t\}} \left(1 - \frac{\delta_{(i)}}{n-i+1}\right), & t \leq Z_{(n)}, \\ 1, & t > Z_{(n)}, \end{cases} = \\ &= \begin{cases} \prod_{\{i: Z_{(i)} \leq t\}} \left(\frac{n-i}{n-i+1}\right)^{\delta_{(i)}}, & t \leq Z_{(n)}, \\ 1, & t > Z_{(n)}. \end{cases} \end{aligned} \quad (5)$$

Поскольку при $x \approx 0$, $e^{-x} \approx 1 - x$, то из формул (4) и (5) легко получаем аппроксимацию:

$$-\log\left(1 - \hat{F}_n^{KM}(t)\right) \approx \Lambda_n^{(1)}(t), \quad t \in R,$$

что показывает близость этих оценок (1) и (4) при больших n .

Автором [2,3] была предложена оценка степенной структуры для $F(t)$ вида

$$F_n(t) = 1 - (1 - H_n(t))^{R_n(t)} = \begin{cases} 0, & t < Z_{(1)}, \\ 1 - \left(\frac{n-j}{n}\right)^{R_n(t)}, & Z_{(j)} \leq t < Z_{(j+1)}, \quad 1 \leq j \leq n-1, \\ 1, & t \geq Z_{(n)}, \end{cases} \quad (6)$$

где $R_n(t) = \Lambda_n^{(1)}(t) \cdot [\Lambda_n(t)]^{-1}$ — оценка дроби $R(t) = \Lambda^{(1)}(t) \cdot [\Lambda(t)]^{-1}$.

2. Численные результаты

В работах [1,2,3,4,7] установлено, что оценки (1), (4) и (6) асимптотически (при $n \rightarrow \infty$) эквивалентны в смысле сходимости к одному и тому же предельному гауссовскому процессу с нулевым средним. Однако, есть и отличительные стороны всех этих трех оценок при небольшом объеме выборки n .

Пусть $G_n^{KM}(t)$, $G_n^{AB}(t)$ и $G_n(t)$ — оценки цензурирующего распределения $G(t)$, получаемые соответственно из формул (1), (4) и (6) заменой индикаторов δ_i на $1 - \delta_i$ при $t \in R$:

$$G_n^{KM}(t) = \begin{cases} 1 - \prod_{\{i: Z_{(i)} \leq t\}} \left(1 - \frac{1 - \delta_{(i)}}{n - i + 1}\right), & t \leq Z_{(n)}, \\ 1, & t > Z_{(n)}, \quad \delta_{(n)} = 0, \\ \text{неопределена,} & t > Z_{(n)}, \quad \delta_{(n)} = 1, \end{cases}$$

$$G_n^{AB}(t) = 1 - \exp\left(-\Lambda_n^{(0)}(t)\right) \text{ и } G_n(t) = 1 - [1 - H_n(t)]^{1 - R_n(t)}.$$

Для степенных оценок (6), т.е. для $F_n(t)$ и $G_n(t)$, справедливо равенство

$$(1 - F_n(t))(1 - G_n(t)) \equiv 1 - H_n(t), \quad t \in R, \quad (7)$$

которое является эмпирическим аналогом равенства

$$(1 - F(t))(1 - G(t)) = 1 - H(t), \quad t \in R, \quad (8)$$

справедливого для распределения минимума в рассматриваемой модели, что означает свойство идентифицируемости с рассматриваемой моделью. Однако, для остальных двух типов оценок имеем соотношения:

I. $(1 - F_n^{KM}(t))(1 - G_n^{KM}(t)) \neq 1 - H_n(t)$, так как при $t \geq Z_{(n)}$, обе оценки $F_n^{KM}(t)$ и $G_n^{KM}(t)$ неопределены.

II. $(1 - F_n^{AB}(t))(1 - G_n^{AB}(t)) = \exp(-\Lambda_n(t)) \neq 1 - H_n(t)$ и $\max\{F_n^{AB}(t), G_n^{AB}(t)\} < 1$.

Таким образом, для случая непрерывных распределений $F(t)$ и $G(t)$ только степенные оценки $F_n(t)$ и $G_n(t)$ обладают свойством идентифицируемости с моделью случайного цензурирования справа.

Для исследования этих трёх оценок используем выборку объёма $n = 97$ из работы [6]. Вариационный ряд $\{Z_{(i)}, \delta_{(i)}, i = 1, \dots, 97\}$, построенный по этим данным, имеет вид:

(777,1), (781,0), (843,0), (866,0), (869,1), (872,1), (876,1), (893,1), (894,1), (895,0), (898,1), (906,0), (907,1), (909,1), (911,1), (911,0), (914,0), (927,1), (932,1), (936,0), (940,0), (942.5,0), (943,0), (945,1), (945,0), (948,1), (951,0), (953,0), (956,0), (957,1), (957,0), (959,0), (960,0), (966,1), (966,0), (969,1), (970,0), (971,1), (972,0), (973,0), (977,0), (983,1), (984,0), (985,1), (989,1), (992.5,1), (993,1), (996,1), (998,1), (1001,0), (1002,0), (1005,0), (1006,0), (1009,1), (1011.5,1), (1012,1), (1012,0), (1013,0), (1015,0), (1016,0), (1018,0), (1022,1), (1023,0), (1025,1), (1027,0), (1029,1), (1031,1), (1031,0), (1031.5,0), (1033,1), (1036,1), (1043,1), (1043,0), (1044,1), (1044,0), (1045,0), (1047,0), (1053,1), (1055,1), (1058,0), (1059,1), (1060,1), (1060,0), (1064,0), (1070,0), (1073,0), (1080,1), (1085,1), (1093,0), (1093.5,1), (1094,1), (1106,0), (1107,0), (1118,0), (1128,1), (1139,1), (1153,0).

Здесь эти данные выражены в месяцах, где “1” — соответствует нецензурированию, а “0” — цензурированию.

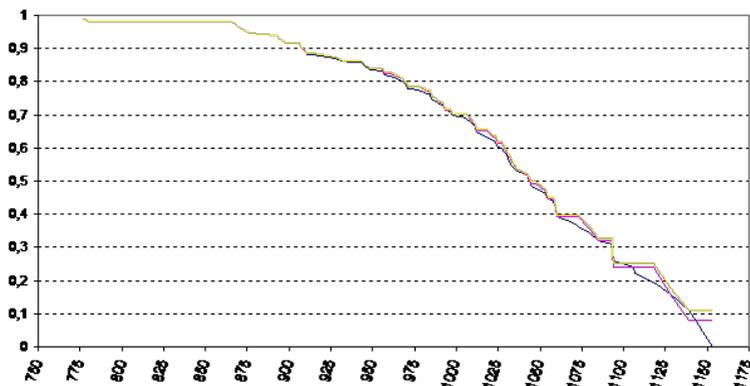


Рис. 1. Оценки $1 - F_{97}^{AB}(t)$, $1 - F_{97}^{KM}(t)$ и $1 - F_{97}(t)$.

Из рисунка 1 видно, что в отличие от экспоненциальных и множительных оценок только степенные оценки определены на всей прямой.

Заключение

Сравниваются три вида оценок: экспоненциальная, множительная и степенная для функции выживания при случайном цензурировании справа. Ранее была установлена асимптотическая эквивалентность этих трех видов оценок при растущем объёме выборки в смысле сходимости к одному и тому же гауссовскому процессу. Для конкретной конечной выборки объёма $n = 97$ показаны некоторые преимущества степенной оценки по сравнению с остальными двумя.

СПИСОК ЛИТЕРАТУРЫ

1. *Абдушукуров А. А.* Статистика неполных наблюдений. // Ташкент: "Университет", 2009. P. 270.
2. *Abdushukurov A. A.* Nonparametric estimation of the survival function from censored data based on relative risk function. Commun. in Stat. - Th. and Meth, 1998. 27 p.
3. *Абдушукуров А. А.* On nonparametric estimation of reliability indices by censored samples. // Theory Probab. Appl. 1999. P. 3–11.
4. *Altshuler B.* Theory for the measurement for competing risks in animal experiments. // Math. Biosciences, 1970. P. 1–11.
5. *Breslow N.* Discussion on professor Cox's paper. // J. Royal Statist. Society, JRSS, 1972. P. 216–217.

6. *Efron B.*, Censored Data and the Bootstrap. // Journal of the American Statistical Association. JASA. 1981. Vol. 76, № 374, P. 312–319.
7. *Csörgő S., Horvath L.* Confidence bands from censored samples. // Canadian Journal of Statistics-revue Canadienne De Statistique. Canadian J. Statist. 1986. Vol. 14, Issue 2. P. 131–144.
8. *Kaplan E., Meier P.* Nonparametric Estimation from Incomplete Observations. // Springer Series in Statistics. Springer Ser. in Stat. 1958. № 282(53). P. 457–481.

Абдурахим Ахмедович Абдушукуров — доктор физико-математических наук, профессор, Московский государственный университет, филиал в г. Ташкенте, (Ташкент, Узбекистан).. E-mail: a_abdushukurov@rambler.ru

Сухроб Ваходирович Бозоров — докторант, Гулистанский государственный университет (Гулистан, Узбекистан).. E-mail: suxrobbek_8912@mail.ru.

КОНСТРУКТИВНЫЙ АЛГОРИТМ УРАВНИВАНИЯ ВЕРОЯТНОСТЕЙ

В. Л. Брызгалов¹, А. В. Войтишек^{2,3}

¹ Лицей № 130 г. Новосибирска, Россия

² Институт вычислительной математики и математической геофизики
СО РАН, г. Новосибирск, Россия

³ Новосибирский государственный университет, г. Новосибирск, Россия

В работе рассмотрена оригинальная технология построения специальной неравномерной сетки, определяющей уравнивание вероятностей для авторского универсального (во всяком случае, для случайных величин, имеющих монотонные плотности на конечных интервалах распределения) экономичного двустороннего алгоритма метода исключения с кусочно-постоянными мажорантой и минорантой. Представлена созданная на основании технологии уравнивания вероятностей и использования двустороннего алгоритма компьютерная система EDSRM (Economic Double-Sided Rejection Method). Эффективность построенного двустороннего алгоритма продемонстрирована на примере моделирования случайной величины, имеющей степенное распределение. **Ключевые слова:** численное моделирование случайных величин, уравнивание вероятностей, двусторонний алгоритм метода исключения с кусочно-постоянными мажорантой и минорантой, построение неравномерных сеток.

Введение

Численное моделирование случайных величин (компьютерная реализация выборочных значений) является ключевым элементом алгоритмов метода Монте-Карло (см., например, [1]). Учитывая относительно медленную сходимость метода Монте-Карло (со скоростью $1/\sqrt{n}$ по числу n используемых выборочных значений – см., например, [1]), такое моделирование должно быть экономичным.

В работах [2, 3] показано, что во многих случаях при численном моделировании выборочного значения ξ_0 случайной величины ξ , распределенной на конечном интервале (отрезке) $[a, b]$, традиционно считающиеся наиболее экономичными формулы метода обратной функции

распределения

$$\xi_0 = F_\xi^{-1}(\alpha_0); \quad \alpha_0 \in U(0, 1), \quad (1)$$

являются трудоемкими. В формуле (1) $F_\xi(x) = \int_{-\infty}^x f_\xi(u) du$ – это функция распределения, а α_0 – стандартное случайное число (выборочное значение случайной величины $\alpha \in U(0, 1)$, равномерно распределенной на интервале $(0, 1)$ – такое значение получается на компьютере с помощью соответствующего генератора – см. раздел 1.1 книги [1]).

В разделе 1 данной работы для монотонной, определенной на конечном интервале $[a, b]$ плотности $f_\xi(u)$; $u \in [a, b]$ сформулирован двусторонний алгоритм метода исключения с кусочно-постоянными мажорантой и минорантой. В разделе 2 подробно описана оригинальная процедура уравнивания вероятностей. В разделе 3 представлена созданная на основании технологии уравнивания вероятностей и использования двустороннего алгоритма компьютерная система EDSRM (Economical Double-Sided Rejection Method) и показана эффективность ее использования на примере моделирования степенного распределения. В разделе «Заключение» сформулированы основные выводы работы и перспективы развития предложенных вычислительных технологий.

1. Двусторонний алгоритм метода исключения

Рассмотрим случайную величину ξ и запишем ее плотность в виде

$$f_\xi(u) = H f(u); \quad u \in [a, b]; \quad -\infty < a < b < +\infty; \quad H = \frac{1}{\int_a^b f(v) dv}; \quad (2)$$

здесь $f(u)$ – непрерывная, монотонная (для определенности – возрастающая) на отрезке $[a, b]$ функция («существенная часть» плотности $f_\xi(u)$), а H – соответствующая нормирующая константа.

Введем на отрезке $[a, b]$ сетку (вообще говоря, неравномерную)

$$a = u_0 < u_1 < u_2 < \dots < u_{M-1} < u_M = b \quad (3)$$

и рассмотрим кусочно-постоянные функции (мажоранту и миноранту функции $f(u)$ из соотношения (2))

$$f^{(1)}(u) \equiv A_i^{(+)} = f(u_i); \quad f^{(2)}(u) \equiv A_i^{(-)} = f(u_{i-1}); \quad u \in \Delta_i = (u_{i-1}; u_i] \quad (4)$$

для всех $i = 1, \dots, M$.

Проведем т. н. **уравнивание вероятностей**, что означает такой выбор сетки (3), чтобы выполнялось соотношение

$$S_i = A_i^{(+)}(u_i - u_{i-1}) = f(u_i)(u_i - u_{i-1}) \equiv S = \text{const}; \quad i = 1, \dots, M \quad (5)$$

(подробности см. в разделе 2). В этом случае можно сформулировать следующий универсальный экономичный двусторонний алгоритм метода исключения (см., например, раздел 1.7.4 книги [1]).

АЛГОРИТМ 1. *Предварительно формируем массив $\mathbf{U}^{(M)} = \{u_0, u_1, u_2, \dots, u_{M-1}, u_M\}$ узлов сетки (3) согласно алгоритмам 2 и 3 (см. далее раздел 2), а также массивы минимумов и максимумов функции $f(u)$ из формулы (4) на элементах разбиения Δ_i : $\mathbf{MIN}^{(M)} = \{A_1^{(-)} = f(u_0), A_2^{(-)} = f(u_1), \dots, A_M^{(-)} = f(u_{M-1})\}$ и $\mathbf{MAX} = \{A_1^{(+)} = f(u_1), A_2^{(+)} = f(u_2), \dots, A_M^{(+)} = f(u_M)\}$.*

В свою очередь, для реализации каждого выборочного значения ξ_0 случайной величины ξ , распределенной согласно монотонно возрастающей плотности (2), реализуем на компьютере следующие действия.

1. Моделируем выборочное значение $\xi_0^{(1)}$ согласно формулам

$$m = [M\alpha_1] + 1; \xi_0^{(1)} = u_{m-1} + (u_m - u_{m-1})(M\alpha_1 - m + 1); \alpha_1 \in U(0, 1), \quad (6)$$

а также значение $\eta_0 = \alpha_2 A_m^{(+)} \in U(0, A_m^{(+)})$; здесь $\alpha_2 \in U(0, 1)$.

Несложно показать, что точка $(\xi_0^{(1)}, \eta_0)$ равномерно распределена в подграфике кусочно-постоянной мажоранты $f^{(1)}(u)$.

2. Сравниваем сначала значение η_0 с соответствующим значением кусочно-постоянной миноранты $f^{(2)}(u)$, т. е. проверяем неравенство $\eta_0 < A_m^{(-)}$. Если оно выполнено, то пара $(\xi_0^{(1)}, \eta_0)$ принадлежит подграфику функции $f^{(2)}(u)$, а значит, и подграфику G функции $f(u)$. Тогда можно положить $\xi_0 = \xi_0^{(1)}$.

3. Если же $\eta_0 \geq A_m^{(-)}$, то проверяем неравенство $\eta_0 < f(\xi_0^{(1)})$. Если оно выполнено, то $(\xi_0^{(1)}, \eta_0) \in G$, и тогда снова можно положить $\xi_0 = \xi_0^{(1)}$. В противном случае (т. е. при $\eta_0 \geq f(\xi_0^{(1)})$) повторяем пункт 1 и т. д. – до получения значения ξ_0 .

2. Уравнивание вероятностей для двустороннего алгоритма

Рекордную экономичность алгоритма 1 определяет применение формул из соотношения (6), что возможно только при таком выборе узлов сетки (3), при котором справедливо соотношение (5).

Руководствуясь идеями из работы [4], отметим, что сетка (3), обеспечивающая выполнение соотношения (5), может быть построена с помощью алгоритма «подбора» (с компьютерной точностью), соответствующего следующей процедуре.

АЛГОРИТМ 2. Определим процедуру-функцию от $u_{M-1} = r$ (наибольшего из левых концов полуинтервалов $(u_{i-1}, u_i]$) – обозначим ее $z(r)$ – как следующую последовательность команд в коде Maple:

$$z(r) : u_{M-1} = r; S = (b - r)f(b); \text{ for } i \text{ from } M - 2$$

$$\text{ by } -1 \text{ to } 1 \text{ do } u_i = u_{i+1} - \frac{S}{f(u_{i+1})} \text{ od}; \text{ return } [S - (u_1 - a)f(u_1)].$$

«Двигаем» (подбираем) параметр r до тех пор, пока с высокой точностью не будем иметь $z(r) \approx 0$.

Отметим, что в работе [4] не описаны точные процедуры получения параметра r , обеспечивающих выполнение соотношения $z(r) \approx 0$ с указанной в [4] точностью $\varepsilon = 10^{-L}$ для $L = 16$ (этот порядок соответствует точности воспроизведения вещественных чисел на ЭВМ – при этом алгоритм «подбора» дает по сути точную сетку (3) для уравнивания вероятностей).

Предлагается следующая процедура приближения.

АЛГОРИТМ 3. Введем $\Delta = \frac{b-a}{M}$, возьмем $r_0 = b - \Delta$ и вычислим $z(r_0)$. Для монотонно возрастающей функции имеем $z(r_0) > 0$, т. е. площадь S «излишне большая» и параметр $r = r_0$ нужно увеличить.

Пусть стандартное представление числа $\Delta = a \times 10^k$; $1 \leq a < 10$ имеет порядок k (это целое – положительное или отрицательное – число).

Возьмем $\varepsilon^{(1)} = 10^k$ и начнем прибавлять эту величину к r_0 до тех пор, пока первый раз не получим точку r_1 такую, что $z(r_1) < 0$.

Теперь возьмем $\varepsilon^{(2)} = 10^{k-1}$ и начнем вычитать эту величину из r_1 до тех пор, пока первый раз не получим точку r_2 такую, что $z(r_2) > 0$ и т. д.

Процедуру продолжаем вплоть до номера итерации s такого, что $\varepsilon^{(s)} < 10^{-L}$ (например, для $L = 16$), и полагаем $r = r_s$.

3. Компьютерная система EDSRM

Универсальный экономичный алгоритм 1 и его аналоги для убывающих и кусочно-монотонных плотностей (2) реализованы нами в компьютерной системе EDSRM (Economical Double-Sided Rejection Method; см. <http://edsrm.andronix1.ru/>), состоящей из библиотеки и удобной для использования исследовательской диалоговой системы.

В библиотеке отдельными модулями представлены: генератор стандартных случайных чисел $\alpha_i \in U(0, 1)$; реализация формулы метода обратной функции распределения (1); реализация экономичного двустороннего алгоритма метода исключения (см. алгоритм 1) для монотонных функций $f(u)$, пропорциональных плотностям вида $f_\xi(u) =$

$Hf(u)$; $u \in [a, b]$ (см. формулу (2)) на ограниченных интервалах (отрезках) распределения $[a, b]$; реализация алгоритма 1 для кусочно-монотонных функций $f(u)$; $u \in [a, b]$; реализация уравнивания вероятностей на заданном отрезке монотонности функции $f(u)$ – см. алгоритмы 2 и 3.

Диалоговая система EDSRM использует описанную библиотеку для сравнительного анализа работы алгоритма 1 и формул метода обратной функции распределения (1) для элементарных монотонных плотностей (2) и позволяет: вносить в систему аналитическое представление функции $f(u)$ из (2) и соответствующие границы отрезка $[a, b]$; вносить формулы метода обратной функции распределения вида (1) для соответствующей плотности (2); проводить сравнение (с построением соответствующих диаграмм) затрат алгоритма 1 и формулы обратной функции распределения (1) для различных монотонных функций $f(u)$ для различных значений параметров n (это количество моделируемых выборочных значений случайной величины $\xi \in [a, b]$) и M (это количество интервалов разбиения для неравномерной сетки $\mathbf{U}^{(M)}$ из алгоритма 1); наблюдать результаты построения кусочно-постоянных мажорант и минорант на неравномерных сетках $\mathbf{U}^{(M)}$, полученных после процедуры уравнивания вероятностей, для различных M – на соответствующих графиках; получать график зависимости времени работы алгоритма 1 от параметра M для заданной функции $f(u)$, $[a, b]$; наблюдать формирование гистограммы для моделируемого распределения и ее приближение (с ростом числа n) к исследуемой плотности распределения.

В качестве достаточно показательного примера эффективности алгоритма 1 отметим, что расчеты, проведенные с помощью системы EDSRM, показали, что двусторонний алгоритм в 1.64 раз эффективнее широко используемой формулы метода обратной функции распределения $\xi_0 = {}^{c+1}\sqrt{\alpha_0}$; $\alpha_0 \in U(0, 1)$ для моделирования степенного распределения с плотностью $f_\xi(u) = (c + 1)u^c$; $u \in [0, 1]$ для всевозможных степеней $c > 0$.

Заключение

В работе рассмотрена оригинальная технология построения специальной неравномерной сетки, определяющей уравнивание вероятностей для авторского универсального (во всяком случае, для случайных величин, имеющих монотонные плотности на конечных интервалах распределения) экономичного двустороннего алгоритма метода исключения с кусочно-постоянными мажорантой и минорантой. Представлена созданная на основании технологии уравнивания вероятностей и ис-

пользования двустороннего алгоритма компьютерная система EDSRM (Econometrical Double-Sided Rejection Method).

Отметим, что относительно несложными для реализации являются возможности построения аналогов алгоритма 1 (включая «зигкурат-конструкции» [4]) для кусочно-монотонных плотностей $f_{\xi}(u)$; $u \in (a, b)$, а также для случайных величин ξ , имеющих неограниченные интервалы распределения (a, b) . Эти возможности частично реализованы нами (в том числе, в системе EDSRM), но требуют отдельного подробного описания.

СПИСОК ЛИТЕРАТУРЫ

1. Михайлов Г. А. Некоторые вопросы теории методов Монте-Карло. Новосибирск: Наука, 1974. 143 с.
2. Рыжов И. А., Трофимов И. А. Исследование экономичности алгоритмов моделирования случайных величин с помощью компьютерной системы NMPUD // Математика: Материалы 59-й Международной науч. студ. конф. 12–23 апреля 2021 г. / Новосиб. гос. ун-т. – Новосибирск: ИПЦ НГУ, 2021. С. 126.
3. Брызгалов В. Л., Войтишек А. В. Анализ трудоемкости формул метода обратной функции распределения для случайных величин с конечным интервалом распределения // Информационные технологии и математической моделирование (ИТММ-2023): Материалы XXII Международной конференции имени А. Ф. Терпугова (4–9 декабря 2023 года), Томск: Издательство Томского государственного университета, 2023. Часть 1. С. 314–321.
4. Marsaglia G., Tsang W. W. The ziggurat method for generating random variables // Journal of Statistical Software. 2000. Vol. 5. Issue 8.

Брызгалов Виктор Леонидович — ученик 11–М класса. E-mail: finerpowerc4acc@gmail.com

Войтишек Антон Вацлавович — доктор физико-математических наук, профессор, главный научный сотрудник, лаборатория стохастических задач ИВМиМГ СО РАН; профессор, кафедра вычислительной математики НГУ. E-mail: vav@osmf.sccc.ru

КОМБИНИРОВАННЫЕ ОЦЕНКИ ФУНКЦИИ РАСПРЕДЕЛЕНИЯ С АПРИОРНОЙ ДОГАДКОЙ О ПАРАМЕТРИЧЕСКОМ КЛАССЕ РАСПРЕДЕЛЕНИЙ

Ю. Г. Дмитриев, Я. В. Кирюшкин

*Национальный исследовательский Томский государственный университет,
г. Томск, Россия*

Рассматривается класс оценок, представляющий собой взвешенные суммы эмпирической функции распределения и некоторой параметрической функции распределения, задаваемой исследователем и называемой априорной догадкой. Оптимальный весовой коэффициент выбирается из критерия минимума среднеквадратической ошибки и определяет оптимальную комбинированную оценку на заданном классе. Строятся адаптивные комбинированные оценки путем замены неизвестного оптимального весового коэффициента на различные статистические оценки. Исследуются асимптотические свойства адаптивных оценок функции распределения, вычисляются асимптотические нормированные среднеквадратические ошибки и проводится их сравнение. **Ключевые слова:** *функция распределения, параметрический класс распределений, априорная догадка, адаптивные оценки, асимптотическое распределение оценки, СКО, комбинированные оценки.*

Введение

Многие задачи обработки экспериментальных данных сводятся к статистическому оцениванию неизвестной функции распределения. Уровень априорной информации о виде функции распределения может быть разнообразным. В работе [1] рассмотрены оценки для параметрического класса функций распределений. В работе [2] предложены оценки с учетом дополнительной информации о вероятностных характеристиках распределений. В данной работе рассмотрена ситуация, когда исследователь делает предположение о виде искомого распределения, однако он в этом не уверен. Такую информацию принято называть априорной догадкой. В данной работе делается предположение о возможном параметрическом классе распределений, строятся оценки и анализируются их свойства.

1. Комбинированная оценка с оптимальным весовым коэффициентом

Пусть X_1, \dots, X_n – независимые наблюдения объема n над случайной величиной X с неизвестной функцией распределения $F(y) = P\{X < y\}$, $y \in R^1$, $F \in \mathbf{F}$ – класс допустимых функции распределения.

Если $\mathbf{F} = \{F(y; \theta), \theta \in \Theta \subseteq R^m\}$, то имеет место параметрический класс распределений, и задача оценивания $F(y)$ сводится к оцениванию неизвестных параметров θ . При отсутствии такой информации класс \mathbf{F} называют непараметрическим и в качестве непараметрической оценки обычно используют эмпирическую функцию распределения (э.ф.р.): $\hat{F}(y) = \frac{1}{n} \sum_{i=1}^n c(y - X_i)$. Известно, что э.ф.р. является несмещенной оценкой, т.е. математическое ожидание $M_F \hat{F}(y) = F(y)$, и $D_F \hat{F}(y) = F(y)(1 - F(y))/n$ – дисперсия.

Функция распределения $F(y; \theta)$ является априорной догадкой, оценивающей неизвестную функцию распределения $F(y)$. Она может совпасть, не совпасть или быть близкой к ней. Задача состоит в построении комбинированной оценки искомой функции распределения, которая одновременно включает непараметрическую оценку $\hat{F}(y)$ и априорную догадку $F(y; \theta)$.

Предположим, что параметр θ задан и значение $F(y; \theta)$ известно. Следуя работе [3], рассмотрим класс оценок функции распределения $F(y)$ в точке y в виде

$$\hat{F}_\lambda = \hat{F}_\lambda(y) = (1 - \lambda)\hat{F}(y) + \lambda F(y; \theta) = \hat{F}(y) - \lambda (\hat{F}(y) - F(y; \theta)), \quad (1)$$

где весовой коэффициент λ выбирается из заданного критерия качества оценки. Пусть критерием качества является среднеквадратическая ошибка (СКО) $S_F^2(\lambda) = M_F (\hat{F}_\lambda(y) - F(y))^2$. В этом случае оптимальное λ_0 , доставляющее минимум СКО, определяется выражением

$$\lambda_0 = \lambda_0(y) = \frac{1}{1 + \frac{n\Delta_F^2(y)}{\sigma_F^2(y)}} = \frac{1}{1 + b_F^2(y)}. \quad (2)$$

Здесь $\Delta_F^2(y) = F(y) - F(y; \theta)$ – величина отклонения априорной догадки $F(y; \theta)$ от истинного значения $F(y)$, $\sigma_F^2(y) = D_F [c(y - X)] = F(y)(1 - F(y))$, а

$$b_n(y) = \frac{\sqrt{n}\Delta_F(y)}{\sigma_F(y)} \quad (3)$$

— нормированное отклонение.

При оптимальном λ_0 минимум СКО в точке y удовлетворяет соотношению $S_F^2(\lambda_0) = D_F \hat{F}(y) - \frac{(D_F \hat{F}(y))^2}{D_F \hat{F}(y) + \Delta_F^2(y)} = (1 - \lambda_0(y)) \frac{\sigma_F^2(y)}{n}$. Оценку $\hat{F}_0 = \hat{F}_{\lambda_0}(y)$ из класса (1) назовем оптимальной. Весовой коэффициент λ_0 изменяется в пределах $0 < \lambda_0 \leq 1$ и показывает, какое влияние на точность оценивания оказывает каждая из величин $\hat{F}(y)$ и $F(y; \theta)$. Его значение зависит от величины нормированного отклонения (3).

2. Адаптивные комбинированные оценки

Практическое применение оценки $\hat{F}_{\lambda_0}(y)$ затруднено в силу незнания оптимального коэффициента λ_0 , который зависит от неизвестной функции распределения F . Здесь мы рассмотрим адаптивные комбинированные оценки

$$\hat{F}_{\hat{\lambda}} = \hat{F}_{\hat{\lambda}}(y) = \hat{F}(y) - \hat{\lambda}(y) \left(\hat{F}(y) - F(y; \theta) \right), \quad (4)$$

построенные путем замены неизвестного λ_0 на какую либо его статистическую оценку $\hat{\lambda}$. Разнообразие возможных оценок приводит к различным адаптивным оценкам (4). Рассмотрим некоторые из них.

1. В качестве первой оценки λ_0 возьмем $\hat{\lambda}_1(y) = \frac{1}{1 + \frac{n\Delta^2}{\sigma^2}} = \frac{1}{1 + b_1^2}$, где

$\hat{\Delta} = \hat{F}(y) - F(y; \theta)$, $\hat{\sigma}^2 = \hat{\sigma}^2(y) = \hat{F}(y) \left(1 - \hat{F}(y) \right)$, $\hat{b}_1 = \frac{\sqrt{n}\hat{\Delta}}{\hat{\sigma}}$ — оценка нормированного отклонения. Полагая в (4) $\hat{\lambda}(y) = \hat{\lambda}_1(y)$, получаем первую адаптивную комбинированную оценку $\hat{F}_1(y) = \hat{F}_{\hat{\lambda}_1}(y)$.

2. Второй оценкой оптимального коэффициента возьмем $\hat{\lambda}_2(y) = \frac{1}{1 + \frac{n\Delta_\theta^2}{\sigma_\theta^2}} = \frac{1}{1 + b_2^2}$, где $\hat{\sigma}_\theta^2 = \hat{\sigma}_\theta^2(y) = F(y; \theta) (1 - F(y; \theta))$, $\hat{b}_2 = \frac{\sqrt{n}\hat{\Delta}}{\hat{\sigma}_\theta}$.

Эта оценка приводит ко второй адаптивной комбинированной оценке $\hat{F}_2(y) = \hat{F}_{\hat{\lambda}_2}(y)$.

3. Заменяем в (2) Δ на $\hat{\Delta}$, а σ^2 на максимальное значение, равное 0,25. Получим третью оценку $\hat{\lambda}_3(y) = \frac{1}{1 + 4n\hat{\Delta}} = \frac{1}{1 + b_3^2}$, где $\hat{b}_3 = 2\sqrt{n}\hat{\Delta}$, и третью адаптивную комбинированную оценку $\hat{F}_3(y) = \hat{F}_{\hat{\lambda}_3}(y)$.

Представляет интерес асимптотическое (при $n \rightarrow \infty$) поведение приведенных оценок.

3. Асимптотические свойства комбинированных оценок

Для изучения асимптотических свойств комбинированных оценок функции распределения $F(y)$ рассмотрим нормированные раз-

ности $\hat{\xi}_j = \frac{\sqrt{n}(\hat{F}_j(y) - F(y))}{\sigma(y)}$, $j = 0, 1, 2, 3$. Обозначим $\hat{\eta} = \hat{\eta}(y) = \frac{\sqrt{n}(\hat{F}(y) - F(y))}{\sigma(y)}$, $\hat{\tau} = \frac{\sigma}{\sigma}$, $\tau_\theta = \frac{\sigma}{\sigma_\theta}$

Теорема 1. Пусть для каждого $F \in \mathbf{F}$ и $-\infty < y < \infty$ выполнено $\lim_{n \rightarrow \infty} b_n(y) = b(y) \in [-\infty, \infty]$.

Тогда при $n \rightarrow \infty$ $\hat{\eta} \Rightarrow \eta$,

1) если $|b(y)| < \infty$, то

$$\begin{aligned} \hat{\xi}_0 \Rightarrow \zeta_0 &= -\frac{b}{1+b^2} + \frac{\eta b^2}{1+b^2}, \hat{\xi}_1 \Rightarrow \zeta_1 = -b + \frac{(\eta+b)^3}{1+(\eta+b)^2}, \\ \hat{\xi}_2 \Rightarrow \zeta_2 &= -b + \frac{(\eta+b)^3}{\frac{1}{\tau_\theta^2} + (\eta+b)^2}, \hat{\xi}_3 \Rightarrow \zeta_3 = -b + \frac{(\eta+b)^3}{\frac{1}{4\sigma^2} + (\eta+b)^2}; \end{aligned}$$

2) при $|b(y)| = \infty$ имеем $\zeta_0 = \zeta_1 = \zeta_2 = \zeta_3 = \eta$.

Здесь знак « \Rightarrow » означает сходимость по распределению, η — стандартная нормальная случайная величина из $N(0, 1)$, $b = b(y)$.

Доказательство. вытекает из теоремы работы [3]. ■

4. Интерпретация результатов

Назовём величины $M\zeta_0^2$, $M\zeta_1^2$, $M\zeta_2^2$, $M\zeta_3^2$ асимптотическими нормированными среднеквадратическими ошибками (АНСКО) соответствующих комбинированных оценок. Приведенная теорема позволяет вычислить АНСКО (проведя усреднение по стандартному нормальному закону) и через их значения сравнить точность оценивания адаптивных комбинированных оценок $\hat{F}_1(y)$, $\hat{F}_2(y)$, $\hat{F}_3(y)$ с оптимальной $\hat{F}_0(y)$ на классе (1).

Отметим, что параметр $b = b(y)$, является интегральной характеристикой АНСКО. Для проведения анализа сделаем замену $F(y) = u$. Для достаточно больших n имеем

$$M\zeta_2^2 = S^2(u, \theta) \approx \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \left(-\beta_n + \frac{(t + \beta_n)^3}{\frac{1}{\pi^2} + (t + \beta_n)^2} \right)^2 e^{-\frac{t^2}{2}} dt, \quad (5)$$

где $\beta_n(u, \theta) = \frac{\sqrt{n}(u - G(u; \theta))}{\sqrt{u - u^2}}$, а $\pi(u, \theta) = \frac{\sqrt{u - u^2}}{\sqrt{G(u; \theta) - G^2(u; \theta)}}$. Выражение для $M\zeta_1^2$ получим, полагая в (5) $\tau_\theta = 1$, а для $M\zeta_3^2$, заменяя τ_θ на 2σ . Полагая в (5) $\tau(u, \theta) = 1$, получаем значение $M\zeta_1^2$. Если принять $\tau(u) = 2\sigma = 2\sqrt{u - u^2}$, получаем значение $M\zeta_3^2$.

Рассмотрим иллюстративные примеры. Возьмем в качестве априорной догадки функцию распределения $G_1(u; \theta) = u^\theta$ и ее симметричный вариант $G_2(u, \theta) = \frac{u^\theta + 1 - (1-u)^\theta}{2}$. Приведем графики асимптотически нормированных СКО описанных выше, представив их как функции от $0 < u < 1$ и $0 < \theta < \infty$. При этом предполагается, что при заданных объемах выборки выполняется нормальное приближение, удовлетворяющее исследователя.

Вычисления $M\zeta_0^2, M\zeta_1^2, M\zeta_2^2, M\zeta_3^2$, при $n = 100$ приведены ниже:

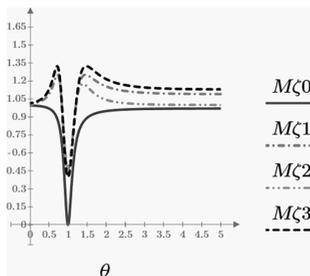


Рис. 1. $G_1(u; \theta)$, $u = 0, 25$

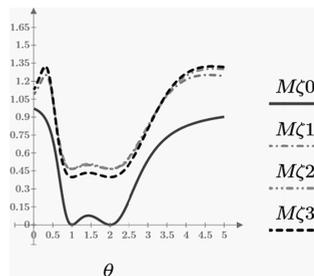


Рис. 2. $G_2(u; \theta)$, $u = 0, 25$

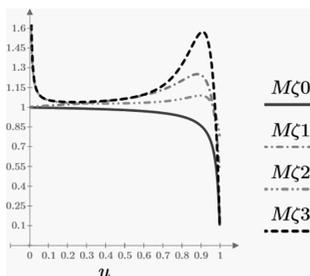


Рис. 3. $G_1(u; \theta)$, $\theta = 0, 25$

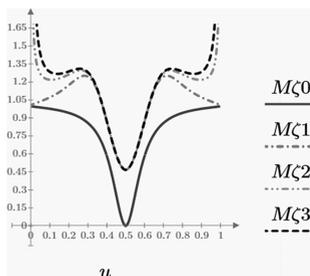


Рис. 4. $G_2(u; \theta)$, $\theta = 0, 25$

Согласно (рис. 1) и (рис. 2) все адаптивные оценки на большей части промежутка θ не точнее обычной эмпирической \hat{F} , но для априорной догадки $G_2(u; \theta)$ (рис. 2) видно, что промежуток, где точность адаптивных оценок выше обычной шире. Самой точной адаптивной оценкой при $G_1(u; \theta)$ является $\hat{F}_2(y)$. Рисунок (рис. 3) и (рис. 4) показывают изменение точности оценки с заданной априорной догадкой при разных вероятностях, здесь самыми точными оценками выступают $\hat{F}_2(y)$

в (рис. 3) и $\hat{F}_1(y)$ в (рис. 4). Снижение точности оценок $\hat{F}_1(y)$, $\hat{F}_2(y)$, $\hat{F}_3(y)$ по сравнению с $\hat{F}_0(y)$ обусловлено наличием погрешности в оценках весового коэффициента λ . Для некоторых значений θ погрешность компенсирует повышение точности от привлечения априорной догадки и делает оценки $\hat{F}_1(y)$, $\hat{F}_2(y)$, $\hat{F}_3(y)$ менее точными по сравнению с эмпирической \hat{F} .

Заключение

В данной работе рассмотрена задача оценивания функции распределения некоторой случайной величины, когда исследователь может предположить принадлежность функции распределения к некому параметрическому классу распределений. Построенные адаптивные оценки исследованы на точность с помощью их асимптотически нормированных СКО, усредненных по стандартному нормальному закону. Это позволило сравнить точность адаптивных оценок $\hat{F}_1(y)$, $\hat{F}_2(y)$, $\hat{F}_3(y)$ с оптимальной $\hat{F}_0(y)$ при заданных априорных догадках.

СПИСОК ЛИТЕРАТУРЫ

1. *Мания Г. М.* Статистическое оценивание распределений вероятностей. Тбилиси: Тбилис. ун-та, 1974.
2. *Дмитриев Ю. Г. Устинов Ю. К.* Статистическое оценивание распределений вероятностей с использованием дополнительной информации, Томск: Томск. ун-та, 1988. 194 с.
3. *Дмитриев Ю. Г. Кошкин Г. М.* Комбинированные оценки линейного функционала с априорной догадкой о параметрическом классе распределений // Вестник Томского государственного университета. Управление, вычислительная техника и информатика, Томск: Томск. ун-та, 2024. № 66. С. 36–43.

Дмитриев Юрий Глебович — доктор физико-математических наук, доцент, профессор кафедры системного анализа и математического моделирования Института прикладной математики и компьютерных наук Национального исследовательского Томского государственного университета (Томск, Россия). E-mail: *dmit@mail.tsu.ru*

Кирюшкин Ярослав Владимирович — студент Института прикладной математики и компьютерных наук Национального исследовательского Томского государственного университета. E-mail: *rran00domm@gmail.com*

КОМПОНЕНТЫ ПОГРЕШНОСТИ ФУНКЦИОНАЛЬНЫХ ВЫЧИСЛИТЕЛЬНЫХ АЛГОРИТМОВ ПРИБЛИЖЕНИЯ ВЕРОЯТНОСТНОЙ ПЛОТНОСТИ ПО ЗАДАННОЙ ВЫБОРКЕ: \mathbb{L}_2 -ПОДХОД

А. В. Войтишек^{1,2}, Н. Х. Шлымбетов²

¹ *Институт вычислительной математики и математической геофизики
СО РАН, г. Новосибирск, Россия*

² *Новосибирский государственный университет, г. Новосибирск, Россия*

В данной работе рассматриваются вопросы построения верхних границ для среднеквадратических погрешностей таких вычислительных конструкций приближения неизвестной плотности по заданной выборке как компьютерные функциональные ядерные и проекционные алгоритмы. **Ключевые слова:** *вычислительный функциональный ядерный алгоритм, вычислительный функциональный проекционный алгоритм, среднеквадратическая погрешность, компонента аппроксимации, компонента смещения, стохастическая компонента, задача условной оптимизации.*

1. Введение: вычислительные функциональные ядерные и проекционные алгоритмы

В наших работах [1–3] отмечено, что при обработке больших данных (например, при применении технологий машинного обучения) актуальным оказывается использование экономичных алгоритмов компьютерного приближения неизвестной плотности $f_{\xi}(\mathbf{x})$ по заданной выборке $\{\xi_1, \dots, \xi_n\}$ с заданным уровнем погрешности $\varepsilon > 0$ на всей компактной области распределения $X \subset \mathbb{R}^d$.

Для такого приближения в [1–3] предлагается использовать одну из следующих вычислительных конструкций. В области X введем сетку $\mathbf{Y}^{(M)} = \{y_1, \dots, y_M\}$.

АЛГОРИТМ 1 (вычислительный функциональный ядерный алгоритм приближения вероятностной плотности [1–3]). *Выберем ядерную функцию $\kappa^{(\mathbf{x})}(\mathbf{z})$ (см., например, [4]) такую, что $f_{\xi}(\mathbf{x}) \approx$*

$\int_X f_{\xi}(\mathbf{z})\varkappa^{(\mathbf{x})}(\mathbf{z}) d\mathbf{z} = \mathbf{E}\varkappa^{(\mathbf{x})}(\xi)$ для всех $\mathbf{x} \in X$. Реализуем на компьютере соответствующие приближения $f_{\xi}(\mathbf{y}_i) \approx \tilde{f}_{\xi}^{(\mathbf{y}_i, ker)}(n) = \frac{1}{n} \sum_{j=1}^n \varkappa^{(\mathbf{y}_i)}(\xi_j)$; $i = 1, \dots, M$ (здесь $\{\xi_1, \dots, \xi_n\}$ – заданная выборка или ее подмножество) и строим окончательную классическую компьютерную аппроксимацию (см., например, главы 2 и 4 книги [5]) плотности $f_{\xi}(\mathbf{x})$:

$$f_{\xi}(\mathbf{x}) \approx L^{(M,n)} \tilde{f}_{\xi}^{(ker)}(\mathbf{x}) = \sum_{i=1}^M w^{(i)} \left[\tilde{f}_{\xi}^{(\mathbf{y}_1, ker)}(n), \dots, \tilde{f}_{\xi}^{(\mathbf{y}_M, ker)}(n) \right] \chi^{(i)}(\mathbf{x}). \quad (1)$$

АЛГОРИТМ 2 (вычислительный функциональный проекционный алгоритм приближения вероятностной плотности [1–3]). Выберем подмножество $\Psi^{(K)} = \{\psi^{(1)}(\mathbf{y}), \dots, \psi^{(K)}(\mathbf{y})\}$ бесконечной ортонормированной системы функций $\Psi^{(\infty)} = \{\psi^{(1)}(\mathbf{y}), \psi^{(2)}(\mathbf{y}), \dots\}$ такой, что $\int_X \psi^{(i)}(\mathbf{y})\psi^{(j)}(\mathbf{y}) d\mathbf{y} = 1$ при $i = j$ и $\int_X \psi^{(i)}(\mathbf{y})\psi^{(j)}(\mathbf{y}) d\mathbf{y} = 0$ при $i \neq j$ и $f_{\xi}(\mathbf{x}) = \sum_{k=1}^{\infty} \left[\int_X f_{\xi}(\mathbf{z})\psi^{(k)}(\mathbf{z}) d\mathbf{z} \right] \psi^{(k)}(\mathbf{x}) = \sum_{k=1}^{\infty} \mathbf{E}\psi^{(k)}(\xi)\psi^{(k)}(\mathbf{x})$ для всех $\mathbf{x} \in X$ (см. § 25 главы 5 книги [6]). Реализуем на компьютере соответствующие приближения $f_{\xi}(\mathbf{y}_i) \approx \tilde{f}_{\xi}^{(\mathbf{y}_i, pr)}(n, K) = \sum_{k=1}^K \left[\frac{1}{n} \sum_{j=1}^n \psi^{(k)}(\xi_j) \right] \psi^{(k)}(\mathbf{y}_i)$; $i = 1, \dots, M$ и строим окончательную аппроксимацию плотности $f_{\xi}(\mathbf{x})$:

$$f_{\xi}(\mathbf{x}) \approx L^{(M,n,K)} \tilde{f}_{\xi}^{(pr)}(\mathbf{x}) = \sum_{i=1}^M w^{(i)} \left[\tilde{f}_{\xi}^{(\mathbf{y}_1, pr)}(n, K), \dots, \tilde{f}_{\xi}^{(\mathbf{y}_M, pr)}(n, K) \right] \chi^{(i)}(\mathbf{x}). \quad (2)$$

В формулах (1), (2) $\chi^{(M)} = \{\chi^{(1)}(\mathbf{x}), \dots, \chi^{(M)}(\mathbf{x})\}$ – заданный набор функций (аппроксимационный базис), определенным образом связанный с введенной в области X сеткой $\mathbf{Y}^{(M)}$, а $\mathbf{W}^{(M)} = \left\{ w^{(i)} \left[f_{\xi}(\mathbf{y}_1), \dots, f_{\xi}(\mathbf{y}_M) \right]; i = 1, \dots, M \right\}$ – набор аппроксимационных коэффициентов.

2. Выбор параметров и функций в алгоритмах 1 и 2

Для обоснованного использования алгоритмов 1 и 2 важной является следующая задача.

ЗАДАЧА 1. Выбрать параметры M, n , вид коэффициентов $\mathbf{W}^{(M)}$, функции $\chi^{(M)}$, $\varkappa^{(\mathbf{x})}(\mathbf{z})$ и/или $\Psi^{(K)}$ таким образом, чтобы обеспечить

выполнение неравенства

$$\delta^{(\mathbb{B})}(M, n) = \left\| f_{\xi} - L^{(M, n)} \tilde{f}_{\xi} \right\|_{\mathbb{B}(X)} \leq \varepsilon, \quad (3)$$

где функция $L^{(M, n)} \tilde{f}_{\xi}(\mathbf{x})$ — это $L^{(M, n)} \tilde{f}_{\xi}^{(ker)}(\mathbf{x})$ из (1) или $L^{(M, n, K)} \tilde{f}_{\xi}^{(pr)}(\mathbf{x})$ из (2), число $\varepsilon > 0$ — заданный уровень погрешности, а $\mathbb{B}(X)$ — подходящее нормированное функциональное пространство.

Отметим, что задача 1 является частной для следующей подробно изученной задачи условной оптимизации (см., например, [3, 7]).

ЗАДАЧА 2. Выбрать параметры M, n , вид коэффициентов $\mathbf{W}^{(M)}$, функции $\chi^{(M)}$, $\varkappa^{(\mathbf{x})}(\mathbf{z})$ и/или $\Psi^{(K)}$ таким образом, чтобы обеспечить выполнение неравенства (3) при наименьших вычислительных затратах $S(M, n)$ функционального алгоритма 1 или 2.

3. \mathbb{C} - и \mathbb{L}_2 -подходы к оценке погрешностей алгоритмов 1 и 2

При решении задачи 1 необходимо выбрать как соответствующее нормированное функциональное пространство $\mathbb{B}(X)$, так и вероятностный смысл выполнения неравенства вида (3) (ведь $\delta^{(\mathbb{B})}(M, n)$ является случайной величиной). Следуя канонам классического численного анализа (см., например, [5]) и теории условной оптимизации (см., например, [3, 7]), рассмотрим пространства $\mathbb{B}(X) = \mathbb{C}(X) \vee \mathbb{L}_2(X)$ и два известных подхода к построению верхней границы $UP^{(\mathbb{B})}(M, n)$ погрешности $\delta^{(\mathbb{B})}(M, n)$:

– « \mathbb{C} -подход», в котором случайная величина

$$\delta^{(\mathbb{C})}(M, n) = \left\| f_{\xi} - L^{(M, n)} \tilde{f}_{\xi} \right\|_{\mathbb{C}(X)} = \sup_{\mathbf{x} \in X} \left| f_{\xi}(\mathbf{x}) - L^{(M, n)} \tilde{f}_{\xi}(\mathbf{x}) \right|$$

ограничивается сверху по вероятности: $\mathbf{P} [\delta^{(\mathbb{C})}(M, n) \leq UP^{(\mathbb{C})}(M, n)] > 1 - p$ для некоторого достаточно малого $p > 0$, и

– « \mathbb{L}_2 -подход», в котором выбирается сходимость в среднем погрешности

$$\delta^{(\mathbb{L}_2)}(M, n) = \left\| f_{\xi} - L^{(M, n)} \tilde{f}_{\xi} \right\|_{\mathbb{L}_2(X)} = \left(\int_X \left[f_{\xi}(\mathbf{x}) - L^{(M, n)} \tilde{f}_{\xi}(\mathbf{x}) \right]^2 d\mathbf{x} \right)^{1/2}$$

к нулю при $M, n \rightarrow \infty$ и строятся оценки сверху $UP^{(\mathbb{L}_2)}(M, n)$ такие, что

$$\left[\mathbf{E} \delta^{(\mathbb{L}_2)}(M, n) \right]^2 \leq UP^{(\mathbb{L}_2)}(M, n). \quad (4)$$

В наших работах [2, 3] для формулировки соображений по выбору аппроксимационного базиса и $\chi^{(M)}$ и коэффициентов $\mathbf{W}^{(M)}$ использован \mathbb{C} -подход: здесь формулы для определения некоторых новых понятий – таких, например, как константа устойчивости $K^{(stab)}$ [2, 3], – являются более компактными. Отметим, что, с учетом очевидного неравенства $\delta^{(\mathbb{L}_2)}(M, n) \leq \sqrt{\text{mes } X} \times \delta^{(\mathbb{C})}(M, n)$, формулировка аналогов соображений из работы [2, 3] для более естественного для приложений \mathbb{L}_2 -подхода является относительно несложной задачей.

4. \mathbb{L}_2 -подход: компоненты погрешности

Следуя работе [3], запишем верхнюю границу погрешности $\delta^{(\mathbb{L}_2)}(M, n)$ из соотношения (4) в виде суммы трех компонент: **компоненты аппроксимации** $\delta_{appr}^{(\mathbb{L}_2)}(M)$, **компоненты смещения** $\delta_{bias}^{(\mathbb{L}_2)}(M)$ и **стохастической компоненты** $\delta_{stoch}^{(\mathbb{L}_2)}(M, n)$.

Для это применим сначала неравенство Коши – Буняковского и теорему Фубини и получим неравенство

$$\left[\mathbf{E} \delta^{(\mathbb{L}_2)}(M, n) \right]^2 \leq \int_X \mathbf{E} \left[f_{\xi}(\mathbf{x}) - L^{(M,n)} \tilde{f}_{\xi}(\mathbf{x}) \right]^2 d\mathbf{x}.$$

Далее заметим, что

$$\begin{aligned} \mathbf{E} \left[f_{\xi}(\mathbf{x}) - L^{(M,n)} \tilde{f}_{\xi}(\mathbf{x}) \right]^2 &= \mathbf{E} \left(\left[f_{\xi}(\mathbf{x}) - L^{(M)} f_{\xi}(\mathbf{x}) \right] + \right. \\ &+ \left. \left[L^{(M)} f_{\xi}(\mathbf{x}) - L^{(M)} \bar{f}_{\xi}(\mathbf{x}) \right] + \left[L^{(M)} \bar{f}_{\xi}(\mathbf{x}) - L^{(M,n)} \tilde{f}_{\xi}(\mathbf{x}) \right] \right)^2 = \\ &= \left(\left[f_{\xi}(\mathbf{x}) - L^{(M)} f_{\xi}(\mathbf{x}) \right] + \left[L^{(M)} f_{\xi}(\mathbf{x}) - L^{(M)} \bar{f}_{\xi}(\mathbf{x}) \right] \right)^2 + \\ &+ \mathbf{E} \left[L^{(M)} \bar{f}_{\xi}(\mathbf{x}) - L^{(M,n)} \tilde{f}_{\xi}(\mathbf{x}) \right]^2; \end{aligned}$$

здесь учтено, что $\mathbf{E} L^{(M,n)} \tilde{f}_{\xi}(\mathbf{x}) = L^{(M)} \bar{f}_{\xi}(\mathbf{x})$, где $L^{(M)} f_{\xi}(\mathbf{x}) = \sum_{i=1}^M w^{(i)} \left[f_{\xi}(\mathbf{y}_1), \dots, f_{\xi}(\mathbf{y}_M) \right] \chi^{(i)}(\mathbf{x})$ – аппроксимация функции $f_{\xi}(\mathbf{x})$ по точным значениям в узлах сетки $\mathbf{Y}^{(M)}$, $L^{(M)} \bar{f}_{\xi}(\mathbf{x}) = L^{(M)} \bar{f}_{\xi}^{(ker)}(\mathbf{x}) = \sum_{i=1}^M w^{(i)} \left[\mathbf{E} \chi^{(\mathbf{y}_1)}(\xi), \dots, \mathbf{E} \chi^{(\mathbf{y}_M)}(\xi) \right] \chi^{(i)}(\mathbf{x})$ для алгоритма 1 и $L^{(M)} \bar{f}_{\xi}(\mathbf{x}) = L^{(M,K)} \bar{f}_{\xi}^{(pr)}(\mathbf{x}) = \sum_{i=1}^M w^{(i)} \left[\sum_{k=1}^K \mathbf{E} \psi^{(k)}(\xi) \psi^{(k)}(\mathbf{y}_1), \dots, \sum_{k=1}^K \mathbf{E} \psi^{(k)}(\xi) \times \psi^{(k)}(\mathbf{y}_M) \right] \chi^{(i)}(\mathbf{x})$ для алгоритма 2.

Используя очевидное неравенство $(a + b)^2 \leq 2a^2 + 2b^2$, получаем

$$\left[\mathbf{E} \delta^{(\mathbb{L}_2)}(M, n) \right]^2 \leq 2 \left[\delta_{appr}^{(\mathbb{L}_2)}(M) \right]^2 + 2 \left[\delta_{bias}^{(\mathbb{L}_2)}(M) \right]^2 + \delta_{stoch}^{(\mathbb{L}_2)}(M, n); \quad (5)$$

$$\text{где } \delta_{appr}^{(\mathbb{L}_2)}(M) = \left\| f\xi - L^{(M)} f\xi \right\|_{\mathbb{L}_2(X)}, \delta_{bias}^{(\mathbb{L}_2)}(M) = \left\| L^{(M)} f\xi - L^{(M)} \bar{f}\xi \right\|_{\mathbb{L}_2(X)}, \\ \delta_{stoch}^{(\mathbb{L}_2)}(M, n) = \int_X \mathbf{D}L^{(M,n)} \tilde{f}\xi(\mathbf{x}) d\mathbf{x}.$$

5. Методы решения задач 1 и 2

Для решения задачи 1 можно предложить следующий метод.

МЕТОД 1. Выбираем параметры M, n , вид коэффициентов $\mathbf{W}^{(M)}$, функции $\chi^{(M)}$, $\varkappa^{(\mathbf{x})}(\mathbf{z})$ и/или $\Psi^{(K)}$ таким образом, чтобы для верхних границ для слагаемых из правой части неравенства (5) выполнялись соотношения: $2 \left[\delta_{appr}^{(\mathbb{L}_2)}(M) \right]^2 \leq UP_{appr}^{(\mathbb{L}_2)}(M) = \frac{\varepsilon^2}{3}$, $2 \left[\delta_{bias}^{(\mathbb{L}_2)}(M) \right]^2 \leq UP_{bias}^{(\mathbb{L}_2)}(M) = \frac{\varepsilon^2}{3}$, $\delta_{stoch}^{(\mathbb{L}_2)}(M, n) \leq UP_{stoch}^{(\mathbb{L}_2)}(M, n) = \frac{\varepsilon^2}{3}$.

Отметим, что в ряде прикладных задач такой выбор параметров и функций дает и приближенное решение задачи 2. Но более точным является следующий способ решения задачи 2 (см., например, [7]).

МЕТОД 2. Для выбранных коэффициентов $\mathbf{W}^{(M)}$, функций $\chi^{(M)}$, $\varkappa^{(\mathbf{x})}(\mathbf{z})$ и/или $\Psi^{(K)}$ строятся верхние границы для слагаемых из правой части неравенства (5), сумма которых приравняется величине ε^2 :

$$UP_{appr}^{(\mathbb{L}_2)}(M) + UP_{bias}^{(\mathbb{L}_2)}(M) + UP_{stoch}^{(\mathbb{L}_2)}(M, n) = \varepsilon^2.$$

Из полученного равенства один из параметров (например, n) выражается через другой: $n = F^{(\varepsilon)}(M)$. Это соотношение подставляется в выражение для затрат $S(M, n)$ (которое тоже зависит от параметров M и n). В результате получается функция $\tilde{S}^{(\mathbb{B}, \varepsilon)}(M)$ одного переменного M , которая исследуется на минимум с помощью известных приемов математического или численного анализа. Найденные значения $M_{min}^{(\mathbb{L}_2)}(\varepsilon) = M_{opt}^{(\mathbb{B})}(\varepsilon), n_{opt}^{(\mathbb{B})}(\varepsilon) = F^{(\varepsilon)} \left[M_{opt}^{(\mathbb{L}_2)}(\varepsilon) \right]$ объявляются **условно-оптимальными параметрами** соответствующего функционального алгоритма.

Полная реализация метода 2 удалась в работе [7] только для многомерного аналога полигона частот (это одновременно частный случай и для алгоритма 1, и для алгоритма 2 [1, 3]).

Заключение

В работе представлены два подхода (методы 1 и 2) к выбору параметров и функций компьютерных функциональных ядерных и проекционных алгоритмов приближения вероятностной плотности по заданной выборке (алгоритмы 1 и 2), основанные на покомпонентном оценивании сверху среднеквадратической погрешности этих вычислительных схем.

Возможности эффективной реализации методов 1 и 2 для алгоритмов 1 и 2 требуют отдельных подробных исследований.

СПИСОК ЛИТЕРАТУРЫ

1. *Войтишек А. В., Шлымбетов Н. Х.* Экономичные алгоритмы приближения вероятностной плотности по заданной выборке // Информационные технологии и математической моделирование (ИТММ–2022): Материалы XXI Международной конференции имени А. Ф. Терпугова (25–29 октября 2022 года). Томск: Издательство Томского государственного университета, 2023. С. 319–324.
2. *Войтишек А. В., Шлымбетов Н. Х.* Требования к аппроксимационным базисам, используемым в вычислительных алгоритмах приближения вероятностной плотности по заданной выборке // Информационные технологии и математической моделирование (ИТММ–2023): Материалы XXII Международной конференции имени А. Ф. Терпугова (4–9 декабря 2023 года). Томск: Издательство Томского государственного университета, 2023. Часть 1. С. 345–350.
3. *Войтишек А. В., Шлымбетов Н. Х.* Выбор аппроксимационных базисов, используемых в компьютерных функциональных алгоритмах приближения вероятностных плотностей по заданной выборке // Сибирский журнал вычислительной математики. 2024. Т. 27, № 2. С. 147–164.
4. *Епанечников В. А.* Непараметрическая оценка многомерной плотности вероятности // Теория вероятностей и ее применения. 1969. Т. 14. Вып. 1. С. 156–161.
5. *Бахвалов Н. С.* Численные методы. М.: Наука, 1975. 632 с.
6. *Ченцов Н. Н.* Статистические решающие правила и оптимальные выводы. М.: Наука, 1972. 520 с.
7. *Булгакова Т. Е.* Оптимизация функциональных вычислительных статистических оценок и алгоритмов: Дис. ... канд. физ.-мат. наук / ИВМиМГ СО РАН. Новосибирск, 2020. 170 с.

Войтишек Антон Вацлавович — доктор физико-математических наук, профессор, главный научный сотрудник, лаборатория стохастических задач ИВМиМГ СО РАН; профессор, кафедра вычислительной математики НГУ. E-mail: vav@osmf.sbcc.ru

Шлымбетов Нурлыбай Хамдуллаевич — аспирант механико-математического факультета НГУ. E-mail: n.shlymbetov@g.nsu.ru

АНАЛИЗ ЭФФЕКТИВНОСТИ КОМБИНИРОВАННОЙ ОЦЕНКИ РЕГРЕССИИ

Ю. Г. Дмитриев, Р. Д. Сорокин

*Национальный исследовательский Томский государственный университет,
г. Томск, Россия*

Рассматривается класс оценок функции регрессии, представляющий собой линейные комбинации двух оценок: исходной оценки функции регрессии и некоторой заданной функции, называемой априорной догадкой. Оптимальная комбинированная оценка определяется через весовой коэффициент, выбираемый из критерия минимума среднеквадратической ошибки (СКО). Анализируется эффективность, определяемая как отношение СКО оптимальной оценки к СКО исходной оценки регрессии, а также поведение оптимального весового коэффициента в зависимости от трех параметров: отклонения априорной догадки от истинной функции регрессии, смещения и дисперсии исходной оценки. Приводятся аналитические выражения и графики, позволяющие судить об эффективности оптимальной оценки и поведении оптимального весового коэффициента. **Ключевые слова:** *функция регрессии, априорная догадка, оценки функции регрессии, среднеквадратическая ошибка, эффективность оценки, комбинированные оценки.*

Введение

Статистическому оцениванию функции регрессии посвящено много работ. В зависимости от располагаемой исходной информации рассматриваются разные модели и методы оценивания этой функции. Работы [1–6] посвящены непараметрическим методам оценивания, в [3,5] предложены параметрические методы, в [5] рассмотрены полупараметрические модели, в [3] предложен метод локальной аппроксимации (МЛА).

Построению оценок, использующих в своей структуре дополнительную (априорную) информацию, посвящены работы [4,7,8]. Проблема повышения точности непараметрических оценок за счет привлечения дополнительной информации рассматривалась в [4]. Выделим работы, в которых дополнительная информация выступает в виде априорной догадки [7,8]. Предполагается, что исследователь на основании своего опыта и знаний может задать конкретную функцию для оценивания

регрессии (возможно с точностью до ряда параметров), однако полной уверенности в правильности задания нет. Предложенные оценки являются взвешенной суммой исходной (задаваемой исследователем) функции регрессии и априорной догадки, и называются комбинированными.

Цель работы заключается в проведении анализа эффективности комбинированной оценки регрессии и оптимального весового коэффициента в зависимости от трех параметров: отклонения априорной догадки от истинной регрессии, смещения и дисперсии исходной оценки.

1. Комбинированные оценки регрессии и эффективность

Функция регрессии определяется как условное математическое ожидание

$$r(x) = M(Y|X = x) = \int_{-\infty}^{+\infty} yf(x, y) dy/g(x), x \in \mathbb{R}^m, \quad (1)$$

где $f(x, y)$ совместная функция плотности случайных величин (X, Y) , $g(x) = \int_{-\infty}^{+\infty} f(x, y) dy$ — функция плотности вектора $X = (X_1, \dots, X_m)$.

Рассмотрим класс комбинированных оценок функции регрессии следующего вида

$$\hat{R}(x; \lambda) = (1 - \lambda) \cdot \hat{r}(x) + \lambda \cdot \varphi(x), \quad (2)$$

где $\hat{r}(x)$ — исходная оценка функции регрессии, построенная по статистическим данным $\{X_i, Y_i\}_{i=1}^n$ объема n , $\varphi(x)$ — априорной догадка, λ — весовой коэффициент, выбираемый из заданного критерия качества оценивания. Существуют разные критерии качества для оценок функции регрессии [5]. В данной работе критерием качества выступает СКО

$$S^2(x; \lambda) = M \left\{ \hat{R}(x; \lambda) - r(x) \right\}^2,$$

$$S^2(x; \lambda) = M \left\{ (1 - \lambda) \cdot \hat{r}(x) + \lambda \cdot \varphi(x) - r(x) \right\}^2. \quad (3)$$

Оптимальное λ , доставляющее минимум СКО, имеет вид

$$\lambda = \lambda_0(x) = \frac{M \{(\hat{r}(x) - r(x)) \cdot (\hat{r}(x) - \varphi(x))\}}{M \{\hat{r}(x) - \varphi(x)\}^2}, \quad (4)$$

при котором

$$S^2(x; \lambda_0) = M \{\hat{r}(x) - r(x)\}^2 - \frac{[M \{(\hat{r}(x) - r(x)) \cdot (\hat{r}(x) - \varphi(x))\}]^2}{M \{\hat{r}(x) - \varphi(x)\}^2}. \quad (5)$$

Формула (5) показывает, что СКО комбинированной оценки при λ_0 не превосходит СКО $S^2(x; 0) = M \{ \hat{r}(x) - r(x) \}^2$ исходной оценки регрессии. Оценку $\hat{R}(x; \lambda_0)$ назовем оптимальной в классе (2) в смысле минимума СКО. Преобразуем (4) и (5) к виду

$$\lambda_0(x) = 1 - \frac{\Delta(x) \cdot (b(x) + \Delta(x))}{D(x) + (b(x) + \Delta(x))^2}, \quad (6)$$

$$S^2(x; \lambda_0) = D\hat{r}(x) + b^2(x) - \frac{(D\hat{r}(x) + b(x) \cdot (b(x) + \Delta(x)))^2}{D\hat{r}(x) + (b(x) + \Delta(x))^2}, \quad (7)$$

где $b(x) = M\hat{r}(x) - r(x)$ — смещение исходной оценки в точке x , $D\hat{r}(x) = M(\hat{r}(x) - Mr(x))^2$ — её дисперсия, $\Delta(x) = r(x) - \varphi(x)$ — отклонение априорной догадки от истинного значения регрессии. Отметим, что смещение оптимальной комбинированной оценки равно $M\hat{R}(x; \lambda_0) - r(x) = b(x) + \Delta(x) = \bar{\Delta}(x)$. В дальнейшем переменную x в формулах будем опускать, $D = D\hat{r}(x)$.

Для сравнения точности оценивания оптимальной комбинированной оценки $\hat{R}(x; \lambda_0)$ с исходной $\hat{r}(x)$ в точке x рассмотрим отношение, которое назовём эффективностью и обозначим через v :

$$\begin{aligned} v = v(x) &= 1 - \frac{M \{ \hat{R}(x; \lambda_0) - r(x) \}^2}{M \{ \hat{r}(x) - r(x) \}^2} = 1 - \frac{S^2(x; \lambda_0)}{S^2(x; 0)} \\ &= 1 - \frac{(D + b \cdot (b + \Delta))^2}{(D + (b + \Delta)^2) \cdot (D + b^2)} = 1 - \frac{(D + b \cdot \bar{\Delta})^2}{(D + \bar{\Delta}^2) \cdot (D + b^2)}. \end{aligned} \quad (8)$$

Формула (8) показывает, что малые значения v соответствуют высокой эффективности оценки $\hat{R}(x; \lambda_0)$ по сравнению с исходной $\hat{r}(x)$. Представляет интерес поведение эффективности $v(x)$ от входящих в нее параметров b, D, Δ .

2. Анализ эффективности

Выявим влияние параметров b, D, Δ на эффективность v и оптимальный весовой коэффициент λ_0 . Рассмотрим различные случаи взаимоотношений этих параметров в (6) и (8).

2.1 Предположим, что $b = 0$ и $D > 0$. Обозначим $\nu = \sqrt{\Delta^2/D}$, тогда

$$\lambda_0 = 1 - \frac{\Delta^2}{D + \Delta^2} = 1 - \frac{\nu^2}{1 + \nu^2} = \frac{1}{1 + \nu^2}, \quad (9)$$

а эффективность

$$v = 1 - \frac{D}{D + \Delta^2} = 1 - \frac{1}{1 + \nu^2} = 1 - \lambda_0. \quad (10)$$

Из (9) следует, что $0 \leq \lambda_0 \leq 1$ для несмещенных оценок $\hat{r}(x)$. При $|\Delta| \rightarrow 0$ ($\nu \rightarrow 0$) коэффициент $\lambda_0 \rightarrow 1$, $v \rightarrow 0$ (см.Рис.1). Влияние априорной догадки на точность оценивания усиливается и оценка $\hat{R}(x; \lambda_0)$ переходит в априорную догадку $\varphi(x)$.

Если $|\Delta| \rightarrow \infty$, то $\lambda_0 \rightarrow 0$, $v \rightarrow 1$ (см.Рис.1). Влияние априорной догадки на точность оценивания ослабевает и $\hat{R}(x; \lambda_0)$ стремится к исходной оценке $\hat{r}(x)$.

Пусть Δ фиксировано. При $D \rightarrow \infty$ коэффициент $\lambda_0 \rightarrow 1$, $v \rightarrow 0$ и предпочтение отдается априорной догадке. Если $D \rightarrow 0$, то $\lambda_0 \rightarrow 0$, $v \rightarrow 1$, а комбинированная оценка переходит к исходной $\hat{r}(x)$.

2.2 Предположим, что в (6) и (8) $\Delta = 0$. Тогда $\lambda_0 = 1$, $v = 0$ (см.Рис.1,2,4), а комбинированная оценка $\hat{R}(x; 1) = \varphi(x)$.

2.3 Обратимся к формуле (6). При $b \neq 0$ возможны случаи, когда $\lambda_0 < 0$, $\lambda_0 > 1$ (см.Рис.2,3). Действительно, из (6) следует, если $\Delta \cdot b > D$, то $\lambda_0 < 0$ и эффективность резко уменьшается $v \rightarrow 1$. Коэффициент $\lambda_0 > 1$, когда b не принадлежит интервалу $(-\Delta, \Delta)$, $\Delta > 0$ и эффективность резко увеличивается $v \rightarrow 0$.

2.4 Рассмотрим случай, когда $b + \Delta = 0$, т.е. смещение комбинированной оценки $\bar{\Delta} = 0$ (см.Рис.4). Тогда из (6) и (8) следует $\lambda_0 = 1$, однако отсюда ещё не следует, что $\Delta = 0$. Эффективность определяется формулой $v = 1 - D/(D + \Delta^2)$.

3. Графическое представление эффективности

Наглядным представлением эффективности является графическое изображение v в зависимости от параметров b, D, Δ .

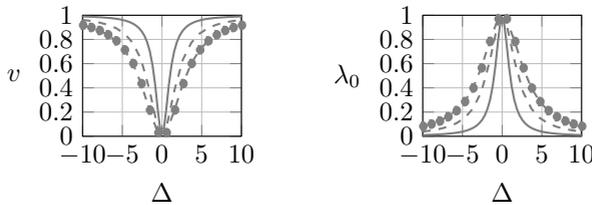


Рис. 1. v и λ_0 от Δ , при $b = 0$ и $D = 1 : -$, $D = 4 : - -$, $D = 9 : - \cdot -$

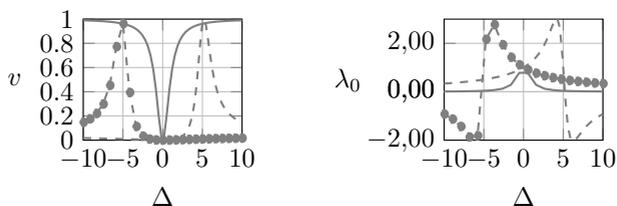


Рис. 2. v и λ_0 от Δ , при $D = 1$ и $b = -5$: - -, $b = 0$: -, $b = 5$: - . -

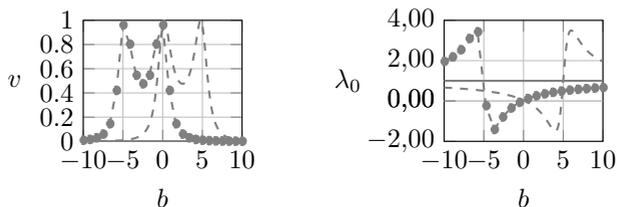


Рис. 3. v и λ_0 от b , при $D = 1$ и $\Delta = -5$: - -, $\Delta = 0$: -, $\Delta = 5$: - . -

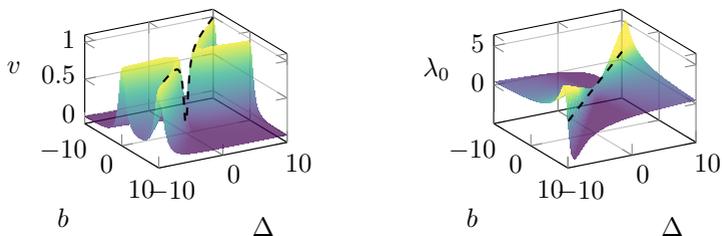


Рис. 4. v и λ_0 от b и Δ , при $\Delta + b = 0$, $v = [0, \dots, 1]$: - -, $\lambda_0 = 1$: - . -

Заключение

Рассмотренный класс оценок неизвестной функции регрессии с оптимальным весовым коэффициентом приводит к уменьшению СКО за счет использования априорной догадки при неизменных исходных данных. Как показал анализ эффективности оценивания, уменьшение СКО может быть довольно существенным в зависимости от соотношений трех параметров: смещения, дисперсии исходной оценки и отклонения априорной догадки от истинной регрессии. Некоторые из этих соотношений могут быть известны априори, что дает возможность практического применения оценок. В других случаях необходимо строить адаптивные оценки путем оценивания оптимального весового коэффициента.

СПИСОК ЛИТЕРАТУРЫ

1. *Надарая Э.А.* Об оценке регрессии // Теория вероятностей и ее применения. 1964. Т. 19. вып. 1. С. 147–149.
2. *Watson G.S.* Smooth regression analysis // Sankhya. Indian J. Statist. 1964. Vol. A26. P. 359–372.
3. *Катковник В.Я.* Непараметрическая идентификация и сглаживание данных: метод локальной аппроксимации. – М.: Главная редакция физико-математической литературы, 1985. 336 с.
4. *Дмитриев Ю.Г., Кошкин Г.М.* Использование дополнительной информации при непараметрическом оценивании функционалов плотности // Автоматика и телемеханика. 1987. № 10. С. 47–59.
5. *Хардле В.* Прикладная непараметрическая регрессия: Пер. с англ. – М., Мир, 1993. 349 с.
6. *Васильев В.А., Добровидов А.В., Кошкин Г.М.* Непараметрическое оценивание функционалов от распределений стационарных последовательностей. – М.: Наука, 2004. – 508 с.
7. *Скрипкин С.В.* Свойства комбинированной оценки регрессии при конечных объемах выборок, Известия Томского политехнического университета, 2008, Т. 313, № 5, С. 10–14.
8. *Yu. G. Dmitriev, Gennady M. Koshkin and Vadim Yu. Lukov.*: Combined Identification Algorithms. Applied Methods of Statistical Analysis. Nonparametric Methods in Cybernetics and System Analysis - AMSA'2017, Krasnoyarsk, Russia, 18-22 September, 2017: Proceedings of the International Workshop. – Novosibirsk: NSTU publisher, 2017. - P. 19–27.

Дмитриев Юрий Глебович — доктор физико-математических наук, доцент, заведующий кафедрой системного анализа и математического моделирования. E-mail: dmit@mail.tsu.ru

Сорокин Роман Дмитриевич — аспирант кафедры системного анализа и математического моделирования института прикладной математики и компьютерных наук НИ ТГУ. E-mail: sorokinroman063@gmail.com

КРАЕВАЯ ЗАДАЧА ДЛЯ УРАВНЕНИЯ ТРЕТЬЕГО ПОРЯДКА СМЕШАННО-СОСТАВНОГО ТИПА В ЧЕТЫРЕХУГОЛЬНОЙ ОБЛАСТИ

М. О. Абулов¹, З. У. Узаков²

¹ *Университет Информационных технологий и менеджмента, Республика
Узбекистан, г Карши, Узбекистан*

² *Каршинский филиал Ташкентского университета информационных
технологий имени Мухаммада аль-Хоразми, г Карши, Узбекистан*

Рассматривается нелокальная краевая задача для уравнения третьего порядка смешанно-составного типа в четырехугольной области. С помощью метода Галеркина при определенных условиях на коэффициенты и правую часть уравнения доказано существование слабо обобщенного решения в пространстве Соболева. При тех же условиях доказана единственность обобщенного решения. **Ключевые слова:** *краевая задача, уравнение смешанно-составного типа, обобщенное решение, уравнение третьего порядка*

Введение

Краевые задачи для неклассических уравнений в частных производных возникают во многих прикладных задачах, в частности, в различных разделах механики, физики, при описании процессов рассеивания и переноса, в геометрии и популяционной генетике, гидродинамике, а также многих других областях. Интерес к нелокальной задаче обусловлен возможностью ее физической интерпретации. Если физический процесс можно описать с помощью дифференциального уравнения, то нелокальные краевые условия являются некоторыми алгебраическими выражениями, связывающими искомое решение и его производные в двух и более точках наблюдения физического процесса. Поэтому изучение нелокальных задач для различных классов дифференциальных уравнений привлекало внимание многих математиков. Библиографию вопроса можно найти, например, в монографиях [7], [10], [13], [9]. Постановке краевых задач для уравнений смешанно-составного типа второго и высокого порядков, исследованию обобщенной разрешимости их в различных пространствах посвящены работы М. С. Салахитдинова [13], [14], Т. Д. Джураева [4], У. О. Рахманова, Л. А. Бобылова, М.

М. Смирнова [5] [6], С. Г. Пяткова [3] [7] и других авторов. В настоящей работе авторы продолжают цикл своих работ [1], [2] по постановке и исследованию краевых задачи для уравнения смешанно-составного и составного типа.

1. Постановка задачи

Зададим область $Q = \{(x, t) : -1 \leq x \leq 1, 0 \leq t \leq T\}$ и рассмотрим уравнение

$$Lu \equiv k(t)u_{ttt} + \mu(x)u_{xxx} + a(x, t)u_{xx} + b(x, t)u_{tt} = f(x, t), \quad (1)$$

где $(t - \frac{T}{2})k(t) > 0$ при $t \neq \frac{T}{2}$, $k(\frac{T}{2}) = 0$, а также $x\mu(x) > 0$ при $x \neq 0$, $\mu(0) = 0$.

Краевая задача. Ставится задача нахождения в области Q решения уравнения (1), удовлетворяющее краевому условию

$$u|_{\partial Q} = 0. \quad (2)$$

Данная постановка отличается от ранее рассмотренной работы [2], в которой была исследована краевая задача:

$$Lu \equiv \frac{\partial}{\partial x} (tu_{xx} + u_{yy} + u_{tt} + au_x + bu) + cu_y + du_t + ju = f(x, y, t),$$

$$u|_{\partial D} = 0, \quad u_x|_{x=0} = 0 \quad (\text{при } t > 0)$$

Будем считать, что коэффициенты уравнения (1) - бесконечно дифференцируемые функции.

2. Решение задачи

Для решения задачи (1), (2) воспользуемся методом Галеркина, чтобы получить априорные оценки и с помощью этих оценок доказать, что решение данной задачи существует и единственно в пространстве $W_2^2(Q)$.

Определение 1. Обозначим через $W_2^2(Q)$ пространство функций, полученное замыканием функций из $C^\infty(Q)$, удовлетворяющих условию (2) по норме [1]

$$\|u\|_{W_2^2(Q)} = \int_Q (u_{tt}^2 + u_{xx}^2 + u_{xt}^2 + u_x^2 + u_t^2 + u^2) dQ.$$

Определение 2. Функцию $u \in W_2^2(Q)$ будем называть слабым обобщенным решением задачи (1) - (2), если для всех $v \in C_0^\infty(Q)$ выполняется тождество

$$\int_Q (-ku_{tt}v_t - k_t u_{tt}v - \mu u_{xx}v_x - \mu_x u_{xx}v + au_{xx}v + bu_{tt}v) dQ = \int_Q fvdQ. \quad (3)$$

Для поставленной задачи сформулируем следующее утверждение.

Теорема. Пусть выполнены условия

$$a(x, t) - \frac{3}{2} |\mu_x| \geq \delta > 0, \quad b(x, t) - \frac{3}{2} |b_t| \geq \delta_1 > 0, \quad (4)$$

тогда для любой функции $f(x, t)$ такой, что $f \in L_2(Q)$, существует единственное решение задачи (1), (2) из $W_2^2(Q)$.

Доказательство. Решение задачи (1), (2) будем искать методом Галеркина

$$u_m(x, t) = \sum_{i=1}^m j_i(t) \varphi_i(x)$$

где базисные функции $\varphi_i(x)$ являются решениями задачи

$$\varphi_i'' = -\lambda_i \varphi_i, \quad \varphi_i(-1) = \varphi_i(1) = 0.$$

а коэффициенты $j_i(t)$ находятся как решения следующих краевых задач для системы обыкновенных дифференциальных уравнений

$$(ku_{mttt}, \varphi_i)_0 + (\mu u_{mxxx}, \varphi_i)_0 + (au_{mxx}, \varphi_i)_0 + (bu_{mtt}, \varphi_i)_0 = (f, \varphi_i)_0, \quad (5)$$

$$j_i(0) = j_i(T) = 0, \quad i = 1, 2, \dots, m, \quad (6)$$

где, например, $(f, \varphi_i)_0 = \int_{-1}^1 f(x) \varphi_i(x) dx$, т.е. скалярное произведение в $L_2(-1, 1)$.

Разрешимость задачи (5) - (6) при фиксированном m вытекает из общей теории обыкновенных дифференциальных уравнений.

В силу граничных условий (2), нетрудно видеть, что для решения верна оценка

$$\int_Q u_m^2 dQ \leq C \int_Q u_{mx}^2 dQ. \quad (7)$$

Выведем равномерные по m оценки для Галеркинских приближений. Для этого умножим (5) на $-j_i(t)$ и, суммируя по i , получим

$$\begin{aligned} & (ku_{mttt}, -u_m)_0 + (\mu u_{mxxx}, -u_m)_0 + \\ & + (au_{mxx}, -u_m)_0 + (bu_{mtt}, -u_m)_0 = (f, -u_m)_0. \end{aligned} \quad (8)$$

Отсюда, интегрируя по t , методом интегрирования по частям, в силу (2), (4), после некоторых преобразований приходим к неравенству

$$\int_Q (u_{mt}^2 + u_{mx}^2 + u_m^2) dQ \leq C. \quad (9)$$

Далее, рассмотрим следующее равенство

$$\begin{aligned} & (u_{mttt}, u_{mxx})_0 + (\mu u_{mxxx}, u_{mxx})_0 + \\ & + (au_{mxx}, u_{mxx})_0 + (bu_{mtt}, u_{mxx})_0 = (f, u_{mxx})_0. \end{aligned} \quad (10)$$

Из тождества (10), в силу (2), (4) и оценки (9), интегрируя по t , и интегрированием по частям, после несложных преобразований вытекает следующая оценка

$$\int_Q (u_{mxx}^2 + u_{mxt}^2) dQ \leq C. \quad (11)$$

Теперь, рассмотрим следующее равенство

$$\begin{aligned} & (u_{mttt}, u_{mtt})_0 + (\mu u_{mxxx}, u_{mtt})_0 + \\ & + (au_{mxx}, u_{mtt})_0 + (bu_{mtt}, u_{mtt})_0 = (f, u_{mtt})_0. \end{aligned} \quad (12)$$

Из тождества (12), в силу (2), (4) и оценок (9), (11), интегрируя по t , и интегрированием по частям, после несложных преобразований вытекает следующая оценка

$$\int_Q (u_{mtt}^2 + u_{mxt}^2) dQ \leq C. \quad (13)$$

Из оценок (9), (11), (13) следует ограниченность последовательности приближенных решений $\{u_m(x, t)\}$ в пространстве $W_2^2(Q)$, можно выбрать подпоследовательность $\{u_{m_k}(x, t)\}$ и перейти к пределу при $m_k \rightarrow \infty$ в системе (5). Нетрудно проверить, что предельная функция принадлежит пространству $W_2^2(Q)$ и удовлетворяет тождеству (3). Поскольку система $\{\varphi_i(x)\}$ плотна в $L_2(-1, 1)$.

Теперь докажем, что решение задачи (1), (2) единственно.

Если u, v — два решения задачи (1), (2), то $w = u - v$ удовлетворяет уравнению

$$kw_{ttt} + \mu(x)w_{xxx} + a(x, t)w_{xx} + b(x, t)w_{tt} = 0.$$

Рассмотрим интеграл

$$\int_Q (kw_{ttt} + \mu(x)w_{xxx} + a(x, t)w_{xx} + b(x, t)w_{tt})w dQ = 0$$

и интегрируя по частям, в силу (2) получим

$$\int_Q (w_t^2 + w_x^2 + w^2) dQ \leq 0.$$

Откуда следует, что $w = 0$ в Q .

Теорема доказана.

3. Заключение

Данная статья посвящена исследованию разрешимости краевой задачи для уравнения третьего порядка смешанно-составного типа в четырехугольной области. В работе доказаны новые теоремы существования и единственности решения краевой задачи (1), (2) в пространстве $W_2^2(Q)$, что позволяет расширить круг решаемых задач теории неклассических уравнений математической физики.

СПИСОК ЛИТЕРАТУРЫ

1. *Абулов М. О.* Краевая задача для смешанно-составного типа третьего порядка // Деп. в ВИНТИ. 1989. № 1436 - В89. С. 5.
2. *Абулов М. О.* Нелокальная задача для одного типа третьего порядка // Математические заметки СВФУ. 2023. № 30.
3. *Бицадзе А. В., Салахитдинов М. С.* К теории уравнений смешанно-составного типа // Сиб. мат. Журнал. 1961. Т. II. № 1. С. 7–19.
4. *Бицадзе А. В.* Об уравнениях смешанно-составного типа // Некоторые проблемы математики и механики (к шестидесятилетию акад. М.А. Лаврентьева), Новосибирск: СО АН СССР, 1961. С. 47–49.
5. *Бобылова Л. А., Смирнов М. М.* Об одной краевой задаче для уравнения смешанно-составного типа 4-го порядка // Известия выс. уч. заведений. Математика. 1972. № 5(120). С. 15–21.
6. *Врагов В. Н.* Об одном уравнении смешанно-составного типа // Дифференциальные уравнения. 1973. Т. IX. № 1. С. 169–171.

7. *Врагов В. Н.* Краевые задачи для неклассических уравнений математической физики. Новосибирск: НГУ, 1983. 84 с.
8. *Джураев Т. Д., Рахманов У.* О корректных краевых задачах для уравнений смешанно-составного типа // Дифференциальные уравнения. 1973. Т. IX. № 1. С. 32–40.
9. *Егоров И. Е., Федоров В. Е.* Неклассические уравнения математической физики высокого порядка. Новосибирск, 1995. 134 с.
10. *Лионс Ж. Л.* Некоторые методы решения нелинейных краевых задач. Москва: Мир. 1972, 408 с.
11. *Лионс Ж. Л., Маджсенес Э.* Неоднородные граничные задачи и их приложения. Москва: Мир. 1971, 372 с.
12. *Пятков С. Г.* Об одном уравнении составного типа. // Дифференциальные уравнения. 1980. Т. 16. № 1. С. 117-123.
13. *Салахитдинов М. С.* Уравнения смешанно-составного типа. Ташкент: Фан, 1974. 154 с.
14. *Салахитдинов М. С., Монов Б. Т.* Краевые задачи для уравнения смешанно-составного типа с одной и двумя линиями вырождения // ДАН Уз ССР. 1972. № 3. С. 19-21.
15. *Соболев С. Л.* Некоторые применения функционального анализа в математической физике. Новосибирск: Наука, 1962. 251 с.

Абулов Мумин Орзикулович — доцент, к.ф.-м.н., Университет Информационных технологий менеджмента. E-mail: abulov1959@mail.ru

Узаков Заир Узакович — доцент, к.ф.-м.н., Исполняющий обязанности профессора кафедры «Программное обеспечение информационных технологий». E-mail: zair90uzakov@gmail.com

Интеллектуальные
системы и
информационные
технологии

ITÔ BRIDGE ANT COLONY OPTIMIZATION (IB-ACO) - A NEW BRIDGE FROM MATHEMATICS TO COMPUTER SCIENCE

A. Taranto

Australian National University, Canberra, Australia

We adapt a Brownian bridge concept into a 2-Dimensional Itô bridge stochastic differential equation (SDE) process, which when adapted to ant colony optimization (ACO) of computer science, allows one to obtain a superior search space capability to traverse loss landscape surfaces (LLS). We call this Itô Bridge ACO (IB-ACO), and is scalable into nonconvex, nonsmooth, dynamically changing, high-dimensional surfaces and has shown to have promising results, moreso than some stochastic gradient descent (SGD) in such complex LLS. **Keywords:** *Itô bridge, ant colony optimization, loss landscape surfaces.*

1. Introduction

Moving away from the exact solutions of convex optimization, and exact solutions of tractible nonconvex problems, of mathematics, to the meta heuristic world of artificial intelligence (AI), machine learning (ML) and deep learning (DL), we face the seemingly insurmountable task of enumerating the gradient at each point of a high-dimensional loss landscape surface (LLS) [1], [2], [3]. This problem has been tackled to date via stochastic gradient descent algorithms (SGD) algorithms such as AdaGrad and RMSprop [5]. More recently, via more advanced SGD approaches such as Riemannian SGD (RSGD). But how can a SGD algorithm find the global minimum (if it exists), or nearest local minimum, of a LLS if it is nonconvex? How can it enumerate a sufficient number of points to then derive a reasonably approximate solution? In fact, determining the convexity of an arbitrary polynomial function is an NP-hard problem [4].

One approach is to leverage the parallel hive mind advantages of swarm intelligence algorithms such as ant colony optimization (ACO), and particle swarm optimization (PSO), artificial bee colony (ABC). In this paper, we adopt an Itô bridge stochastic differential equation (SDE) approach [8] to form a novel SGD algorithm.

Supported by an Australian RTS Scholarship, and an Australian Defence Innovation Network (DIN) Scholarship

2. Methodology

For a point $x \in \mathbb{R}^2$, let \mathbb{P}^x denote the law of X given initial datum $X_0 = x$, and let \mathbb{E}^x denote expectation with respect to \mathbb{P}^x . Let $X_1, X_2 : [0, \infty) \times \Omega \rightarrow \mathbb{R}^2$ defined on a probability space $(\Omega, \mathcal{F}, \{\mathcal{F}\}_{t \geq 0}, \mathbb{P})$ be two 1-Dimensional Itô diffusions satisfying the conditions given in the definition of the 1-Dimensional Itô process.

$$\begin{aligned} dX_1(t) &= f_1(X_t, t) dt + g_{1,1}(X_t, t) dW_1(t) + g_{1,2}(X_t, t) dW_2(t) \\ dX_2(t) &= f_2(X_t, t) dt + g_{2,1}(X_t, t) dW_1(t) + g_{2,2}(X_t, t) dW_2(t) \end{aligned} \tag{1}$$

The 2-Dimensional Itô bridge $\hat{X} = \{\hat{X}_1(t), \hat{X}_2(t)\}$ is an SDE of the form,

$$\begin{aligned} \hat{X}_1(t) &= X_1(t) - \frac{t}{T} X_1(T), \quad t \in [0, T], \\ \hat{X}_2(t) &= X_2(t) - \frac{t}{T} X_2(T), \quad t \in [0, T], \end{aligned} \tag{2}$$

where by substituting (1) into (2) gives,

$$\begin{aligned} \hat{X}_1(t) &= f_1(X_t, t) dt + g_{1,1}(X_t, t) dW_1(t) + g_{1,2}(X_t, t) dW_2(t) \\ &\quad - \frac{t}{T} \left(f_1(X_T, T) dt + g_{1,1}(X_T, T) dW_1(T) + g_{1,2}(X_T, T) dW_2(T) \right), \quad t \in [0, T], \\ \hat{X}_2(t) &= f_2(X_t, t) dt + g_{2,1}(X_t, t) dW_1(t) + g_{2,2}(X_t, t) dW_2(t) \\ &\quad - \frac{t}{T} \left(f_2(X_T, T) dt + g_{2,1}(X_T, T) dW_1(T) + g_{2,2}(X_T, T) dW_2(T) \right), \quad t \in [0, T]. \end{aligned} \tag{3}$$

This methodology was implemented within Python 3 environment, whose pseudocode is found below.

3. Implementation

To achieve a baseline comparison for IB-ACO, the original continuous ACO (CACO) algorithm [6] was implemented as part of our simulations, where CACO's pseudocode is outlined in Algorithm 1.

Algorithm 1. CACO Algorithm [6] [7]

```

INPUT: Number of ants
t ← 0
Initialize A(t)
Evaluate A(t)
while ( not Termination_Condition ) do
    t = t + 1
    Add_Trail A(t)
    Send_Ants A(t)
    Evaluate A(t)
    Evaporate A(t)
OUTPUT: Metaheuristic sub-optimal path to food sources
    
```

To highlight the corresponding differences and benefits of the IB-ACO over CACO, the IB-ACO pseudocode is outlined in Algorithm 2.

Algorithm 2. Itô Bridge - ACO (IB-ACO) Algorithm

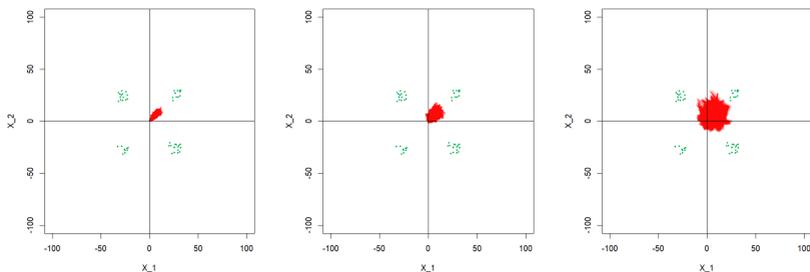
```

INPUT: Set_Algorithm_Parameters()
i, j, k ← 0
for (i = 1 to Total_Ants_In_Colony) do
  Ant s0 ← Create_sub_colony_and_release_agent
  for (j = 1 to Total_Path_Steps) do
    Calculate_Next_Step( $\mu$ ,  $\sigma$ )
    if (Path_Crossed_A_New_Food_Source) then
      Add_New_Food_Source_to_Found_Food_Table()
    if (Food_Still_Exists_In_Found_Food_Table) then
      Send_Ant_Directly_To_Known_Food_Source()
      Consume_One_Unit_Of_Food()
      Send_Ant_Directly_To_Nest()
      k = i + 1
    j = k + 1
  i = j + 1
  Update_Pheromone_On_Arc()
OUTPUT: Sopt “candidate” to be best found solution  $x \in I$ 

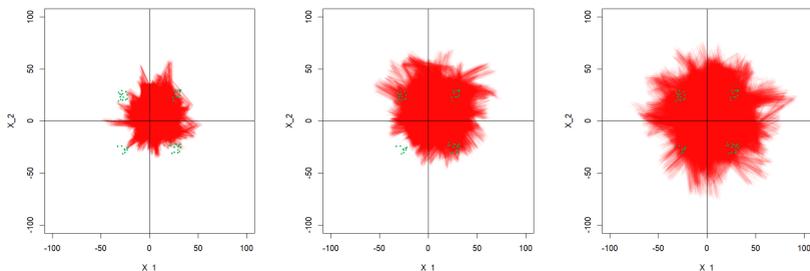
```

4. Results

Algorithm 2 was simulated to demonstrate how the search space was surveyed for given amounts of food at certain distances from the nest/origin, as shown in Figure 1.



(a). 1,000 ants, $\sigma = 0.05$ (b). 1,000 ants, $\sigma = 0.1$ (c). 1,000 ants, $\sigma = 0.25$



(d). 1,000 ants, $\sigma = 0.5$ (e). 1,000 ants, $\sigma = 0.75$ (f). 1,000 ants, $\sigma = 1$

Figure 1. Simulating IB-ACO for Search Space Coverage

The food is randomly assigned in all scenarios with $\text{Min} = 20$ and $\text{Max} = 30$. For clarity, the ants have been made transparent, to allow one to see the food and the pheromones and paths.

IB-ACO: Rather than increasing the number of ants to reach the food, IB-ACO requires the variance σ be increased. At first, the paths remain in the first quadrant due to low variance, but as this is increased, the coverage or reach is increased, later extending beyond the location of the food.

Figure 1 shows that unlike CACO, IB-ACO can keep extending beyond the food sources without necessarily having to increase the number of ants, only the variance parameter σ is required to be regulated. Initially, with small σ , all 1,000 IB-ACO paths are constrained within the first quadrant, showing that the random component does not significantly impact the bridge mechanism of IB-ACO. As variance is successively increased, the ant paths extend further and further from the nest, but are ultimately constrained back to the nest due to the dominant impact of the bridge.

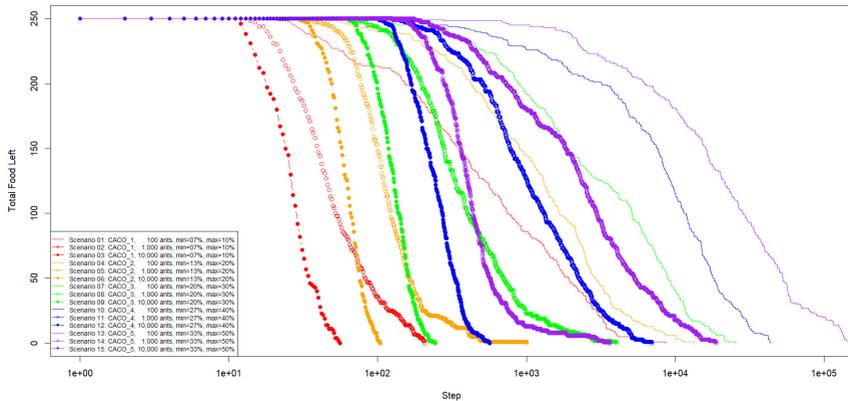


Figure 2. Loss Rate Comparison of CACO (The loss rate for the various CACO algorithm variations (CACO₁, ..., CACO₅) shows that the further the food is, the longer it takes to fully consume it. CACO₆, ..., CACO₈ were not plotted since they didn't have valid results for all ant and distance value combinations. It also shows that the less ants there are, then the longer it takes to fully consume the food. The lines closest to the left involve a solid circle for 10,000 ants, a hollow circle for 1,000 ants, and finally no circle for 100 ants.

Figure 2 shows that the less ants there are, then the longer it takes to fully consume the food. It also shows that there is an initial plateau before the ants accumulate sufficient pheromone deposits to promote efficient trail.

5. Conclusion

This paper has found a novel way to apply SDEs to obtain a scalable way to survey or traverse the immense LLS arising from AI, ML and DL. Rather than finding exact solutions to curated convex surfaces found in mathematics, we have enhanced traditional (continuous) ACO algorithms by introducing an Itô bridge system of SDEs to be able to extend the ant agents search space reach. This has all been achieved without having to increase the number of ants, which is a major limitation of ACO algorithms. Our approach to ACO, i.e. Itô bridge ACO (IB-ACO), not only exploits this system of SDEs, but also includes a number of sophisticated memory address lookups to ensure that the nest or hub can keep track of where all the food is. Ultimately, the global minimum or nearest local minimum can be equated to a food source. This means that IB-ACO can be used as a novel

algorithm to for numerous constraint-based optimization problems found in large scale, nonconvex and dynamically changing ML applications and LLS.

REFERENCES

1. *H. Li, Z. Xu, G. Taylor, C. Studer and T. Goldstein* Visualizing the loss landscape of neural nets. *Advances in neural information processing systems*. 2008. Vol. 31. P. 1–11. https://proceedings.neurips.cc/paper_files/paper/2018/file/a41b3bb3e6b050b6c9067c67f663b915-Paper.pdf.
2. *M. Geiger, L. Petrini and M. Wyart* Landscape and training regimes in deep learning. *Physics Reports*. 2021. Vol. 924. P. 1–18. <https://www.sciencedirect.com/science/article/pii/S0370157321001290>.
3. *R. Bain, M. Tokarev, H. Kothari and R. Damineni* LossPlot: A Better Way to Visualize Loss Landscapes. *arXiv preprint, arXiv:2111.15133v1*, 2021. P. 1–5. <https://arxiv.org/pdf/2111.15133>.
4. *A.A. Ahmadi, A. Olshevsky, P.A. Parrilo and J.N. Tsitsiklis*, NP-hardness of deciding convexity of quartic polynomials and related problems. *Mathematical programming*, 2013. Vol. 137. P. 453–476. <https://arxiv.org/pdf/1012.1908>.
5. *S. Ruder*, An overview of gradient descent optimization algorithms. *arXiv preprint*, 2016. [arXiv:1609.04747](https://arxiv.org/abs/1609.04747).
6. *G. Bilchev and I.C. Parmee* The ant colony metaphor for searching continuous design spaces, *Evolutionary Computing: AISB Workshop Sheffield*, 1995. P. 25–39.
7. *M. Dorigo, V. Maniezzo and A. Colorni* The Ant System: An autocatalytic optimizing process. *Dipartimento di Elettronica, Politecnico di Milano*. 1991.
8. *B. D’Auria and A. Ferriero* The optimal stopping time for a class of Itô diffusion bridges. *arXiv Preprint*, 2019. P. 1–9. <https://arxiv.org/pdf/1909.02916.pdf>.

Dr. Aldo Taranto — PhD, Researcher, School of Computing, College of Engineering, Computing and Cybernetics, ANU. E-mail: Aldo.Taranto@anu.edu.au

ANALYSIS OF THE EFFECTIVENESS OF PIPELINING OF APPLICATION PROCESSING

A. Korsun¹, S. Sushchenko^{1,2}, A. Shkurkin¹

¹*Tomsk State University, Tomsk, Russia*

²*Tomsk Branch of the Federal Research Center for Information and Computing Technologies, Tomsk, Russia*

A pipeline model of the process of executing a heterogeneous flow of commands by a homogeneous processor is considered. The functional dependence of the optimal pipeline length on the execution time of individual application commands, the size of the application, the overhead of writing and reading intermediate results at each stage of the pipeline and the heterogeneous probabilities of rebooting the pipeline in case of branch prediction failures in various sections of the program code is obtained analytically. The analysis of application execution time by the processor pipeline optimal in length is carried out. Conditions have been found under which pipelining provides acceleration of calculations. The continuity of the conducted research to the known results is shown. **Keywords:** *processor pipeline, pipeline length, optimization, heterogeneous command flow, application execution time, application size, probability of pipeline restart.*

Introduction

The most important parameter of the processor of a computer system, which determines its performance, is the number of phases of the processor pipeline. The classic processor pipeline [1] has five phases – reading the command, decoding the command, reading the operands, executing the command and writing the result. To increase performance, each of the classic stages is divided into separate phases, which makes it possible to increase the operating frequency of the processor and, as a result, performance. As a rule, the phases of reading commands and operands, as well as the phase of writing results, are accelerated due to the advanced command sampling, the organization of multi-level associative caching of addressable objects with a high non-blocking coefficient of hierarchical memory, which ensures the execution of a large number of non-blocking access transactions to a multi-level memory subsystem, and the possibility of write-back [2, 3]. The speed of the instruction decoding phase they are usually accelerated by organizing parallel operation of several pipelined decrypting

devices of varying complexity. The high speed of the instruction execution phase when processing preconditions are met (availability of operands and necessary actuators, absence of true interdependencies) is achieved using disordered and speculative instruction execution, several independent actuators of various subsets of the processor instruction system, decomposed in turn into several independent stages [1, 3]. Also, the acceleration of these phases is achieved by increasing the volume of register memory to overcome the problems of "false" interdependencies using the architectural technique "register renaming which provides mapping of a limited number of general-purpose architectural registers to a significantly larger number of physical registers [1, 3]. In addition, acceleration of the execution stage with "true" interdependencies of sequential application commands is achieved using the architectural technique of "data traversal and promotion which reduces the overhead of writing the results of the previous command and reading operands for the next [1, 3]. In well-known works, results have been obtained only for a homogeneous flow of processed application commands and an equally probable restart of the pipeline if all conditional program transitions of the application are incorrectly predicted. In addition, these publications do not contain a study of the effect of optimizing the operational characteristics of the conveyor processing process. In this paper, the optimization of the processor pipeline and the analysis of its performance indicators are performed in conditions of heterogeneous duration of execution of individual application commands and heterogeneous overhead costs associated with the execution of various commands on the pipeline, as well as heterogeneous probabilities of rebooting the pipeline in case of branch prediction failures in various sections of applications.

1. The processor model

A central processor with dynamic command execution [1] is modeled as a homogeneous pipeline with a number of stages $D \geq 1$, processing a heterogeneous stream of application commands with the same speed at different stages of the pipeline. The heterogeneity of the command flow implies a different complexity of individual application commands. We assume that the processor is running an application consisting of $N \geq 1$ commands. We believe that the execution time of each application command is different and, when executed as a single stage, is $T_n, n = \overline{1, N}$. Usually, the pipeline stages are implemented as independent actuators [1, 3]. Because of this, when executing a command at certain stages, intermediate results are stored in the service buffer fields of the pipeline control structure, and before processing at the next stage, they are read from the control buffer. We assume that the sum of the times of read-

ing and writing the intermediate results of each stage of the pipeline is determined by the type of command, is proportional to the duration of its execution and is $(N - 1)p$, reboot time $- (D - 1)(N - 1)p \left(\frac{T}{D} + t\right)$, and the average application execution time is determined by the ratio [4] $T(D, N) = \{(D + N - 1) + (D - 1)(N - 1)p\} \left(\frac{T}{D} + t\right)$. In the case of a heterogeneous command flow, the distribution of the number of reboots also follows a binomial distribution, and the time of a separate reboot of the pipeline is determined by the delay in performing $D-1$ stages of processing that command, after which, due to an incorrect branch prediction, a reboot was required. The average execution time of an application that is heterogeneous in terms of the composition of commands by a processor with dynamic execution will take the form [5]

$$T(D, N) = (D - 1) \left(\frac{T_m}{D} + t_m + \sum_{n=2}^N p_n \left(\frac{\bar{T}_n}{D} + \bar{t}_n \right) \right) + \sum_{n=1}^N \left(\frac{T_n}{D} + t_n \right),$$

$$T_m = \max_{n=1, N} T_n, t_m = \max_{n=1, N} t_n.$$

2. Optimization of the length of the processor pipeline

Let's determine the time gain from the pipelined execution of the application compared to its execution by a single-stage device:

$$\Delta(D, N) = T(1, N) - T(D, N) =$$

$$= \frac{D-1}{D} \left[\sum_{n=1, n \neq m}^N T_n - Dt_m - \sum_{n=2}^N p_n (\bar{T}_n + D\bar{t}_n) \right].$$

From here, it is easy to see that the gain is positive under interval constraints on the length of the pipeline

$$1 < D < \left(\sum_{n=1, n \neq m}^N T_n - \sum_{n=2}^N p_n \bar{T}_n \right) / \left(t_m + \sum_{n=2}^N p_n \bar{t}_n \right).$$

In addition, the functional dependence of the gain on D has an asymptote $\Delta_A(D, N) = -(D - 1) \left(t_m + \sum_{n=2}^N p_n \bar{t}_n \right) + \sum_{n=1, n \neq m}^N T_n - \sum_{n=2}^N p_n \bar{T}_n$ and when interval conditions are met, positivity is a unimodal function of the length of the conveyor. For a uniform flow of commands ($T_n = T, t_n = t, n = 1, N$) and an equally likely reboot of the pipeline after each command ($p_n = p, n = 1, N$) the winnings are converted to $\Delta(D, N) = \frac{D-1}{D} [T(N - 1)(1 - p) - Dt(1 + p(N - 1))]$, and the condition

of the positivity of the win takes the form $1 < D < \frac{T(N-1)(1-p)}{t(1+p(N-1))}$. The gain has a linear dependence on the size of the application with a slope angle coefficient equal to $\frac{D-1}{D} [T(1-p) - Dtp]$. A non-zero gain is achieved for applications of the size $N > 1 + \frac{Dt}{T(1-p) - Dtp}$. The same gain for pipelines with different lengths equal to $D > 1$ and $D + d, d \geq 1$ it is achieved when $N = 1 + \frac{t(D+d)}{T(1-p) - t(D+d)p}$. By $D = T(1-p)/tp$ the gain is invariant to the size of the application N and is determined by the ratio $\Delta(D, N) = t - T(1-p)/p$. The range of values of the dependence of the gain on the probability of rebooting the pipeline p varies from $\Delta(D, N) = \frac{D-1}{D} [T(N-1) - Dt]$ at $p = 0$ to $\Delta(D, N) = -t(D-1)N$ at $p = 1$. A positive gain is achieved for the probability of a reboot $p < \frac{T(N-1) - Dt}{(N-1)(T+Dt)}$. With an unlimited increase in the size of the application, this inequality takes the form $p < \frac{T}{T+Dt}$. From the ratio for the win, we find the optimal length of the conveyor that maximizes the win

$$D_o = \sqrt{\left(\sum_{n=1, n \neq m}^N T_n - \sum_{n=2}^N p_n \bar{T}_n \right) / \left(t_m + \sum_{n=2}^N p_n \bar{t}_n \right)}$$
. A uniform flow of commands and an equally likely pipeline restart at any point in the application leads to a ratio for the optimal length of the view $D_o = \sqrt{\frac{T(N-1)(1-p)}{t(1+p(N-1))}}$. From here, it is easy to see that with an unlimited application size ($N \rightarrow \infty$), the optimal length of the pipeline processing a homogeneous stream of commands takes the well-known form $D_o = \sqrt{\frac{T(1-p)}{tp}}$.

The optimal gain is twice the difference between the arithmetic mean and geometric mean of the cumulative application execution time without the most time-consuming command $\sum_{n=1, n \neq m}^N T_n - \sum_{n=2}^N p_n \bar{T}_n$ and overhead

$$t_m + \sum_{n=2}^N p_n \bar{t}_n:$$

$$\Delta(D_o, N) = 2 \left[\frac{\sum_{n=1, n \neq m}^N T_n - \sum_{n=2}^N p_n \bar{T}_n + t_m + \sum_{n=2}^N p_n \bar{t}_n}{2} - \sqrt{\left(\sum_{n=1, n \neq m}^N T_n - \sum_{n=2}^N p_n \bar{T}_n \right) \left(t_m + \sum_{n=2}^N p_n \bar{t}_n \right)} \right]$$

From this it can be seen that the zero value of the optimal gain is achieved when the terms of the arithmetic mean and geomet-

ric mean are equal $\sum_{n=1, n \neq m}^N T_n - \sum_{n=2}^N p_n \bar{T}_n = t_m + \sum_{n=2}^N p_n \bar{t}_n$. For a homogeneous flow of commands and an equally probable restart of the processor pipeline, the optimal gain will be $\Delta(D_o, N) = 2 \left[\frac{T(N-1)(1-p) + t(1+p(N-1))}{2} - \sqrt{T(N-1)(1-p)t(1+p(N-1))} \right]$.

Zero gain is achieved with equality $T(N-1)(1-p) = t(1+p(N-1))$. Then the size of the application must exceed the minimum value $N_{\min} = 1 + \frac{t}{T(1-p)-tp}$, that ensures $\Delta(D_o, N) = 0$. In addition, the probability of a reboot must satisfy the constraint $p < \frac{T(N-1)-t}{(N-1)(T+t)}$. The dependence of the optimal areas of the pipeline length on the size of the application has the form of a saturation curve. The optimal number of processor pipeline phases is largely determined by the probability of a reboot.

REFERENCES

1. *Tanenbaum E., Austin T.* Computer architecture. 6-th edition St. Petersburg: St. Petersburg, 2021. 816 p.
2. *Sushchenko M.S., Sushchenko S.P.* Mathematical models of hierarchical memory of computing systems Novosibirsk: Nauka; Tomsk: Publishing House of NTL, 2012. 148 p.
3. *Sushchenko S.P.* Architecture of computing systems. Textbook Tomsk: Publishing house "SKK-Press 2006. 198 p.
4. *Sushchenko S.P.* Parametric optimization of a packet switching network // Automation and computer engineering. 1985. № 2. P. 43–49.
5. *Sushchenko S.P.* Analysis of end-to-end message delay in a multi-link virtual channel // Automation and computer engineering. 1989. № 3. P. 52–64.

Korsun Alexander — Assistant of the Department of Applied Informatics. E-mail: aleksnfs15@gmail.com

Sushchenko Sergey — Doctor of Technical Science, Head of the Department of Applied Informatics. E-mail: ssp.inf.tsu@gmail.com

Shkurkin Alexey — Candidate of Technical Science, Associate Professor of Department of Applied Informatics. E-mail: shkurkin@mail.tsu.ru

A METHOD FOR GENERATING RANDOM NUMBERS ON INTERNET OF THINGS DEVICES AND ANALYZING THE RESULTS

M. M. Nurullaev

Bukhara Engineering Technological Institute, Bukhara, Uzbekistan

This article presents a novel method for generating random numbers tailored specifically for IoT (Internet of Things) devices. It integrates entropy from both hardware-based sources, such as ring oscillators, and software algorithms, including cryptographic hash functions, to enhance randomness and security. The Hybrid Cryptographic Key Generator (HCKG) algorithm is introduced, which comprises three phases: Initialization, Generation, and Reseeding. The effectiveness of the proposed method is evaluated through NIST SP 800-22 and Dieharder statistical tests, showing that the generated random numbers possess high randomness quality. This study underscores the importance of robust random number generation for securing IoT devices, addressing unique challenges in their resource-limited settings. **Keywords:** *IoT Security, Random Number Generation, Entropy Sources, Statistical Tests.*

Introduction

The proliferation of Internet of Things (IoT) devices has brought forth new challenges in ensuring secure communications and data integrity. One critical aspect of IoT security is the generation of high-quality random numbers, which are fundamental for encryption, authentication, and other cryptographic processes. Traditional methods of random number generation (RNG) often rely on hardware or software techniques that may not be suitable for the constrained environments of IoT devices. This article explores a novel method for generating random numbers specifically tailored for IoT devices, examining both its implementation and the quality of the generated numbers. The study of random number generation (RNG) and its application in ensuring the security of IoT (Internet of Things) devices has gained significant attention in recent years. The increasing deployment of IoT devices in various sectors, such as healthcare, smart homes, industrial automation, and transportation, has highlighted the critical need for robust security mechanisms. The generation of high-quality random numbers is fundamental to these security mechanisms, as they underpin cryptographic processes, secure communications, and authentication protocols. Li et al. [1] proposed

a hybrid chaotic map and genetic algorithm-based pseudorandom number generator specifically designed for IoT security, highlighting the importance of integrating different methodologies to enhance randomness. Mukherjee et al. [2] developed a lightweight encryption scheme for IoT devices using RNG, addressing the need for efficient and secure random number generation in resource-constrained environments. Zhou, Wang, and Yu [3] explored the use of blockchain technology to create an efficient and secure random number generator for IoT devices, demonstrating innovative approaches to enhancing security. In [4], a detailed classification of the applications of random number generators in IoT is provided.

1. Hardware and Software Environment

IoT devices are characterized by limited computational power, memory, and energy resources. The proposed method leverages a hybrid approach combining hardware-based noise sources with software algorithms to enhance entropy [5]. The hardware component includes a ring oscillator, a widely used source of entropy in constrained environments due to its simplicity and low power consumption. The software component involves a post-processing algorithm based on cryptographic hash functions to ensure uniformity and unpredictability of the generated numbers.

2. Post-Processing Algorithm

The raw entropy collected from the ring oscillator is subjected to a post-processing algorithm to remove any bias and ensure high-quality randomness. We employ the O'zDSt 1106:2009 cryptographic hash function for this purpose [6]. By feeding the raw entropy into the hash function, we produce a uniformly distributed output that meets the requirements for cryptographic applications [7], [8], [9]. The hash function's output serves as the final random number.

3. The Hybrid Cryptographic Key Generator (HCKG) algorithm for IoT Devices

The Hybrid Cryptographic Key Generator (HCKG) algorithm is designed to generate high-quality random numbers for securing IoT devices. This algorithm combines entropy from both True Random Number Generators (TRNGs) and system events to ensure the randomness and unpredictability of the generated keys [10]. The HCKG algorithm comprises three main phases: Initialization, Generation, and Reseeding. Below is a detailed representation of the algorithm.

1) Initialization Phase

1. Gather Entropy:
 - Collect entropy from TRNG, which may include sources such as thermal noise, quantum phenomena, or other physical processes.
 - Collect entropy from system events, such as keyboard inputs, mouse movements, disk I/O timings, network activity, and environmental sensor data.
 2. Set Nonce:
 - Generate a unique nonce value to ensure that each initialization is unique.
 3. Set Personalization String:
 - Optionally, set a personalization string provided by the user to customize the seed material.
 4. Create Seed Material:
 - Combine the collected entropy, nonce, and personalization string to form the seed material.
 - `seed_material = entropy_input||nonce||personalization_string`
 5. Compute Initial V:
 - Compute the initial value V by hashing the seed material.
 - $V = \text{Hash}(\text{seed_material})$
 6. Compute Initial C:
 - Compute the initial value C by hashing the concatenation of V and the nonce.
 - $C = \text{Hash}(V||\text{nonce})$
 7. Set Reseed Counter:
 - Initialize the reseed counter to 1.
 - `reseed_counter = 1`
- 2) Generation Phase
1. Check Reseed Condition:
 - Before generating random bits, check if the reseed counter exceeds a predefined reseed interval.
 - If the reseed counter is greater than the reseed interval, perform the Reseeding phase.
 2. Generate Random Bits:
 - Initialize an empty output string to store the random bits.
 - For n iterations (where n is the number of random bits needed):
 - Update V by hashing the concatenation of V and C .
 - Append the new V value to the output string.
 - $V = \text{Hash}(V||C)$
 - Increment the reseed counter by 1.

- `reseed_counter` += 1
- 3. Return Output:
 - Return the first n bits of the output string as the generated random bits.
- 3) Reseeding Phase
 1. Gather New Entropy:
 - Collect new entropy from TRNG and system events to refresh the seed material.
 2. Set Additional Input:
 - Optionally, set additional input provided by the user to further enhance the seed material.
 3. Create New Seed Material:
 - Combine the new entropy and additional input to form the new seed material.
 - `seed_material` = `new_entropy_input`||`additional_input`
 4. Update V :
 - Update the value of V by hashing the concatenation of the old V value and the new seed material.
 - $V = \text{Hash}(V||\text{seed_material})$
 5. Update C :
 - Update the value of C by hashing the concatenation of the updated V value and the new entropy input.
 - $C = \text{Hash}(V||\text{new_entropy_input})$
 6. Reset Reseed Counter:
 - Reset the reseed counter to 1.
 - `reseed_counter` = 1

The HCKG algorithm ensures high-quality random number generation by integrating entropy from both TRNGs and system events, making it suitable for securing IoT devices. The Initialization phase sets up the initial state, the Generation phase produces random bits, and the Reseeding phase refreshes the entropy to maintain security. By following these steps, the HCKG algorithm provides a robust and reliable solution for generating cryptographic keys and random numbers.

4. Entropy Assessment

To evaluate the quality of the generated random numbers, we conducted a series of tests based on the NIST Statistical Test Suite [11] and the Diehard Tests [12]. These tests assess various aspects of randomness, including frequency, serial correlation, and uniform distribution. Our results demonstrate that the proposed method produces numbers that pass all the tests

with high confidence levels, indicating strong randomness. Given the constrained nature of IoT devices, it is crucial to assess the computational overhead of the proposed method. The performance metrics, including power consumption, processing time, and memory usage, were recorded. The results show that the method is efficient and suitable for real-time applications in IoT environments. Using the algorithm HCKG (Hybrid Cryptographic Key Generator), 1000 random numbers were created and statistical tests were carried out to assess their degree of randomness. In this, random numbers were generated in Python, and then NIST SP 800-22 and Dieharder tests were used to assess the degree of randomness. These tests provide powerful statistical methods for estimating random number randomness levels. Through these processes, it is possible to verify the quality and safety of the random numbers generated. Traditional RNG methods on IoT devices often rely solely on software techniques, which can suffer from predictability and lack of entropy in constrained environments. Our hybrid approach combines the strengths of hardware and software methods, resulting in superior entropy and randomness. Compared to purely hardware-based methods, our approach provides an additional layer of security through cryptographic post-processing. The ability to generate high-quality random numbers is fundamental for various security protocols in IoT networks. Our method enhances the overall security posture of IoT devices by providing reliable randomness for key generation, encryption, and authentication processes. This contributes to the robustness and resilience of IoT systems against cyber threats.

Conclusion

In conclusion, we have presented a novel method for generating random numbers on IoT devices that combines hardware-based entropy with cryptographic post-processing. Our approach addresses the unique challenges posed by the constrained environments of IoT devices, providing high-quality randomness essential for secure communications and data integrity. NIST SP 800-22 and Dieharder test results show that random numbers generated by the HCKG algorithm provide a high degree of randomness. The results in both test sets show good randomness in most cases, taking values between 0 and 1. These results show that the random bits generated by the random number generator HCKG are safe and random enough for cryptographic purposes.

REFERENCES

1. *Li X., Cheng X.* A Hybrid Chaotic Map and Genetic Algorithm-Based Pseudorandom Number Generator for IoT Security // IEEE Internet of Things Journal. 2021. № 8(7). P. 5521-5530.
2. *Mukherjee S., Pal A., Ghosh S., Misra S* A Lightweight Encryption Scheme for IoT Devices Using RNG // IEEE Transactions on Network and Service Management. 2021. № 18(2). P. 1475-1487.
3. *Zhou R., Wang D., Yu S* An Efficient and Secure Random Number Generator for IoT Devices Using Blockchain Technology // Future Generation Computer Systems. 2021. № 114. P. 226-239.
4. *Seyhan K., Akleylek S.* Classification of random number generator applications in IoT: A comprehensive taxonomy // Journal of Information Security and Applications. 2022. № 71. P. 103365. <https://doi.org/10.1016/j.jisa.2022.103365>
5. *Alaba F. A., Othman M., Hashem I. A., Alotaibi F* Internet of Things security: A survey // Journal of Network and Computer Applications. 2021. № 170. P. 102813. <https://doi.org/10.1016/j.jnca.2020.102813>
6. *Nurullaev M. M.* Embedding data hashing algorithms into a cryptographic information security tool // Science and innovation. 2023. V. 2. № 3. P. 584-587. doi:10.5281/zenodo.785760
7. *Nurullaev M. M.* Functions and their mechanisms for generating cryptographic keys and random numbers // AIP Conference Proceedings, 2024. <https://doi.org/10.1063/5.0181797>
8. *Nurullaev M. M.* Generating random numbers for a cryptographic key based on smartphone sensors // AIP Conference Proceedings. – AIP Publishing, 2024. <https://doi.org/10.1063/5.0199570>
9. *Mukhammadovich N. M., Djuraevich A. R.* Working with cryptographic key information // International Journal of Electrical and Computer Engineering. 2023. V. 13. № 1. P. 911-913.
10. *Kaur K., Kumar N., Kaur S.* Internet of Things: A Review on Technologies, Protocols, Applications, Security, and Privacy // Journal of Network and Computer Applications. 2020. № 148. P. 102481.
11. *Rukhin A., et al.* A statistical test suite for the validation of random number generators and pseudo random number generators for cryptographic application // NIST Special Publication 800-22. 2010.
12. *Marsaglia G.* The Marsaglia Random Number CDROM including the Diehard Battery of Tests of Randomness // Florida State University. 1995.

EVALUATION OF THE ROBUSTNESS OF E_BLIND/D_LC WATERMARKING ALGORITHM

V. Anzhin, V. Trenkaev

Tomsk State University, Tomsk, Russian Federation

The robustness of a watermarking algorithm reflects its resilience against various signal processing operations. This paper considers a methodology for testing watermark robustness by estimating watermark detection rates before and after signal processing and evaluating the distortions introduced by these manipulations. The methodology is applied to assess the robustness of E_BLIND/D_LC watermarking algorithm. Robustness evaluations are conducted for a range of image filters used in the testing process.

Keywords: *watermark, robustness, perceptual distortion measure.*

Introduction

Digital watermarking is one of the tools used for protecting multimedia data against piracy. Digital watermarks can embed information about the authorship of transmitted data. In the case of a content leak, they can be used to prove authorship and restrict the distribution of pirated copies. When each recipient can be identified and it is possible to transmit multimedia data to each recipient individually, the watermarks can be used to find which of recipients illegally redistribute received data [1].

The intruder can attempt to remove or modify embedded watermarks. Watermark robustness [2] refers to the resilience of the watermark against various signal processing operations that can be applied to the watermarked multimedia content. These operations include signal transformations such as image filtering, cropping, recompression into different formats, etc. The study evaluates the robustness of the E_BLIND/D_LC [3] watermarking algorithm against the range of signal transformations.

1. E_BLIND/D_LC watermarking algorithm

E_BLIND/D_LC algorithm is designed for embedding a 1-bit message m into the original image C_o . An image is a two-dimensional array of pixels, each pixel defined by color values. The result of algorithm is the watermarked image C_w . The reference pattern W_r with the same size as the original image C_o is used for embedding. W_m is selected based on W_r according to the embedding bit m . α value is used to scale the reference

pattern for enabling the tradeoff between imperceptibility and robustness. The resulting watermarked image C_w is calculated by applying the scaled watermarking pattern W_a to the original image C_o .

$$W_m = \begin{cases} W_r, & \text{if } m=1 \\ -W_r, & \text{if } m=0 \end{cases} \quad (1)$$

$$W_a = \alpha * W_m \quad (2)$$

$$C_w = C_o + W_a \quad (3)$$

Scaling W_m by α in (2) involves multiplying values of every pixel in image W_m to α . Applying the scaled watermarking pattern W_a to original image C_o in (3) implies pixel by pixel adding.

For detection the watermark in the image C the linear correlation between C and reference pattern W_r is used:

$$z(C, W_r) = 1/N * C * W_r = 1/N * \sum_{x,y} C[x,y] * W_r[x,y] \quad (4)$$

In the equation (4) N is the number of pixels in image, $C[x,y]$ and $W_r[x,y]$ are the pixel values at the location x,y in the image C and reference pattern W_r .

To determine the watermark is present in the image C the following equation with selected threshold value θ is used:

$$m = \begin{cases} 1, & \text{if } z(C, W_r) > \theta \\ 0, & \text{if } z(C, W_r) < -\theta \\ \text{no-watermark,} & \text{if } -\theta \leq z(C, W_r) \leq \theta \end{cases}$$

2. Robustness testing methodology

For robustness testing E.BLIND/D.LC algorithm has been implemented in C++. The source code of the implementation is available at [4]. The image dataset Flickr8k [5] is used for evaluating. The detection threshold θ is fixed at $\theta = 0.01$ in E.BLIND/D.LC algorithm implementation. There are no false positive detections in the Flickr8K database for every $\alpha \geq 0.005$ with this detection threshold.

Robustness testing methodology is the following:

- 1) Random reference pattern is generated and embedded using range of α values into each image in the Flickr8K database.
- 2) The probability of detection at a given α value is determined by the number of successful detections in images with embedded watermarks.

- 3) Images with embedded watermarks are processed using an image filters.
- 4) The number of successful post-filter detections describes the robustness of the watermark to that particular image filter.

Filters from FFmpeg [6] are used in the evaluation. The list of filters used in evaluation includes the following:

- 2D DCT filtering using `detdnoiz` with sigma values of 5 and 10.
- Blurring with the unsharp filter using parameters `unsharp = 3:3:-0.25:3:3:-0.25`.
- Sharpening with the unsharp filter using parameters `unsharp = 3:3:0.25:3:3:0.25`.
- Scaling to 75% of the original image height.
- Cropping 25% of the image by height.
- JPEG transcoding with quantizer values of 1, 25, and 45.

The objective full-reference video quality metrics VMAF [7] and PSNR are used for estimating watermark imperceptibility and changes in quality after filters applied. The scripts developed for watermark robustness testing can be accessed by the link [8].

3. Testing results

Table 1

Probability of watermark detection for different α values (rows) and filters applied after watermarking (columns)

α	W/o filtering	Denoise 10	Blurring	Sharping	JPEG q=25
0.005	0.021	0.000	0.003	0.054	0.000
0.010	0.742	0.000	0.346	0.920	0.000
0.020	0.999	0.016	0.987	0.999	0.000
0.030	1.000	0.109	0.999	1.000	0.000
0.050	1.000	0.403	1.000	1.000	0.014
0.100	1.000	0.883	1.000	1.000	0.258
0.150	1.000	0.998	1.000	1.000	0.634
0.300	1.000	1.000	1.000	1.000	0.995

Tables (see Table 1 and Table 2) contain part of the testing results which describes the probabilities of watermark detection and VMAF quality metrics for different α values and filters, calculated on the Flickr8K database. The first column contains α values used for embedding. The subsequent col-

Table 2

VMAF values after embedding the watermark for different α values (rows) and filters applied after watermarking (columns)

α	W/o filtering	Denoise 10	Blurring	Sharping	JPEG q=25
0.005	96.993	96.809	99.951	91.383	73.286
0.010	96.971	96.816	99.927	91.369	73.292
0.020	96.808	96.823	99.824	91.342	73.289
0.030	96.532	96.782	99.589	91.267	73.268
0.050	95.969	96.548	98.801	90.985	73.203
0.100	93.593	95.082	95.917	89.284	72.750
0.150	90.123	92.616	92.286	86.275	71.930
0.300	76.647	80.571	79.060	73.499	67.719

umn names describe filters applied after watermark embedding. The value at the intersection of a specific α value row and a filter column in Table 1 represents the probability of detecting the watermark embedded with that α value after the image has been processed by the corresponding filter. In Table 2 the respective value represents the VMAF metric between the original image and the image with the watermark embedded at given α value, followed by filter processing. Full tables of evaluation results can be found at the link [8]. The Figure 1 shows the probability of watermark detection according to α value used on embedding and the type of filter applied after watermarking. Figure 2 describes the quality of the processed image as calculated by the VMAF metric.

A notable result is the improvement in the effectiveness of E_BLIND/D_LC algorithm, in terms of watermark detection probability, through sharpening filtering for small α values. For example, as shown in Table 1, embedding with $\alpha=0.01$ without applying any filters gives a detection probability of 0.742, but applying a sharpening filter increases the detection probability to 0.92.

Blurring filtering decreases detection probability for small α values and increases image quality as calculated by the VMAF metric for all α values used in the experiment. This means that using blurring for E_BLIND/D_LC watermark removal does not have penalties in the resulting image quality.

DCT denoising, as well as blurring, increases image quality as measured by the VMAF metric for all α values except the smallest ones. Also DCT denoising is more effective at decreasing the probability of watermark detec-

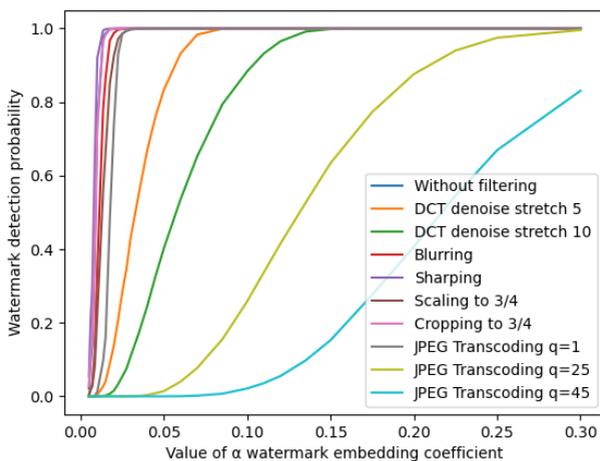


Figure 1. Probability of watermark detection according to α value used for embedding and the type of filter applied after watermarking

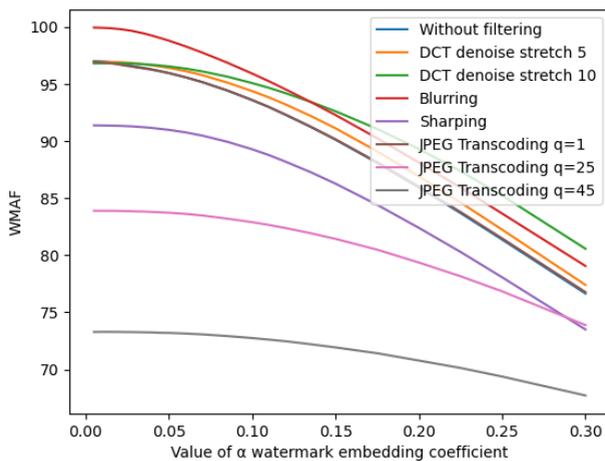


Figure 2. VMAF quality metric for the resulting image based on α value used for embedding and the type of filter applied after watermarking

tion compared to the blurring filter with the filters coefficients selected for testing. As a result, DCT filtering is effective at removing E_BLIND/D_LC watermark but may slightly decrease image quality for small α values.

4. Conclusion

The methodology for testing watermark robustness is considered, based on estimating watermark detection probability before and after signal processing and evaluating the distortions introduced by these manipulations. This methodology was applied to assess the robustness of E_BLIND/D_LC watermarking algorithm. The robustness of E_BLIND/D_LC was estimated, with results showing improvements in detection rates after applying a sharpening filter, and enhancements in objective quality combined with reductions in detection probability after blurring and DCT denoising filtering.

REFERENCES

1. *Wagner N.* Fingerprinting // 1983 IEEE Symposium on Security and Privacy, Oakland: IEEE Computer Society, 1983. P. 18-22.
2. *Asikuzzaman M., Pickering MR.* An overview of digital video watermarking. // IEEE Transactions on Circuits and Systems for Video Technology., 2017., Vol.28 № 9, P. 2131–2153.
3. *Cox I.* Digital watermarking and steganography. Burlington: Morgan kaufmann, 2007. 593 p.
4. github.com/anjin-viktor/watermarking_performance — E_BLIND/D_LC watermarking algorithm implementation. 2024.
5. *Rashtchian C., Young P., Hodosh M., Hockenmaier J.* Collecting image annotations using amazon’s mechanical turk // Proceedings of the NAACL HLT 2010 workshop on creating speech and language data with Amazon’s Mechanical Turk, Los Angeles, 2010. P. 139–147.
6. ffmpeg.org — A complete, cross-platform solution to record, convert and stream audio and video. 2024.
7. *Li Z.* Toward a practical perceptual video quality metric // The Netflix Tech Blog. 2021. Vol. 6. № 2. P. 2.
8. github.com/anjin-viktor/watermark_testing — Watermarking testing framework implementation. 2024.

Viktor Anzhin — Ph.D. student, Institute of Applied Mathematics and Computer Science, TSU. E-mail: viktor.anjin@gmail.com

Vadim Trenkaev — Ph.D. (Technical), Associate Professor, Institute of Applied Mathematics and Computer Science, TSU. E-mail: kziiktvn@gmail.com

РАЗРАБОТКА ИНТЕРФЕЙСОВ КЛИЕНТ-СЕРВЕРНОГО ВЗАИМОДЕЙСТВИЯ ВЕБ-СЕРВИСА SIMQ

Э. А. Оруджов, А. Н. Моисеев

*Национальный исследовательский Томский государственный университет,
г. Томск, Россия*

SimQ — программный комплекс, используемый для моделирования систем массового обслуживания. Данный комплекс разрабатывается на базе ТГУ для дальнейшего применения в образовательных и научных целях. В статье приведено общее описание устройства комплекса, перечислены виды систем обслуживания, доступных для моделирования в настоящий момент, а также описано текущее состояние процесса разработки. **Ключевые слова:** *системы массового обслуживания, имитационное моделирование, дискретно-событийное моделирование, агентное моделирование.*

Введение

С развитием общества и технологического прогресса появляются различные системы основная цель которых заключается в обработке поступающих запросов и требований. Системы могут варьироваться от небольших ремонтных мастерских или билетных касс до масштабных индустриальных комплексов. Независимо от размера и специализации, для всех таких систем крайне важна оптимизация рабочих процессов. Однако, чтобы внедрить необходимые изменения и улучшения, требуются точные данные, которые позволят сделать обоснованные выводы о текущем состоянии и возможных направлениях развития системы. Одним из наиболее эффективных способов получения таких данных является моделирование поведения этих систем [1, 2].

Для выполнения задач моделирования таких систем на базе ТГУ была начата разработка программного комплекса SimQ [3]. С его помощью можно выполнять имитационное моделирование систем массового обслуживания (СМО), устанавливая их параметры и функциональные возможности, собирать статистические данные, делать определенные выводы. В данной статье описан текущий прогресс разработки комплекса и план дальнейшего процесса разработки.

1. Архитектура MVP-версии проекта

Программный комплекс SimQ подразумевался как единое решение, позволяющее конечным пользователям применять инструмент SimQ для решения типовых задач теории массового обслуживания. В ходе процесса разработки было решено сфокусироваться на разработке проекта в виде веб-сервиса, вследствие чего текущей задачей работы является реализация минимальной жизнеспособной (MVP) версии комплекса SimQ, для которой будет реализован единственный способ доступа: путём обращения к серверу комплекса через веб-интерфейс в браузере пользователя. Концепция MVP-приложения отличается от представленной в ранней публикации [3]. Актуальная концепция отражена на рисунке 1.

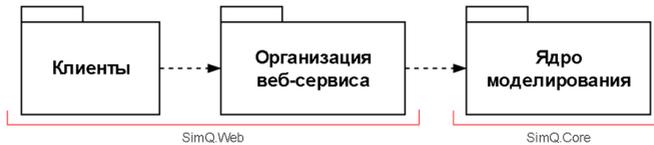


Рис. 1. Концепция организации комплекса SimQ

На диаграмме указаны три основных компонента, необходимые для реализации MVP-версии комплекса. Первые два компонента («Клиенты» и «Организация веб-сервиса») обеспечат работу комплекса с пользователями. Логика ранее представленного компонента по обеспечению многопоточковой работы [3] теперь входит в компонент «Организации веб-сервиса».

Компонент «Ядро моделирования» обеспечивает основную задачу комплекса – моделирование СМО, занимается сбором и обработкой результатов, полученных в ходе моделирования, а также предоставляет возможность хранения задач и результатов для дальнейшего их использования.

2. Анализ моделируемых задач

Ядро моделирования уже реализовано, и на данный момент взаимодействие с компонентом сводится к ручной компиляции проекта с изменением исходного кода для установки параметров моделируемой задачи. Процесс моделирования выполняется с помощью дискретно-событийного подхода [4], а построение систем массового обслуживания основано на агентном моделировании.

В ходе анализа задач, которые могут быть построены при помощи комплекса, были определены следующие наиболее распространённые виды систем массового обслуживания:

- классические системы;
- RQ-системы;
- циклические системы;
- системы с конечным числом источников.

Классические СМО. Данный вид систем представляет собой простейшие модели СМО. Элементами, из которых состоят данные системы, являются: входящие потоки заявок, очереди (буферы) и блоки приборов, обрабатывающие поступающие заявки. Примеры таких систем представлены на рисунке 2 (системы включают: *а* – входящий поток, ограниченный буфер, один прибор; *б* – входящий поток, неограниченный буфер, блок приборов с ограниченным числом серверов; *в* – входящий поток, блок приборов с неограниченным числом серверов).

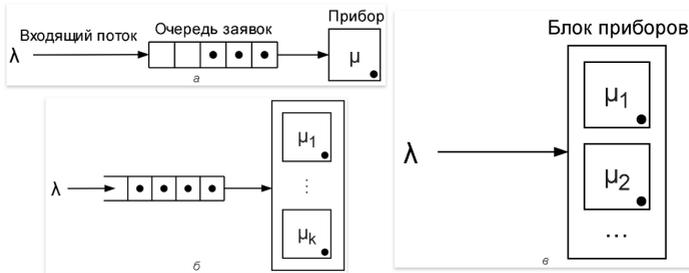


Рис. 2. Примеры классических СМО

RQ-системы. Ключевой особенностью систем с повторными вызовами – RQ-систем (от Retrial Queuing) является наличие орбит. Орбита – объект системы, хранящий в себе коллекцию заявок, который будет предпринимать попытки отправки заявки на обработку в блок приборов с некоторой периодичностью.

Система, представленная в качестве примера на рисунке 3, содержит: входящий поток заявок, орбиту и прибор, обрабатывающий заявки. Каждая прибывшая в систему заявка пытается пройти на прибор. В случае если прибор в текущий момент занят, то заявка отправляется на орбиту и ожидает её отправки на обработку.

Циклические СМО (системы поллинга). Для данных систем характерно наличие нескольких очередей, обработкой которых занимается единственный блок приборов. Блок приборов переключается меж-

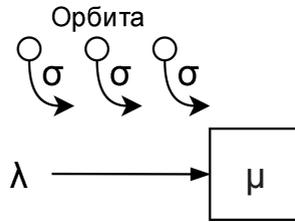


Рис. 3. Пример RQ-системы

ду очередями для обработки заявок, которые в них накопились. Выбор следующей очереди и длительность её обработки также могут быть заданы правилом.

Пример подобной системы продемонстрирован на рисунке 4. В такой системе имеется четыре источника, заявки которых попадают в разные очереди. Выбор текущего обрабатываемого хранилища заявок выполняется блоком приборов по очереди раз в L единиц времени.

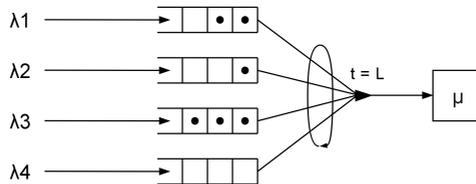


Рис. 4. Пример циклической СМО

Системы с конечным числом источников заявок. В таких системах поток входящих заявок зависит от всех источников заявок. Для источников вводится понятие «занятости», характеризующее возможность или невозможность введения новых заявок в систему. После поступления новой заявки, источник, из которого она пришла, будет считаться «занятым», из-за чего добавление новых заявок этим источником будет невозможным. Источник считается занятым до тех пор, пока его заявка находится в системе.

Вариант такой системы показан на рисунке 5. В данном примере имеется четыре источника заявок. Заявки, поступающие из разных источников, помечены соответствующими геометрическими фигурами. На текущем снимке системы занятыми оказались второй, третий и четвёртый источники. Заявки, созданные вторым и третьим источником, находятся в очереди, в то время как заявка четвёртого источника нахо-

дится в обработке. Источник под номером 1 не занят, и готов к добавлению новой заявки в систему.

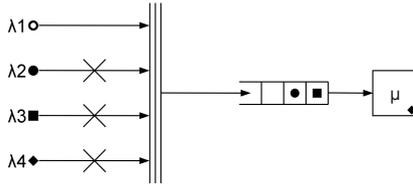


Рис. 5. Пример СМО с конечным числом источников

Описанные выше виды систем наиболее часто встречаются в задачах моделирования. В связи с этим в ядре моделирования были построены модели агентов с уникальной логикой поведения, использование которых необходимо для построения перечисленных видов систем.

3. Проектирование клиент-серверного API

Веб-сервис планируется разместить на сервере с доступом к сети, через которую будет осуществлён доступ пользователей, и вычислительной мощности которого будет достаточно для обеспечения основных функций проекта. Взаимодействие с сервисом будет происходить по протоколу HTTP, в связи с чем было описано API для взаимодействия клиента с сервисом.

Клиентом выступает веб-интерфейс приложения, отображаемый в веб-браузере пользователя. При установке соединения с сервером и отправке первых определённых HTTP-запросов, пользователь получит все необходимые данные и статические файлы, которые обеспечат дальнейшую работу веб-интерфейса и взаимодействие с API сервиса.

Основная задача комплекса SimQ — предоставить инструмент, с помощью которого будет возможно изучать различные системы массового обслуживания и их поведение в зависимости от параметров задаваемых пользователем. Исходя из данного требования, были определены четыре ключевые возможности, которые должны присутствовать в MVP-версии проекта:

1. Создать собственную модель агента СМО.
2. Сконструировать СМО.
3. Запустить процесс моделирования СМО.
4. Посмотреть результаты моделирования СМО.

Данные действия охватывают потребность в моделировании вышеописанных видов систем. На основании списка действий были выявле-

ны сущности предметной области, с которыми будет взаимодействовать пользователь по проектируемому интерфейсу:

- Задача моделирования – набор параметризованных агентов и связей между ними, представляющий изучаемую систему массового обслуживания;
- Процесс моделирования – запущенный пользователем вычислительный процесс, в ходе которого выполняется моделирование заранее созданной задачи моделирования;
- Результат моделирования – набор информации, получаемый по окончании процесса моделирования;
- Агент моделирования – объект, используемый для построения систем массового обслуживания.

Исходя из описанного списка сущностей, были определены доступные для пользователей операции над объектами данных сущностей и спроектированы в виде HTTP-запросов. Список запросов представлен на рисунке 6.

Tasks	Problems	Agents	Results
GET Получить список процессов моделирования	GET Получить список задач моделирования	GET Получить список агентов моделирования	GET Получить список результатов моделирования
POST Запустить процесс моделирования	POST Добавить задачу моделирования	POST Добавить новый тип агента	GET Получить подробную информацию о результате
GET Получить подробную информацию процесса моделирования	GET Получить подробную информацию о задаче моделирования	DEL Удалить тип агента	DEL Удалить результат моделирования
POST Остановить моделирование	DEL Удалить задачу моделирования		

Рис. 6. Список реализованных запросов

В ходе работы было описано четырнадцать HTTP-запросов в архитектурном стиле REST [5]. Следование архитектурному стилю позволит систематизировать процесс проработки интерфейсов, обеспечит масштабируемость и лёгкость интеграции в проект, а открытое описание интерфейсов клиент-серверного взаимодействия позволит создавать клиентов различных представлений.

Для документации API было необходимо подобрать спецификацию, следуемую выбранному архитектурному стилю, по которой будет описан итоговый файл. В качестве такой спецификации был выбран OpenAPI. Спецификация позволяет чётко определить структуры запросов, параметров и ответов. Наличие структурированного файла

с описанием интерфейсов также открывает возможности для добавления генеративных механизмов в проект.

Заключение

В результате работы был проведён анализ предметной области, на основании которого были определены наиболее распространённые виды систем массового обслуживания. Возможность построения и выполнения моделирования данных систем была реализована в пакете ядра моделирования. В рамках задачи по обеспечению работы программного комплекса в виде веб-приложения был спроектирован интерфейс прикладного взаимодействия между клиентами и сервером.

Дальнейший путь разработки программного комплекса SimQ заключается в проектировании и реализации серверной и клиентской составляющих проекта по заявленным интерфейсам, а также в проектировании веб-интерфейса пользователей.

СПИСОК ЛИТЕРАТУРЫ

1. *Adan I., Resign J.* Queueing Systems. Eindhoven, Netherlands: Department of Mathematics and Computing Science Eindhoven University of Technology, 2015. 182 с.
2. *Солнышкина И. В.* Теория систем массового обслуживания : учебное пособие. Комсомольск-на-Амуре: ФГБОУ ВПО «КНАГТУ», 2015. 76 с.
3. *Оруджов Э. А., Моисеев А. Н.* Разработка архитектурного прототипа подсистемы моделирования программного комплекса SimQ // Материалы IX-й Международной научной конференции "Математическое и программное обеспечение информационных, технических и экономических систем Томск: Изд. Дом Том. гос. ун-та, 2022. С. 289–294.
4. *Илюхина Н. А., Комарцева О. О.* Дискретно-событийное моделирование в управлении экономической системы муниципального образования // Современные наукоемкие технологии. 2015. № 7. С. 77–80.
5. <https://www.ibm.com/topics/rest-apis> — What is a REST API? | IBM. 2024.

Оруджов Эмиль Азер оглы — студент магистратуры Томского Государственного Университета. E-mail: k1merpro@gmail.com

Моисеев Александр Николаевич — д-р физ.-мат. наук, доцент, заведующий кафедрой, Институт прикладной математики и компьютерных наук. E-mail: moiseev.tsu@gmail.com

ПРИМЕНЕНИЕ ТЕОРИИ МАССОВОГО ОБСЛУЖИВАНИЯ В СПИРАЛЬНОМ ЖИЗНЕННОМ ЦИКЛЕ ОНЛАЙН-ШКОЛЫ ДЛЯ ОПТИМИЗАЦИИ ПРОДУКТОВОЙ МЕТРИКИ ЗАНЯТОСТИ ПРЕПОДАВАТЕЛЕЙ

А. А. Шайхутдинов

*Московский государственный технический университет
имени Н.Э. Баумана, г. Москва, Россия*

В статье рассматривается применение имитационного моделирования систем массового обслуживания в контексте спирального жизненного цикла программных продуктов. Проведение серии экспериментов с пессимистичными и оптимистичными сценариями позволяет выявлять скрытые риски предполагаемых гипотез и разрабатывать новые механики, которые помогают нивелировать или минимизировать эти риски. По итогам моделирования предложена новая продуктовая механика, внедрение которой решит проблему низкой явки учеников на бесплатные пробные уроки преподавателей, что, в свою очередь, улучшит показатели бизнеса для онлайн-школы.

Ключевые слова: *имитационное моделирование, теория массового обслуживания, система массового обслуживания, Global Purpose Simulation System, нелинейный нестационарный процесс, оптимальное управление.*

Введение

Основная функция онлайн-сервиса подбора репетиторов «Тетрика» – облегчение взаимодействия между «преподавателем» и «учеником».

Базовая транзакция сервиса включает отправку заявки учеником (или его представителем) на лендинге [1], назначение времени пробного занятия и передачу информации преподавателю, который бронирует свободное время в своем расписании.

Спиральная модель жизненного цикла программного обеспечения (рис. 1) предполагает продуктовые инкременты, которые приносят дополнительную ценность продукту.

1. Анализ исходных требований

В онлайн-сервисе подбора репетиторов «Тетрика» [1] ученики занимаются с преподавателями один на один – на каждом занятии при-

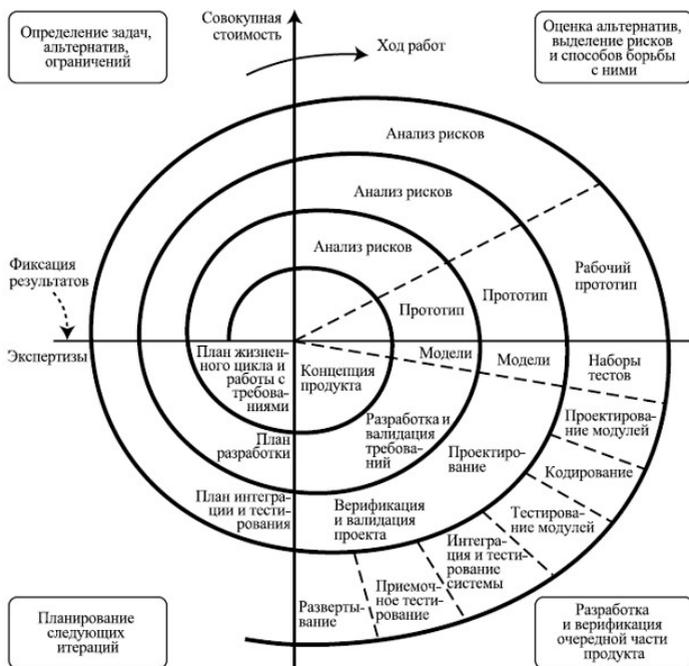


Рис. 1. Спиральная модель жизненного цикла программного продукта

существует только один ученик и один репетитор – невозможно, чтобы два ученика занимались с одним репетитором одновременно [2]. После подачи заявки ученик получает звонок от сотрудника отдела продаж, который с помощью специального интерфейса (рис. 2) бронирует ему временной слот пробного урока [3]. Каждый преподаватель специализируется на определённом предмете и цели обучения, т.е. не может обучать учеников другой дисциплине или по другой цели [4].

В этом процессе существует проблемное место – т.к. пробный урок в «Тетрике» бесплатный, часть учеников не являются на занятие. Явка на него достигает лишь 40%.

Задача: разработать механику, которая оптимизирует продуктовую метрику занятости, увеличив её до близкого к 100% значения и освободив часть преподавателей от бесполезной нагрузки.

Математическая постановка задачи оптимизации состоит в определении наибольшего значения целевой функции f_{UTIL} занятости преподавателей: $f_{UTIL} \rightarrow max$.

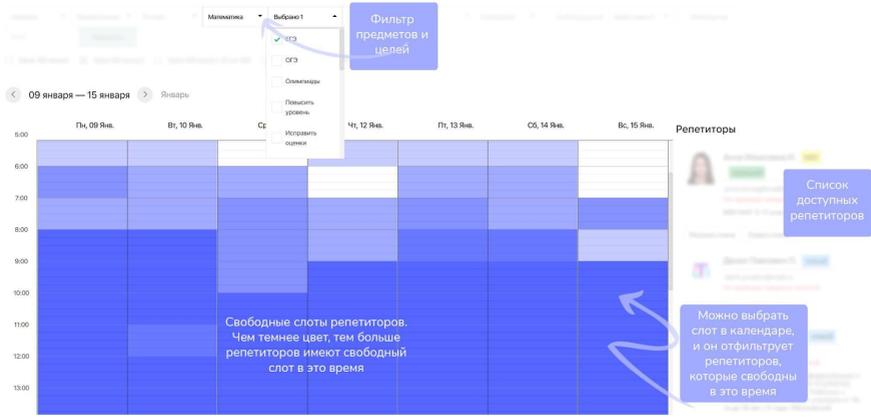


Рис. 2. Интерфейс, используемый менеджерами отдела продаж для подбора подходящих репетиторов и записи учеников на урок

2. Имитационное моделирование пробных уроков

Для данной задачи рассмотрим имитационную модель системы массового обслуживания [5] для 400 пробных уроков подготовки к ЕГЭ по информатике, проводимых одновременно (рис. 3).

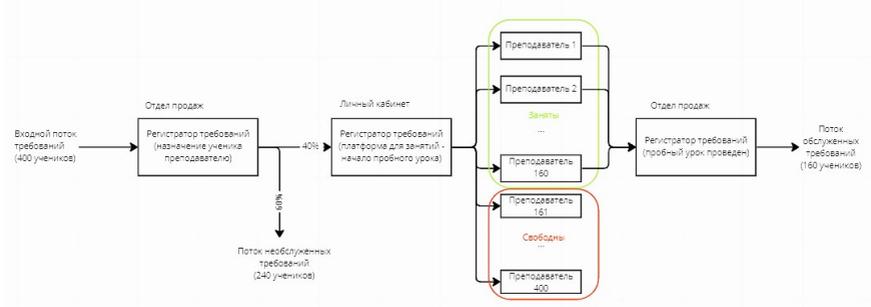


Рис. 3. СМО для явки на пробный урок 40%

Для имитационного моделирования систем массового обслуживания использовались инструменты GPSS (Global Purpose Simulation System) и MATLAB [6], что позволило увидеть загруженность преподавателей напрямую в зависимости от доступной мощности потока «горящих заявок». Имитационная модель (см. табл. 2):

Таблица 1

Имитационная модель GPSS

GPSS блок	Комментарий
TEACHERS STORAGE 400	преподаватели
generate ,,400	вх. поток 400 учеников одновременно
queue ASSIGNREG	назначение ученика преподавателю
transfer 0.6,,absence	вероятность неявки 60%
depart ASSIGNREG	
lesson queue LESSONREG	начало пробного урока
enter TEACHERS,1	ученик занимает преподавателя
depart LESSONREG	
advance 60,3	длительность урока 60+3 минуты
leave TEACHERS,1	ученик освобождает преподавателя
queue FINISH	пробный урок проведен
depart FINISH	
terminate	
absence queue ABSENCEREG	фиксация факта неявки
depart ASSIGNREG	
depart ABSENCEREG	
terminate	
generate ,,240	вх. поток 240 "горячих" учеников
queue ASSIGNREG	назначение ученика преподавателю
depart ASSIGNREG	
transfer ,lesson	начало пробного урока

QUEUE	MAX	CONT.	ENTRY	ENTRY (0)	AVE.CONT.	AVE.TIME	AVE. (-0)	RETRY
ASSIGNREG	1	0	640	640	0.000	0.000	0.000	0
LESSONREG	2	0	402	400	1.901	0.284	57.034	0
ABSENCEREG	1	0	238	238	0.000	0.000	0.000	0
FINISH	1	0	199	199	0.000	0.000	0.000	0

STORAGE	CAP.	REM.	MIN.	MAX.	ENTRIES	AVL.	AVE.C.	UTIL.	RETRY	DELAY
TEACHERS	400	197	0	400	402	1	394.995	0.987	0	0

Рис. 4. Результаты имитационного моделирования

Результаты имитационного моделирования 60 минут представлены на рисунке 4, где $max(f_{UTIL}(x)) = 0,987$ (занятость преподавателей 98,7%) при мощности потока «горящих заявок» $x = 240$. На рисунке 5

показан оптимальный случай. Для практической реализации формирования потока требований второго типа можно применить генетический алгоритм [7] или нелинейный нестационарный процесс [8].

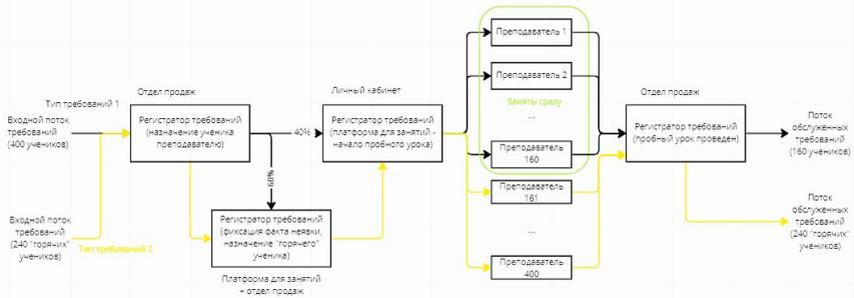


Рис. 5. СМО с дополнительным входным потоком «горящих заявок»

Вывод: назначение новых учеников из «горящих заявок» при неявке обеспечит повышение метрики занятости преподавателей.

Заключение

По результатам анализа требований и экспериментов по моделированию систем массового обслуживания для решения задачи была предложена механика назначения учеников по «горящим заявкам» в момент фиксации неявки. Онлайн-школе «Тетрика» рекомендовано провести А/В тестирование этой механики, предварительно промоделировав её с помощью предложенного метода имитационного моделирования [9]. После внедрения предложенной механики продуктовый инкремент будет завершён, таким образом, продукт с добавленной ценностью сможет продолжить движение по спирали своего жизненного цикла.

СПИСОК ЛИТЕРАТУРЫ

1. *tetrika-school.ru* — Онлайн-школа Тетрика – подготовка к ЕГЭ, ОГЭ, ВПР и олимпиадам с репетитором. 2024.
2. Гаврилова Е. Ю., Кузнецов В. И., Попов А. Ю. Алгоритм балансировки нагрузки в распределённых системах, основанный на универсальной величине вычислительной сложности задач // Искусственный интеллект в автоматизированных системах управления и обработки данных : Сборник статей Всероссийской научной конференции. В 2-х томах, Москва, 27–28 апреля 2022 года, М.: Московский государственный технический университет имени Н.Э. Баумана (национальный исследовательский университет), 2022. Т. 1. С. 301–309.

3. *Minitaeva A. M.* Analysis of the multi-criteria decision-making problem under conditions of heterogeneous interval uncertainty // Proceedings of the 4th International Conference on Control Systems, Mathematical Modeling, Automation and Energy Efficiency (SUMMA), Lipetsk: IEEE, 2022. P. 42–45.
4. *Minitaeva A. M.* Multimodel Approach to Forecasting Nonlinear Nonstationary Processes in Optimal Control Problems // 2024 6th International Youth Conference on Radio Electronics, Electrical and Power Engineering (REEPE), Moscow: IEEE, 2024. DOI 10.1109/REEPE60449.2024.10479796
5. *Гаврилова Е. Ю.* Имитационная модель виртуальной инфраструктуры для исследования задач балансировки нагрузки в распределённой системе на базе одноплатных компьютеров Raspberry Pi // Искусственный интеллект в автоматизированных системах управления и обработки данных : Сборник статей Всероссийской научной конференции. В 2-х томах, Москва, 27–28 апреля 2022 года, М.: Московский государственный технический университет имени Н.Э. Баумана (национальный исследовательский университет), 2022. Т. 1. С. 294–300.
6. *Бычков Б. И., Сюзев В. В., Гуренко В. В.* Применение MATLAB для исследования алгоритмов имитации сигналов // Будущее машиностроения России : сборник докладов: в 2 томах, Москва, 22–25 сентября 2020 года / Союз машиностроителей России, Московский государственный технический университет имени Н.Э. Баумана (национальный исследовательский университет). Том 2., М.: Издательство МГТУ им. Н.Э. Баумана, 2020. С. 95–98.
7. *Минитаева А. М., Векшин Р. Д., Шатилов А. А.* Анализ различных видов генетических алгоритмов в задачах оптимизации // Технологии инженерных и информационных систем. 2022. № 1. С. 21–34.
8. *Минитаева А. М., Шайтура С. В.* Системный анализ и разработка методики моделирования нелинейных нестационарных процессов в системе поддержки принятия решений // Информационно-технологический вестник. 2023. Т. 37. № 3. С. 54–64.
9. *Шайхутдинов А. А.* Применение методов имитационного моделирования в спиральном жизненном цикле программных продуктов для улучшения продуктовых метрик на примере онлайн сервиса подбора репетиторов // Искусственный интеллект в автоматизированных системах управления и обработки данных : Сборник статей II Всероссийской научной конференции: в 5 томах, Москва, 27–28 апреля 2023 года, М.: КДУ, Добросвет, 2023. С. 237–244.

ЛИНГВИСТИЧЕСКАЯ ОБРАБОТКА НАУЧНО-ТЕХНИЧЕСКОЙ ДОКУМЕНТАЦИИ С ЦЕЛЬЮ ПОИСКА И КАТЕГОРИЗАЦИИ

О. В. Блинова, Е. В. Панкратова

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

Рассматриваются методы, модели и технологии, применяемые для поиска и обработки информации в текстовых данных применительно к обработке научно-технической документации. Использование теоретико-множественных, векторных и вероятностных моделей имеют ряд недостатков при работе с рассматриваемыми потоками данных, связанных со сложностью алгоритмической реализации специфических особенностей области. Целесообразно применять дополнительные инструменты фильтрации, группировки данных, учет внутренней структуры документа. Перспективным представляется использование методов обработки и представления данных, используемых в хранилищах больших данных и больших языковых моделей.

Ключевые слова: *модели информационного поиска, обработка текстов, интеллектуальный поиск.*

Введение

По мере увеличения объемов баз данных научно-технической документации возникает проблема организации результативного поиска и обработки информации. В современных поисковых системах поиск все же осуществляется в первую очередь по содержащимся в запросе словам, а не по смысловой нагрузке запроса. Однако желаемым результатом для пользователя (его информационной надобностью) является совокупность научно-технических документов, содержание которых соответствует его запросу именно по смыслу. Есть два основных пути достижения этой цели: использование лингвистической и алгоритмической обработки текста или использование больших языковых моделей. Научно-техническая информация имеет ряд особенностей, важных при обработке: наличие определенной внутренней структуры у публикаций, особые требования к фильтрации и группировке данных, срав-

Исследование выполнено в рамках научной программы Национального центра физики и математики, направление № 9 «Искусственный интеллект и большие данные в технических, промышленных, природных и социальных системах

нительно узкий поиск с высокими требованиями к точности и качеству результатов. Указанные особенности требуют особых подходов и уникальных алгоритмов и методов обработки.

1. Модели извлечения информации из текстовых документов

Для осуществления поиска по словам или других операций по обработке содержания текста необходимо предварительно подготовить текст, приведя все слова в начальную форму и проделав еще ряд операций. [1]. В результате обработки поискового запроса необходимо найти различные информационные документы, которые будут содержать как минимум все варианты написания словосочетания из информационного запроса, а зачастую необходим смысловой анализ терминов и поиск синонимов, учет омонимов и т.д. Современные поисковые системы ранжируют результаты поиска на основе вычисленной релевантности. Для увеличения степени пертинентности этих результатов для конкретного пользователя используются технологии персонализированного поиска. Обработка научно-технической документации отличается использованием строгих и специфичных терминов, особыми языковыми конструкциями, большим количеством терминов и формул. При обработке таких данных помимо языковых моделей целесообразно использование экспертных знаний в виде библиотек терминов, составление семантических облаков научных областей, применение определенных критериев для кластеризации и индексации материала. Рассмотрим три основных класса моделей получения информации с помощью поисковых запросов.

Теоретико-множественные модели поиска информации

Это модели поиска информации, основанные на теории множеств. Логическая модель поиска информации подразумевает обработку любого запроса в форме предикатов с использованием логических операций AND, OR и NOT. При этом документы делятся только на две категории: релевантные (предикат принимает значение «True») и нерелевантные (предикат принимает значение «False»). Зачастую этого недостаточно для соответствия сложному пользовательскому запросу и необходимо введение дополнительных критериев оценки релевантности. Стоит понимать что составление логических конструкций в запросе может вызывать затруднения у пользователя и повышать вероятность ошибок обработки. Простые же запросы влекут за собой генерацию слишком большого объема результирующей выборки. Для решения этой проблемы целесообразно применять как интерфейсные решения, упрощающие составление логических запросов, так и современные логические модели поиска: модель нечетких множеств и метод TF-IDF, с помощью которых можно не только выявить наличие искомого термина в до-

кументе, но и оценить степень связности элементов запроса, позицию слова запроса относительно других его элементов, расстояние между отдельными терминами запроса.

Наиболее применимым и результативным методом оценки релевантности документа запросу на сегодняшний день является метод TF-IDF, который заключается в нахождении произведения двух параметров TF и IDF. Здесь TF (term frequency — частота слова) — отношение количества вхождений какого-либо слова в документ к общему количеству слов в нем; IDF (inverse document frequency — обратная частота документа) — частотная инверсия, с которой заданное слово встречается в базах данных электронных документов. Таким образом, максимальный вес слово получает в том случае, когда оно максимально часто встречается в анализируемом документе и крайне редко в иных документах. Для оценки эффективности информационного поиска применяют два основных статистических показателя: точность (P) (доля релевантных электронных документов в количественном объёме найденных документов) и полноту (V) (доля выявленных релевантных электронных документов среди всех релевантных электронных документов). Однако применение данного метода ставит пользователя перед выбором: приоритетным является полнота информационного поиска или же его точность, поскольку увеличение одного параметра влечет за собой уменьшение другого. Для решения этой проблемы вводится переменная баланса $x \in [0; 1]$ и пользователь может задать ее значения, руководствуясь своими субъективными предпочтениями. Далее вычисляется среднее гармоническое взвешенное (H) полноты и точности полученной в результате информационного поиска выборки [2]: $H = PV[xV + (1 - x)P]^{(-1)}$.

Векторная модель информационного поиска

В этой модели электронный документ, который изначально представляет из себя неупорядоченное множество слов, представляется в виде вектора весов слов; общая совокупность документов — векторного пространства. Чтобы определить вес слова в документе («важность» слова для идентификации данного текста), можно подсчитать частоту вхождения этого слова в текст — чем чаще слово встречается в документе, тем больший у него будет вес. Если слово вовсе не встречается в документе, то его вес в этом документе равен нулю. Далее, выписав по порядку веса всех слов для конкретного текстового документа, получим вектор, который и будет представлением данного документа в векторном пространстве (совокупности всех документов в агрегаторе научно-технических документов). Существует несколько стандартных

способов задания функции взвешивания для определения веса слова в текстовом документе:

- булевский вес: равен 1, если слово встречается в документе, и 0 в противном случае;
- TF: вес определяется как функция от количества вхождений слова в документ;
- TF-IDF — метод, о котором шла речь ранее.

Помимо определения степени соответствия документа поисковому информационному запросу с помощью векторной модели можно решать задачу подбора документов, составления научных карт, семантических облаков, других визуальных и интерактивных решений для навигации по совокупности документов.

Вероятностная модель информационного поиска

При использовании рассматриваемой модели электронные документы и поисковые запросы анализируются подобно векторной модели поиска информации, но дополнительно вводятся следующие параметры: P_{jnr} — вероятность нерелевантности электронного документа j ; P_{jr} — вероятность его релевантности; R_{nr} и R_{nor} — потери, характеризующие соответственно получение нерелевантных электронных документов и неполучение релевантных. Значение функции $F(j) = P_{jr}/P_{jnr} - R_{nor}/R_{nr}$ характеризует нерелевантность или релевантность анализируемого электронного документа [2].

Несмотря на то, что самыми простыми в реализации являются теоретико-множественные модели, на практике чаще всего в современных поисковых системах отдается предпочтение векторным моделям, т. к. их эффективность выше и шире круг применения.

2. Алгоритмы поиска для очень больших и слабо структурированных объемов данных

Объемы цифровой и оцифрованной научно-технической информации стремительно увеличиваются, помимо документации существуют большие массивы аудио, видео информации (записи выступлений, лекций и т. д.). Для обработки таких массивов крайне затруднительно использовать методы, предполагающие полную последовательную обработку. Для решения задач подобного масштаба целесообразно обращаться к методам big data, позволяющим быстро и качественно обработать огромные массивы разнородной слабоструктурированной информации. Для таких данных обычно используются специальные алгоритмы машинного обучения, основанные на существующих критериях релевантности, но без явного алгоритма поиска. Для работы с разно-

родными данными применяется формирование «озера данных». Озеро данных (data lake) — это хранилище практически исходных данных, как неструктурированных, так и частично структурированных. Данные собираются из различных источников, не модифицируются и не преобразуются в какой-либо формат, но озера данных содержат информацию о способах разбиения разнородных данных, об источниках, исполнителях задач сбора данных, метаданных индексирования и т. д. Для анализа этих данных, придания им однородности требуется длительная предварительная подготовка, очистка и форматирование [3]. Преимуществом хранилищ больших данных является возможность обучать алгоритмы на большом количестве материала. Для задач конкретного текстового поиска по большим массивам данных требуется предварительная группировка информации: метапрофилирование информации, кластеризация, классификация, регрессионный анализ. К основным алгоритмам кластеризации относят метод k -средних; пространственную кластеризацию для приложений с шумами — DBSCAN; алгоритм кластеризации OPTICS; метод главных компонент. Для решения задач классификации применяются методы опорных векторов, наивный байесовский классификатор, методы k -ближайших соседей, нейронные сети, деревья принятия решений. Для реализации алгоритмов обработки больших данных используют инструменты экосистем apache hadoop или greenplum, поддерживающие обработку распределенных хранилищ, параллельные вычисления и еще множество технологий. Объединенные вычислительные ресурсы позволяют применять для поиска и обработки большие языковые модели (LLM), позволяющие производить интеллектуальную обработку текста, но крайне требовательные к вычислительным ресурсам.

Заключение

При работе с научной информацией часто требуются более сложные и интеллектуальные методы обработки текстов, чем при работе пользователей в универсальных поисковых системах. Существует потребность в новых инструментах для работы с научной документацией, учитывающих специфику области. Особый интерес представляют графические и интерактивные представления данных. Динамическое формирование графических представлений по запросу, а не по заранее известным данным — задача сложная и не решенная в полной мере, особенно для работы с научной информацией. Необходимо адаптировать существующие технологии для обработки распределенных хранилищ больших данных под обработку научно-технической информации, в том числе разнородной. Широкое применение автоматизированных средств обра-

ботки и представления данных может качественно изменить работу научного сотрудника, сократив время на решение рутинных задач, и акцентировать внимание на наиболее актуальных задачах и направлениях [4].

СПИСОК ЛИТЕРАТУРЫ

1. *Ерохин В. В.* Инновационная деятельность в повышении конкурентоспособности продукции // Вестник образовательного консорциума среднерусский университет. Информационные технологии. 2015. № 5. С. 25–28.
2. *Ерохин В. В.* Аспекты применения в информационном поиске математических моделей // Вестник современных исследований. 2018. № 11.3 (26). С. 254–258.
3. *Нещерет М. Ю.* Цифровая библиография: библиотеки в поисках инновационных инструментов библиографической деятельности // Научные и технические библиотеки. 2021. Т. 1 № 7. С. 33–50.
4. *Парамонов И. Ю.; Смагин В. А.; Косых Н. Е.; Хомоненко А. Д.* Методы и модели исследования сложных систем и обработки больших данных. СПб: Лань, 2020. 236 с.

Блинова Ольга Викторовна — н.с., лаб. № 17. E-mail: blinova_olga_v@mail.ru

Панкратова Екатерина Владимировна — к.ф. -м.н., с.н.с., лаб. № 17.
E-mail: pankatya86@gmail.com

ПРИМЕНЕНИЕ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ ДЛЯ ОБНАРУЖЕНИЯ РАКА МОЛОЧНОЙ ЖЕЛЕЗЫ ПО УРОВНЮ ЭКСПРЕССИИ ГЕНОВ

Е. Е. Голденок^{1,2}, Н. А. Лукьянова^{1,2}, Я. С. Кухто¹

¹ *Красноярский государственный медицинский университет имени профессора В. Ф. Войно-Ясенецкого МЗ РФ, г. Красноярск, Россия*

² *Сибирский федеральный университет, г. Красноярск, Россия*

В работе представлен сравнительный анализ четырех моделей классического машинного обучения (k-ближайших соседей, наивный Байес, метод опорных векторов, случайный лес) и адаптивной нейро-нечеткой системы вывода ANFIS для бинарной классификации пациентов по уровню экспрессии генов на ДНК-чипах, используемых в изучении патогенетических основ рака молочной железы. Уровни экспрессии каждого гена представляют собой стандартизированные показатели интенсивности флуоресценции гибризованных цепей ДНК. Все модели классификации протестированы на трех непересекающихся подмножествах информативных признаков, отобранных тремя методами: случайный лес, рекурсивное удаление признаков, экспертная оценка. Сравнительный анализ качества моделей показал, что наибольшая точность классификации достигнута при использовании ансамблевого метода случайный лес на основе показателей экспрессии пяти генов, образующих подмножество признаков, отобранных методом рекурсивного удаления признаков. Предлагаемая модель позволит значительно ускорить обработку результатов ДНК-чипа и может быть использована в качестве инструмента для первичной диагностики – выявления наличия или отсутствия онкопатологии у пациента.

Ключевые слова: *машинное обучение, рак молочной железы, классификация, профили экспрессии генов.*

Введение

Генетические тесты, такие как ДНК-чипы позволяют проводить мониторинг генной экспрессии, определять функциональное значение генов, изучать патогенетические основы различных заболеваний, в том

Работа поддержана Красноярским математическим центром, финансируемым Минобрнауки РФ (Соглашение 075-02-2024-1429).

числе канцерогенез. Опухоль развивается в результате повреждения генетического материала, не приводящего к появлению летальных мутаций [1]. Генами-мишенями для патогенных факторов являются про-тоонкогены, гены-супрессоры опухолевого роста, гены, регулирующие клеточный апоптоз, гены репарации ДНК [2, 3]. Раннее лечение рака повышает вероятность излечения, снижает летальность и вероятность рецидива. ДНК-чипы используются во многих диагностических тест-системах РМЖ [4, 5]. Преимуществом ДНК-чипов перед традиционными методами диагностики является возможность одновременного анализа множества образцов ткани с использованием минимальных количеств исследуемого материала и реагентов. Данные экспрессии генов ДНК-микрочипа отражают состояние клетки на молекулярном уровне через стандартизированные показатели интенсивности флюоресценции гибридизованных цепей кДНК и имеют большую перспективу в качестве инструмента для постановки диагноза РМЖ [6], в том числе с использованием методов машинного обучения [7, 8, 9].

Цель исследования: сравнительный анализ моделей классического машинного обучения и адаптивной нейро-нечеткой системы вывода (ANFIS) для бинарной классификации РМЖ по уровню экспрессии генов.

1. Материалы и методы исследования

Материалом для исследования является база данных профиля экспрессии генов рака молочной железы программы Атлас генома рака (TCGA). База данных ДНК-микрочипов состоит из количественных показателей экспрессии 17814 генов по каждому из 590 образцов ткани молочной железы, в числе которых 61 образец нормальной ткани и 529 образцов с РМЖ.

Дизайн исследования включает 4 этапа. На первом этапе произведена редобработка данных ДНК-микрочипов: замена пропусков медианой, решена проблемы дисбаланса классов методов SMOTE, класс здоровых дорощен до класса больных. На втором этапе сокращена размерность признакового пространства тремя методами (случайный лес (RF), рекурсивное исключение признаков (RFE), экспертная оценка). Третий этап состоит из проектирования моделей бинарной классификации методами машинного обучения (k-ближайших соседей (kNN), наивный Байес (NB), случайный лес (RF), метод опорных векторов (SVM)) и адаптивной сетью на основе системы нечеткого вывода ANFIS. Четвертый этап включает оценку качества моделей с помощью различных метрик (Accuracy, Precision, Recall, F1) и выбор наилучших моделей бинарной классификации.

2. Результаты

База данных состоит из двух несбалансированных классов, образцов ткани без патологии в 9,67 раз меньше, чем наблюдений с РМЖ, что является типичным для медицинских исследований. Общее количество наблюдений после предобработки методом передискретизации меньшинства (SMOTE) составило 1058, по 529 в каждом классе. В обучающую выборку включено 741 наблюдение, в тестовую — 317 наблюдений. Предобработка данных и проектирование моделей машинного обучения произведено в среде Python (версии 3.10.12). Для построения сети ANFIS использовался MATLAB (версии 8.3.0.532). В системах нечеткого логического вывода большое число входов не только порождает увеличение количества логических правил, но и является ограничением для применения сети ANFIS, обусловленными малой обучающей выборкой и большим числом настраиваемых параметров сети. Поэтому в данном исследовании для всех проектируемых в работе моделей бинарной классификации, как классических методов машинного обучения так и для ANFIS в качестве входов из десятков тысяч признаков выбраны всего пять информативных признаков — пять генов. Значениями этих признаков являются уровни экспрессии генов. В результате снижения размерности признакового пространства методами случайного леса (RF), рекурсивного удаления признаков (RFE) и экспертной оценки найдены три непересекающихся подмножества признаков. Методом рекурсивного исключения признаков отобраны пять признаков — гены: COL10A1, PARP1 WISP1, MMP11, IGSF10, а методом случайного леса — другие пять признаков — гены: HOOK3, SYP, C6orf57, RAB1F, HRASLS. Время отбора признаков алгоритмом RFE составило 310 минут, 30 минут занял отбор методом RF. Экспрессия генов, отобранных в данном исследовании, имеет связь с РМЖ, что подтверждают работы [11, 12]. Задача классификации решена следующими методами машинного обучения: k-ближайших соседей (kNN), где количество соседей составило 5, наивного Байеса (NB), опорных векторов (SVM), случайным лесом (RF), состоящим из 100 деревьев, адаптивной сетью (ANFIS). Спроектированная в данной работе модель ANFIS, реализующая алгоритм нечеткого вывода Тагаки-Сугено, имеет пять входов и два выхода: класс 1 — человек болен РМЖ, класс 0 — человек здоров, функции принадлежности имеют треугольную форму, база правил содержит 243 нечетких правила, выполняющих агрегирование в соответствии с операцией t-нормы. Наилучшие результаты классификации показала модель, построенная по пяти генам (COL10A1, PARP1 WISP1, MMP11, IGSF10), отобранным методом рекурсивного удаления призна-

ков. Увеличение уровня экспрессии в группе пациентов с РМЖ наблюдается для COL10A1, PARP1, WISP1, MMP11 ($p < 0,05$ по критерию Манна-Уитни). При этом отмечается статистически значимое снижение экспрессии для гена IGSF10 ($p < 0,05$). На основе результатов исследований Genomic Data Commons (GDC) Национального института рака (NCI), принятых в качестве экспертной оценки, также отобраны 5 генов из исследуемой базы данных наиболее часто подвергающихся мутациям при раке молочной железы: TP53, TTN, GATA3, MAP3K1, HMCN1. Значения экспрессии генов TTN, GATA3, MAP3K1, HMCN1 в группах больных и здоровых имеют статистически значимые различия ($p < 0,05$). Среди отобранных встроенной функцией в методе случайного леса генов (HOOK3, SYP, C6orf57, RAB1F, HRASLS) статистически значимые различия в группах пациентов с РМЖ и здоровых наблюдаются в показателях экспрессии только одного гена RAB1F ($p < 0,05$). Бинарная классификация проведена на тестовой выборке, включающей 317 наблюдений с показателями экспрессии генов, отобранных тремя методами (RF, RFE и экспертной оценки). В таблице 1 представлены усредненные оценки качества моделей классификации, полученные в результате 10-кратной кросс-валидации.

Модель RF показала наилучшие результаты (accuracy = 99,8%, precision = 99,9%, recall = 99%, $f1 = 99,3\%$) на подмножестве признаков, отобранных методом RFE. Стоит отметить, что методы RF и SVM показали наилучшие результаты классификации на подмножестве признаков, отобранных RF. На подмножестве признаков, отобранных методом экспертной оценки, по метрикам классификации лидирует метод RF. Худший результат классификации получен при использовании метода NB (accuracy = 88,2%, precision = 88,8%, recall = 89%, $f1 = 88,2\%$) на подмножестве признаков, отобранным методом экспертной оценки.

3. Заключение

Сравнительный анализ качества разработанных моделей бинарной классификации показал, что модели случайного леса (RF) на основе всех трех подмножеств признаков, отобранных на первом этапе исследования, являются наиболее эффективными. С клинической точки зрения, отобранные группы признаков являются релевантными для диагностики РМЖ, что подтверждает потенциал этих генов как биомаркеров для диагностики данного заболевания. Выявленная в результате исследования повышенная экспрессия генов COL10A1, PARP1, MMP11, WISP1, RAB1F и сниженная экспрессия генов IGSF10, TTN, HMCN1, MAP3K1, GATA3 в тканях при раке молочной железы позволяет сократить время анализа результатов ДНК-чипирования, своевременно

Таблица 1

Сравнение классификационных моделей, построенных на основе показателей экспрессии генов, отобранных методами RF, RFE, экспертной оценки

Методы классификации	Accuracy(%)	Precision(%)	Recall(%)	F_1
RF (HOOK3, SYP, C6orf57, RAB1F, HRASLS)				
NB	95,4	98,1	96,1	95,4
RF	97,7	99,8	96,9	97,7
SVM	97,7	99,7	96,9	97,7
KNN	97,7	97,4	96,1	97,6
ANFIS	96,2	96,3	96,3	96,2
RFE (COL10A1, PARP1 WISP1, MMP11, IGSF10)				
NB	98,1	99,8	99,0	98,1
RF	99,8	99,9	99,0	99,3
SVM	99,0	99,9	99,0	99,0
KNN	99,3	99,3	98,6	99,3
ANFIS	98,4	98,4	98,4	96,2
Экспертный метод (TP53, TTN, GATA3, MAP3K1, HMCN1)				
NB	88,2	88,8	89,0	88,2
RF	96,1	94,2	94,6	94,3
SVM	92,5	92,8	93,3	92,9
KNN	94,1	93,3	90,1	93,8
ANFIS	91,5	91,5	91,6	91,5

начать лечение и снизить летальность от РМЖ. В дальнейшем планируется исследовать влияние различных способов устранения дисбаланса классов на качество разработанных моделей, а также провести бинарную классификацию, снизив размерность признакового пространства методом главных компонент.

СПИСОК ЛИТЕРАТУРЫ

1. Павлова Н. В. [и др.]. Современное представление о факторах риска и механизмах развития рака молочной железы // Успехи молекулярной онкологии. 2023. Т. 10. № 3. С. 15–23.
2. Аскандирова А. Б. [и др.]. Роль эпигенетических исследований в диагностике и лечении рака молочной железы // Онкология и радиология Казахстана. 2019. № 2(52). С. 39–44.

3. *Eliyatkın N. [et al.]*. Molecular classification of breast carcinoma: from traditional, old-fashioned way to a new age, and a new way // The journal of breast health. 2015. Vol. 11. № 2. P. 59.
4. *Ahmad F. K., Yusof Y., Yusoff N.* Filter-based gene selection method for tissues classification on large scale gene expression data // International Journal of Engineering and Technology. 2018. Vol. 7. № 2. P. 68–71.
5. *Torre L. A. [et al.]*. Global cancer statistics // CA: a cancer journal for clinicians. 2015. Vol. 65. № 2. P. 87–108.
6. *Тельшьева Т. В.* Медико-генетическое консультирование, как основа профилактики наследственных болезней // Вестник научных конференций. 2019. № 41. С. 106–108.
7. *Bhargava N.* Detection of Malicious Executables Using Rule Based Classification Algorithms // Technology and Knowledge Management: proceeding of the First International Conference on Information Technology and Knowledge Management. 2018. P. 35–38.
8. *Rajeshwari J.* Dermatology disease prediction based on firefly optimization of ANFIS classifier // AIMS Electronics and Electrical Engineering. 2022. Vol. 6. № 1. P. 61–80.
9. *Mazzotta A. [et al.]*. Nuclear PARP1 expression and its prognostic significance in breast cancer patients // Tumor Biol. 2016. Vol. 37. P. 6143–6153.
10. *Борисов В. В., Круглов В. В., Федюлов А. С.* Нечеткие модели и сети. 2-е изд., стереотип. М.: Гор. линия-Телеком, 2012. 284 с.
11. *Saxena N. [et al.]*. Differential expression of WISP-1 and WISP-2 genes in normal and transformed human breast cell lines // Mol Cell Biochemistry. 2001. Vol. 228. P. 99–10.
12. *Васильченко Д. В. [и др.]*. Роль факторов транскрипции GATA3, FOXA1, ELF5 в патогенезе и прогнозе рака молочной железы // Сибирский онкологический журнал. 2020. Vol. 19. P. 146–155.

Голденок Елена Евгеньевна — к.ф.-м.н., доцент, доцент кафедры медицинской кибернетики и информатики, КрасГМУ, доцент кафедры высшей и прикладной математики, СФУ. E-mail: egoldenok@gmail.com

Лукиянова Наталья Александровна — к.ф.-м.н., доцент кафедры медицинской кибернетики и информатики, КрасГМУ, доцент кафедры высшей математики 2, СФУ. E-mail: nalukyjanovakrsk@gmail.com

Кухто Яна Сергеевна — студент 6 курса специальности 30.05.03 Медицинская кибернетика, КрасГМУ. E-mail: yanakuhto7@gmail.com

ИССЛЕДОВАНИЕ МЕТОДОВ РАСПОЗНАВАНИЯ ЖЕСТОВ РУКИ НА ИЗОБРАЖЕНИЯХ

И. В. Баранова, А. А. Судакова

Сибирский федеральный университет, г. Красноярск, Россия

Целью данной работы является исследование и реализация методов распознавания жестов руки на изображениях. Решение указанной задачи распознавания состоит из двух частей: обнаружения на изображении руки и классификации жеста обнаруженной руки. Предлагаются модификации нейронных сетей для детекции и классификации жестов. Разработано программное обеспечение на языке Python, реализующее работу методов обработки и фильтрации изображений, алгоритмов обнаружения руки и классификации жестов. Проводятся вычислительные эксперименты по оценке работы различных методов детекции и классификации жестов по точности и качеству распознавания. **Ключевые слова:** *распознавание изображений, распознавание жестов, нейронная сеть, классификатор, машинное обучение.*

Введение

Разработка систем компьютерного зрения является одним из наиболее востребованных и актуальных направлений в области анализа данных, искусственного интеллекта и машинного обучения. Основным предназначением указанных систем является распознавание на изображениях объектов различной природы и классификация обнаруженных объектов. Одним из возможных направлений применения таких систем является распознавание жестов руки на изображении. Поставленная задача распознавания является важной и практически значимой, поскольку алгоритмы распознавания жестов могут быть использованы для управления мультимедийными системами [1], для реализации алгоритмов работы с дополненной реальностью, для синхронного перевода с языка глухонемых [2] и ряда других задач анализа и интерпретации данных.

Работа поддержана Красноярским математическим центром, финансируемым Минобрнауки РФ (Соглашение 075-02-2024-1429).

1. Постановка задачи распознавания образов

Задача распознавания жеста руки заключается в обнаружении руки человека и последующей классификации жеста обнаруженной руки. Распознавание выполняется на растровых изображениях в формате RGB. В данной работе будет рассматриваться задача распознавания статических жестов руки на изображении. Под статическим жестом понимается фиксированное положение частей человеческого тела в определённый момент времени, то есть поза.

Решение задачи распознавания объекта на изображениях выполняется в несколько этапов [3]:

- 1) предварительная обработка и фильтрация — выполняется удаление шумов, фильтрация и выделение областей, подлежащих распознаванию.
- 2) обработка результатов фильтрации — формируются признаковые описания выделенных на предыдущем этапе областей.
- 3) обучение — используются различные методы машинного обучения для создания на основе обучающей выборки классификаторов.
- 4) классификация — по признаковым описаниям, построенным ранее, принимается решение о том, к какому заранее известному классу необходимо отнести выделенные элементы.

2. Математическая постановка задачи

Пусть имеется множество объектов $X = x_1, \dots, x_n$, где $x_i \in X$ — объект распознавания. Задано F — пространство признаков, тогда функция $f(x) : X \rightarrow F$ ставит в соответствие каждому объекту x точку $f(x)$. Здесь вектор $f(x)$ является образом объекта x .

Пусть в пространстве признаков F определено множество классов $C = \{C_1, \dots, C_m\}$, $C_i \in F$, $i = 1, \dots, m$, $C_i \cap C_j = \emptyset$ при $i \neq j$. Функция классификации $g(x) : X \rightarrow C$ — неизвестная функция, разбивающая пространство образов F на m непересекающихся классов C_i . Решающее правило $\hat{g}(f) : F \rightarrow C$ — функция, являющаяся оценкой для $g(x)$ на основании образа объекта, то есть $\hat{g}(f) = \hat{g}(f(x))$. Пусть задано множество $\hat{X} = \{x_j\}$, $j = 1, \dots, k$, $k < n$, $\hat{X} \subset X$, для которого задана информация о функциях $f(x)$ и $g(x)$, т.е. $f_j = f(x_j)$ и $g_j = f(g_j)$, для $x_j \in \hat{X}$, $j = 1, \dots, k$. Множество (g_j, f_j) называется множеством прецедентов.

Требуется построить такое решающее правило (*алгоритм классификации*) $\hat{g}(f)$, чтобы $Q(\hat{g}, C) = \min_f Q(g, C)$, где $Q(g, C)$ — выбранный критерий качества классификации. В данной работе в качестве критерия используются четыре метрики (точность, полнота, усредненный индекс Жаккара и мера F1) [4].

Точность (Precision) — доля объектов, действительно принадлежащих данному классу, относительно всех объектов, которые алгоритм отнес к этому классу: $Precision = \frac{TP}{TP+FP}$, где TP — истинно положительные значения, FP — ложно положительные значения.

Полнота (Recall) — доля объектов класса, действительно относящихся к этому классу, ко всем объектам, которые нашел алгоритм: $Recall = \frac{TP}{TP+FN}$, где FN — ложно отрицательные значения.

Индекс Жаккара или метрика IoU (Intersection over Union) — метрика степени пересечения между двумя ограничивающими рамками:

$$IoU = \frac{TP}{TP+FP+FN}.$$

Мера F1 (F1-Score) — мера, сочетающая точность и полноту:

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall}.$$

3. Методы предварительной обработки и фильтрации изображения

Как уже было сказано ранее, перед проведением распознавания выполняется предварительная обработка изображения (для устранения шумов и дефектов изображения) и фильтрация необходимых областей. В данной работе для этих целей используются следующие методы: фильтр Кувахары, цветовая таблица поиска (CLTable) и выделение контуров (границ) объектов. Для выделения контуров на изображении применяется свёрточный оператор Кэнни [4].

Фильтр Кувахары — метод обработки изображений, который используется для сглаживания изображений и уменьшения шума. Он вычисляет локальное среднее значение и локальную дисперсию для четырех квадратных областей вокруг каждого пикселя и выбирает результат с минимальной дисперсией, что помогает сохранить детали изображения и сгладить шумы.

4. Методы обнаружения руки на изображении

В работе для решения задачи обнаружения (детекции) руки на изображении используется два подхода: классификатор AdaBoost [5] на основе признаков Хаара и сверточные нейронные сети с архитектурами Single Shot Detector (SSD) [6] и RetinaNet [6].

Single Shot Detector (SSD) — свёрточная нейронная сеть, разработанная для решения задачи выделения объектов и их классификации, которая работает путем предсказания рамок объектов и их соответствующих классов на различных масштабах изображения.

RetinaNet — однопроходная нейронная сеть модель для обнаружения объектов, хорошо работающая с плотными объектами небольших масштабов.

5. Методы классификации жестов руки

Для решения задачи классификации жестов руки в данной работе предлагаются две модифицированные свёрточные нейронные сети распознавания на основе архитектур MobileNetV3 [7] и ResNeXt50 [7]. Реализованный в нашей работе метод AdaBoost применяется и для решения описанной задачи классификации.

ResNeXt — свёрточная нейронная сеть, использующая несколько «групп» слоев с функциями невязки вместо многих отдельных путей [7]. В данной работе будет рассматриваться нейронная сеть данной архитектуры, состоящая из 50 слоев.

MobileNet — класс свёрточных нейронных сетей, разработанных для решения задачи классификации. MobileNet состоит из свёрточных блоков, которые помогают выделять признаки из исходного изображения.

6. Вычислительные эксперименты по сравнению методов и программный комплекс

В работе были проведены вычислительные эксперименты по оцениванию эффективности предложенного метода распознавания и классификации жестов по точности и качеству распознавания, а также их сравнению с алгоритмом AdaBoost. Сравнение и оценка работы алгоритмов выполнялись на основе датасета жестов HaGRID от Сбербанка. Датасет содержит изображения 18 различных классов (т.е. жестов). Каждый класс содержит порядка 15000 изображений в обучающей выборке и 2000 в тестовой. В данной работе будут использоваться 5 классов из-за отсутствия достаточных вычислительных ресурсов, поскольку обучение нейронных сетей распознавания и классификации на каждом из классов занимает существенное время. В данной работе рассматривались следующие жесты: кулак (fist), ок (ok), один (one), два (peace) и стоп (stop).

Для решения поставленной в работе задачи распознавания был разработан программный комплекс на языке Python, реализующий работу методов обработки и фильтрации изображений, а также перечисленных выше алгоритмов обнаружения руки и классификации жестов. На рис.1 приведен пример работы программного комплекса, распознавание выполняется с помощью нейросети MobileNet V3.

В табл. 1 представлены результаты проведенного вычислительного эксперимента по сравнению двух предложенных нейросетей классификации и классификатора AdaBoost с различными фильтрами предварительной обработки.

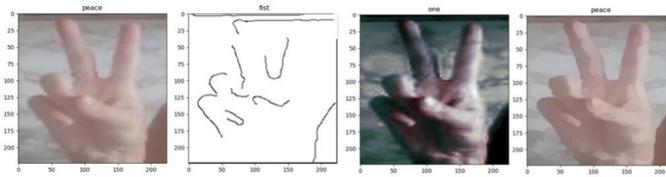


Рис. 1. Пример классификации жеста с помощью нейросети MobileNet V3

Таблица 1

Результат вычислительных экспериментов по сравнению методов классификации жестов

Метод	Мера F1	Время обучения (мин)
Классификатор AdaBoost	0,82	35
ResNeXt50	0,992	176
Кэнни + ResNeXt50	0,9512	187
CLTable + ResNeXt50	0,9756	170
Фильтр Кувахары + ResNeXt50	0,9856	188
MobileNet V3	0,8476	40
Кэнни + MobileNet V3	0,536	41
CLTable + MobileNet V3	0,8269	35
Фильтр Кувахары + MobileNet V3	0,972	32

Из результатов таблицы 1 следует, что ResNeXt50 выигрывает по точности распознавания (мера F1) по сравнению с MobileNet V3 и классификатором AdaBoost. Для нейросети MobileNet V3 применение фильтра Кувахары позволяет повысить точность распознавания. Также нами в работе [5] были выполнены вычислительные эксперименты по сравнению методов обнаружения руки на изображении. Проведенное сравнение показало, что предложенная нами нейронная сеть обнаружения руки выигрывает по точности и мере F1 по сравнению с Single Shot Detector и классификатором AdaBoost.

Заключение

В работе решена задача распознавания жестов руки на изображениях. Предложены модификации нейронных сетей для детекции и классификации жестов. Разработано программное обеспечение на языке Python, реализующее работу методов обработки и фильтрации изоб-

ражений, алгоритмов обнаружения руки и классификации жестов. Были проведены вычислительные эксперименты по оценке работы предложенных методов детекции и классификации жестов по точности и качеству распознавания. Также было проведено исследование влияния на работу предложенных алгоритмов ряда методов предобработки и фильтрации: фильтра Кувахары, цветowych таблиц поиска и сверточного оператора выделения контуров Кэнни. Предложенные методы показали высокую эффективность и точность распознавания.

СПИСОК ЛИТЕРАТУРЫ

1. *Чудновский М. М.* Алгоритм распознавания жестов руки человека на видеопоследовательности в режиме реального времени для реализации интерфейсов человеко-машинного взаимодействия // Вестник СибГАУ. 2014. № 3 (55). С. 162–167.
2. *Jayanthi P., Sathia Bhama P. R. K., Swetha K., Subash S. A.* Real Time Static and Dynamic Sign Language Recognition using Deep Learning // Journal of Scientific & Industrial Research. 2022. Vol. 81. P. 1186–1194.
3. *Лепский А. Е., Броневиц А. Е.* Математические методы распознавания образов. Таганрог: Изд-во ТТИ ЮФУ, 2009. 155 с.
4. *Баранова И. В., Гилли С. В.* Гибридный алгоритм распознавания строений на спутниковых снимках на основе метода жука и алгоритма исключения областей // Вестник ЮУрГУ. Серия: Вычислительная математика и информатика. 2024. Т. 13. № 2. С. 56–76.
5. *Судакова А. А., Баранова И. В.* Сравнение алгоритмов распознавания жестов руки на изображении // Системы управления, информационные технологии и математическое моделирование. 2024. С. 509–514.
6. *Kaiming H, Xiangyu Z., Shaoqing R., Jian S. D.* Deep Residual Learning for Image Recognition // IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2016. P. 770–778.
7. *Howard A, Sandler M., Chu G.* Aggregated Residual Transformations for Deep Neural Networks // IEEE/CVF International Conference on Computer Vision (ICCV). 2019. P. 1314–1324.

Баранова Ирина Владимировна — кандидат физико-математических наук, доцент, доцент, базовая кафедра вычислительных и информационных технологий Института математики и фундаментальной информатики Сибирского федерального университета. E-mail: irinabar@yandex.ru

Судакова Арина Анатольевна — студент, Институт математики и фундаментальной информатики Сибирского федерального университета. E-mail: sarinacfg@yahoo.com

ПРИМЕНЕНИЕ СВЕРТОЧНЫХ НЕЙРОННЫХ СЕТЕЙ ДЛЯ КЛАССИФИКАЦИИ МРТ ЗДОРОВЫХ КОЛЕННЫХ СУСТАВОВ И С ПРИЗНАКАМИ ОСТЕОАРТРИТА

Н. А. Лукьянова^{1,2}, Е. Е. Голденок^{1,2}, К. Е. Черняк¹

¹ *Красноярский государственный медицинский университет имени профессора В. Ф. Войно-Ясенецкого МЗ РФ, г. Красноярск, Россия*

² *Сибирский федеральный университет, г. Красноярск, Россия*

Цель исследования: разработка моделей сверточных нейронных сетей для повышения качества классификации МРТ снимков здоровых коленных суставов и с остеоартритом. Реализованы и обучены 5 моделей сверточных нейронных сетей на основе базовых архитектур EfficientNetB5, InceptionResNetV2, MobileNetV2, MobileNet, VGG16. Материалом исследования является объединенная база снимков, включающая набор цифровых изображений из электронного открытого репозитория Kaggle по остеоартриту (4590 снимков) и сформированный набор снимков коленных суставов пациентов Красноярского края (1904 снимков): 2690 снимков — с наличием признаков остеоартрита, 3804 — здоровые коленные суставы. На этапе предобработки данных была решена проблема недостаточности входных данных и дисбаланса классов за счет аугментации, увеличив набор до 40 000 изображений. Проведен сравнительный анализ нейросетевых моделей по доле правильных классификаций. Наилучший результат на тестовой выборке в 8 000 снимков показала модель с архитектурой EfficientNetB5 (accuracy = 95,6%), она также имеет высокие значения специфичности и чувствительности (Se = 95,5%, Sp = 95,5%), ее можно использовать в качестве скрининговой. Полученные результаты свидетельствуют о потенциале использования нейронных сетей для автоматизации диагностики признаков остеоартрита по МРТ коленного сустава.

Ключевые слова: *сверточная нейронная сеть, классификация, остеоартрит коленного сустава.*

Авторы благодарят ООО «Лечебно-диагностический центр Международного института биологических систем» города Красноярска, за предоставление обезличенных снимков коленных суставов и врача-рентгенолога Бородинского Леонида Михайловича за экспертную оценку МРТ изображений. Работа поддержана Красноярским математическим центром, финансируемым Минобрнауки РФ (Соглашение 075-02-2024-1429).

Введение

Дегенеративно–дистрофические поражения суставов имеют широкую распространенность, представляя серьезную медико-социальную проблему во многих странах мира, включая Россию. По данным [1] остеоартритом (ОА) страдают более 10% взрослого населения. Показатели заболеваемости продолжают демонстрировать тенденцию к росту [2], в связи с чем особенно актуальным представляется поиск новых методов наиболее эффективной диагностики. Постановкой диагноза занимаются высококвалифицированные специалисты, но с развитием информационных технологий, в том числе методов машинного обучения [3–5], существует возможность автоматизированного скрининга ОА по МРТ коленного сустава. Поэтому востребованы модели на базе сверточных нейронных сетей (СНС) способных работать в условиях ограниченных вычислительных ресурсов и обладающих высокой точностью распознавания повреждения коленных суставов.

Цель исследования: разработка моделей сверточных нейронных сетей для повышения качества классификации МРТ изображений с наличием или отсутствием признаков остеоартрита.

1. Материалы и методы исследования

Материалом исследования является объединенная база снимков из открытой базы Kaggle [6] «Knee Osteoarthritis Dataset with Severity Grading» (4590 снимков) и сформированная база снимков коленных суставов пациентов Красноярского края (1904 снимка). Набор реальных МРТ-изображений был предоставлен для исследования частной медицинской организацией ООО «ЛДЦ МИБС» г. Красноярска. Все изображения были предварительно деперсонализированы врачом-рентгенологом, который выступил в качестве эксперта при разметке снимков на два класса: класс Healthy — изображение здорового колена; класс Severe — тяжелое повреждение колена, включающее крупные остеофиты, значительное сужение суставов и тяжелый склероз. В процессе работы над базой снимков [6] были исключены папки с изображениями среднего и минимального повреждения коленных суставов.

Объединение набора реальных снимков колен пациентов Красноярского края и базы снимков с Kaggle составила в итоге 6494 снимка:

- класс Healthy: 3804 изображения здоровых коленных суставов (551 снимок колен пациентов Красноярского края; 3253 снимка с Kaggle);
- класс Severe: 2680 изображений поврежденных коленных суставов (1353 снимка колен пациентов Красноярского края; 1337 снимков с Kaggle).

Дизайн исследования состоит из 4-х этапов:

Этап 1. Проведение предобработки реальных МРТ снимков коленных суставов пациентов Красноярского края путем изменения DICOM формата изображений в JPEG и проведение масштабирования до размера снимков с Kaggle в 224×224 пикселей.

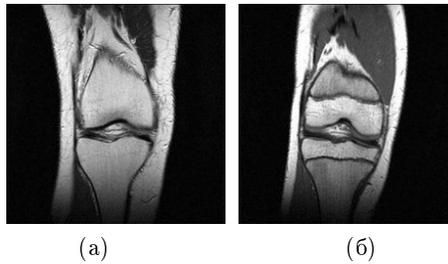


Рис. 1. Примеры обезличенных изображений суставов: (а) — неповрежденный коленный сустав, (б) — поврежденный коленный сустав

Этап 2. Формирование базы данных путем объединения базы реальных снимков с базой из Kaggle. Решение проблемы недостаточности входных данных и дисбаланса классов с помощью аугментации (табл. 1). В результате итоговый набор составляет 40 000 изображений. Разделение набора снимков на три выборки: тренировочная выборка (70%) — 28 000 снимков (по 14 000 в каждом классе); валидационная выборка (10%) — 4 000 снимков (по 2 000 в каждом классе); тестовая выборка (20%) — 8 000 снимков (по 4 000 в каждом классе).

Таблица 1

Решение проблемы дисбаланса классов для изображений коленных суставов пациентов Красноярского края совместно с БД Kaggle

	Health	Severe	Всего
Исходные данные	3 804	2 690	6 494
Зеркальная трансформация (вертикальная и горизонтальная)	+11 412	+8 070	+19 482
Аугментация	+4 785	+9 240	+14 024
Всего	20 000	20 000	40 000

Этап 3. Реализация и обучение пяти классификационных моделей сверточных нейронных сетей с различными архитектурами

(EfficientNetB5, InceptionResNetV2, MobileNetV2, MobileNet, VGG16), которые уже предобучены на базе ImageNet — наборе данных из более чем 14 миллионов изображений, принадлежащих к 1000 классам.

Этап 4. Сравнительный анализ результатов нейросетевых моделей по классификации здоровых и поврежденных коленных суставов. При проведении оценки качества классификации использовались метрики аккуратность (accuracy), точность (precision), полнота (recall), а также показатели, соответствующие клинической информативности: чувствительность (sensitivity), специфичность (specificity).

В нашей работе матрица ошибок имеет структуру (табл. 2).

Таблица 2

Матрица ошибок

Предсказанные СНС классы	Реальные классы	
	Класс Severe	Класс Health
Класс Severe	TP	FP
Класс Health	FN	TN

Поскольку в данном исследовании была решена проблема дисбаланса классов, то это позволило рассматривать долю правильных классификаций снимков моделью, как метрику качества работы алгоритма:

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + FN + TN} \cdot 100\%.$$

Показатели клинической информативности, чувствительность и специфичность, помогают определить, насколько хорошо модель способна выявить наличие или отсутствие заболевания. Чувствительность диагностики модели (Recall для класса Severe) — вероятность того, что поврежденные колени будут классифицированы моделью именно как больные: $Se = \frac{TP}{TP+FN} \cdot 100\%$. Специфичность диагностики модели (Recall для класса Health) — вероятность того, что здоровые будут классифицированы именно как здоровые: $Sp = \frac{TN}{FP+TN} \cdot 100\%$.

Стоит отметить, что метрики алгоритма классификации и показатели клинической информативности имеют некоторое сходство, но они предназначены для разных целей и работают в разных контекстах. Метрики классификации измеряют производительность модели, в то время как показатели клинической информативности помогают оценить значимость результатов диагностики или прогнозирования в медицинском контексте.

2. Результаты и обсуждения

Для построения нейронной сети были использованы открытые библиотеки Keras и TensorFlow. Программная реализация нейронных сетей осуществлена на языке Python 3.9.1 с использованием программы Visual Studio Code 2019г. Обучены 5 сверточных нейронных сетей с известными базовыми архитектурами на основе трансферного подхода, а именно EfficientNetB5, InceptionResNetV2, MobileNet, MobileNetV2, VGG16. Применена аугментация данных и метод регуляризации dropout с параметром 0,4 для преодоления эффекта переобучения нейронных сетей. С помощью метода compile() из пакета Keras связывали модели с оптимизатором Adam и метрикой accuracy, в качестве функции потерь использовали категориальную кросс-энтропию. Пространственное объединение осуществляется с помощью maxpooling. Все скрытые слои снабжены функцией активации ReLU, в последнем слое — сигмоидальная функция. Обучение сетей запускали с ограничением 10 эпох. После завершения обучения происходило сохранение весов и образа модели. Наилучшие результаты на обучающей выборке показали модели EfficientNetB5 (accuracy = 96,26%) и InceptionResNetV2 (accuracy = 96,28%). Модель EfficientNetB5 в течение всего процесса обучения показывала самый устойчивый результат, не переобучалась и показала самое низкое значение функции ошибки, равное 0,111. Модель InceptionResNetV2 имеет loss=0,15, MobileNet — loss=0,136, MobileNetV2 — loss=0,135. Самое высокое значение функции ошибки, равное 0,43, показала модель VGG16. Результаты классификации моделей представлены в таблице 3.

Таблица 3
Распределение моделей по метрикам классификации на тестовой выборке

Название модели	Accuracy (%)	Se (%)	Sp (%)
EfficientNetB5	95,55%	95,5%	95,5%
InceptionResNetV2	94,39%	92,3%	96,4%
MobileNet	93,19%	91,3%	95,1%
MobileNetV2	91,80%	91,2%	92,3%
VGG16	91,51%	94,6%	88,4%

Сравнительный анализ результатов моделей на тестовой выборке показал, что лучшей моделью классификации МРТ снимков по признаку отсутствия или наличия ОА является модель EfficientNetB5. Наихудший результат у модели с базовой архитектурой VGG16.

Заключение

В ходе данной работы были обучены и построены 5 моделей СНС позволяющие достаточно хорошо решать задачу классификации МРТ изображений коленных суставов для распознавания наличия ОА. Наилучший результат показала модель с архитектурой EfficientNetB5, ее можно использовать в качестве вспомогательного инструмента в скрининговой диагностике ОА, путем интегрирования в соответствующие медицинские информационные системы.

СПИСОК ЛИТЕРАТУРЫ

1. Балабанова Р. М., Дубинина Т. В. Динамика пятилетней заболеваемости болезнями костно-мышечной системы и их распространенности среди взрослого населения России за 2013–2017 гг // Современная ревматология. 2019. № 13(4). С. 11–17.
2. Жугрова Е. С., Беляева И. Б., Самигуллина Р. Р. Взгляд на остеоартрит с позиции доказательной медицины и собственного опыта // РМЖ. Медицинское обозрение. 2023. № 7(3). С. 167–173.
3. Зеленина Л. И., Хаймина Л. Э., Деменкова Е. А. [и др.]. Сверточные нейронные сети в задаче классификации медицинских изображений // Современные наукоемкие технологии. 2021. № 9. С. 68–73.
4. Лукьянова Н. А., Потылицина С. Е. Применение сверточных нейронных сетей в задаче распознавания пневмонии // Информационные технологии и математическое моделирование (ИТММ-2022). Томск: НИ Томский гос. университет, 2023. С. 277–282.
5. Мамедов Т. Х. [и др.]. Применение сверточных нейронных сетей для распознавания диабетической ретинопатии на цифровых изображениях // Сибирское медицинское обозрение. 2022. № 1. С. 83–87.
6. <https://www.kaggle.com/datasets/shashwatwork/knee-osteoarthritis-dataset-with-severity> — Открытый репозиторий Kaggle: Knee Osteoarthritis Dataset with Severity Grading. 2018.

Лукьянова Наталья Александровна — к.ф.-м.н., доцент кафедры медицинской кибернетики и информатики, КрасГМУ, доцент кафедры высшей математики 2, СФУ. E-mail: nalukyjanovakrsk@gmail.com

Голденок Елена Евгеньевна — к.ф.-м.н., доцент, доцент кафедры медицинской кибернетики и информатики, КрасГМУ, доцент кафедры высшей и прикладной математики, СФУ. E-mail: egoldenok@gmail.com

Черняк Кирилл Евгеньевич — студент 6 курса специальности 30.05.03 Медицинская кибернетика, КрасГМУ. E-mail: blackdragon2000@mail.ru

ИНТЕГРАЦИЯ МАТЕМАТИЧЕСКИХ И ИНТЕЛЛЕКТУАЛЬНЫХ ТЕХНОЛОГИЙ В ЦИФРОВОЙ ПЛАТФОРМЕ ДЛЯ ПОДДЕРЖКИ ПРИНЯТИЯ СТРАТЕГИЧЕСКИХ РЕШЕНИЙ ПО РАЗВИТИЮ ЭНЕРГЕТИКИ РОССИИ

Л. В. Массель, А. Г. Массель

*Институт систем энергетики им. Л.А. Мелентьева СО РАН, г. Иркутск,
Россия*

Предлагается разработка цифровой платформы для поддержки принятия стратегических решений по развитию энергетики России. Разработку предлагается выполнять на основе экосистемы знаний, которая в свою очередь, рассматривается как ядро ИТ-инфраструктуры системных исследований в энергетике, разработанной в ИСЭМ СО РАН под руководством авторов в рамках проекта по госзаказу РФ «Методология построения ИТ-инфраструктуры для разработки интеллектуальных систем управления развитием и функционированием систем энергетики». **Ключевые слова:** *цифровая платформа, экосистема знаний, ИТ-инфраструктура системных исследований в энергетике, интеллектуальные и математические технологии.*

Введение

Актуальность решения обозначенной проблемы обусловлена в первую очередь значимостью энергетики как критической инфраструктуры, устойчивое развитие которой является залогом процветания государства.

К настоящему времени в России сформировалась комплексная иерархическая система управления текущей и перспективной деятельностью топливно-энергетического комплекса (ТЭК) страны и его отраслей: электроэнергетической, угольной, нефтяной, газовой теплоснабжения. Каждая отрасль, в силу определенной специфики, имеет свою отличающуюся систему управления. Система управления включает государственный и корпоративный уровни.

Результаты получены в рамках проекта по госзаданию ИСЭМ СО РАН «Методология построения ИТ-инфраструктуры для разработки интеллектуальных систем управления развитием и функционированием систем энергетики», № госрегистрации АААА-А21-12101209007-7.

В данной системе существуют определенные проблемы. В частности, требует расширения законодательная база энергетики. Кроме того, требуют большего согласования решения, формируемые на верхнем уровне государственного управления и на корпоративном уровне энергокомпаний, которые принимают решения, в первую очередь, исходя из своих коммерческих интересов. Наряду с этим налицо недостаточная скоординированность перспектив развития ТЭК отдельных регионов и стратегических решений по развитию ТЭК страны. В целом, формируется нерациональная структура топливно-энергетических балансов отдельных регионов с соответствующим сдерживанием темпов их развития.

К возникновению указанных проблем фактически привели результаты решений, принимаемых в рамках существующей системы управления ТЭК и ее отраслей (обосновываемых с использованием ныне имеющегося в распоряжении субъектов системы управления инструментария). Очевидно, что для решения и предупреждения возникающих проблем необходимы подготовка и принятие научно обоснованных стратегических решений по устойчивому развитию энергетики, с учетом требований энергетической безопасности, экологических и социально-экономических факторов.

Для выработки таких решений необходимо привлекать как математические методы, так и методы искусственного интеллекта (ИИ) и инновационные методы управления знаниями. В свою очередь, актуальность предлагаемого проекта обусловлена как необходимостью развития методов управления знаниями и методов ИИ, так и необходимостью их интеграции с математическими методами для успешного решения поставленной проблемы.

В связи с развитием тенденции цифровой экономики стали говорить об экономике, основанной на знаниях (knowledge based economy), или проще, но менее корректно – экономике знаний (knowledge economy). В то же время при проектировании цифровой платформы поддержки принятия стратегических решений в энергетике (ЦП ППСР), которая должна объединять информационную, модельно-программную и интеллектуально-аналитическую системы, как правило, основное внимание уделяется первым двум. Авторы предлагают обратить внимание на разработку интеллектуально-аналитической системы и использовать для поддержки принятия решений в исследованиях и управлении развитием энергосистем экосистему знаний в энергетике.

В качестве научного прототипа ЦП ППСР рассматривается разрабатываемая под руководством авторов ИТ-инфраструктура для построения интеллектуальных систем управления развитием и функцио-

нированием систем энергетики (далее для краткости будем называть ее ИТ-инфраструктура системных исследований в энергетике). ИТ-инфраструктура разрабатывается в рамках проекта по госзаказу ИС-ЭМ СО РАН «Методология построения ИТ-инфраструктуры для разработки интеллектуальных систем управления развитием и функционированием систем энергетики» (2021-2025).

1. Основные определения

Экосистема. Термин «биологические экосистемы», как сообщества взаимодействующих организмов, расположенных в определенной географической среде, первоначально был введен британским ботаником Артуром Тэнсли в 1930-х годах [1]. Он отражает непрерывную совместную эволюцию организмов, которые приспосабливаются к внешним изменениям и нарушениям, ощущаемым в окружающей их среде. За рубежом термин «экосистема» в последнее время использовался как метафора и часто фигурировал в литературе по промышленному маркетингу [2, 3].

Цифровая экосистема. По мере развития тенденции цифровой экономики в РФ получил широкое распространение термин «Цифровая экосистема», под которой понимают сеть взаимосвязанных цифровых технологий, платформ и услуг, взаимодействующих друг с другом для создания ценности для бизнеса и потребителей. Также «Цифровую экосистему» определяют как цифровое пространство, в котором «бесшовно» функционирует множество сервисов одной компании или нескольких участников-партнёров. Примеры цифровых экосистем: Apple, Microsoft, Google, Tesla, Amazon, в России: Сбер, Яндекс, МТС, Магнит и другие громкие бренды. Авторами экосистема знаний рассматривается как развитие и специализация цифровой экосистемы [4].

Цифровая платформа. Согласно решению Высшего Евразийского экономического совета от 11.10.2017 № 12 «Об основных направлениях реализации цифровой повестки Евразийского экономического союза до 2025 года», под цифровой платформой (ЦП) понимают систему средств, поддерживающую использование цифровых процессов, ресурсов и сервисов значительным количеством субъектов цифровой экосистемы и обеспечивающую возможность их бесшовного взаимодействия. Иначе говоря, понятие ЦП для экосистемы знаний включает не только компьютерные и математические модели, но и базы данных и знаний, информационную и интеллектуальную подсистемы, подсистему визуализации и необходимые сервисы.

2. Экосистемы знаний

Экосистемы знаний рассматриваются как инновационный подход к управлению знаниями. Идея «Экосистемы знаний» представляет собой подход к управлению знаниями, направленный на развитие взаимодействий между участниками обмена (агентами), упрощение процесса принятия решений и стимулирование инноваций, благодаря эволюции сотрудничества между агентами [5]. Это направление активно развивается за рубежом (например, в [6] выполнен обзор 80 источников). В России проблематика экосистем знаний проработана гораздо хуже, что подтверждается, в том числе, расхождениями в терминологии: например, используются термины «знаниевые экосистемы», «экосистема управления знаниями» [7, 8]. Важность проблемы управления знаниями подчеркивается в работе [9]. В ней отмечается, что существующие подходы к управлению знаниями (УЗ) в российских компаниях носят пока фрагментарный характер, фокусируясь лишь на отдельных аспектах УЗ (например, на информационных технологиях).

Экосистемы знаний определяют, как «организации, состоящие из различных акторов, объединенных совместным поиском ценных знаний, и в то же время обладающих независимой деятельностью за пределами экосистемы знаний» [10]. Одним из ключевых является понятие «артефакты». В контексте экосистемы артефакты относятся к продуктам и услугам, входам и выходам (включая материальные и нематериальные ресурсы), которые совместно создаются всеми участниками как выход на уровне экосистемы [11, 12]. Субъектов, взаимодействующих с экосистемой знаний, можно разделить на две категории: разработчики (участники) и пользователи (бенефициары). Эти две категории не обязательно исключают друг друга, и вкладчики (участники) могут стать членами-бенефициарами и наоборот [13].

Экосистемы знаний отличаются от первоначальной концепции бизнес-экосистемы по нескольким параметрам: 1) в основном географически сгруппированные или локализованные вокруг конкретной «горячей точки», экосистемы знаний вращаются вокруг ключевого актора или «опорного владельца»; 2) компании располагаются в этих конкретных географических горячих точках и вокруг них, чтобы развивать и обмениваться неявными знаниями; краеугольным актором или игроком часто выступает университет или исследовательская организация; 3) основная деятельность экосистемы знаний сосредоточена на разработке и создании общей базы знаний, при этом основное внимание уделяется совместному исследованию, а не использованию знаний [14].

Актуальность решения проблемы построения экосистем знаний в России обусловлена в первую очередь возрастанием в современном мире роли управления знаниями (УЗ), в связи со сложностями идентификации актуальной и ценной для принятия решений информации из-за возрастания ее объемов; быстрой потери актуальности знаний; а также осознания того, что ошибки и игнорирование проблематики УЗ может привести к фатальным последствиям [9]. В научных организациях эта проблема также полностью не осознана, например, невидимые знания ученых и получаемые научные результаты, в частности, в исследованиях энергетики, не трансформируются своевременно в прагматические знания, которые могут быть применены в энергетических организациях, а это невозможно сделать без разработки теоретических основ, методов и средств управления знаниями. Предлагается подход, развиваемый коллективом под руководством авторов, в котором ИТ-инфраструктура системных исследований в энергетике, ядром которой является экосистема знаний, рассматривается как научный прототип ЦП ППСР.

3. ИТ-инфраструктура системных исследований энергетики как основа экосистемы знаний в области энергетики

Отличие экосистем знаний в бизнесе и науке состоит в том, что в первом случае основное – накопление знаний как стратегического ресурса и их обработка для обеспечения конкурентного преимущества. Для экосистемы знаний в науке главное – накопление и обработка знаний для получения новых знаний, в перспективе – применение накопленных знаний для практических целей, с учетом необходимости конкурентных преимуществ. Архитектура ИТ-инфраструктуры включает четыре уровня: 1) математические, семантические и информационные модели, интегрируемые с помощью авторского Языка управления знаниями; 2) Инструментальная среда, включающая: интеллектуальную ИТ-среду семантического моделирования; Хранилище данных и знаний; онтологический портал, объединяющий онтологическое пространство знаний в энергетике; инструментальную среду для построения цифровых двойников энергетических объектов и систем; 3) цифровые двойники (ЦД) агрегатов, объектов и систем и умные ЦД энергетических систем и ТЭК в целом. Уровни построения ЦД соответствуют уровням управления: оперативное, тактическое и стратегическое (уровень умных ЦД). Подробно архитектура ИТ-инфраструктуры представлена в [15]. Экосистема знаний в энергетике на основе онтологического портала рассматривается как ядро ИТ-инфраструктуры.

4. Цифровая платформа поддержки принятия стратегических решений в энергетике (ЦП ППСР)

Цель проекта – разработка цифровой платформы для анализа условий функционирования ТЭК при взаимосвязанной работе энергетических отраслей, с учетом взаимосвязей энергетики и экономики, и поддержки обоснования и принятия управленческих (стратегических) решений в целях обеспечения рационального и надежного топливо- и энергоснабжения потребителей энергоресурсов в перспективе.

Основные задачи, решаемые для достижения цели: 1) построение информационной модели ТЭК, включая разработку онтологических моделей энергетических систем и ТЭК в целом, необходимых моделей данных и структуры взаимосвязанных баз данных, используемых математическими моделями; 2) определение системы необходимых математических моделей, включая модели энергетических отраслей страны, модели ТЭК, модели взаимосвязей энергетики и экономики и их взаимодействия в рамках единого ТЭК страны; 3) определение требований к web-сервисам, реализующим математические модели, и интерфейсам разных категорий пользователей (ЛПР – лиц, принимающих решения); экспертов, обеспечивающих обоснование решений; и лиц, обеспечивающих подготовку информации для экспертов); 4) подготовка информации для принятия управленческих решений в областях: оценки состояния ТЭК (области, региона, федерального округа, страны) как в стабильных условиях, так и в условиях предкризисных и кризисных ситуаций; прогнозирования направлений развития ТЭК и необходимых мероприятий, с учетом требований энергетической безопасности; обоснования решений по модернизации и/или строительству новых объектов ТЭК; 5) адаптация и обеспечение возможности использования интеллектуальных систем (экспертных систем, систем семантического моделирования, систем машинного обучения) для подготовки и обоснования управленческих решений; 6) 3D-визуализация предлагаемых решений, в том числе с помощью интерактивной геоинформационной системы (ГИС).

Заключение

Предложен проект построения цифровой платформы для поддержки принятия стратегических решений (ЦП ППСР) по развитию энергетики. В качестве основы этой ЦП предлагается рассматривать ИТ-инфраструктуру системных исследований в энергетике, ядром которой, в свою очередь, является экосистема знаний, интегрирующая типовые компоненты СППР: информационную систему (включающую ГИС),

модельно-программную систему и интеллектуально-аналитическую систему. Новизна проекта построения ЦП ППСР по развитию энергетики определяется, в том числе, инновационным подходом к управлению знаниями в области энергетики (концепцией экосистемы знаний в энергетике, обеспечивающей интеграцию информационных, интеллектуальных и математических технологий). Имеется существенный задел, выполненный в рамках проекта по госзаказу ИСЭМ СО РАН «Методология построения ИТ-инфраструктуры для разработки интеллектуальных систем управления развитием и функционированием систем энергетики» (2021-2025).

СПИСОК ЛИТЕРАТУРЫ

1. *Kelly E.* Ecosystems Come of Age; Business Trends // Industry Report; Deloitte University Press: London, UK, 2015. P. 1–17
2. *Velu C.* Knowledge Management Capabilities of Lead Firms in Innovation Ecosystems. // AMS Rev. 2015. № 5. P. 123–141
3. *Zhang J. Z., Watson G. F.* Marketing Ecosystem: An Outside-in View for Sustainable Advantage. // Ind. Mark. Manag. 2020. № 88. P. 287–304
4. *Массель Л. В.* Экосистема знаний как развитие и специализация цифровой экосистемы. // Труды Международного научно-технического конгресса «Интеллектуальные системы и информационные технологии – 2023». , 2023. № 2. С. 155–164
5. *David A. Bray.* Knowledge Ecosystems: A Theoretical Lens for Organizations Confronting Hyperturbulent Environments // 2007.
6. <https://www.mdpi.com/2071-1050/12/18/7372>. — Jeandri Robertson. Competition in Knowledge Ecosystems: A Theory Elaboration Approach Using a Case Study. 2020.
7. *Абузярова М. И.*, Знаниевые экосистемы как доминирующий подход формирования новых моделей управления. // Экономика, предпринимательство и право, 2021. № 12. С. 2259–2669.
8. *Шарипов, Ф. Ф.* Ф. Ф., Экосистемы управления знаниями в отраслях промышленности. // Экономика. Налоги, Право, 2020. № 13. С. 54–60.
9. *Гаврилова Т. А., Алсуфьев А. И., Кокоулина Л. О.*, Управление знаниями с российским акцентом: победы и поражения // Инновации. 2017. Т. 219. № 1. С. 59–69.
10. *Bathelt, H.; Cohendet P.*, The Creation of Knowledge: Local Building, Global Accessing and Economic Development // Toward an Agenda. J. Econ. Geogr. 2014. № 14. Pp. 869–882.
11. *Scaringella, L.; Radziwon A.* Innovation, Entrepreneurial, Knowledge, and Business Ecosystems: Old Wine in New Bottles? // Technol. Forecast. Soc. Chang. 2018. № 136. P. 59–87.

12. *Valkokari K.* Business, Innovation, and Knowledge Ecosystems: How They Differ and How to Survive and Thrive within Them. // *Technol. Innov. Manag. Rev.* 2015. № 5. P. 17–24.
13. *Granstrand O., Holgersson M.* Innovation Ecosystems: A Conceptual Review and a New Definition. // *Technovation* 2020. P. 552–560.
14. *Trischler J., Johnson M., Kristensson P.* A Service Ecosystem Perspective on the Diffusion of Sustainability-Oriented User Innovations. // *J. Bus. Res.* 2020. № 116. Pp. 552–560.
15. *Массель Л. В., Массель А. Г.*, Построение экосистемы знаний на основе ИТ-инфраструктуры системных исследований в энергетике // *Вестник Югорского университета.* 2023. № 4. С. 78–87.

Массель Людмила Васильевна — д.т.н, профессор, г.н.с., зав. отделом, отдел систем искусственного интеллекта в энергетике. E-mail: massel@isem.irk.ru

Массель Алексей Геннадьевич — к.т.н., с.н.с., отдел систем искусственного интеллекта в энергетике. E-mail: amassel@isem.irk.ru

ПОДГОТОВКА ДАННЫХ ДАТАСЕТА «VEHICLE INSURANCE DATA 2018» ДЛЯ ПРЕДСКАЗАНИЯ СТРАХОВОЙ ПРЕМИИ

М. И. Морозов

Югорский Государственный Университет, г. Ханты-Мансийск, Россия

Перед подачей исходных данных в алгоритмы машинного обучения и методы построения ансамблей деревьев решений важным процессом является анализ и предобработка данных. В данной статье мы рассматриваем процесс предобработки данных, используя конкретный набор данных Vehicle Insurance Data от страховой организации Эфиопии, содержащий информацию о страховании транспортных средств за 2011 - 2018 год. Набор данных включает в себя от 10 ключевых признаков, которые будут использоваться для предсказания страховой премии на основе методов построения деревьев решений и моделей нейронных сетей. Мы подробно рассматриваем этапы очистки данных, обработки пропущенных значений, кодирования категориальных переменных, нормализации, масштабирования данных, обработки временных данных с целью повышения точности работы будущей модели. **Ключевые слова:** *страхование, предобработка данных, машинное обучение, страховая премия.*

Введение

В эпоху больших данных и машинного обучения качественная подготовка данных играет ключевую роль в успешном решении различных аналитических задач. Одной из таких задач является предсказание страховой премии, что имеет особенное значение для страховых компаний. В данной статье рассматривается процесс подготовки данных датасета «Vehicle Insurance Data 2018» [1] с целью дальнейшего построения модели машинного обучения для предсказания страховой премии.

Предсказание страховой премии является критически важной задачей для страховых компаний, так как от точности этих предсказаний зависит их прибыльность и конкурентоспособность. Правильное предсказание позволяет устанавливать адекватные цены на страховые полисы, обеспечивая баланс между привлекательностью предложений для клиентов и финансовой устойчивостью компании.

1. Обзор датасета

Датасет «Vehicle Insurance Data 2018» содержит информацию о страховых полисах, включающий обезличенные данные клиентов, характеристики транспортных средств (далее ТС) и детали страховых договоров. Этот набор данных представляет собой ценный ресурс для анализа и построения прогнозных моделей, поскольку он охватывает широкий спектр факторов, влияющих на расчет страховой премии. В необработанном виде он содержит 802036 записей за время работы Эфиопской страховой организации с 2011 по 2018. Датасет содержит 12 ключевых признаков, которые предстоит проанализировать в данной работе (см. табл. 1).

Таблица 1

Описание признаков для набора данных «Vehicle Insurance Data 2018»

№	Наименование	Тип	Описание
0	SEX	int64	Пол клиента
1	INSR_TYPE	int64	Тип страхования
2	INSURED_VALUE	float64	Стоимость ТС
3	PREMIUM	float64	Страховая премия
4	PROD_YEAR	float64	Год производства ТС
5	SEATS_NUM	float64	Кол-во сидений в ТС
6	CARRYING_CAPACITY	float64	Грузоподъемность ТС
7	TYPE_VEHICLE	object	Тип ТС
8	CCM_TON	float64	Вес в тоннах для ТС.
9	MAKE	object	Марка ТС
10	USAGE	object	Цель использования ТС
11	CLAIM_PAID	float64	Выплата по убыткам

Статистические данные датасета содержат среднее значение m , медиану M и среднеквадратическое отклонение σ для каждого из признаков датасета.

На основе распределения данных в датасете можно судить о необходимости в следующих этапах работы:

1. Провести маппинг числовых признаков «SEX» и «INSR_TYPE» в соответствии с данными из справочника по страхованию [2].

2. Необходимость перевести некатегориальные данные «object» в числовой тип данных для вторичных признаков «INSR_END», «INSR_BEGIN», «EFFECTIVE_YR».

3. В виду сильной разницы между медианой M и средним значением m , а также высокого среднеквадратического отклонения σ , для каждого и из признаков можно судить о наличии выбросов в данных, соответственно, необходимо провести удаление выбросов для признаков «INSURED_VALUE», «PREMIUM» «SEATS_NUM», «CARRYING_CAPACITY».

2. Очистка Данных

Качество данных напрямую влияет на производительность и точность моделей машинного обучения. Неполные и неточные данные могут привести к созданию моделей, которые делают неверные прогнозы или принимают ошибочные решения. Очистка данных, включает в себя удаление дубликатов, исправление ошибок и заполнение пропусков, что позволяет повысить надежность исходных данных и как следствие улучшить качество модели. Работа с данными проводилась на языке программирования Python с использованием библиотек «pandas» [3] и «numpy» [4] в среде разработки «Jupyter Notebook». Исходный файл с проектом обработки данных можно найти в Github проекте автора статьи [5].

Первым этапом очистки данных является преобразование всех возможных нечисловых значений в числовой формат. Делается это с целью последующей подачи обработанных данных в алгоритмы машинного обучения. В данном случае, было выполнено преобразование строкового значения дат с шаблоном строки «%d-%b-%y» в числовой формат «datetime» для двух признаков «INSR_BEGIN» и «INSR_END». Выполнено преобразование признака «EFFECTIVE_YR» из строкового значения в числовой формат «float64».

Следующим этапом избавляемся от дубликатов и отрицательных значений. Так как мы знаем, что признак «EFFECTIVE_YR» отображает год вступления страхового полиса в силу, мы убираем отрицательные значения для этого признака путем возведения числа в модуль.

В случае обработки пропусков, за пропуск мы считаем значение признака равного «NULL» (в признаке нет данных) или значение признака равного нулю. С целью упростить обработку пропусков, мы удалим «NULL» записи для признаков, в которых количество «NULL» значений меньше 600. Для текущего набора данных такими признаками оказались «EFFECTIVE_YR», «PREMIUM», «PROD_YEAR», «SEATS_NUM», «CCM_TON», «MAKE». Значения для признака «CLAIM_PAID» заменяются на значение равное нулю, так как данный признак отражает выплату по договору страхования, а отсутствие данных означает что выплаты не было.

Для признаков, значения которых равны нулю мы заменяем значения на медианные, ориентируясь на косвенные признаки. Так, для признака «INSURED_VALUE» мы берем медианные значения ориентируясь на признаки «TYPE_VEHICLE», «MAKE», «PROD_YEAR»

Похожие действия проделываются для признаков: «SEATS_NUM», «CCM_TON» и «CARRYING_CAPACITY». Для целевого признака «PREMIUM» содержание нулевых значений в данных не несет для нас никакой информативной ценности, такие значения удаляем.

В качестве метода определения выбросов для каждого числового признака была использована Z-оценка [6] с использованием квантилей Q1 (значение, ниже которого находятся 25% данных) и Q3 (значение, ниже которого находятся 75% данных). На основе этих квантилей вычислялся интервальный размах IQR и определялись нижние и верхние границы для определения и удаления выбросов. Для признаков «INSURED_VALUE» и «PREMIUM» Z-оценка проводилась без привязки к сторонним признакам.

В случае признаков «SEATS_NUM», «CARRYING_CAPACITY», «CCM_TON» удаление выбросов проводилось путем группировки значений по признаку типа автомобиля «TYPE_VEHICLE», так как каждый тип автомобиля (автомобиль, автобус, мотоцикл и т.д.) имеет свои средние значения по количеству мест, грузоподъемности и весу транспортного средства.

После очистки данных было проведено вычисление корреляции между целевым признаком «PREMIUM» и вторичными признаками (см. таблицу 2). Как можно видеть, в большинстве случаев корреляция уменьшается, что в дальнейшем поможет повысить устойчивость и обобщаемость модели машинного обучения.

Таблица 2

Корреляция между признаками до и после удаления выбросов

Коррелируемый признак	Корреляция до удаления выбросов	Корреляция после удаления выбросов
INSURED_VALUE	0,574	0,387
SEATS_NUM	0,269	0,180
CARRYING_CAPACITY	0,067	0,052
CCM_TON	0,352	0,352

3. Преобразование Данных

Различные типы данных могут иметь разные масштабы и единицы измерения. Нормализация и стандартизация данных помогают привести их к единому виду, что облегчает обучение модели. Это особенно важно для алгоритмов, чувствительных к масштабированию признаков, к примеру: градиентный спуск.

В качестве метода кодирования категориальных признаков был выбран метод — One-Hot Encoding [6, с. 205], в силу небольшого количества имеющихся категорий в данных. Как инструмент кодирования информации, был выбран метод «OneHotEncoder» из библиотеки Python «sklearn». Для всех категориальных признаков было проведено One-Hot кодирование, что увеличило размерность признаков с 13 до 563 (размерность столбцов в наборе данных).

Окончательным этапом преобразования текущих данных является нормализация. В качестве метода нормализации данных был выбран «MinMaxScaler» [7, с. 130]. В качестве инструмента нормализации, была выбрана функция «MinMaxScaler» из библиотеки Python «sklearn».

Однако, данное масштабирование выполнено только для обучения одной модели машинного обучения и может не подойти для таких алгоритмов как деревья решений или случайные леса.

4. Визуализация Данных

В качестве инструмента визуализации данных была выбрана библиотека Python «seaborn», содержащая множество методов отображения данных. На рисунке 1 видно распределение данных целевого признака до и после удаления выбросов. Визуальное распределение отображает однородность подготовленных данных, а также отсутствие обработанных выбросов

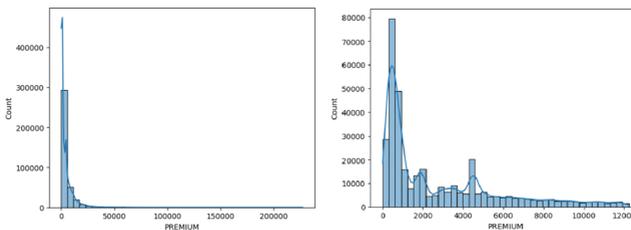


Рис. 1. Распределение целевого признака страховой премии до и после обработки данных

В процессе построения корреляционной матрицы для числовых признаков была выявлена необходимость удаления признаков с корреляцией выше значения 0,75 с целью улучшения обобщенности данных. Из таких признаков были выявлены и удалены: «INSR_BEGIN», «INSR_END», «OBJECT_ID». На рисунке 2 можно видеть матрицу корреляции после удаления лишних признаков.

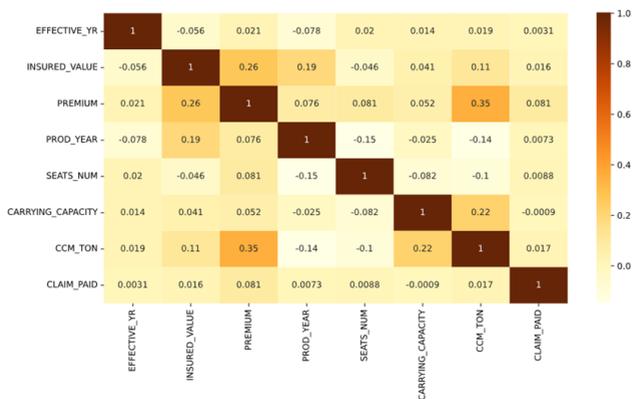


Рис. 2. Матрица корреляции после удаления лишних признаков

Заключение

В процессе данной работы была произведена обработка набора данных «Vehicle Insurance Data 2018». Были выполнены основные этапы анализа и обработки данных: приведение данных к числовому формату, удаление дубликатов, обработка пропусков, удаление выбросов, кодирование категориальных переменных, масштабирование числовых признаков. В ходе анализа данных были выявлены и удалены признаки с высокой корреляцией, что предположительно поможет улучшить обобщаемость модели.

В ходе процесса удаления дубликатов было удалено 293537 дубликатов. В процессе заполнения пропущенных данных был использован метод заполнения медианными значениями, что позволило заполнить пропуски в 243873 экземплярах данных. С помощью метода Z-оценки было произведено удаление 249944 выбросов для всех признаков данных. Визуализация данных до и после обработки наглядно демонстрирует улучшение распределения признаков, что подтверждает эффективность примененных методов.

СПИСОК ЛИТЕРАТУРЫ

1. <https://www.kaggle.com/datasets/imtkaggleteam/vehicle-insurance-data> — Vehicle Insurance Data. 2024.
2. *Terefe E. M.* Tree-Based Machine Learning Methods For Vehicle Insurance Claims Size Prediction // 2023.
3. <https://pandas.pydata.org/docs/> — Pandas documentation. 2024.
4. <https://numpy.org/doc/> — NumPy Documentation. 2024.
5. https://github.com/MM758/insurance_dataset — insurance_dataset. 2024.
6. *Дорофеев В. С., Волосатова Т. М.*, Ансамблирование методов обнаружения выбросов при подготовке обучающей выборки данных // Моделирование, оптимизация и информационные технологии. 2022. Т. 10. № 3(38). С. 22–33.
7. *Muller A., Guido S.* Introduction to Machine Learning with Python: A Guide for Data Scientists. Beijing Boston Farnham Sebastopol Tokyo: O'Reilly Media, 2016. 398 с.

Морозов Михаил Ильич — Аспирант. E-mail: mikhailmorozov99@bk.ru

АДАПТАЦИЯ СИСТЕМ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА ПОД КУЛЬТУРНЫЙ, НРАВСТВЕННЫЙ, ЭТИЧЕСКИЙ И КОГНИТИВНЫЙ ПРОФИЛИ

О. В. Блинова, Ю. В. Таратухина

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

Развитие систем искусственного интеллекта (ИИ) требует тщательного рассмотрения способов обеспечения безопасности и комфорта пользователей. Необходимо понимать, что невозможно подобрать универсальное решение, пригодное во всех странах и социальных группах. Культурные, социальные, этические, религиозные и законодательные нормы действуют в разных сообществах, и их нельзя разделять только по географическому признаку. В данной работе акцентируется необходимость гибкой адаптации систем ИИ под различные сообщества, рассмотрены возможные проблемы и пути решения данного вопроса. Разработка различных профилей взаимодействия ИИ с пользователями является междисциплинарной проблемой, требующей совместной работы как технических специальностей, так и в области лингвистики, этики, религии, юриспруденции и других областей. **Ключевые слова:** *этика искусственного интеллекта, большие языковые модели, культурный профиль искусственного интеллекта, локальная адаптация интеллектуальных систем*

Введение

На сегодняшний день и в перспективе ближайшего будущего можно утверждать, что искусственный интеллект (ИИ) становится полноценным «партнером» человека в деловой и личной коммуникациях. Он используется и применяется во многих сферах: образовательной, медицинской, военной, научной, творческой, в сфере первичного отбора персонала и мониторинга, в сфере психологической диагностики, в сфере отбора и адаптации персонала в международных компаниях, направленных на то, как согласно не только твердым навыкам и компетенциям, но и индивидуально-психологическому и культурному профилю

Исследование выполнено в рамках научной программы Национального центра физики и математики, направление № 9 «Искусственный интеллект и большие данные в технических, промышленных, природных и социальных системах

подобрать кандидату наиболее релевантную вакансию. В психологии и психотерапии на сегодняшний день получает распространение использование электронных помощников в психотерапевтических практиках, практиках улучшения навыков коммуникации и стратегии совладания со стрессом (копинге), индивидуальном коучинге. Более того, ИИ, который научится разбираться в кросс-культурной психологии и специфике кросс-культурных коммуникаций, сможет осуществлять помощь и поддержку в адаптации сотрудников из разных стран и улучшении коммуникаций в кросс-культурных командах. С помощью различного рода приложений, онлайн ассимиляторов и рекомендательных сервисов можно ускорить процессы адаптации к «новым когнитивным реальностям». В ближайшем будущем множество рутинных трудовых функций будет передано искусственному интеллекту и роботам и человек вынужден будет с ними взаимодействовать, то возникнет вопрос выстраивания эффективного взаимодействия человеческого и искусственного интеллектов. Крайне важно обеспечить возможность комфортного и безопасного взаимодействия, взаимопонимание и формирование положительных впечатлений от общения с ИИ, а для этого важно обеспечить гибкость и возможность индивидуальной настройки коммуникации.

1. Возможность обучения ИИ «пониманию» культурных норм

Как человеческий, так и искусственный интеллект развивается и проявляется благодаря обучению. С точки зрения этики, безусловно, должен встать вопрос о том, что искусственный интеллект может восприниматься как полноценный партнер по коммуникации. Отсюда встают вопросы о юридическом статусе искусственного интеллекта и формировании этических норм в данной коммуникации.

Очевидно, что есть необходимость интегрирования человеческих культурных, социальных, религиозных и т.д. норм в системы ИИ. Задача эта крайне сложная, так как некоторые области могут быть крайне острыми и чувствительными для пользователя. Первое, что необходимо делать – это встраивать некоторую систему правил, инструкций и ограничений в работу ИИ, для обеспечения нейтрально-доброжелательного общения, базового уважения к собеседнику независимо от его убеждений. Эта система правил должна разрабатываться и курироваться учеными-юристами, специалистами по этике, психологии, религии, социологии и т.д. и создаваться с учетом возможностей и ограничений технических способов реализации.

Существует также обширный ряд инженерных проблем, особенно с учетом того, что поведение многих интеллектуальных систем «непро-

зрачно» и с трудом поддается корректировке, поиск и исправление нежелательных эффектов может быть затруднено.

Следует повышать информированность и грамотность разработчиков и пользователей о необходимости учета основных культурно-языковых и этических профилях с целью обучения ИИ настройкам эффективной и этической коммуникации (развивать «культурный интеллект ИИ»), развивать системы модерации и контроля действий пользователя для избегания нарушений закона или разжигания конфликтов.

Культурный интеллект ИИ – способность «понимать» и распознавать каким образом мыслят представители разных культурных групп, предсказывать их поведение в целях выстраивания эффективных коммуникаций. Безусловно, при воспитании ИИ важен как этический (общий, универсальный), так и эмический (культурно-специфичный) аспект.

Основной задачей разработчиков в контексте этик-подхода является «научить» ИИ общаться исходя из общечеловеческих ценностей и морально-этических принципов. То есть, в некотором смысле, быть образцом поведения, даже если человек в коммуникации допускает грубость и не совсем этичное поведение. Для примера возьмем Алису. Она «научилась» достаточно этично вести себя с разными пользователями и весьма корректно реагировать на не совсем этичные реплики. Однако в кросс-культурных коммуникациях для достижения уровня этического общения требуются дополнительные компетенции разработчиков, учет различных культурных профилей, где помимо общей безопасности систем ИИ с точки зрения потенциально возможного вреда для человека и универсальных этических норм будут учитываться еще и культурные параметры.

Одним из возможных решений этой проблемы является проведение в некоторых системах мини-тестирования на основе принадлежности к определенному культурному профилю. Например, система может спросить о культурном происхождении человека и его личных убеждениях для наиболее релевантной коммуникации в дальнейшем. Здесь крайне важно обеспечить безопасность данных о пользователе и избежать возможности утечки «досье» на пользователя, содержащего не только персональные данные, но и личностные особенности и убеждения.

2. Способы реализации гибкой настройки систем ИИ под особенности пользователя

Выявление и учет дифференцированных культурно-когнитивных параметров пользователей не является новой задачей, например, уже довольно давно существуют исследования и работы об адаптации интер-

фейсов, учебных материалов и т.д. под различные культурные особенности. Например, могут быть использованы инструменты учета влияния культурных особенностей на специфику коммуникации, основу которых составляют подходы Г. Хофстеде, Э. Холла [1], Р Нисбетта, М. Холодной. [2]. Г. Триандиса, [3]. Н. Лебедевой [5]. и др. Опираясь на вышеперечисленные концепции, можно выделить ряд параметров, которые лежат в основе анализа культурной специфики поведения, мышления и деятельности и которые будут обуславливать собственно специфику культурно-когнитивного профиля личности, коммуникации в целом и этической коммуникации в частности: восприятие времени, пространства и как следствие, специфика деятельности; специфика представления информации, специфика мышления и внимания; специфика социальных коммуникаций; доминирующие ценности. [4].

В любой коммуникации можно условно выделить когнитивный и операционный компоненты. Когнитивный компонент будет включать в себя специфику когнитивного стиля, специфику принятия решений и работы с информацией, специфику распределения внимания, специфику использования творческих подходов (креативности).

Операционный компонент – отношение ко времени, к инструкциям, к целеполаганию, место в социуме, отношение к внешней среде, к командной работе. Также будет значимым эмоциональный компонент, как некий контекст, детерминирующий когнитивный и операционный компоненты (контекстуальные параметры).

Используя данные инструменты можно построить модель культурно-когнитивного профиля личности и, при необходимости, группы людей. Далее, на основе выявленных личностных признаков возможно создать несколько профилей ИИ, соответствующих основным культурным группам, между которыми ИИ сможет переключаться с появлением информации о пользователе, с которым ведется взаимодействие. Набор этих профилей и способы выбора между ними зависят от сферы применения системы ИИ, масштаба, особенностей конкретной системы и решаемых с ее помощью задач.

Таким образом, необходимо учитывать культуру, в которой формируется индивид как когнитивную и семиотическую системы. Любая культура представляет собой определенный набор установок (так называемых культурных фреймов или схем, где четко описано ценности, нормы, табу, включающие в себя отношение к возрасту (эйджизм) и полу (гендерные нормы), отношение к статусу (имеет значение или «все равны»), расовые стереотипы и предрассудки, отношение ко времени

коммуникации, способы обработки информации и принятия решений и т.п.

В идеале, речь идет о локализации работы ИИ в зависимости от принадлежности индивида к культурной группе. То есть речь идет о выработке общечеловеческих ценностей, которые впоследствии будут интегрированы в ядро интеллектуальных информационных систем (ИИС) и будут определять «моральный облик ИИ». Речь идет о том, чтобы научить в итоге ИИС различать этические и неэтические действия. То есть помимо общечеловеческого этического и морального кодекса имеет смысл добавить еще и культурно-специфический (по основным культурным группам) свод этических правил и норм.

Можно критически относиться к этическому и ценностному кодексу других стран и культур, однако, необходимо понимать их генезис и причины появления. Здесь не должно в конечном счете быть дихотомии между универсализмом и культурной специфичностью. Поскольку языковые модели обучаются на большом объеме данных, то следует предположить, что в зависимости от культуры и страны в ИИ будет так или иначе воплощаться культура (языковая, нравственная и этическая) его создателей. Человечество в любом случае будет сталкиваться с кросс-культурными различиями и до конца данные барьеры непреодолимы.

Научить ИИ этическому общению в рамках этнокультурных и религиозных особенностей возможно, задав ему основные алгоритмы и скрипты коммуникации, характерные для разных культур: базовые ценности, специфика дискурса, табу, религиозные табу, стереотипы и предубеждения и т.п.

3. Возможность цифровой эмпатии и более гибкой адаптации к индивидуальным особенностям

Очень интересным в данном контексте будет выглядеть вопрос о «цифровой эмпатии» и цифровом эмоциональном интеллекте. Поскольку в настоящее время ИИ используется в разных сферах, то его необходимо учить не только общеэтическим нормам, нормам коммуникации в кросс-культурном контексте, межпрофессиональной коммуникации, но и учету гораздо большему количеству личностных особенностей: возраста, опыт, специальности, типа темперамента, особенностей социализации, психологическим и интеллектуальным факторам.

Здесь же может быть реализована возможность инклюзии, учета врожденных ограничений или вызванных заболеваниями. Опыт составления портрета личности и настройки способов коммуникации в зависимости от этого уже есть у психологов, психиатров, специалистов спецслужб и т.д. Но вот возможность учета этих параметров ИИ, с обеспече-

ние безопасности пользователя, недопущению влияния на его мнение и давления, соблюдения принципов «не навреди» - задача сложная, требующая дальнейших исследований и экспериментов.

Заключение

Системы ИИ развиваются стремительно и очевидно, что степень их участия в жизни обычных людей будет увеличиваться. И для обеспечения успешной и положительно окрашенной коммуникации необходимо обучать и ИИ способам коммуникации с людьми и у людей формировать понимание основных принципов и основ функционирования технических интеллектуальных систем и их отличий от живого общения с людьми.

Современное общество сейчас движется от идеи общего «выравнивания» и усреднения к уважению к различиям между людьми и культурами, возможности гибкой адаптации среды под человека для максимального раскрытия его способностей и талантов. И системы искусственного интеллекта должны быть гибкими, адаптивными, эргатическими, индивидуально настраиваемыми.

Создание (проектирование) релевантных этнокультурных профилей ИИ позволит существенно продвинуться в решении проблемы этической коммуникации с учетом разницы культурных и личностных особенностей, опыт создания таких профилей поможет разработке и внедрению других адаптивных механизмов с упором на другие личностные особенности.

Это сложная многоэтапная работа, но актуальная и целесообразная, так как позволит избежать большого количества проблем, а, в некоторых случаях, и судебных исков.

СПИСОК ЛИТЕРАТУРЫ

1. *Hall E. T.* Understanding Cultural Differences: Germans, French and Americans. // BIntercultural Press. 1990.
2. *Холодная М. А.* Психология интеллекта. Парадоксы исследования 3-е изд., пер. и доп. Учебное пособие для бакалавриата и магистратуры // Litres, 2018, 392 с.
3. *Triandis H. S.* Culture and Social Behavior. //Mcgraw-Hill Book Company. 2015.
4. *Таратухина Ю. В.; Цыганова Л. А.; Ткаленко Д. Э.* Межкультурные коммуникации в информационном обществе. // Изд. дом НИУ ВШЭ, 2019 . 23 с.
5. *Лебедева Н. М.; Бульцева М. А.* Взаимосвязь опыта межкультурных взаимодействий, аккультурационных ожиданий и креативности у российских

студентов. // Культурно-историческая психология. 2019. Т. 15. № 3. С. 51-59

Блинова Ольга Викторовна — н.с., лаб № 17. E-mail: *blinova_olga_v@mail.ru*

Таратухина Юлия Валерьевна — к.ф.н., с.н.с., лаб № 17. E-mail: *jvt6@ipu.ru*

ПРАКТИКА ПРИМЕНЕНИЯ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА ДЛЯ ДИАГНОСТИКИ СОСТОЯНИЯ ИЗОЛЯЦИИ СИЛОВОГО ТРАНСФОРМАТОРА ПРИ ОГРАНИЧЕННОМ НАБОРЕ ДАННЫХ

С. Ю. Петрова, А. Ю. Швец

Севастопольский государственный университет, г. Севастополь, Россия

Одним из ключевых препятствий на пути успешного применения искусственного интеллекта в задаче прогнозирования дефектов и предаварийного состояния электросетевого оборудования является недостаточное количество данных, характеризующих такого рода состояния. Исследование посвящено оценке возможности применения моделей машинного обучения k-NN, SVM, XGBoost, RandomForest, Prototypical Networks и Matching Networks к задаче классификации аномальных состояний силового трансформатора. Работа проводилась с ограниченным количеством исходных маркированных данных, а именно, для обучения модели были использованы 27 образцов, описывающих аномальное технологическое состояние трансформаторов по 86 параметрам. Результаты сравнения показали, что Prototypical Networks демонстрируют более сбалансированное и устойчивое поведение на ограниченном обучающем наборе данных. Успешное выполнение экспериментов продемонстрировало практическую применимость методов One-Shot и возможность их использования в реальных промышленных сценариях.

Ключевые слова: *машинное обучение, неразрушающий контроль оборудования, ограниченный набор исходных данных.*

Введение

Силовой трансформатор относится к ключевому элементу электроэнергетических систем, так как это дорогостоящее, сложное в ремонте и обслуживании оборудование, отказоустойчивость которого непосредственно влияет на надежность поставки электроэнергии. Состояние изо-

Работа поддержана Российским научным фондом в рамках программы: «Проведение фундаментальных научных исследований и поисковых научных исследований малыми индивидуальными научными группами» (региональный конкурс) № 24-21-2001 «Искусственный интеллект в диагностике неисправностей и оценке состояния электросетевого оборудования»...

ляции силового трансформатора является важным параметром, от него зависит надежность и безопасность эксплуатации.

Важно отметить, что традиционные методы диагностики трансформаторов производят ограниченное количество данных, характеризующих аномальные состояния оборудования, которые можно было бы использовать в промышленных интеллектуальных агентах. Это является несомненной проблемой, так как агенты искусственного интеллекта, способные принимать решения с необходимой точностью, строятся на алгоритмах машинного обучения, для обучения которых требуются тысячи экземпляров дефектных состояний, маркированных экспертами.

Таким образом, возникает необходимость оценить существующие методы машинного обучения и выработать рекомендации по их применению для разработки агентов искусственного интеллекта, диагностирующих состояние изоляции силового трансформатора при ограниченном наборе исходных данных.

1. Описание исходных данных

Обычно, маркировка данных проводится вручную и для сложных предметных областей к этой работе привлекаются эксперты, обладающие обширными знаниями соответствующей области. Такой процесс аннотирования данных является дорогостоящим, занимает много времени и чреват ошибками.

Исследование проводилось на ограниченном объёме исходных маркированных данных, а именно, для обучения моделей использовалось 27 экземпляров, описывающих технологическое состояние силовых трансформаторов с помощью 86 параметров. Данные были получены в результате измерения электрических параметров изоляционных промежутков, измерения сопротивлений постоянному току, потерь холостого хода, измерения физико-химических показателей трансформаторного масла, а также обследования и визуального осмотра конструктивных элементов силовых масляных трансформаторов классом напряжения 10/0,4 кВ (см. табл. 1) У данных экземпляров был выявлен 31 тип дефектных состояний (см. табл. 2).

Анализ данных показал, что помимо ограниченности выборки, необходимо решать такие проблемы, как:

- Наличие пропусков в данных.
- Дисбаланс классов, а именно, существенное различие в количестве данных, описывающих аномальные состояния, и данных, описывающих нормальную работу трансформатора.
- Данные разного масштаба и типа.

Таблица 1

Список параметров технологического состояния трансформатора

Параметр	ТР1	ТР2
<i>Внешний осмотр вводов</i>		
Неудовлетворительное состояние уплотнений	2	2
<i>Измерение электрических параметров изоляционных промежутков</i>		
Сопротивление изоляции ВН-К, МОм	735	336
Тангенс угла диэлектрических потерь ВН-К, %	0.85	1.15
<i>Измерение сопротивлений постоянному току</i>		
ΔR_{\min} ВН (по всем ответвл.), %	0.4	19.5
Сопротивление обмотки НН фазы а, Ом	0.139	0.003
Сопротивление обмотки НН фазы с, Ом	0.242	0.034
ΔR_{\max} НН, %	109	1244
<i>Измерение потерь ХХ</i>		
Потери трехфазным методом, Вт	1097	1378
<i>Измерение физико-химических показателей трансформаторного масла</i>		
Пробивное напряжение масла, кВ	52.6	47.2
Влагосодержание масла, г/т	17	27
Кислотное число мг КОН/г	0.02	0.02
Температура вспышки, градусов	140	130
Тангенс угла диэлектрических потерь при 90°	4.13	3.29

Все вышеперечисленные проблемы были устранены посредством тщательной предобработки данных, и внесением дополнительных функций в реализуемые алгоритмы машинного обучения.

2. Экспериментальная часть

В работе исследовались алгоритмы способные классифицировать состояние изоляции силового трансформатора. К таким алгоритмам относятся, например, *SVM* [1, 2], *RandomForest* [3], *XGBoost* [5] и *k-NN* [6]. Также была рассмотрена возможность использования более сложных моделей, решающих проблему малого количества исходных данных, а именно разновидность трансферного обучения, алгоритмы класса One-Shot Learning: *Prototypical Networks* [7] и *Matching Networks* [8]. Оба алгоритма позволяют модели обучаться распознавать новые классы на основе всего одного обучающего примера данных.

Таблица 2

Сокращенный список дефектов элементов силового трансформатора

Компонент	Наименование дефекта
Магнитопровод	Повреждение
Обмотки	Увлажнение изоляции
Масло	Низкий уровень масла
Масло	Увлажнение
Масло	Загрязнение
Вводы	Повреждение уплотнений
Вводы	Повреждение внешних контактов
ПБН	Повреждение контактов
ПБН	Отсутствие неповрежденных ответвлений
Радиаторы	Неудовлетворительное состояние поверхности
...	...

Цель экспериментов – определить смогут ли данные модели, обученные на ограниченном наборе маркированных экземпляров, эффективно прогнозировать состояние изоляции силового трансформатора.

Оценка точности прогнозирования модели выполнялась с помощью с помощью базовых метрик. Наиболее информативной метрикой для решаемой задачи была *Weighted Average*. Данная метрика хорошо работает с несбалансированными классами, так как учитывает различную степень важности чисел в наборе данных и дает адекватное представление о производительности модели. Метрика рассчитывается по формулам (1) и (2).

$$WeightedF1 = \frac{\sum F1_i W_i}{\sum W_i}, \quad (1)$$

$$W_i = \frac{n - i}{N}. \quad (2)$$

где N – общее количество экземпляров; n – количество экземпляров в классе. Алгоритм расчёта следующий: для каждого класса состояний создается задача бинарной классификации, в которой рассматриваемый класс обозначается как положительный класс, а остальные классы – как единый отрицательный класс. Данные по оценке качества и производительности каждой модели, сведены в таблицу (см. табл. 3).

Таблица 3

Результаты оценки точности моделей

XGBoost				
(Training Accuracy: 80.0%, Model Accuracy Score: 20.00%)				
	precision	recall	f1-score	support
Micro avg	0.95	0.97	0.96	63
Macro avg	0.46	0.48	0.47	63
Weighted avg	0.94	0.97	0.95	63
Samples avg	0.95	0.97	0.96	63
RandomForest				
(Training Accuracy: 100.0%, Model Accuracy Score: 16.67%)				
	precision	recall	f1-score	support
Micro avg	0.91	0.96	0.93	74
Macro avg	0.44	0.47	0.45	74
Weighted avg	0.91	0.96	0.93	74
Samples avg	0.91	0.96	0.93	74
k-NN				
(Training Accuracy: 20.0%, Model Accuracy Score: 0.0%)				
	precision	recall	f1-score	support
Micro avg	0.84	0.80	0.82	61
Macro avg	0.36	0.35	0.35	61
Weighted avg	0.81	0.80	0.79	61
Samples avg	0.85	0.80	0.82	61
Prototypical Networks				
(Training Accuracy: 86,7%, Model Accuracy Score: 65,3%, Loss: 0,35%)				
	precision	recall	f1-score	support
Macro avg	0,83	0,83	0,83	2500
Weighted avg	0,83	0,83	0,83	2500
Matching Networks				
(Training Accuracy: 92,5%, Model Accuracy Score: 39,4%, Loss:0,46%)				
	precision	recall	f1-score	support
Macro avg	0,70	0,71	0,70	7500
Weighted avg	0,70	0,70	0,70	7500

3. Обсуждение результатов

Модель, обученная с помощью алгоритма RandomForest сильно переобучена, так как точность на обучающей выборке составляет 100%, в

то время как точность на тестовой выборке очень низкая 16.67%. Это указывает на то, что модель плохо обобщает на новых данных.

Модель, обученная с помощью метода k-NN, не справляется с задачей классификации на тестовой выборке (0.0% точности). Это может быть связано с тем, что алгоритм k-NN требует тщательного подбора параметров и возможно, большего объема данных для адекватной работы.

XGBoost имеет самые высокие значения micro avg (precision 0.95, recall 0.97, f1-score 0.96) и weighted avg (precision 0.94, recall 0.97, f1-score 0.95), что указывает на его лучшую производительность для часто встречающихся классов по сравнению с другими алгоритмами. XGBoost также показал немного лучшие результаты по macro avg (precision 0.46, recall 0.48, f1-score 0.47) по сравнению с RandomForest, что указывает на лучшую работу с редкими классами.

Алгоритм классификации Matching Networks продемонстрировал высокую способность к запоминанию обучающего набора данных, что подтверждается высоким значением Training Accuracy (92.5%). Однако, при оценке модели на новых данных, Model Accuracy Score оказалась низкой (39%). Это свидетельствует о недостаточной обобщающей способности модели, что, вероятно, связано со способностью запоминать даже шум в данных, который может приводить к недооценке обобщающей способности модели при оценке на новых данных. Это особенно заметно в условиях ограниченного и несбалансированного набора данных, а также при наличии различий в масштабе признаков. Для улучшения производительности Matching Networks необходимо проводить более тщательное управление переобучением, возможно, с использованием регуляризации или аугментации данных.

Модель Prototypical Networks демонстрирует более сбалансированное и устойчивое поведение на данных. И хотя Training Accuracy у Prototypical Network (86,7%) несколько ниже чем у Matching Networks (92,5%), модель Prototypical Network показывает значительно лучшее обобщение на новых данных, так как Model Accuracy Score у Prototypical Network 65,3% значительно выше 39,4% у Matching Networks, что указывает на меньшее переобучение Prototypical Network.

Заключение

Таким образом, эксперимент подтвердил жизнеспособность использования искусственного интеллекта для диагностики состояния изоляции силовых трансформаторов. Исследование может быть полезным для специалистов в области электроэнергетики, которые занимаются проектированием, эксплуатацией и ремонтом силовых трансформато-

ров. Результаты исследования могут быть использованы для оптимизации процессов обслуживания и ремонта трансформаторов, что в свою очередь позволит снизить затраты на их эксплуатацию и ремонт.

Несмотря на удовлетворительные результаты обучения, данная тематика имеет значительный потенциал для дальнейшего развития.

СПИСОК ЛИТЕРАТУРЫ

1. *Yang S., Xin H., Yang B.* Implicit Regularization in Over-Parameterized Support Vector Machine // arXiv:2310.17124v1 [math.ST], 2023.
2. *Haoxing L., Renjie X., Haoshen Q., Yizhou L.* Research on Dangerous Flight Weather Prediction based on Machine Learning // arXiv:2406.12298 [math.ST], 2024.
3. *Rivera-Lopez R., Canul-Reich J., Mezura-Montes E., Cruz-Chávez M.A.* Induction of decision trees as classification models through metaheuristics // Swarm Evol Comput, 2022.
4. *Zhao J., Liu X., Yan Q., Li B., Shao M., Peng H., Sun L.* Automatically predicting cyber attack preference with attributed heterogeneous attention networks and transductive learning // Comput Secur., 2021.
5. *Holloway E., Marks R.* High Dimensional Human Guided Machine Learning // arXiv:1609.00904 [cs.AI], 2016.
6. *Shekkizhar S., Ortega A.* Revisiting local neighborhood methods in machine learning // IEEE data science and learning workshop (DSLW), 2021. p. 1–6.
7. *Snell J., Swersky K., Zemel R.* Prototypical Networks for Few-shot Learning // arXiv:1703.05175v2 [cs.LG], 2017.
8. *Vinyals O., Blundell C., Lillicrap T., Kavukcuoglu K., Wierstra D.* Matching Networks for One Shot Learning // arXiv:1606.04080 [cs.LG], 2016.

Петрова Светлана Юрьевна — к.т.н., доцент, заведующий кафедрой Интеллектуальные сети энергоснабжения, Институт Ядерной энергии и промышленности, Севастопольский государственный университет. E-mail: sypetrova@sevsu.ru

Швец Антон Юрьевич — научный сотрудник, кафедра Интеллектуальные сети энергоснабжения, институт Ядерной энергии и промышленности, Севастопольский государственный университет. E-mail: shvetsanton01@gmail.com

О ПЕРЕОЦЕНКЕ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В ГОСУДАРСТВЕННОМ РЕГУЛИРОВАНИИ ПОДГОТОВКИ КАДРОВ БУХГАЛТЕРОВ В СОВРЕМЕННОЙ РОССИИ

А. Ложникова, Н. Балясова, Н. Кичко

*Национальный исследовательский Томский государственный университет,
г. Томск, Россия*

Авторы критикуют распространённый в России подход, прежде всего в правительственных кругах и в сфере подготовки кадров высшей квалификации и высшего образования, переоценивающий роль искусственного интеллекта в современной бухгалтерской деятельности, и недооценивающий труд специалиста-бухгалтера с высшим образованием и/или с ученой степенью. Для обоснования ценности профессионального бухгалтерского суждения в сфере бухгалтерского управленческого учета приводится ряд примеров, в том числе из лучших зарубежных практик, о необходимости обоснования специалистом-бухгалтером непростого выбора из ряда альтернативных решений. Фокусируется внимание на сохранении западными аудиторскими компаниями своей ниши на российском рынке аудиторских услуг на фоне целенаправленного снижения государством уровня подготовки отечественных специалистов-бухгалтеров.

Ключевые слова: *бухгалтерский управленческий учёт, бухгалтеры-специалисты с высшим образованием (бакалавры, магистры), бухгалтеры-кадры высшей квалификации (кандидаты и доктора наук по научной специальности 08.00.12 «Бухгалтерский учет, статистика»), искусственный интеллект (ИИ), западные аудиторские компании.*

Введение

В 2020 году министр финансов РФ А. Силуанов назвал бухгалтерскую профессию вымирающей ввиду применения новых технологий и искусственного интеллекта. Добавим к этому исключение с 2021 года научной специальности 08.00.12 «Бухгалтерский учет, статистика» из перечня специальностей подготовки кадров высшей квалификации, по которым присуждаются ученые степени. Таким образом, в последние годы в России, прежде всего в правительственных кругах и в сфере подготовки кадров высшей квалификации и высшего образования ак-

тивно распространяется мнение о том, что искусственный интеллект (далее – ИИ) в самом ближайшем будущем вытеснит специалистов по бухгалтерскому учёту с рынка труда.

Посмотрим на ситуацию с другой стороны. Современный бухгалтерский управленческий и финансовый учёт в силу своей высокой значимости для задач управления организацией/бизнесом определяется во многом как деятельность специалистов с высшим образованием (бакалавры, магистры) и высшей квалификации (с учёной степенью по прежней научной специальности «Бухгалтерский учёт, статистика»). Деятельность в сфере бухгалтерского учёта нацелена на создание и сопровождение информационной системы, которая не просто измеряет, но обрабатывает и передает информацию о хозяйствующем субъекте разного рода пользователям. Ошибочно думать, что бухгалтеры компаний заняты только сбором и несложной обработкой информации. В сферу деятельности бухгалтерского управленческого учёта включаются все сложные процессы: прогнозирования, планирования, выполнения показателей и контроля. Сам современный термин «бухгалтерский управленческий учёт» как нельзя лучше отражает его важнейшую роль в управлении компанией и диктует высокий уровень требований к квалификации занятых здесь сотрудников.

Кроме того, бухгалтеры выполняют критично важную для коллективов работников коммуникативную функцию. Немалая часть рабочего времени бухгалтеров тратится на непосредственное общение с сотрудниками компаний и с их клиентами, на передачу, пояснения и интерпретацию бухгалтерских показателей разным пользователям, чтобы предоставляемая им информация была максимально доступна для понимания.

Подчеркнём, что в учётной политике современных компаний очень много развилочек, что требует учёта факторов разной природы, технологической, технической, финансовой и социальной. Так, ярким примером в контексте текущей повестки обеспечения технологического суверенитета РФ является развилка для технологических компаний – когда именно списать расходы на НИОКР и сформировать, таким образом, прибыль или убыток в текущем или будущем отчетном периоде. Другим примером развилки являются вопросы распределения прибыли. Например, на годовом собрании-2021 акционеров Тесла Илон Маск заявил, что компания никогда не платила, не платит и не собирается платить дивиденды, поскольку лучше остальных понимает как развиваться дальше и как финансировать это развитие [1]. Новостью прошлого года стало то, что компания Alphabet – владелец Google впервые в

истории выплатит дивиденды. «Объявленные компанией Alphabet выплаты дивидендов и обратный выкуп в дополнение к стабильным показателям прибыли — это не только поток свежего воздуха для технологического рынка в целом, но и очень разумная стратегия для гиганта поисковых систем, переживающего трудное время года» [2]. Именно в поле зрения специалистов-бухгалтеров находится структура формирования обобщающего финансового результата компаний. Например, чистая прибыль уже упомянутой выше американской Тесла определяется регулятивными ZEV-кредитами в большей степени, чем эффективностью операционной деятельности. Этот факт подтолкнул компанию к принятию сложного решения о снижении цен на производимую продукцию. Анализ структуры формирования чистой прибыли российских экспортёров показывает большую значимость курсовой ренты, чем прибыли от продаж несмотря на многолетнюю недооцененность российского рубля по зарубежным оценкам [3]. Продолжая рассуждения о выполнении бухгалтерами коммуникативной функции в коллективе работников, отметим важную роль специалистов бухгалтерского управленческого учета в достижении взаимопонимания между администрацией и владельцами компании. Всё вышеперечисленное позволяет определить, к примеру, функцию управления в части ценообразования как искусство. Под силу ли данное искусство искусственному интеллекту? Сомневаемся в этом, поскольку большая часть данных, которыми оперируют специалисты в области бухгалтерского управленческого учёта составляет коммерческую тайну и не обнаружится ни в каком виде. Это относится к размерам надбавок розничной торговли и оптового посредника, нормам прибыли производителей, оценкам текущих и капитальных затрат, а также к большинству показателей в физических единицах измерения.

Заключение

В контрасте с российскими подходами государственного регулирования подготовки кадров бухгалтеров в современной России принципиальной особенностью международных стандартов бухгалтерского учета является поступательное расширение функций и ответственности бухгалтеров. Как следствие этого мы наблюдаем прочное закрепление западных аудиторских компаний на российском рынке аудиторских услуг, прежде всего для отечественных крупнейших компаний. Сегодня, в 2024 году, на фоне рекордного количества санкций в отношении нашей страны особенно отчетливо заметно отсутствие запрета на оказание аудиторских услуг западными компаниями российским. Например, «бывшее российское подразделение Ernst & Young (EY) сменило название на ООО «Центр аудиторских технологий и решений – аудиторские услу-

ги». Это первые наследники «большой четверки», сменившие название» [4]. Мы обратились к официальному ресурсу бухгалтерской финансовой отчётности БФО [5], который публикует аудиторские заключения российских компаний, в том числе самых крупнейших (ПАО Лукойл, ПАО НК Роснефть, ПАО Сбер) и отмечаем сохранение позиций прежних западных аудиторов в новейшей истории.

СПИСОК ЛИТЕРАТУРЫ

1. Илон Маск на ежегодном собрании акционеров tesla 2021 [HTTPS://YANDEX.RU/VIDEO/PREVIEW/9665919122782219311](https://yandex.ru/video/preview/9665919122782219311)
2. Компания – владелец Google впервые в истории выплатит дивиденды | РБК Инвестиции <https://quote.rbc.ru/news/article/662b5a859a7947cd2b0813e9>
3. Our Big Mac index shows how burger prices differ across borders <https://www.economist.com/big-mac-index>
4. Российское подразделение EY первым из аудиторов «большой четверки» сменило название – Коммерсантъ <https://www.kommersant.ru/doc/5305491>
5. Ресурс БФО <https://bo.nalog.ru/search?query=ПАО+Сбербанк&page=1>

Ложникова Анна Владимировна — доктор экономических наук, профессор кафедры экономики Национального исследовательского Томского государственного университета. E-mail: tfg@mail.ru

Балясова Нина Юрьевна — независимый исследователь. E-mail: nn0503@mail.ru

Кичко Наталья Ивановна — владелец и директор аудиторской компании ООО «СТКМ-аудит». E-mail: nn7202@gmail.com

О ВОЗМОЖНОСТЯХ МЕТОДА ВЫДЕЛЕНИЯ СТРУКТУР ПРИ ОБРАБОТКЕ ДАННЫХ ПРОИЗВОЛЬНОЙ ПРИРОДЫ

С. Г. Катаев¹, С. С. Катаева²

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

² *Национальный исследовательский Томский государственный университет, г. Томск, Россия*

Обсуждаются возможности применения алгоритмов метода выделения структур при обработке эмпирических данных различной природы. Этот метод может быть отнесен к кластерным методам, и он позволяет отыскивать в массивах многомерных данных латентные структуры, фактически осуществляя процедуру классификации объектов. Постановка и решение большого количества практических задач очень близка к задачам именно классификации: распределить все множество объектов, характеризующихся общей системой признаков, на небольшое число классов, руководствуясь поставленной целью и выбранными критериями. Рассмотрены примеры решаемых задач.

Ключевые слова: *кластерный анализ, структура, графы, компонента связности, поток случайных событий.*

Введение

Данная работа посвящена демонстрации возможностей применения **метода выделения структур (МВС)** в различных модификациях к исследованию геофизических данных.

Понятие структура подразумевает наличие множества элементов, связанных между собой. В геофизических задачах сигналы могут иметь разную физическую природу. Одной из основных задач геофизических исследований является обнаружение скрытых объектов, обладающих отличными от фона физическими свойствами (электрическими, магнитными, упругими и т.д.). Обычно данные получают на некоторой пространственной сетке (x, y) (не обязательно равномерной) или на отдельных трассах (профилях). В любом случае каждой пространственной точке ставится в соответствие свой сигнал $S(x, y, t)$. Если сигналы в двух точках мало отличаются друг от друга, то можно считать, что в них и примерно одинаковое измеряемое свойство среды.

Иными словами, из данных исследований надо найти такие сигналы, которые близки между собой, но отличаются от сигналов в других

точках. В этом плане, постановка геофизических задач очень близка к задачам классификации: распределить все множество объектов, характеризующихся общей системой признаков, на небольшое число классов, руководствуясь выбранным критерием.

1. Алгоритмы Метода Выделения Структур (МВС)

Методы многомерного анализа, к которым принадлежит МВС, сейчас широко развиваются, и этот процесс не может быть отнесен только к математической или прикладной статистике [1]. Эти методы применяются при решении большого количества практических задач [2,3]. МВС был разработан для нахождения явных и латентных структур в числовом массиве данных любой природы и включает в себя 3 алгоритма: APFLOW, BINREL, MRS [4]. Для использования любого из алгоритмов МВС данные должны быть представлены в виде матрицы Объект-Признак и матрица расстояний D_{ij} между каждой парой объектов с номерами i и j , выбрав метрику, как правило, Евклида или Хемминга.

Для описания работы алгоритмов МВС удобно использовать язык теории графов. Вершины графа – объекты, рёбра – попарные расстояния D_{ij} между объектами. Алгоритмы МВС в матрице D удаляют часть ребер, получая для исходного полного графа несвязный подграф, каждая связанная часть которого называется компонентой связности. Этот подграф объявляется структурой, которая называется «естественной». МВС предусматривает и отыскания структур, обладающих определенными, заранее заданными критериями. В этом случае, компонента связности становится отдельной структурой, если свойства объектов, ей принадлежащих, удовлетворяют этим критериям. Критерии – это набор требований к выделенному кластеру (отдельной структуре), которые могут формулироваться как с использованием основных, так и вспомогательных признаков. Например, объекты выделенной компоненты связности должны быть близки как в пространстве признаков, так и в реальном пространстве.

Так, алгоритм **APFLOW** разработан для решения «поточковых» задач, т.е. когда имеется множество объектов или событий, каким-то образом упорядоченных в пространстве или во времени, и характеризующихся только одним признаком (одномерные данные). Алгоритм разбивает весь поток на участки, называемые «интервалами стационарности» (ИС) по критерию близости значения этого признака. Этот алгоритм был первоначально разработан для аппроксимации потока случайных событий подходящим МС-поток, поэтому в адаптированном варианте сохранилась изначальная терминология.

BINREL эффективен при обработке пространственных многомерных данных, особенно, в задачах отыскания малоконтрастных объектов или отделения объектов от фона.

MRS – классический дивизимный алгоритм кластерного анализа. Этот алгоритм позволяет разбить все множество объектов на небольшое число подмножеств. Подмножество фиксируется в качестве класса, если элементы его связаны между собой сильнее, нежели с другими объектами.

Примеры задач, решаемых МВС

Пример 1. Исследование отраженного электромагнитного сигнала от поверхности материалов. Решалась задача о возможности использования электромагнитного сигнала в качестве детектора, реагирующего на состояние объектов, перемещающихся по ленте конвейера. Схожие по постановке задачи были рассмотрены в [5]. Лента с объектами движется с определенной скоростью, над лентой закреплен радар, излучающий электромагнитный сигнал определенной частоты. Отраженный от материала, находящегося на ленте, сигнал периодически записывается. Проведенные экспериментальные исследования со средами, отличающимися друг от друга диэлектрическими параметрами, показали, что в качестве информационного признака, реагирующего на состояние объекта, может выступать максимальное или минимальное значение отраженного сигнала, т.е. одномерный поток случайных данных. На рис.1 приведены поведение максимальных значений отраженного сигнала для модельной среды, (песок, бетон, сухие и влажные опилки), вместе с результатами работы APFLOW [6].

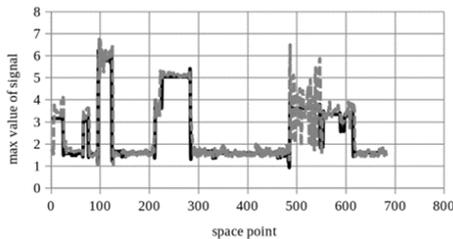


Рис. 1. Поведение максимального значения отраженного сигнала от разных поверхностей (пунктирная линия) и рассчитанные средние значения ИС (сплошная линия).

Пример 2. Гамма-каротаж. Задача, решаемая методом гамма-каротажа, заключается в нахождении литологических границ в разрезе с привязкой их по глубине. В этом методе гамма-кванты из окружающей среды поступают на медленно опускающийся с постоянной скоростью в скважину прибор с детектором, который фиксирует суммарное количество импульсов $N(t)$ за единицу времени. Поток гамма-квантов, достигающий детектора в скважине, представляет собой случайный процесс. И в каждой точке отсчета (на своей глубине) фиксируется случайная величина – интенсивность $N(t, H)$. На рис.2 представлены данные модельного эксперимента, имитирующего опускание датчика в скважину и полученные алгоритмом APFLOW средние значения найденных ИС.

При фиксированном положении прибора на определенной глубине, $N(t)$ будет распределена по закону Пуассона. Для каждой породы при одних и тех же условиях измерения характерна своя интенсивность излучения гамма-квантов, и на этом факте построена методика выделения слоев и нахождения их границ. Исследования показали, что алгоритм надежно выделяет границы ИС в том случае, когда средние значения интенсивности соседних слоев различаются на величину порядка 10% [7].

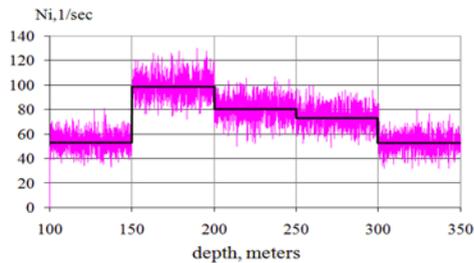


Рис. 2. Экспериментальные интенсивности и рассчитанные средние значения ИС (черная линия).

Пример 3. Гамма-спектрометрия. Методы, основанные на изучении потоков естественной радиации, применяются не только в скважинах, но и других геофизических работах. Широко используются гамма-спектрометрические методы в экологических исследованиях, в частности, при поиске зон повышенного радиационного фона искусственного или естественного происхождения. Приборы, называемые гамма-спектрометром, фиксируют гамма-кванты в каждом узком энергетическом канале. Несколько тысяч таких каналов охватывают широкий диапазон энергий фотонов (до 4 МэВ). В этом случае, мы имеем дело

с многомерными данными и можно применять, наряду с APFLOW, и другие алгоритмы. При реальных скоростях перемещения датчика, особенно с использованием самолетов или БПЛА, количество фотонов, попавших в конкретный энергетический канал очень мало. В этом случае имеет смысл группировать исходные интенсивности в небольшое число «интегральных» каналов. В наших исследованиях мы использовали 5 таких каналов. В каждом интегральном канале $N(t)$ получалось суммирование по 750-900 идущим подряд аппаратурным каналам. Причем, границы каналов подбирались таким образом, чтобы выделенные области энергии включали в себя пики, соответствующие гамма излучению атомов Cs , K , U и Th . В этом случае каждой пространственной точке (объект) исследуемых профилей присваивался упорядоченный набор из 5 чисел (признаков). Такой подход был применен к решению задачи обнаружения возможных радиационных аномалий на полигоне твердых бытовых отходов. Использовались все алгоритмы MBC, причем MRS и BINREL применялись в многомерном варианте, с учетом сразу всех 5 признаков, а APFLOW находил ИС отдельно для каждого канала. Анализ результатов показал, что использование всех алгоритмов дает возможность более корректно выделять аномальные зоны (рис.3).

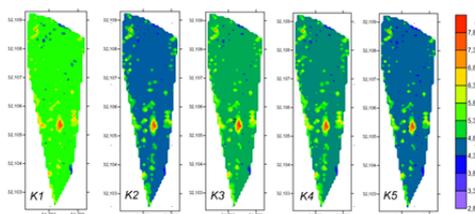


Рис. 3. Распределение ИС на исследуемой территории для каждого интегрального канала.

Пример 4. Исследование классовой динамики. Для нахождения закономерностей поведения метеорологических полей также можно использовать MBC, поскольку, по сути, цель одна и та же – найти территории, обладающие одинаковым поведением измеряемого параметра. Однако структуры метеорологических данных в отличие от геофизических исследований, обладают высокой изменчивостью [8]. Для выявления устойчивых закономерностей в поведении метеопараметров полезно использовать среднемесячные данные (температуры, влажности, давления, количество осадков, концентрации парниковых газов и т.д.) за большой период времени [9]. В последнее 10-15 лет наблюда-

ется усиление интереса исследователей к использованию для анализа геофизических полей процедуры классификации методами кластерного анализа [10,11]. Это объясняется, в частности тем, что в отличие от других методов многомерной статистики, результаты разбиения на кластеры хорошо интерпретируемы. Классификация, т.е. распределение пунктов измерения (метеостанций) по классам, дает возможность находить общие территории для различных систем признаков. Под общими территориями для k наборов признаков понимаются территории, на которых находятся станции, одновременно попадающие в класс номер m_1 для первого набора, в класс номер m_2 для второго и т.д. и, наконец, в класс номер m_k для k -ого набора. Для описания динамики пространственного распределения классов внутри года предлагается процедура, использующая результаты классификации отдельно за каждый календарный месяц. Каждая пространственная точка (станция) в каждый месяц принадлежит к определенному классу, т.е. за год ей будет соответствовать последовательность из 12 чисел – номеров классов. Назовем эту последовательность *годовой траекторией* станции, а группу станций с одинаковой траекторией – «траекторным классом». Ясно, что ареал, на котором располагается определенный траекторный класс, будет демонстрировать, в целом, одинаковое поведение метеопараметра в течение всего года или сезона. На рис. 4 в качестве примера приведены траекторные классы для «зимних» месяцев среднемесячной температуры территории Северного полушария. Осуществив классификацию территории по нескольким параметрам, можно выделить области с одинаковым поведением уже не одного, а целой группы параметров. Анализ расположения таких областей может служить базисом для последующего анализа с целью выявления факторов, влияющих на климатическую систему Земли.



Рис. 4. Траекторные классы для зимних месяцев.

Заключение

Приведенные примеры показывают широкие возможности метода выделения структур при решении разнообразных практических задач.

СПИСОК ЛИТЕРАТУРЫ

1. *Ким Дж. О.* Факторный, дискриминантный и кластерный анализ: пер с англ. под ред. И. С. Енюкова. М.: Финансы и статистика. 1989. 215 с
2. *Pal M.* Random forest classifier for remote sensing classification // *International Journal of Remote Sensing*. 2005. Vol.26. P. 217–222.
3. *Maulik U., Bandyopadhyay S.* Genetic Algorithm-Based Clustering Technique // *Pattern Recognition*. 2000. Vol. 33. P. 1455–1465.
4. *Катаев С. Г.* Алгоритмические и программные средства аналитического и неразрушающего контроля параметров природной среды на основе метода выделения структур: дис. на соискание ученой степени докт. техн. наук. Томск: 2013. 355 с.
5. *Greco M. S., Abramovich Y., Ovarlez J. P., Li H., Yang X.* Introduction to the issue on advanced signal processing techniques for radar applications // *IEEE Journal of Selected Topics in Signal Processing*. 2015. Vol. 9(8). P. 1363–1365.
6. *Antonyuk N. A., Kataev S. G., Kataeva S. S., Klein Wolterink M., Miller A. A.* On obtaining information about the structure and quality of investigated objects revealed from streams of data through the method for identifying structures. // *Journal of Physics: Conference Series*. 2020. Vol.1680. P(012002).
7. *Катаев С. Г., Катаева С. С., Сенчина Н. П., Миллер А. А., Ермолин Е. Ю.* Использование метода выделения структур для анализа данных спектрометрического гамма-каротажа // *Сборник материалов 18-ой научно-практической конференции и выставки*. Москва: 2022. С. 312–320.
8. *Хромов С. П., Петросяну М.* А Метеорология и климатология. М.: Изд-во МГУ. 2004. 582 с.
9. *Pérez I. A., Sánchez M. L., García M. Á., Torre B.* A classification of CO_2 concentrations based on a binary meteorological six variable system // *Agricultural and Forest Meteorology* 2009. Vol. 149. № 10. P. 1686–1692.
10. *Turcotte B., Morse B.* A global river ice classification model // *Journal of Hydrology* 2013. Vol. 507, №12. P. 134–148.
11. *Jacobeit J.* Classifications in climate research // *Physics and Chemistry of the Earth, Parts A/B/C*. 2010. Vol. 35. № 9–12, P. 411–421.

Катаев Сергей Григорьевич — д-н, профессор, Томский государственный педагогический университет. E-mail: sgkataev2010@yandex.ru

Катаева София Семёновна — к-н, доцент, Национальный исследовательский Томский государственный университет. E-mail: sskataeva@yandex.ru

ПРОТОТИП ИНТЕЛЛЕКТУАЛЬНОЙ РЕКОМЕНДАТЕЛЬНОЙ СИСТЕМЫ ПОДДЕРЖКИ ФОРМИРОВАНИЯ ТРЕБОВАНИЙ ДЛЯ ВАКАНСИЙ ПО АКТУАЛЬНЫМ ТРЕБОВАНИЯМ РЫНКА ТРУДА

И. Е. Николаев

Челябинский государственный университет, г. Челябинск, Россия

В статье приводится описание прототипа интеллектуальной рекомендательной системы поддержки формирования списка требований для вакансии по актуальным требованиям рынка труда. Кратко описываются методы, лежащие в основе прототипа, описывается модель прототипа системы и процесс его разработки. Приводятся результаты оценки эффективности использования прототипа системы на реальных данных. **Ключевые слова:** интеллектуальные рекомендательные системы, рынок труда, требования вакансии, машинное обучение, нейросетевые модели.

Введение

В настоящее время, с ростом числа вакансий на рынке труда, линейным руководителям и специалистам по подбору персонала в процессе формирования списка требований для вакансий становится все сложнее проводить анализ существующих тенденций востребованности отдельных знаний и навыков на рынке труда. Учитывая разнообразие требований на рынке труда, возникает необходимость в разработке интеллектуальных рекомендательных систем, способных предложить удобный инструмент анализа и создания оптимального списка требований для вакансии.

В данной статье приводится описание и апробация прототипа интеллектуальной рекомендательной системы поддержки формирования списка требований для вакансий, с учетом актуальных тенденций рынка труда на уровне анализа отдельных сущностей требований, таких как знания и навыки/компетенции. Разработка и апробация такой системы позволит не только ускорить процесс формирования списка требований для вакансий, но и в целом позволит повысить эффективность процесса подбора персонала.

1. Методология

В основе прототипа системы лежат ранее разработанные модели, методы и алгоритмы. Модель и метод извлечения отдельных сущностей требований представлены в работе [1] (общая схема метода представлена на рисунке 1). Интеллектуальный метод формирования списка требований профиля должности представлен в работе [2]. (общая схема метода представлена на рисунке 2).

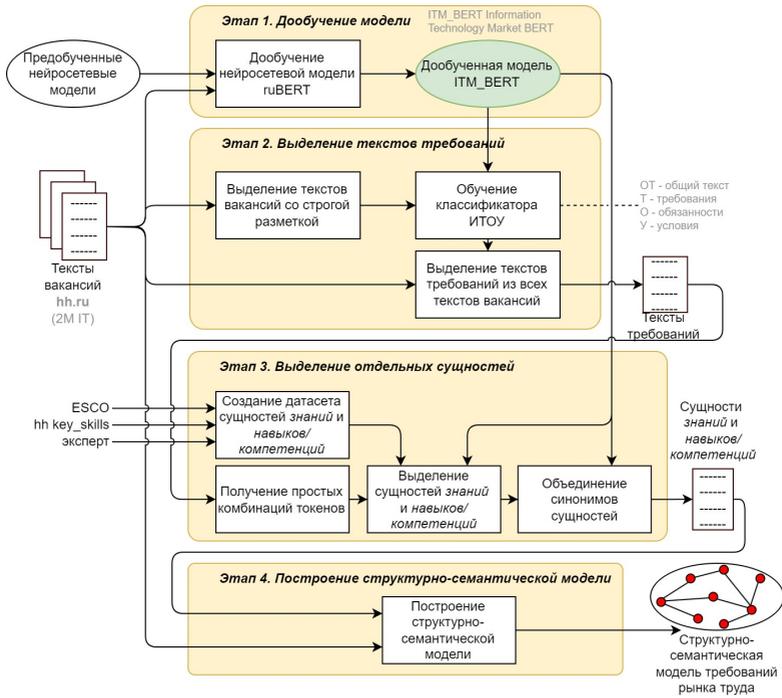


Рис. 1. Общая схема метода извлечения отдельных сущностей требований



Рис. 2. Общая схема формирования списка требований для вакансии

2. Разработка прототипа

Программная реализация прототипа системы выполнена на языке Python 3.9 с использованием веб-фреймворка Flask, в качестве СУБД использовалась MySQL. Для обучения нейросетевых моделей BERT [3] и классификаторов на основе нейронных сетей использовалась библиотека HuggingFace [4], для остальных алгоритмов классификации и кластеризации, а также для оценки качества проводимых экспериментов использовалась библиотека scikit-learn. Для извлечения сущностей из текстов вакансий использовалась открытая библиотека Yargy. Для построения графов использовалась библиотека NetworkX, а для их визуализации библиотека plotly. Структура прототипа представлена на рисунке 3.

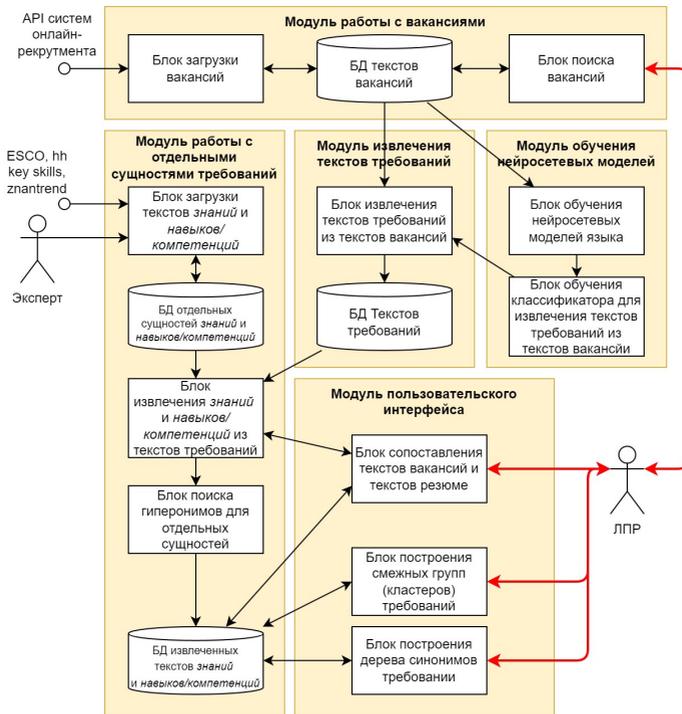


Рис. 3. Структура прототипа системы поддержки формирования списка требований вакансии по актуальным требованиям рынка труда

В качестве ЛПР (лица принимающего решения) могут выступать: линейные и функциональные руководители служб, отделов, структур-

ных подразделений; специалисты по подбору персонала; эксперты из профессионального сообщества.

3. Использование прототипа

Прототип системы позволяет строить несколько типов диаграмм для анализа отдельных сущностей требований и связей между ними:

- Карта смежных сущностей требований с учетом их совместной встречаемости (см. рисунок 4);
- Дерево синонимов отдельных сущностей требований (см. рисунок 5).

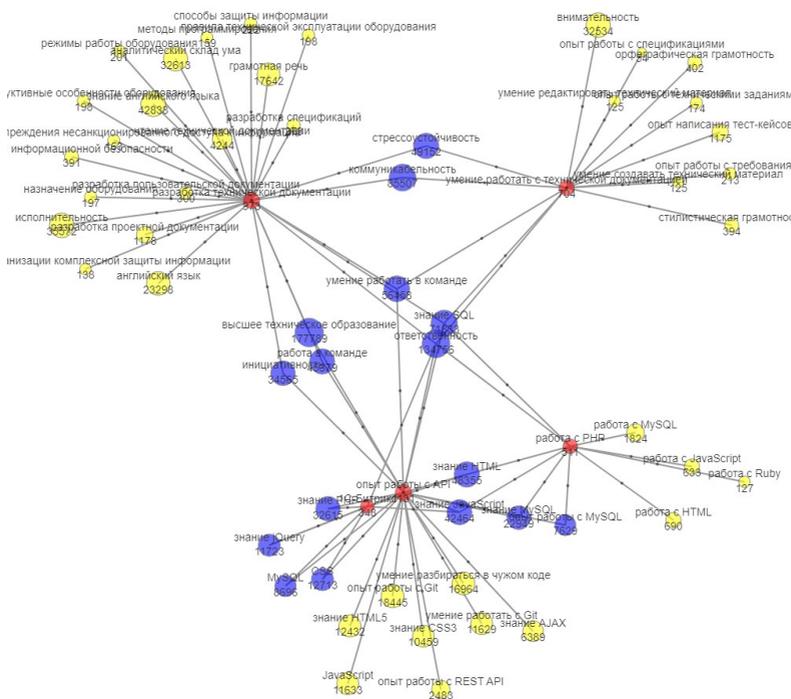


Рис. 4. Пример карты совместной встречаемости смежных сущностей требований для профессии «web-разработчик»

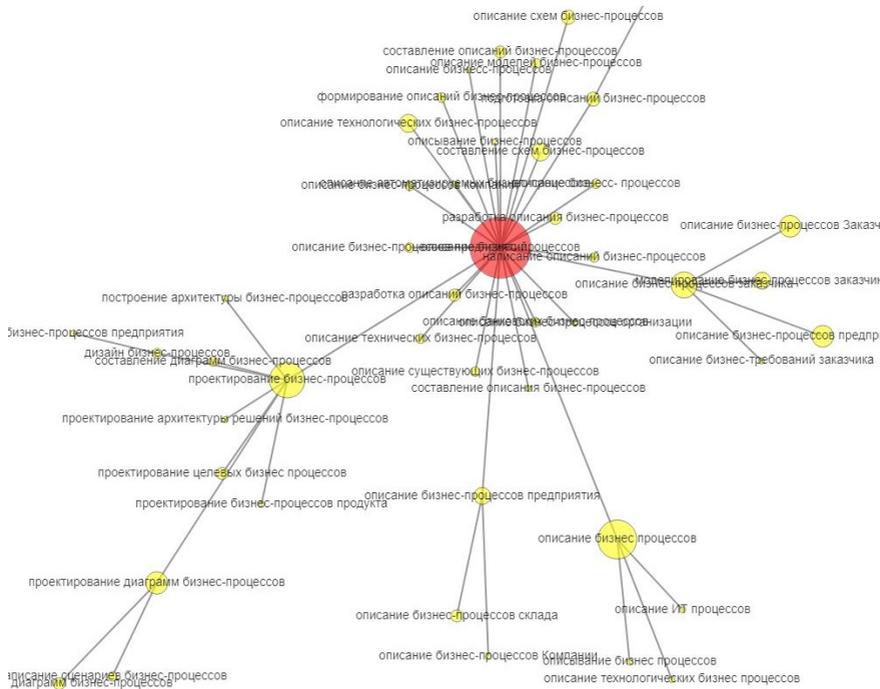


Рис. 5. Пример дерева синонимов для навыка/компетенции «описание бизнес-процессов»

4. Результаты

Для оценки эффективности прототипа системы были использованы следующие метрики: оценка изменения среднего времени формирования описания вакансии; оценка изменения среднего количества соискателей; оценка изменения среднего времени закрытия одной вакансии.

Для вычисления перечисленных метрик на базе двух предприятий ЮНИИТ (г. Ханты-Мансийск) и ООО Фирма «Интерсвязь» (г. Челябинск) в период с 2022 по 2024 год была собрана информация о подготовке 87 вакансий до внедрения системы, и о 57 вакансиях после внедрения системы. Результаты оценки трех метрик представлены в таблице 1.

Эффективность управленческих решений, принятых с использованием прототипа системы, заключается в повышении степени соответствия списка формируемых требований проекта вакансии с реальными требованиями рынка труда, и тем самым повышается эффективность процесса кадрового обеспечения.

Таблица 1

Результаты оценки эффективности использования прототипа

Метрика	Было	Стало	%
Время подготовки вакансии (часов)	4	1,5	↓62,5%
Кол-во соискателей на 1 вакансию (шт.)	3,7	4,3	↑16,2%
Время закрытия 1 вакансии (дней)	17,7	15,1	↓14,7%

Заключение

Разработан прототип интеллектуальной рекомендательной системы поддержки формирования требований в проектах вакансий. Проведена апробация прототипа в нескольких организациях. По результатам экспертной оценки, отмечается высокая релевантность предлагаемых системой рекомендаций вариантов отдельных сущностей требований. Проведена комплексная оценка эффективности использования интеллектуальной рекомендательной системы в процессе подбора персонала по нескольким метрикам.

СПИСОК ЛИТЕРАТУРЫ

1. Николаев И. Е. Метод извлечения знаний и навыков/компетенций из текстов требований вакансий // Онтология проектирования. 2023. Т. 13. № 2(48). С. 282–293.
2. Николаев И. Е. Интеллектуальный метод формирования списка требований профиля должности на основе нейросетевых моделей языка с использованием таксономии ESCO и корпуса онлайн-вакансий // Бизнес-информатика. 2023. Т. 17. № 2. С. 71–84.
3. Devlin J., Chang M.W., Lee K., Toutanova K. Bert: Pre-training of deep bidirectional transformers for language understanding // arXiv preprint arXiv:1810.04805. 2018.
4. Библиотека transformers от Hugging Face. URL: <https://huggingface.co/docs/transformers/index>

Николаев Иван Евгеньевич — старший преподаватель кафедры информационных технологий и экономической информатики института информационных технологий Челябинского государственного университета. E-mail: ivan_nikolaev@csu.ru

АБСТРАГИРОВАНИЕ ДАННЫХ В КОНТЕКСТЕ МАШИННОГО ОБУЧЕНИЯ: МЕТОДЫ, ПРИМЕНЕНИЯ И ПЕРСПЕКТИВЫ

М. Е. Исмагулов

Югорский государственный университет, г. Ханты-Мансийск, Россия

В данной обзорной статье рассматривается стратегия абстрагирования данных, какой смысл она несет в контексте машинного обучения и мультимодальной обработке данных, рассматриваются методы, подходы и алгоритмы, а также какие положительные эффекты несет в себе эта стратегия. В работе рассматриваются теоретические основы абстрагирования, её связь с теорией информации. Особое внимание уделяется различным методам абстрагирования, включая выделение признаков, понижение размерности, кластеризации и использованию автоэнкодеров. Статья исследует применение этих методов в ключевых областях машинного обучения, таких как компьютерное зрение, обработка естественного языка и т. д. **Ключевые слова:** абстрагирование данных, машинное обучение, выделение признаков, понижение размерности, глубокое обучение, интерпретируемость моделей

Введение

Абстрагирование — когнитивный процесс и методологический прием, заключающийся в выделении существенных свойств и связей предмета и отвлечении от других, несущественных в данном контексте свойств. Это позволяет создать обобщенное представление о предмете или явлении, выявить его ключевые характеристики и закономерности [1].

В контексте машинного обучения, абстрагирование это процесс и метод преобразования сырых данных в более компактные и информативные представления, выделяющие ключевые признаки и паттерны, релевантные для решения конкретной задачи. Включает в себя различные техники сокращения размерности, выделения признаков и создания высокоуровневых представлений данных, что позволяет повысить эффективность и обобщающую способность моделей машинного обучения [2].

1. Значимость темы в современных исследованиях машинного обучения

На современном этапе развития машинного обучения особое внимание уделяется мультимодальной обработке данных – процессу, при котором системы искусственного интеллекта способны одновременно обрабатывать и интегрировать информацию из различных модальностей, таких как текст, изображения, звук или видео [3].

Стратегия абстрагирования особенно важна для оптимизации вычислительных ресурсов и повышения обобщающей способности модели [4]. Абстрагирование применяется как на этапе предобработки данных и обучения моделей, так и при их непосредственном функционировании, позволяя создавать высокоуровневые представления, инвариантные к незначительным изменениям во входных данных [5].

2. Теоретические основы абстрагирования данных и связь с теорией информации

Абстрагирование в машинном обучении основывается на выделении существенных признаков, игнорировании несущественных деталей [6], создании иерархических представлений данных [7], достижении инвариантности к незначительным изменениям [5] и эффективной компрессии информации [8].

Теория информации играет ключевую роль в абстрагировании. Информационная энтропия Шеннона уменьшает избыточность данных, сохраняя их информативность [9]. Принцип минимальной длины описания (MDL) Риссанена помогает выбрать оптимальный уровень абстракции [10]. Взаимная информация оценивает значимость признаков [11], а теорема кодирования источника определяет пределы сжатия данных, связанного с сохранением важной информации [12].

3. Выделение признаков и понижение размерности как методы абстрагирования

Выделение признаков — ключевой аспект абстрагирования в машинном обучении, представляющий процесс преобразования сырых данных в информативные характеристики [11]. Методы варьируются от простых статистических подходов до сложных алгоритмов глубокого обучения [4]. Например, сверточные нейронные сети (CNN) автоматически извлекают иерархию признаков от простых краев до сложных объектов [12].

Выделение признаков связано с уменьшением размерности. Методы, такие как анализ главных компонент (PCA) и t-SNE, представляют

данные в пространстве меньшей размерности, сохраняя важную информацию [13]. В глубоком обучении автоэнкодеры создают сжатые представления данных [14].

Понижение размерности — это метод машинного обучения, направленный на уменьшение количества признаков в данных при сохранении максимального объема информации.

4. Кластеризация и сегментация данных

Абстрагирование кадров в видео, как показано на изображении (Рис. 1), можно рассматривать как форму сегментации данных. В этом случае, видеопоток сегментируется на отдельные кадры, которые затем классифицируются как «хорошие» (зеленые) или «плохие» (красные) на основе определенных критериев. Этот процесс по сути является бинарной кластеризацией, где каждый кадр относится к одному из двух кластеров.

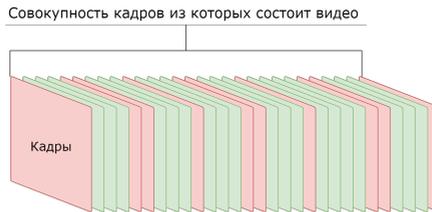


Рис. 1. Абстрагирование кадров видеоконтента

Эти критерии могут включать:

1. Наличие артефактов: Кадры, содержащие визуальные искажения, пиксельные ошибки или другие технические дефекты, классифицируются как "плохие" (красные).
2. Вмешательство в кадр: Например, когда преподаватель пересекает область презентации, загорая важную информацию на слайде, такой кадр также отмечается как "плохой" (красный).

Абстрагирование по сценам представляет собой высокоуровневую сегментацию, анализируя последовательность кадров для выявления значительных изменений в содержании или визуальных характеристиках, указывающих на смену сцены. В свою очередь абстрагирование по кадрам фокусируется на мелком уровне детализации, выделяя ключевые кадры, которые наилучшим образом представляют содержание видео или сцены.

5. Применение абстрагирования в компьютерном зрении

Изображение (Рис. 2), представленное в качестве примера, демонстрирует сложный процесс выделения признаков (Feature extraction) и классификации объектов в контексте современных методов компьютерного зрения и машинного обучения.

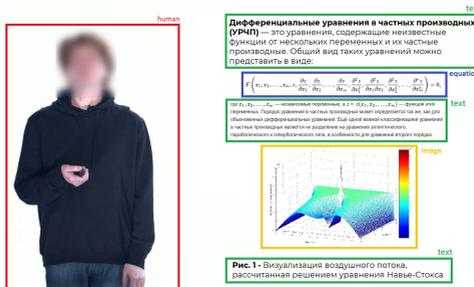


Рис. 2. Выделение (Извлечение) признаков и их классификация

Центральным элементом изображения является человек в темной одежде, что представляет классическую задачу обнаружения и классификации объектов. Рядом расположены текстовые блоки с математическими формулами и описаниями дифференциальных уравнений в частных производных, демонстрируя способность систем распознавать и классифицировать научный текст. В нижней части изображена трехмерная визуализация, иллюстрирующая возможности машинного обучения в анализе сложных графических данных и научных визуализаций.

Для такого анализа можно использовать различные архитектуры нейронных сетей и алгоритмы. YOLO и SSD подходят для быстрого обнаружения и локализации объектов, а ResNet и EfficientNet — для

точной классификации, включая распознавание человеческих фигур и графических элементов. Для сегментации изображений применяются U-Net и Mask R-CNN, создающие точные маски для объектов. Распознавание текста осуществляется с помощью OCR-систем, таких как Tesseract, работающих даже со сложными математическими формулами.

6. Мультимодальное обучение

Абстрагирование в мультимодальном обучении представляет собой комплексный подход к анализу и интерпретации данных из различных источников восприятия, таких как текст, изображения, аудио и видео. В предыдущих главах мы рассмотрели процессы абстрагирования отдельных модальностей.

Мультимодальное обучение идет дальше, объединяя эти отдельные абстракции для создания более полного и контекстно-богатого представления информации. Важно отметить, что мультимодальность может подразумевать не только параллельную обработку различных типов данных, но и итоговую компиляцию абстракций отдельных модальностей.

Заключение

В данном обзоре были рассмотрены ключевые аспекты абстрагирования данных в контексте машинного обучения. Проанализированы теоретические основы абстрагирования, связь с теорией информации, а также различные методы, включая выделение признаков, понижение размерности и кластеризация. Особое внимание уделено применению абстрагирования в таких областях, как компьютерное зрение и мультимодальное обучение, что демонстрирует универсальность и важность этого подхода для современных систем искусственного интеллекта.

Результаты обзора указывают на значительный потенциал методов абстрагирования для повышения эффективности и интерпретируемости моделей машинного обучения.

СПИСОК ЛИТЕРАТУРЫ

1. Горский Д. П. Вопросы абстракции и образование понятий. М.: Издательство Академии наук СССР, 1961. 351 с.
2. DeepAI.org — What is Feature Extraction? // DeepAI. URL: <https://deepai.org/machine-learningglossary-and-terms/feature-extraction> (дата обращения: 21.06.2024).

3. *Baltrusaitis T. Ahuja C. Morency L. P.* Multimodal Machine Learning: A Survey and Taxonomy // IEEE Transactions on Pattern Analysis and Machine Intelligence., 2019. Vol. 41, № 2. P. 423-443.
4. *Courville A. Bengio Y. Vincent P.* Representation Learning: A Review and New Perspectives // IEEE Transactions on Pattern Analysis and Machine Intelligence., 2013. Vol. 35, № 8. P. 1798-1828.
5. *LeCun Y. Bengio Y. Hinton G.* Deep learning // Nature., 2015. Vol. 521, № 7553. P. 436-444.
6. *Goodfellow I. Bengio Y. Courville A.* Deep Learning. Cambridge: MIT Press, 2016. 800 p.
7. *Bengio Y.* Learning deep architectures for AI // Foundations and Trends in Machine Learning., 2009. Vol. 2, № 1. P. 1-127.
8. *MacKay D. J. C.* Information Theory, Inference, and Learning Algorithms. Cambridge: Cambridge University Press, 2003. 628 p.
9. *Cover T. M. Thomas J. A.* Elements of Information Theory. 2nd ed. Hoboken: Wiley-Interscience, 2006. 776 p.
10. *Rissanen J.* Modeling by shortest data description // Automatica., 1978. Vol. 14, № 5. P. 465-471.
11. *Guyon I. Elisseeff A.* An introduction to variable and feature selection // Journal of Machine Learning Research., 2003. Vol. 3. P. 1157-1182.
12. *Krizhevsky A. Sutskever I. Hinton G. E.* ImageNet classification with deep convolutional neural networks // Advances in Neural Information Processing Systems., 2012. Vol. 25. P. 1097-1105.
13. *Van der Maaten L. Hinton G.* Visualizing data using t-SNE // Journal of Machine Learning Research., 2008. Vol. 9. P. 2579-2605.
14. *Vincent P. Larochelle H. Bengio Y. Manzagol P. A.* Extracting and composing robust features with denoising autoencoders // Proceedings of the 25th International Conference on Machine Learning., 2008. P. 1096-1103.

Исмагулов Милан Ерикович — аспирант 1 года обучения. E-mail: *milan.erikovich@gmail.com*

СОДЕРЖАНИЕ

Математическая теория телеграфика и теория массового обслуживания	5
<i>Tóth Á., Sztrik J.</i> Investigation of M/G/1//N system with impatient customers, unreliable primary and a backup server	6
<i>Dudin A., Dudina O.</i> Tandem of a self-service queueing system and a finite capacity queueing system processing in which is described by the Markov chain.....	12
<i>Nazarin A., Sopin E.</i> A Computational Algorithm Development for Stationary Characteristics of Resource Loss System With Resource-Dependent Service Time	19
<i>Binumon Joseph, Jose K. P.</i> A k-out-of-n system with single server taking multiple vacation and external service under non-preemption	25
<i>Bobina J. Mattam, K. P. Jose.</i> A Perishable Inventory System with Unreliable Server	31
<i>Tsitsiashvili G.</i> Synthesis of a new Markov process based on the characteristics of other two processes	37
<i>Joy Jijo, Jose K. P.</i> Analysis of a Discrete-Time Perishable Inventory Model with Positive Service and Lead Times Under Modified N-Policy Framework	44
<i>Lizyura O.</i> Two-Way Communication MAP/GI/1 Retrial Queue with Multiple Types of Outgoing Calls	51
<i>Astafiev S.</i> Multi-server Job Model: lumping and numerical analysis	58
<i>Manaeva V., Zhbankova E., Markova E.</i> Age of Information analysis for wireless systems with time-correlated traffic arrival process	64
<i>Sakai Y., Phung-Duc T.</i> Analysis of a Batch Arrival Queue with Power Saving Mode	70
<i>Власкина А. С., Бурицева С. А., Кочеткова И. А.</i> Анализ модели с относительным приоритетом перераспределения ресурса между двумя сегментами сети с эластичным трафиком.....	76
<i>Назаров А. А., Измайлова Я. Е.</i> Асимптотический анализ RQ-системы M M 1 с двумя орбитами и высокой интенсивностью обращения заявок со второй орбиты.....	80
<i>Зорин А. В.</i> О некоторых соотношениях для времени разгрузки очереди при циклическом управлении формируемыми в случайной среде потоками	85

<i>Бабаева Б. Э., Лисовская Е. Ю.</i> Модель сети массового обслуживания с тремя узлами для расчета требуемой вычислительной мощности сети Интернета вещей	91
<i>Королев Д. А., Моисеева С. П., А. Н. Моисеев</i> Асимптотический анализ системы массового обслуживания с ожидающими отрицательными заявками	98
<i>Николаева Д. Ю., Фёдорова Е. А.</i> Асимптотический анализ RQ-системы с s-настойчивыми заявками и обратной связью в условиях большой задержки.....	105
<i>Николаев Д. И., Гайдамака Ю. В.</i> Анализ временных характеристик поллинговой системы с двумя очередями и пуассоновскими входящими потоками 2-го рода	112
<i>Копать Д. Я.</i> Анализ G-сети с контрольными и карантинными очередями и нетерпеливыми положительными и отрицательными заявками	118
<i>Невенченко Е. А., Шепилов С. С., Моисеева С. П.</i> Асимптотический анализ многопоточной неоднородной СМО в условии предельно редких изменений состояний случайной марковской среды	126
<i>Гайдамака Е. А., Милехин А. А., Самуйлов К. Е.</i> Анализ пикового возраста информации в сети интегрированного доступа и транзита ..	133
<i>Гебриал И. Е. З., Леонтьева К. А., Хажимов А. А., Мутханна А. С. А., Кочеткова И. А.</i> Натурное моделирование нарезки сети 5G на базе комплексной интеграции SDN и SDR.....	141
<i>Ву Х. З., Жарков М. Л., Казаков А. Л., Лемперт А. А.</i> О применении сетей массового обслуживания для моделирования работы морского контейнерного терминала.....	147
<i>Сергеева Н. В., Тананко И. Е., Станкевич Е. П.</i> Моделирование системы хранения и транспортировки грузов сетью массового обслуживания с групповым обслуживанием требований	153
<i>Хаустова Н. Г., Моисеева С. П., Пакулова Е. А.</i> Алгоритм вычисления стационарного распределения вероятностей системы с приоритетами для произвольного числа мест в очереди	159
<i>Мелюшикова Н. П.</i> Асимптотический анализ RQ-системы с катастрофами в условии высокой интенсивности входящего потока	166
<i>Пауль С. В., Назаров А. А., Лапатин И. Л., Ву Х. А.</i> Область применимости асимптотических результатов при исследовании маркированного ММРР	171
<i>Русилко Т. В., Сальников Д. А.</i> G-сеть с доходами как модель кластерной системы	178

<i>Вавилов В. А., Назаров А. А.</i> Асимптотический анализ системы $M(t) M(t) _{\infty}$	184
<i>Цодиков В. Л., Зорин А. В.</i> Об условиях существования стационарного распределения в СМО с циклическим управлением периодическими конфликтными потоками	190
<i>Иголкин В. А., Румянцев А. С.</i> Тестирование долгой памяти в нагрузке односерверной системы	196
<i>Лапатин И. Л., Романов В. В.</i> Имитационный анализ характеристик выходящего потока бесконечнолинейной системы с деградацией скорости обслуживания	202
<i>Лавриненко М. Д., Шкленник М. А.</i> Исследование бесконечнолинейной СМО с проверкой качества обработки при нестационарном режиме работы.	208
<i>Кузнецов Н. А., Шкленник М. А.</i> Математическая модель коммерческого банка, использующего страхование кредитных продуктов	214
<i>Постнова О. С., Тананко И. Е., Рогачко Е. С.</i> Анализ системы сборки фрагментов требований, обслуженных в ациклической сети массового обслуживания	220
<i>Прокудина Ю. А., Моисеев А. Н.</i> Концепция архитектуры приложения для построения аналитической аппроксимации распределения вероятностей числа заявок в СМО на основе имитационного моделирования	226
Математическое моделирование	232
<i>Samatov B., Turgunboeva M.</i> Pursuit problem in a differential game with inertial players under integral constraints on controls	233
<i>Ibragimov M.</i> On Automorphisms of a Genetic Algebra Associated by a Non-Volterra Operator	239
<i>Uteuliev N., Seyrov A., Qutlimuratov Yu., Begilov B., Madreymova Z.</i> Stochastic Two-Criteria Model For Optimising The Location Of Agricultural Production Under Conditions Of Soil Saturation	244
<i>Starodubov R.</i> Implementation of genetic algorithm optimization approach for sigma-delta MEMS accelerometer design.....	250
<i>Ivannikov V.</i> The development of resonant MEMS accelerometer based on system level approach.....	256
<i>Солдатенко А. А., Семенова Д. В.</i> Построение сравнительных рядов для задачи ольфакторного анализа.....	263
<i>Героев А. С., Гергет О. М.</i> Влияние метаокружений на подбор гиперпараметров системы обучения с подкреплением.....	269

<i>Халиуллина А. Р.</i> Статистические характеристики углового распределения яркости для различных моделей атмосферы	275
<i>Кириллин А. М., Куликов В. Р.</i> Об аппроксимации множества точек поверхностью на основе приближенного расстояния из ряда Тейлора	281
<i>Булавчук А. М., Семенова Д. В., Солдатенко А. А.</i> О модификациях алгоритма генерации сетевого графика проектов для задачи RCPSP	287
<i>Подолин Д. А., Рассадин А. Э.</i> О пространственно-неоднородном распространении болезни	293
<i>Ибрагимова Э. И.</i> Обзор методов генерации знаковых графов	299
<i>Гендрина И. Ю.</i> Моделирование отраженной солнечной радиации с учетом отражения от подстилающей поверхности	306
<i>Лукьянова Н. А., Кульгаева А. И., Мишина Н. В.</i> Анализ кардиокомплексов групп больных с внезапной сердечной смертью и людей с нелетальной аритмией по записям холтеровского мониторинга	313
<i>Дьякович М. П.</i> Моделирование развития профессиональных заболеваний от физических факторов	319
<i>Бештоков М. Х.</i> Разностная схема повышенного порядка точности для модифицированного уравнения влагопереноса дробного порядка	325
<i>Гермидер О. В., Попов В. Н.</i> Математическое моделирование изгиба изотропных нанопластин в рамках нелокальной теории градиента деформации	331
<i>Тисленко Т. И.</i> Моделирование интенсивностей транспортных потоков для модуля валидации комплекса MARLIN24	336
Прикладной вероятностный и статистический анализ	342
<i>Abdulkazanova Z.</i> On Fixed Point Of A Mapping Related To Ising Model On \mathbb{Q}_p^2	343
<i>Mimitaeva A.</i> Method of Factorization of Simultaneous Equations for Unbiased Estimation of Their Coefficients	347
<i>Kitaeva A., Cao Y.</i> A Power-Law Adjustable Coefficient Dynamic Pricing Model Considering Shortages	353
<i>Itomov A. A., Murtazaev M. S., Tukhtaev E. E.</i> Local Limit Theorems for the Noncritical Galton-Watson Branching Systems	359
<i>Nurmukhamedova N., Berdimuradov M.</i> Comparative Analysis of Unknown Parameter Estimation of the Exponential Distributions with Right-Censored Data in Incomplete Statistical Models	363

<i>Imomov A. A., Ibrohimova Y.</i> On a Limiting Structure of Discrete-Time Stochastic Branching Systems with the Eventually Awaiting Death Moment	371
<i>Rahmatullaev M., Armedov O.</i> p-adic quasi Gibbs measures for the four states p-adic SOS model on the Cayley tree of order two	378
<i>Imomov A. A., Nazarov Z. A.</i> On Application of Karamata Slowly Varying Functions in the Theory of Noncritical Markov Branching Systems	384
<i>Rudić B., Pichler-Scheder M., Efrosinin D.</i> Maximum A Posteriori Predecessors in State Observation Models	391
<i>Sánchez-Benavides M. A., Saucedo-Zendejo F. R.</i> Meshfree Least Squares Modelling Applied To The Time-Dependent Schrödinger Equation	396
<i>Rahmatullaev M., Rasulova M.</i> Periodic Gibbs measures for the Potts-SOS model with an external field	402
<i>Abdushukurov A., Nurmukhamedova N.</i> Fisher Information Under Random Censoring On The Right	407
<i>Маркович Н. М., Рыжов М. С.</i> Модулярность графов, эволюционирующих моделью кластерного присоединения	413
<i>Jokhadze V., Purtukhia O.</i> Explicit integral representation of some path-dependent Brownian functionals	420
<i>Asadov Q. U., Sabirov K. K.</i> Modeling Reflectionless Transmission Of The Kink Soliton Of Klein-Gordon Equation On Networks	426
<i>Muradov R., Ziyoyitdinova M.</i> On Estimation of a Mixture of Distributions with Incomplete Data	430
<i>Dalabaev U., Ikramova M.</i> Numerical Modeling of Fluid Flow Based on the Rakhmatulin Model Taking Into Account the Forchheimer Equation	434
<i>Расулов А. С., Раимова Г. М.</i> Применения ветвящихся процессов для решения краевых задач для полулинейного уравнения Гельмгольца	440
<i>Войтишек А. В., Кирильчик Д. Е.</i> Исследование алгоритмов численного моделирования двумерных случайных векторов с зависимыми компонентами	446
<i>Мурадов Р. С., Душатов Н. Т.</i> Оценивание некоторых функционалов функции распределения по зависимым неполным наблюдениям в теории массового обслуживания	452
<i>Абдушукуров А. А., Бозоров С. Б.</i> Сравнение оценок функции выживания с учётом их свойства идентифицируемости с моделью неполных наблюдений	456

<i>Брызгалов В. Л., Войтишек А. В.</i> Конструктивный алгоритм уравнивания вероятностей	463
<i>Дмитриев Ю. Г., Киришжигин Я. В.</i> Комбинированные оценки функции распределения с априорной догадкой о параметрическом классе распределений.....	469
<i>Войтишек А. В., Шлымбетов Н. Х.</i> Компоненты погрешности функциональных вычислительных алгоритмов приближения вероятностной плотности по заданной выборке: L_2 -подход	475
<i>Дмитриев Ю. Г., Сорокин Р. Д.</i> Анализ эффективности комбинированной оценки регрессии	481
<i>Абулов М. О., Узаков З. У.</i> Краевая задача для уравнения третьего порядка смешанно-составного типа в четырехугольной области	487
Интеллектуальные системы и информационные технологии	493
<i>Taranto A.</i> Itô Bridge Ant Colony Optimization (IB-ACO) - A New Bridge from Mathematics to Computer Science.....	494
<i>Korsun A., Sushchenko S., Shkurkin A.</i> Analysis of the Effectiveness of Pipelining of Application Processing.....	500
<i>Nurullaev M. M.</i> A Method for Generating Random Numbers on Internet of Things Devices and Analyzing the Results.....	505
<i>Anzhin V., Trenkaev V.</i> Evaluation of the robustness of E_BLIND/D_LC watermarking algorithm.....	511
<i>Оруджов Э. А., Мусеев А. Н.</i> Разработка интерфейсов клиент-серверного взаимодействия веб-сервиса SimQ	517
<i>Шайхутдинов А. А.</i> Применение теории массового обслуживания в спиральном жизненном цикле онлайн-школы для оптимизации продуктовой метрики занятости преподавателей.....	524
<i>Блинова О. В., Панкратова Е. В.</i> Лингвистическая обработка научно-технической документации с целью поиска и категоризации.....	530
<i>Голденюк Е. Е., Лукьянова Н. А., Кусто Я. С.</i> Применение методов машинного обучения для обнаружения рака молочной железы по уровню экспрессии генов	536
<i>Баранова И. В., Судакова А. А.</i> Исследование методов распознавания жестов руки на изображениях	542
<i>Лукьянова Н. А., Голденюк Е. Е., Черняк К. Е.</i> Применение сверточных нейронных сетей для классификации МРТ здоровых коленных суставов и с признаками остеоартрита	548

<i>Массель Л. В., Массель А. Г.</i> Интеграция математических и интеллектуальных технологий в цифровой платформе для поддержки принятия стратегических решений по развитию энергетики России.....	554
<i>Морозов М. И.</i> Подготовка данных датасета «Vehicle Insurance Data 2018» для предсказания страховой премии.....	562
<i>Блинова О. В., Таратухина Ю. В.</i> Адаптация систем искусственного интеллекта под культурный, нравственный, этический и когнитивный профили	569
<i>Петрова С. Ю., Швец А. Ю.</i> Практика применения искусственного интеллекта для диагностики состояния изоляции силового трансформатора при ограниченном наборе данных.....	576
<i>Ложникова А., Балясова Н., Кичко Н.</i> О переоценке искусственного интеллекта в государственном регулировании подготовки кадров бухгалтеров в современной России	583
<i>Катаев С. Г., Катаева С. С.</i> О возможностях метода выделения структур при обработке данных произвольной природы.....	587
<i>Николаев И. Е.</i> Прототип интеллектуальной рекомендательной системы поддержки формирования требований для вакансий по актуальным требованиям рынка труда.....	594
<i>Исмагулов М.</i> Абстрагирование данных в контексте машинного обучения: методы, применения и перспективы	600

Научное издание

**ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ
И МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ (ИТММ-2024)**

МАТЕРИАЛЫ

XXIII Международной конференции имени А. Ф. Терпугова
20–26 октября 2024 г.

Редактор В.Г. Лихачева
Компьютерная верстка О.Д. Лизюра
Дизайн обложки Л.Д. Кривцовой

Отпечатано на оборудовании
Издательства Томского государственного университета
634050, г. Томск, пр. Ленина, 36.
Тел. 8+(382-2)–52-98-49
Сайт: <http://publish.tsu.ru>
E-mail: rio.tsu@mail.ru

Подписано к печати 14.10.2024 г.
Формат 60 × 84¹/₁₆. Бумага для офисной техники.
Гарнитура Times.
Печ. л. 38,4. Усл.-печ. л. 35,7. Тираж 500 экз. Заказ № 6025.

ISBN 978-5-907890-08-4



9 785907 890084 >