DISCRETE AND CONTINUOUS MODELS AND APPLIED COMPUTATIONAL SCIENCE

Volume 32 Number 3 (2024)

Founded in 1993

Founder: PEOPLES' FRIENDSHIP UNIVERSITY OF RUSSIA NAMED AFTER PATRICE LUMUMBA

DOI: 10.22363/2658-4670-2024-32-3

Edition registered by the Federal Service for Supervision of Communications, Information Technology and Mass Media

Registration Certificate: ПИ № ФС 77-76317, 19.07.2019

ISSN 2658-7149 (Online); 2658-4670 (Print)
4 issues per year.
Language: English.

Publisher

Peoples' Friendship University of Russia named after Patrice Lumumba (RUDN University).

Indexed by

- Scopus (https://www.scopus.com),
- Ulrich's Periodicals Directory (http://www.ulrichsweb.com),
- Directory of Open Access Journals (DOAJ) (https://doaj.org),
- Russian Index of Science Citation (https://elibrary.ru),
- CyberLeninka (https://cyberleninka.ru).

Aim and Scope

Discrete and Continuous Models and Applied Computational Science arose in 2019 as a continuation of RUDN Journal of Mathematics, Information Sciences and Physics. RUDN Journal of Mathematics, Information Sciences and Physics arose in 2006 as a merger and continuation of the series "Physics", "Mathematics", "Applied Mathematics and Computer Science", "Applied Mathematics and Computer Mathematics".

Discussed issues affecting modern problems of physics, mathematics, queuing theory, the Teletraffic theory, computer science, software and databases development.

It's an international journal regarding both the editorial board and contributing authors as well as research and topics of publications. Its authors are leading researchers possessing PhD and PhDr degrees, and PhD and MA students from Russia and abroad. Articles are indexed in the Russian and foreign databases. Each paper is reviewed by at least two reviewers, the composition of which includes PhDs, are well known in their circles. Author's part of the magazine includes both young scientists, graduate students and talented students, who publish their works, and famous giants of world science.

The Journal is published in accordance with the policies of COPE (Committee on Publication Ethics). The editors are open to thematic issue initiatives with guest editors. Further information regarding notes for contributors, subscription, and back volumes is available at http://journals.rudn.ru/miph

E-mail: miphj@rudn.ru, dcm@sci.pfu.edu.ru

Editorial board

Computer Design: *Anna V. Korolkova*, *Dmitry S. Kulyabov*

English Text Editors: *Nikolay E. Nikolaev*, *Ivan S. Zaryadov*, *Konstantin P. Lovetskiy*

Address of editorial board:
3 Ordzhonikidze St, 115419 Moscow, Russia
+7 (495) 955-07-16, e-mail: publishing@rudn.ru

Editorial office:
+7 (495) 952-02-50, miphj@rudn.ru, dcm@rudn.su,
site: http://journals.rudn.ru/miph

Paper size 70×108/16. Offset paper. Offset printing. Typeface “Adobe Source”.
Conventional printed sheet 7.52. Printing run 500 copies. Open price. The order 1053.
PEOPLES’ FRIENDSHIP UNIVERSITY OF RUSSIA NAMED AFTER PATRICE LUMUMBA
6 Miklukho-Maklaya St, Moscow, 117198, Russian Federation

Printed at RUDN Publishing House:
3 Ordzhonikidze St, Moscow, 115419, Russian Federation,
+7 (495) 955-08-61; e-mail: publishing@rudn.ru

Discrete & Continuous Models
& Applied Computational Science
ISSN 2658-7149 (Online), 2658-4670 (Print)

2024, 32 (3)
http://journals.rudn.ru/miph

Contents

Editorial

Computer science

Modeling and simulation

Physics and astronomy

Letters

Discrete & Continuous Models & Applied Computational Science
ISSN 2658-7149 (Online), 2658-4670 (Print)

2024, 32 (3) 255–259
http://journals.rudn.ru/miph

Editorial

DOI: 10.22363/2658-4670-2024-32-3-255–259 EDN: DKMJJS

Journal rubrics

Dmitry S. Kulyabov[1, 2], Leonid A. Sevastianov[1, 2]

[1] RUDN University, 6 Miklukho-Maklaya St, Moscow, 117198, Russian Federation

[2] Joint Institute for Nuclear Research, 6 Joliot-Curie St, Dubna, 141980, Russian Federation

Abstract. We describe introduced in the journal the rubric system.

Key words and phrases: rubrics, physics, computer science, modeling

For citation: Kulyabov, D. S., Sevastianov, L. A. Journal rubrics. *Discrete and Continuous Models and Applied Computational Science* **32** (3), 255–259. doi: 10.22363/2658-4670-2024-32-3-255–259. edn: DKMJJS (2024).

1. Rubrics of the journal

The editorial board has decided to return the explicitly allocated rubrics in the journal. The journal contains subject and organizational rubrics. The subject rubrics are:

- informatics (computer science);
- mathematical modeling;
- physics.

Organizational rubrics are:

- editorials;
- letters.

2. Subject rubrics

The journal publishes articles on computer science, mathematical modeling, and physics. All these scientific fields are closely related, and the connecting core is mathematical modeling.

2.1. Computer science

Computer science research articles cover a wide range of topics related to modern information processing technologies and techniques:

- networks and database management systems;
- artificial intelligence and machine learning;
- software engineering;
- cybersecurity;
- software methods development;
- theory of computing;
- information technology.

2.2. Physics

Physics research articles cover a wide range of topics, each aiming to advance understanding of different aspects of nature (Nature-1) and to develop new technologies (Nature-2):

- particle physics;
- cosmology and astrophysics;
- nuclear physics;
- condensed state physics;
- nonlinear physics;
- laser physics and optics;
- quantum mechanics;
- field theory;
- computational physics.

2.3. Mathematical modeling

Mathematical modeling emerged at the intersection of several fields of science:

- applied theoretical physics;
- mathematical physics;
- computational physics.

The structure of the field of mathematical modeling is given as the triad Model–Algorithm–Program [1, 2]. Modeling as a discipline encompasses different types of modeling approaches [3]. From our point of view, these approaches can be schematically described in a unified way. The research framework consists of operational and theoretical parts. The operational parts are represented by the procedures of system preparation and measurement. It is also possible to describe the operational parts as input and output data.

The theoretical part consists of two layers: model layer and implementation layer. The modeling layer is the main layer and defines the actual model under study. The realization layer describes the specific structure of the system evolution. Depending on the type of realization, different types of models can be obtained:

- realization—mathematical expressions: analytical mathematical models;
- implementation—analog system: physical model;
- realization—algorithm: simulation models;
- implementation—behavior approximation: surrogate model.

Each type of model has its own area of applicability, advantages and disadvantages. The use of the full range of models allows the deepest and most comprehensive study of the system being modeled.

2.3.1. Analytical modeling

The most rigorous study is based, usually, on an analytical mathematical model. In this case, the model layer is realized through mathematical expressions describing the evolution of the system.

2.3.2. Physical modeling

The obtained mathematical model should be compared with experimental data, verify it. For this purpose, a model analogous to a physical or technical system can be created. A physical model can also be a virtual model. For example, you can build a model of a data network using images of operating systems of routers and switches.

2.3.3. Simulation Modeling

With the development of computer technology, it became possible to specify a model realization not in the form of a mathematical description, but in the form of some algorithm. This type of models was called *simulation models*, and the approach itself was called *simulation modeling*. A simulation model plays a dual role.

- A simulation model, debugged and tested on experimental data and a physical model, can itself serve the purpose of verification of a mathematical model.
- On the other hand, the simulation model allows to investigate the behavior of the simulated system under different variants of input data more effectively than the mathematical model.

2.3.4. Statistical modeling

This type of modeling includes models that are implemented through *machine learning* methods. It can be divided into several approaches (e.g., surrogate modeling; data-driven modeling).

Surrogate modeling Other names: approximation models, response surface models, metamodels, black box models. In this approach, the modeling layer is known and even has an implementation (most often in the form of an analytical model). For many real-world problems, modeling can take quite a long time. Extremely difficult (rather routine) problems such as solution optimization, solution space exploration, sensitivity analysis and "what if" analysis become impossible because they require thousands or millions of simulation evaluations. To simplify the study, models are built that mimic the behavior of the original model as closely as possible while being computationally cheap. Surrogate models are built using a data-driven approach. The exact inner workings of the simulation code are not assumed to be known (or even understood), only the input-output (cooking-measuring) behavior is important. The model is built by simulating the response to a limited number (sometimes quite large) of selected data points. The scientific goal of surrogate modeling is to create a surrogate that is as accurate as possible, using as few modeling estimates as possible.

Data-driven modeling Used when there is no model per se and the nature of the true function is unknown a priori, so it is unclear which surrogate model will be most accurate. Furthermore, it is not clear how to obtain the most reliable estimates of the accuracy of this surrogate. In this case, the model layer is replaced by the researcher's guesses.

3. Organizational rubrics

The following organizational rubrics are suggested for the journal:

- editorial;
- letters.

In our view, editorials fulfill several tasks. In these articles, the editorial board:

- gives advice and recommendations on the proper design and structuring of scientific articles (in particular, for natural science articles, the structure of the IMRAD (Introduction, Methods, Results and Discussion) article [4, 5] is supposed to be used, which facilitates the perception and understanding of the research);
- communicates the requirements of the journal; indicates the specific requirements of journals for article layout, both semantic and syntactic;

– gives direction to improve the quality of articles, emphasizes the elements of research; the editorial board tries to strengthen the scientific value of the papers;
– provides semantic navigation through the current issue.

Editorial articles inform the authors about the current goal-setting of the journal.

The column *Letters to the Editor* is intended for publication of short messages containing comments, explanations or alternative points of view concerning both previously published articles and different thoughts that the editorial board does not consider it necessary to place in the corresponding subject rubrics. The main purposes of such letters are:

– comments and clarifications on certain aspects of a previously published article, if some points remained unclear or insufficiently disclosed;
– alternative or additional opinions on the issues discussed in the published articles;
– constructive criticism of the data or conclusions presented in the article;
– expansion of scientific discourse.

Author Contributions: The contributions of the authors are equal. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: No new data were created or analysed during this study. Data sharing is not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Samarskii, A. A. & Mikhailov, A. P. *Principles of Mathematical Modelling. Ideas, Methods, Examples* 360 pp. (CRC Press, Dec. 20, 2001).
2. Chetverushkin, B. N. & Mikhailov, A. P. Triad of Samarskii. The 100th anniversary of academician A.A. Samarskii. *Herald of the Russian Academy of Sciences* **89,** 187–193. doi:10.31857/s0869-5873892187-193 (Mar. 2019).
3. Korolkova, A. V., Kulyabov, D. S. & Hnatič, M. *Practical Application of the Multi-model Approach in the Study of Complex Systems* in *Distributed Computer and Communication Networks. DCCN 2020* (eds Vishnevskiy, V. M., Samouylov, K. E. & Kozyrev, D. V.) 526–537 (Springer Nature Switzerland AG, Cham, 2020). doi:10.1007/978-3-030-66471-8_40.
4. Wu, J. Improving the writing of research papers: IMRAD and beyond. *Landscape Ecology* **26,** 1345–1349. doi:10.1007/s10980-011-9674-3 (Nov. 2011).
5. Brain, L. Structure of the scientific paper. *British Medical Journal* **2,** 868–869. doi:10.1136/bmj.2.5466.868 (Oct. 1965).

Information about the authors

Dmitry S. Kulyabov (Russian Federation)—Professor, Doctor of Sciences in Physics and Mathematics, Professor of Department of Probability Theory and Cyber Security of RUDN University; Senior Researcher of Laboratory of Information Technologies, Joint Institute for Nuclear Research (e-mail: kulyabov-ds@rudn.ru, phone: +7 (495) 952-02-50, ORCID: 0000-0002-0877-7063, ResearcherID: I-3183-2013, Scopus Author ID: 35194130800)

Leonid A. Sevastianov (Russian Federation)—Professor, Doctor of Sciences in Physics and Mathematics, Professor of Department of Computational Mathematics and Artificial Intelligence of RUDN University (e-mail: sevastianov-la@rudn.ru, phone: +7 (495) 955-07-83, ORCID: 0000-0002-1856-4643, ResearcherID: B-8497-2016, Scopus Author ID: 8783969400)

DOI: 10.22363/2658-4670-2024-32-3-255–259 EDN: DKMJJS

Рубрики журнала

Д. С. Кулябов[1,2], Л. А. Севастьянов[1,2]

[1] Российский университет дружбы народов, ул. Миклухо-Маклая, д. 6, Москва, 117198, Российская Федерация

[2] Объединённый институт ядерных исследований, ул. Жолио-Кюри, д. 6, Дубна, 141980, Российская Федерация

Аннотация. Описывается вводимая в журнале система рубрик.

Ключевые слова: рубрики, физика, информатика, моделирование

Discrete & Continuous Models & Applied Computational Science
ISSN 2658-7149 (Online), 2658-4670 (Print)

2024, 32 (3) 260–270
http://journals.rudn.ru/miph

Computer science

Research article

UDC 519.21
DOI: 10.22363/2658-4670-2024-32-3-260–270
EDN: EUVTRK

Distribution of the peak age of information in a two-node transmission group modeled by a system with a group flow and a phase-type service time

Sergey I. Matyushenko, Konstantin E. Samouylov

RUDN University, 6 Miklukho-Maklaya St, Moscow, 117198, Russian Federation

(received: September 16, 2023; revised: September 30, 2023; accepted: October 1, 2023)

Abstract. This article continues the cycle of works by the authors devoted to the problem of the age of information (AoI), a metric used in information systems for monitoring and managing remote sources of information from the control center. The theoretical analysis of information transmission systems requires a quantitative assessment of the "freshness" of information delivered to the control center. The process of transferring information from peripheral sources to the center is usually modeled using queuing systems. In this paper, a queuing system with phase-type distributions is used to estimate the maximum value of the information age, called the peak age. This takes into account the special requirement of the transmission protocol, which consists in the fact that information enters the system in groups of random size. For this case, an expression is obtained for the Laplace–Stieltjes transformation of the stationary distribution function of the peak age of information and its average value. Based on the results of analytical modeling, a numerical study of the dependence of the average value of the peak age of information on the system load was carried out. The correctness of the expressions obtained was verified by comparing the analytical results with the results of simulation modeling.

Key words and phrases: information age, peak information age, queuing system, phase type distribution, group flow

For citation: Matyushenko, S. I., Samouylov, K. E. Distribution of the peak age of information in a two-node transmission group modeled by a system with a group flow and a phase-type service time. *Discrete and Continuous Models and Applied Computational Science* **32** (3), 260–270. doi: 10.22363/2658-4670-2024-32-3-260–270. edn: EUVTRK (2024).

1. Introduction

The problem of timely delivery of information to the control and management center arises in various spheres of human activity: in energy systems, in the industrial Internet of things, in the field of autonomous transport, in video surveillance systems, etc. [1–3]. In 2011, to quantify the freshness of information received by the control and monitoring center, the Age of Information (AoI) metric was proposed, which is a function of the time between the generation of updates at the sending node and the delay in their delivery over the network to the control and monitoring center (recipient node) [4–13]. The most convenient device for studying the problem of information age is the device of queuing systems and networks. An overview of the works in which the analysis of the age of information is proposed to be carried out using this device can be found, for example, in [14]. It should be noted that

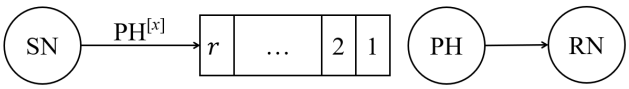

Figure 1. A two-node GT

most specialists limit themselves to simple models, for example, with an exponential distribution of time between the moments of generation of updates at the sending node and an exponential or deterministic distribution of the duration of update processing at the receiving node [15–17]. However, the simplest models of queuing systems allow us to obtain only a rough estimate of the age of information, since single-parameter distributions do not make it possible to take into account all the features of the protocols of modern dispatch control and data collection systems. In this paper, the process of transferring information from the sending node to the receiving node is modeled using a queuing system with phase-type distributions, the choice of phase parameters of which allows flexibly modeling complex dependencies that arise in modern data transmission systems.

2. Description of the model

Let's consider a group of information transmission (GT) consisting of a sender node (SN), a recipient node (RN) and a communication channel between them (Fig. 1).

Transmission from the SN to the RN is carried out by groups of packets of random length over a single communication channel. If the channel is busy, groups of packages line up in a queue with a limited number of waiting places. If there are no places in the queue, the group is lost and no longer has an impact on the information transfer process. A group is considered transferred if the last packet of this group is transmitted. By the peak age of the Z_n information transmitted by the n-th group, we will understand sum

$$Z_n = \hat{G}_{n+1} + W_{n+1}, \tag{1}$$

where $\hat{G}_{n+1}$ is the duration of the generation of the group following the n-th group, which ended with the first successful joining of the group to the transfer queue; W_{n+1} is the time of transfer of the $(n+1)$-st group from the SN to the RN.

We will model the transmission of information over the communication channel, using a single-line queuing system (QS) with a finite capacity accumulator r, $1 \le r < \infty$.

Applications for the system are received in groups. The flow of groups of applications is recurrent with a distribution function $A(t)$, $t \ge 0$, of the phase interval [18]

$$A(t) = 1 - \boldsymbol{\alpha}^T \mathrm{e}^{\boldsymbol{\Lambda} t}\mathbf{1}, \quad \boldsymbol{\alpha}^T\mathbf{1} = 1, \tag{2}$$

with an irreducible *PH* representation $(\boldsymbol{\alpha}, \boldsymbol{\Lambda})$ of order l.

Each group receives a random number of applications η with a given probability in advance:

$$\alpha_k = \mathrm{P}\{\eta = k\}, \quad k = \overline{1, r+1}.$$

At the same time

$$\sum_{k=1}^{r+1} \alpha_k = 1.$$

If, upon receipt of an application in the system, the number of applications in a group exceeds the number of available places in it, then the available places are filled, and the remainder of the applications from the new group is lost. A group is considered accepted into the system if at least one application from its membership is accepted. In the opposite case, the group is lost and the generation of the next group immediately begins.

Applications are served one at a time. Their service times do not depend on the aggregate and do not depend on the duration of generation and the number of applications in the system and have a common PH of $B(t)$ phase type:

$$B(t) = 1 - \boldsymbol{\beta}^T \mathrm{e}^{\boldsymbol{M}t}\mathbf{1}, \quad t \geq 0, \quad \boldsymbol{\beta}^T\mathbf{1} = 1, \tag{3}$$

with an irreducible *PH* representation $(\boldsymbol{\beta}, \boldsymbol{M})$ of the order of m.

We will assume that groups of applications are served in the order they are received, and applications within the group are served in a random order. In other words, an arbitrary application may end up in the right place in a group of k applications with the same probability of $1/k$.

The QS in question will be encoded in Kendall's notation as $PH^{[x]}/PH/1/r$.

3. Mathematical model

This system was considered in [19]. In this work, the functioning of the system was described by a homogeneous Markov process (MP) $\{X(t),\ t \geq 0\}$ over a set of states

$$\mathscr{X} = \bigcup_{n=0}^{r+1} \mathscr{X}_n,$$

where $\mathscr{X}_0 = \{(i,0),\ i = \overline{1,l}\}$, $\mathscr{X}_n = \{(i,n,j),\ i = \overline{1,l},\ j = \overline{1,m}\}$, $n = \overline{1,r+1}$.

Here and further, the expression $X(t) = (i,0)$ means that at time t the system is empty, and the generation process is going through phase i; $X(t) = (i,n,j)$ means that at time t there are n applications in the system, the generation process is going through phase i, and the maintenance process is phase j.

Since all states (MP) $\{X(t),\ t \geq 0\}$ communicate with each other, it is ergodic.

The limiting probabilities

$$p_{i0} = \lim_{t\to\infty} \mathrm{P}\{X(t) = (i,0)\},$$

$$p_{inj} = \lim_{t\to\infty} \mathrm{P}\{X(t) = (i,n,j)\},$$

exist, are strictly positive, do not depend on the initial distribution, coincide with stationary and they are the only solution to the system of equilibrium equations (SEE) with a normalization condition.

If you want to get acquainted with the conclusion of the SEE, we recommend that you refer to the materials [19]. We will limit ourselves to reproducing the main result of this work related to the development of a recurrent matrix algorithm for calculating stationary probabilities of system states. To do this, we introduce vectors:

$$\boldsymbol{p}_0^T = (p_{10}, p_{20}, ..., p_{l0}),$$

$$\boldsymbol{p}_n^T = (p_{1n1}, ..., p_{1nm}, p_{2n1}, ..., p_{2nm}, ..., p_{lnm})$$

and let's introduce the notation:

$$\boldsymbol{\lambda} = -\boldsymbol{\Lambda}\mathbf{1},$$

$$\boldsymbol{\mu} = -\boldsymbol{M}\mathbf{1},$$

$$A_k = \begin{cases} 1, \ k = 0, \\ 1 - \sum_{s=1}^{k} a_s, \end{cases}$$

where $1 - \sum_{s=1}^{k} a_s$ is the probability that the group size will be greater than k, $k = \overline{1, r+1}$.

$\boldsymbol{\Lambda} \oplus \boldsymbol{M} = \boldsymbol{\Lambda} \otimes \boldsymbol{I} + \boldsymbol{I} \otimes \boldsymbol{M}$ is the Kronecker sum $\boldsymbol{\Lambda}$ and $\boldsymbol{M}$, the symbol $\otimes$ denotes the tensor product of matrices; $\lambda = -(\boldsymbol{\alpha}^T \boldsymbol{\Lambda}^{-1} \mathbf{1})^{-1}$ is the intensity of receipt of groups of applications; $\mu = -(\boldsymbol{\beta}^T \boldsymbol{m}^{-1} \mathbf{1})^{-1}$ is the intensity of service of applications.

$$\widetilde{\boldsymbol{\Lambda}} = -(\boldsymbol{\Lambda} \oplus \boldsymbol{M}) + \mathbf{1}\boldsymbol{\alpha}^T \otimes \boldsymbol{M},$$

$$\widetilde{\boldsymbol{M}} = -(\boldsymbol{\Lambda} \oplus \boldsymbol{M}) + \boldsymbol{\Lambda} \otimes \mathbf{1}\boldsymbol{\beta}^T,$$

$$\widetilde{\widetilde{\boldsymbol{M}}} = \boldsymbol{\Lambda} \oplus \boldsymbol{M} + \boldsymbol{\lambda}\boldsymbol{\alpha}^T \otimes \boldsymbol{I},$$

$$\boldsymbol{F}_1 = -(\boldsymbol{\Lambda} \otimes \boldsymbol{\beta}^T),$$

$$\boldsymbol{F}_2 = \boldsymbol{\lambda}\boldsymbol{\alpha}^T \otimes \boldsymbol{\beta}^T,$$

$$\boldsymbol{F}_3 = \widetilde{\boldsymbol{\Lambda}} - (\boldsymbol{\lambda}\boldsymbol{\alpha}^T \otimes \boldsymbol{I}),$$

$$\boldsymbol{F}_4 = \boldsymbol{\lambda}\boldsymbol{\alpha}^T \otimes \boldsymbol{I},$$

$$\boldsymbol{F}_5 = \boldsymbol{\lambda}\boldsymbol{\alpha}^T \otimes \mathbf{1}\boldsymbol{\beta}^T,$$

$$\boldsymbol{W}_1 = \boldsymbol{F}_1 \widetilde{\boldsymbol{M}}^{-1},$$

$$\boldsymbol{W}_n = \left[A_{n-1}\boldsymbol{F}_2 + \sum_{k=1}^{n-1} \boldsymbol{W}_k (a_{n-k}\boldsymbol{F}_4 + A_{n-k}\boldsymbol{F}_5) + \boldsymbol{W}_{n-1}\boldsymbol{F}_3 \right] \widetilde{\boldsymbol{M}}^{-1},$$

$$\boldsymbol{W}_{r+1} = -\left(a_{r+1}\boldsymbol{F}_2 + \sum_{k=1}^{r} A_{r-k}\boldsymbol{W}_k \boldsymbol{F}_4 \right) \widetilde{\widetilde{\boldsymbol{M}}}^{-1},$$

$$\boldsymbol{v} = \mathbf{1} + \sum_{k=1}^{r+1} \boldsymbol{W}_k (\mathbf{1} \otimes \mathbf{1}),$$

$$\boldsymbol{Z} = \left[-\boldsymbol{W}_r \boldsymbol{F}_3 + \boldsymbol{W}_{r+1} (\widetilde{\boldsymbol{M}} + \widetilde{\widetilde{\boldsymbol{M}}}) \right] (\boldsymbol{I} \otimes \mathbf{1}).$$

The following theorem is proved in [19]:

Theorem 1. *The stationary probability distribution* $\{\boldsymbol{p}_n, n = \overline{0, r+1}\}$ *for QS* $PH^{[x]}/PH/1/r$ *is determined by the expressions:*

$$\boldsymbol{p}_n^T = \boldsymbol{p}_0^T \boldsymbol{W}_n, \quad n = \overline{1, r+1}, \tag{4}$$

where $\boldsymbol{p}_0$ *is the only solution to the system of equations:*

$$\boldsymbol{p}_0^T \boldsymbol{Z} = \mathbf{0}^T, \tag{5}$$

$$\boldsymbol{p}_0^T \boldsymbol{v} = 1. \tag{6}$$

4. Laplace–Stiltjes transformation of the peak age of information

First, we will find the distribution of the time spent by the group of applications in the system under consideration. To do this, we construct a Markov chain (CM) $\{X_k^- = X(t_k - 0),\ k \geq 1\}$ nested at the moments $t_k - 0$ of the arrival of groups of appearances, over the set of states

$$\mathscr{X}_A^- = \{(0), (n, j),\ n = \overline{1, r+1},\ j = \overline{1, m}\}.$$

We denote by q_0 and q_{nj} the stationary probabilities of the states of the CM (0) and (n, j) and introduce the vectors:

$$\boldsymbol{q}_n^T = (q_{n1}, ..., q_{nm}).$$

From [20] we get:

$$q_0 = \frac{1}{\lambda}\boldsymbol{p}_0^T\boldsymbol{\lambda}, \tag{7}$$

$$\boldsymbol{q}_n^T = \frac{1}{\lambda}\boldsymbol{p}_n^T(\boldsymbol{\lambda} \otimes \boldsymbol{I}), \quad n = \overline{1, r+1}. \tag{8}$$

Letq_{xk} be $x \in \mathscr{X}_A^-$ there is a stationary probability that at the moment $t - 0$ before the group arrived, the system was in state x, and a group of k applications entered the system.

Next, we introduce the vectors:

$$\boldsymbol{q}_{nk}^T = (q_{(n,1),k}, ..., q_{(n,m),k}).$$

It is obvious that

$$q_{0k} = q_0 a_k, \quad k = \overline{1, r+1}, \tag{9}$$

$$\boldsymbol{q}_{nk}^T = \boldsymbol{q}_n^T a_k, \quad k = \overline{1, r+1}, \tag{10}$$

$$\boldsymbol{q}_{n,r+1-n}^T = \boldsymbol{q}_n^T A_{r-n}, \quad n = \overline{1, r}. \tag{11}$$

Let us denote by $W(t|(0), k)$, $k = \overline{1, r+1}$, the conditional distribution function (CDF) of the time spent in the system by a group of k applications, provided that at the time $t - 0$ the group was received in the system 0 applications.

Through $W(t|(n, j), k)$, $k = \overline{1, r+1-n}$ is the CDF of the time spent in the system by a group of k applications, provided that at the time $t - 0$ of the group's receipt, there were n applications in the system, and the application was serviced on the device phase j, $j = \overline{1, m}$.

The Laplace–Stieltjes transform (LST) CDF $W(t|(0), k)$ and $W(t|(n, j), k)$ are denoted by $w_0(s|k)$ and $w_n(s|j, k)$ and we introduce vectors

$$\boldsymbol{w}_n^T(s|k) = (w_n(s|1, k), ..., w_n(s|m, k)).$$

Note that if at the moment $t - 0$ of the group's receipt the system was in the state (0), then the first application from the group immediately goes to service, and the entire group of size k is serviced after the last service is completed (k-th) applications from this group. Therefore

$$w_0(s|k) = \beta^k(s),$$

where $\beta(s)$ is the LST of the CDF $B(t)$, is determined by the formula

$$\beta(s) = \boldsymbol{\beta}^T(s\boldsymbol{I} - \boldsymbol{M})^{-1}\boldsymbol{\mu}. \tag{12}$$

If the incoming group finds the system in the state (n, j), $n = \overline{1,r}$, then the service of the first application from the group will begin after the service of the application on the device is completed, $n-1$ application from queues and all k applications of this group will be served

Thus, we have:

$$\boldsymbol{w}_n(s|k) = (s\boldsymbol{I} - \boldsymbol{M})^{-1}\boldsymbol{\mu}(\beta(s))^{n-1+k}. \tag{13}$$

And finally, it should be taken into account that if a group finds the system in one of the states $(r+1, j)$, then the entire group is lost. Consequently, in accordance with (8), the probability of loss of the group π is determined through the probabilities of the states $(r+1, j)$ of the CM nested at the moments $t-0$ of the application groups as follows:

$$\pi = \frac{1}{\lambda}\boldsymbol{p}_{r+1}^T(\boldsymbol{\lambda} \otimes \mathbf{1}). \tag{14}$$

Summing up our reasoning and applying the formula of complete certainty, we come to the following result:

Theorem 2. *LST of the residence time of a group of applications accepted in the QS $PH^{[x]}/PH/1/r$ is defined by the expression:*

$$w(s) = \frac{1}{1-\pi}\left[\sum_{k=1}^{r+1} a_k q_0 \beta^k(s) + \sum_{n=1}^{r}\left[\sum_{k=1}^{r+1-n} a_k \boldsymbol{q}_n^T(s\boldsymbol{I} - \boldsymbol{M})^{-1}\boldsymbol{\mu}(\beta(s))^{n-1+k} + \right.\right.$$
$$\left.\left. + A_{r+1-n}\boldsymbol{q}_n^T(s\boldsymbol{I} - \boldsymbol{M})^{-1}\boldsymbol{\mu}(\beta(s))^r\right]\right]. \tag{15}$$

As already noted above, the peak age of the Z_n of the n-th group of turnouts is equal to the sum of two terms $\hat{G}_{n+1}$ – the duration of generation of groups following the nth and completed the first successful connection to the queue and W_{n+1} is the time spent by the $(n+1)$-th group in the system. It is obvious that

$$\hat{G}_{n+1} = \frac{1}{1-\pi}G_{n+1},$$

where G_{n+1} is the duration of generation of the $(n+1)$-th group with CDF $A(t)$ and LST

$$\alpha(s) = \boldsymbol{\alpha}^T(s\boldsymbol{I} - \boldsymbol{\Lambda})^{-1}\boldsymbol{\lambda}. \tag{16}$$

Considering the independence of $\hat{G}_{n+1}$ and W_{n+1}, their independence from n and the fact that the LST of $\frac{1}{1-\pi}G$ is equal to $\alpha\left(\frac{s}{1-\pi}\right)$, we come to the following result.

Theorem 3. *The LST of the peak age of information transmitted by a group of applications in the $PH^{[x]}/PH/1/r$ system is determined by the expression:*

$$z(s) = \alpha\left(\frac{s}{1-\pi}\right)w(s), \tag{17}$$

where $w(s)$ and $\alpha\left(\frac{s}{1-\pi}\right)$ calculated according to (15)–(16)

5. LST of the peak age information for QS $M^{[x]}/M/1/r$

As an example, let us consider a special case – QS with a group Poisson flow of applications of intensity λ and exponential service of intensity μ.

In this case, the parameters of the initial system will be determined by the expressions:

$$\boldsymbol{\Lambda} = (-\lambda), \quad \boldsymbol{\lambda} = (\lambda), \quad \boldsymbol{\alpha} = (1), \quad \alpha(s) = \frac{\lambda}{\lambda + s}, \tag{18}$$

$$\boldsymbol{M} = (-\mu), \quad \boldsymbol{\mu} = (\mu), \quad \boldsymbol{\beta} = (1), \quad \beta(s) = \frac{\mu}{\mu + s}, \tag{19}$$

and the algorithm of Theorem 1 is expressed in the form of a recurrent formula:

$$p = \rho \sum_{k=0}^{n-1} p_k A_{n-k-1}, \quad n = \overline{1, r+1}, \tag{20}$$

where $\rho = \lambda/\mu$ and p_n are the stationary probabilities that there is n applications. The probability p_0 is determined from the normalization condition.

In accordance with (7),(8) and (14), we obtain:

$$q_n = p_n, \quad n = \overline{0, r+1},$$

$$\pi = p_{r+1}. \tag{21}$$

And finally, based on (15) and (17), taking into account (18)–(21), we come to the following result:

Theorem 4. *LST of the peak age of information transmitted the group of applications in the $M^{[x]}/M/1/r$ system is defined by the expression:*

$$z(s) = \alpha\left(\frac{s}{1 - p_{r+1}}\right) w(s),$$

where $\alpha(s) = \dfrac{\lambda}{\lambda + s}$;

$$w(s) = \frac{1}{1 - p_{r+1}} \left[\sum_{k=0}^{r+1} a_k p_0 \left[\frac{\mu}{\mu + s}\right]^k + \sum_{n=1}^{r} \sum_{k=1}^{r+1} a_k p_n \times \right.$$
$$\left. \times \left[u(r+2-n-k) \left[\frac{\mu}{\mu + s}\right]^{n+k} + u(k-(r+1-n)) \left[\frac{\mu}{\mu + s}\right]^{r+1} \right] \right]. \tag{22}$$

$$u(x) = \begin{cases} 1, & x > 0, \\ 0, & x \le 0, \end{cases} \qquad \textit{is the Heaviside function.}$$

Corollary 1. *The average value MZ of the peak age of information transmitted by a group of applications in the $M^{[x]}/M/1/r$ system is determined by the expression:*

$$MZ = \frac{1}{1 - p_{r+1}} \left[\frac{1}{\lambda} + \sum_{k=1}^{r+1} p_0 a_k \frac{k}{\mu} + \sum_{n=1}^{r} p_n \sum_{k=1}^{r+1} a_k \left[u(r+2-n-k) \frac{n+k}{\mu} + \right. \right.$$
$$\left. \left. + u(k-(r+1-n)) \frac{r+1}{\mu} \right] \right]. \tag{23}$$

The proof is based on the application of the formula:

$$MZ = -z'(0).$$

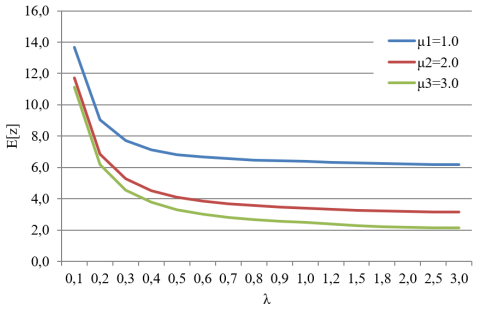

Figure 2. Dependence of the average peak age of information on the intensity of the incoming flow λ at fixed values of μ

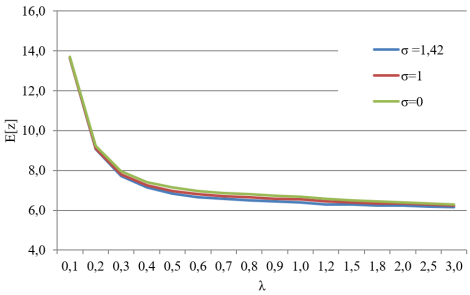

Figure 3. The dependence of the average peak age of information on the intensity of the incoming flow λ for different distributions of the number of applications in the group

6. The numerical results

We conducted a numerical study of the peak age of information for QS $M^{[x]}/M/1/r$ the results of which are shown in fig. 2 and fig. 3.

Figure 2 shows the dependence of the average peak increase on the intensity of the flow λ for three different values of the service intensity μ: $\mu = 1.0$, $\mu = 2.0$, $\mu = 3.0.0$ and storage capacity $r = 4$. At the same time, the distribution of the number of applications in the group was set by the formula:

$$\mathrm{P}\{\eta = k\} = \frac{1}{5}, \quad k = \overline{1,5}. \tag{24}$$

Note that the variant with $\mu = 1.0$ was calculated in two ways: analytically and imitatively. As can be seen from Figure 2, with the growth of λ, there is a decrease in the peak age. Moreover, the larger the μ, the lower the average age, which is quite expected, because with large λ and μ, the transfer of information from the SN and RN occurs more often and faster.

Figure 3 reflects the dependence of the average peak age of information on the intensity of the flow λ for $\mu = 1.0$ and $r = 4$ for different variations of the distribution of the number of applications in the group. So in option 1, the distribution of the number of applications in the group is given by the formula (24). In option 2:

$$\mathrm{P}\{\eta = k\} = \begin{cases} \frac{1}{2}, & k = 2,4; \\ 0, & k = 1,3,5. \end{cases}$$

In option 3, the number of applications in the group is constant and equal to 3. Obviously, for all three options, the average number of applications in the group is the same and equal to 3, but the standard deviation is different: $\sigma_1 = \sqrt{2}$; $\sigma_2 = 1$; $\sigma_3 = 0$. Despite the fairly close values of the average peak age for all three variants, there is a well-defined dependence: the greater the standard deviation, the lower the value of the average peak age. What is the reason for this pattern, it is not yet clear to us. Perhaps this is due to the loss of applications due to a limited storage device. In any case, this issue remains a topic for discussion.

7. Conclusion

As a result of the conducted research, we managed to obtain an expression for the Laplace–Stieltjes transformation of the stationary distribution function of the peak age of information transmitted from a peripheral source to the control center, modeling the transmission process using a queuing system with a group flow of applications and with phase-type distributions. This study allows us to obtain fairly accurate estimates of the age of information for real technical systems, due to the versatility of phase-type distributions.

Author Contributions: Conceptualization, S.I. Matyushenko and K.E. Samouylov; methodology, S.I. Matyushenko; validation, S.I. Matyushenko; writing—original draft preparation, S.I. Matyushenko; writing—review and editing, S.I. Matyushenko and K.E. Samouylov; visualization, S.I. Matyushenko; supervision, K.E. Samouylov; project administration, K.E. Samouylov. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Data sharing is not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Sultan, A. *Ultra Reliable and Low Latency Communications. 3GPP* 2023.
2. Kamoun, F. & Kleinrock, L. Analysis of Shared Finite Storage in a Computer Network Node Environment Under General Traffic Conditions. *IEEE Transactions on Communications* **28,** 992–1003. doi:10.1109/TCOM.1980.1094756 (1980).
3. Baskett, F., Chandy, K. M., Muntz, R. R. & Palacios, F. G. Open, Closed, and Mixed Networks of Queues with Different Classes of Customers. *Journal of the ACM* **22,** 248–260. doi:10.1145/321879.321887 (1975).
4. Bedewy, A. M., Sun, Y. & Shroff, N. B. *Age-optimal information updates in multihop networks* in *2017 IEEE International Symposium on Information Theory (ISIT)* (2017), 576–580. doi:10.1109/ISIT.2017.8006593.
5. Bojan, T. M., Kumar, U. R. & Bojan, V. M. *An internet of things based intelligent transportation system* in *2014 IEEE International Conference on Vehicular Electronics and Safety* (2014), 174–179. doi:10.1109/ICVES.2014.7063743.
6. Hu, C. & Dong, Y. Age of information of two-way data exchanging systems with power-splitting. *Journal of Communications and Networks* **21,** 295–306. doi:10.1109/JCN.2019.000037 (2019).
7. Costa, M., Codreanu, M. & Ephremides, A. *Age of information with packet management* in *IEEE International Symposium on Information Theory (ISIT)* (2014), 1583–1587.
8. Chiariotti, F., Vikhrova, O., Soret, B. & Popovski, P. Peak Age of Information Distribution for Edge Computing With Wireless Links. *IEEE Transactions on Communications* **69,** 3176–3191. doi:10.1109/TCOMM.2021.3053038 (2021).
9. Kadota, I., Uysal-Biyikoglu, E., Singh, R. & Modiano, E. *Minimizing the Age of Information in broadcast wireless networks* in *2016 54th Annual Allerton Conference on Communication, Control, and Computing (Allerton)* (2016), 844–851. doi:10.1109/ALLERTON.2016.7852321.
10. Kosta, A., Pappas, N., Ephremides, A. & Angelakis, V. *Non-linear Age of Information in a Discrete Time Queue: Stationary Distribution and Average Performance Analysis* in *ICC 2020 — 2020 IEEE International Conference on Communications (ICC)* (2020), 1–6. doi:10.1109/ICC40277.2020.9148775.

11. Talak, R., Karaman, S. & Modiano, E. Improving Age of Information in Wireless Networks With Perfect Channel State Information. *IEEE/ACM Transactions on Networking* **28,** 1765–1778. doi:10.1109/TNET.2020.2996237 (2020).
12. Tripathi, V., Talak, R. & Modiano, E. *Age of information for discrete time queues* 2019.
13. Wijerathna Basnayaka, C. M., Jayakody, D. N. K., Ponnimbaduge Perera, T. D. & Vidal Ribeiro, M. *Age of Information in an URLLC-enabled Decode-and-Forward Wireless Communication System* in *2021 IEEE 93rd Vehicular Technology Conference (VTC2021-Spring)* (2021), 1–6. doi:10.1109/VTC2021-Spring51267.2021.9449007.
14. Zhbankova, E., Khakimov, A., Markova, E. & Gaidamaka, Y. The Age of Information in Wireless Cellular Systems: Gaps, Open Problems, and Research Challenges. *Sensors* **23.** doi:10.3390/s23198238 (2023).
15. Kaul, S., Yates, R. & Gruteser, M. *Real-time status: How often should one update?* in *2012 Proceedings IEEE INFOCOM* (2012), 2731–2735. doi:10.1109/INFCOM.2012.6195689.
16. Costa, M., Codreanu, M. & Ephremides, A. On the Age of Information in Status Update Systems With Packet Management. *IEEE Transactions on Information Theory* **62,** 1897–1910. doi:10.1109/TIT.2016.2533395 (2016).
17. Kaul, S. K., Yates, R. D. & Gruteser, M. *Status updates through queues* in *2012 46th Annual Conference on Information Sciences and Systems (CISS)* (2012), 1–6. doi:10.1109/CISS.2012.6310931.
18. Bocharov, P. P. & Pechinkin, A. V. *Queueing Theory [Teoriya massovogo obsluzhivaniya]* in Russian. 529 pp. (Izd-vo RUDN, Moscow, 1995).
19. Bocharov, P. P. & Yakutina, S. V. Stationary queue distribution in a finite capacity service system with group flow and phase-type service time [Stacionarnoe raspredelenie ocheredi v sisteme obslugivaniya konechnoy emkosty s gruppovim potokom i vremenem obslugivaniya fazovogo tipa]. *Avtomatika i Telemekhanika.* in Russian, 106–119 (1994).
20. Naumov, V. A. *O predelnih veroyatnostyah polumarkovskogo processa [On the limiting probabilities of a semi-Markov process]* in Russian. 35–39 (Universitet drugbi narodov, Moscow, 1975).

Information about the authors

Sergey I. Matyushenko—Candidate of Physical and Mathematical Sciences, Assistant professor of Department of Probability Theory and Cyber Security of RUDN University (e-mail: matyushenko-si@rudn.ru, ORCID: 0000-0001-8247-8988)

Konstantin E. Samouylov—Professor, Doctor of Technical Sciences, Head of the Department of Probability Theory and Cyber Security of RUDN University (e-mail: samuylov-ke@rudn.ru, ORCID: 0000-0002-6368-9680)

УДК 519.21
DOI: 10.22363/2658-4670-2024-32-3-260–270 EDN: EUVTRK

Распределение пикового возраста информации в двухузловой группе передачи, моделируемой системой обслуживания с групповым потоком и обслуживанием фазового типа

С. И. Матюшенко, К. Е. Самуйлов

Российский университет дружбы народов им. П. Лумумбы, ул. Миклухо-Маклая, д. 6, Москва, 117198, Российская Федерация

Аннотация. Данная статья продолжает цикл работ авторов, посвященных проблеме возраста информации (AoI) – метрики, используемой в информационных системах для мониторинга и управления удаленными источниками информации со стороны центра управления. Теоретический анализ систем передачи информации требует количественной оценки «свежести» информации, доставляемой в центр управления. Процесс передачи информации от периферийных источников к центру обычно моделируется с помощью систем массового обслуживания. В данной работе для оценки максимального значения возраста информации, называемого пиковым возрастом, используется система массового обслуживания с распределениями фазового типа. При этом учитывается специальное требование протокола передачи, состоящее в том, что информация в систему поступает группами случайного размера. Для данного случая получено выражение для преобразования Лапласа–Стилтьеса стационарной функции распределения пикового возраста информации и его среднего значения. По результатам аналитического моделирования проведено численное исследование зависимости среднего значения пикового возраста информации от загрузки системы. Корректность полученных выражений проверена путем сравнения аналитических результатов с результатами имитационного моделирования.

Ключевые слова: возраст информации, пиковый возраст информации, система массового обслуживания, распределение фазового типа, групповой поток

Discrete & Continuous Models
& Applied Computational Science
ISSN 2658-7149 (Online), 2658-4670 (Print)

2024, 32 (3) 271–282
http://journals.rudn.ru/miph

Computer science

Research article

UDC 519.21
DOI: 10.22363/2658-4670-2024-32-3-271–282
EDN: BAUGIT

Analysis of a queuing system of a single capacity with phase-type distributions and queue updating

Sergey I. Matyushenko, Konstantin E. Samouylov, Nikolai Yu. Gritsenko

RUDN University, 6 Miklukho-Maklaya St, Moscow, 117198, Russian Federation

(received: September 14, 2023; revised: September 30, 2023; accepted: October 1, 2023)

Abstract. In this paper, we study a queuing system with a single-capacity storage device and queue updating. An update is understood as the following mechanism: an application that enters the system and finds another application in the drive destroys it, taking its place in the drive. It should be noted that systems with one or another update mechanism have long attracted the attention of researchers, since they have important applied significance. Recently, interest in systems of this kind has grown in connection with the tasks of assessing and managing the age of information. A system with a queue update mechanism similar to the one we are considering has already been studied earlier in the works of other authors. However, in these works we were talking about the simplest version of the system with Poisson flow and exponential maintenance. In this paper, we consider a phase-type flow and maintenance system. As a result of our research, we developed a recurrent matrix algorithm for calculating the stationary distribution of states of a Markov process describing the stochastic behavior of the system in question, and obtained expressions for the main indicators of its performance.

Key words and phrases: queuing system, phase-type distribution, queue update mechanism

For citation: Matyushenko, S. I., Samouylov, K. E., Gritsenko, N. Y. Analysis of a queuing system of a single capacity with phase-type distributions and queue updating. *Discrete and Continuous Models and Applied Computational Science* **32** (3), 271–282. doi: 10.22363/2658-4670-2024-32-3-271–282. edn: BAUGIT (2024).

1. Introduction

The tasks related to the assessment and management of information age, which were initiated in [1–13], revived interest in the study of systems with various kinds of updating mechanisms. One of these systems is a system with a queue update mechanism [14], the essence of which is that an application entering the system and finding another application in the drive "kills" it and takes its place in the drive. This ensures that the information transmitted by the application is updated as quickly as possible, which is extremely important for real technical systems implementing service complexes for which the time factor plays the most important role. A system with this queue update mechanism was considered in [7, 15, 16]. However, the authors of these papers considered a system with Poisson flow and exponential maintenance, which, according to Kendall's notation, is usually encoded as $M/M/1/1$. It is known that such models of queuing systems allow us to obtain only rough estimates of the characteristics of real technical systems, since single-parameter distributions do not make it possible to take into account all the features of the protocols of modern dispatch

control and data collection systems, random multiple access from several remote sender nodes, multistep information transmission routes, etc. Therefore, in this paper we have followed the path of generalization, assuming that the time intervals between the receipt of applications and the duration of their service are random variables with phase-type distributions. This circumstance will allow us to subsequently use the universality of phase-type distributions to build more accurate models of real technical systems.

2. Description of the model

A single-line queuing system (QS) with a single-capacity storage device is considered, which receives a recurrent flow of applications with a phase-type distribution function (DF) $A(t)$:

$$A(t) = 1 - \boldsymbol{\alpha}^T e^{\boldsymbol{\Lambda} t}\mathbf{1}, \quad t \geq 0, \quad \boldsymbol{\alpha}^T\mathbf{1} = 1,$$

admitting an irreducible PH representation $(\boldsymbol{\alpha}, \boldsymbol{\Lambda})$ of order l [17]. The duration of the application service has a phase type DF $B(t)$ with an irreducible PH representation of the order m:

$$B(t) = 1 - \boldsymbol{\beta}^T e^{\mathbf{M} t}\mathbf{1}, \quad t \geq 0, \quad \boldsymbol{\beta}^T\mathbf{1} = 1.$$

Consider the QS with queue update. This means that an application that enters the system and finds the drive busy displaces the application from the drive and takes its place. The repressed application leaves the system and does not return to it anymore. In accordance with Kendall's notation, the QS in question will be encoded as $PH/PH/1/1$ with queue update (Fig. 1).

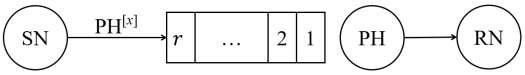

Figure 1. QS $PH/PH/1/1$ with queue update

3. Mathematical model

Based on the probabilistic interpretation of the PH distributions, the functioning of the QS under consideration is described by a homogeneous Markov process (MP) $\{X(t),\ t \geq 0\}$ over the state space

$$\mathcal{X} = \bigcup_{k=0}^{2} \mathcal{X}_k,$$

where $\mathcal{X}_0 = \{(i,0), i = \overline{1,l}\}$, $\mathcal{X}_k = \{(k,i,j), i = \overline{1,l}, j = \overline{1,m}\}$, $k = 1,2$.

Here $X(t) = (i,0)$ if at time t the system is empty and the process of generating a new application is in phase i. In turn, equality $X(t) = (k,i,j)$ means that in the system of k applications, the process of generating a new application is in phase i, and maintenance is in phase j.

It follows from the irreducibility of PH distributions [17] that all states of the process $\{X(t),\ t \geq 0\}$ are reported, the process itself is ergodic, and the limiting probabilities

$$p_{i0} = \lim_{t\to\infty} P\{X(t) = (i,0)\},$$
$$p_{ikj} = \lim_{t\to\infty} P\{X(t) = (i,k,j)\},$$

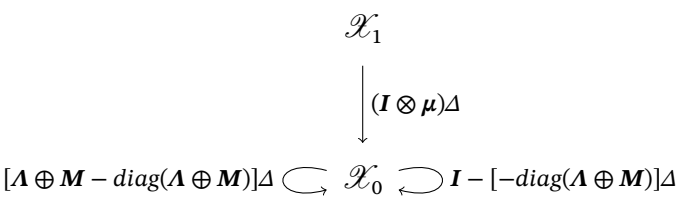

Figure 2. Diagram of transitions of MP $X(t)$ to states $\mathscr{X}_0$ for Δ

strictly positive, independent of the initial distribution, and consistent with stationary probabilities.

Let's introduce vectors:

$$\boldsymbol{p}_0^T = (p_{10}, \dots, p_{l0}),$$
$$\boldsymbol{p}_k^T = (p_{1k1}, \dots, p_{1km}, \dots, p_{lk1}, \dots, p_{lkm}), \quad k = 1, 2.$$

Stationary probabilities $\{\boldsymbol{p}_k,\ k = 0, 1, 2\}$ are the only solution to the system of equilibrium equations (SEE):

$$\mathbf{0}^T = \boldsymbol{p}_0^T\boldsymbol{\Lambda} + \boldsymbol{p}_1^T(\boldsymbol{I}\otimes\boldsymbol{\mu}), \tag{1}$$

$$\mathbf{0}^T = \boldsymbol{p}_0^T(\boldsymbol{\lambda}\boldsymbol{\alpha}^T\otimes\boldsymbol{\beta}^T) + \boldsymbol{p}_1^T(\boldsymbol{\Lambda}\oplus\boldsymbol{M}) + \boldsymbol{p}_2^T(\boldsymbol{I}\otimes\boldsymbol{\mu}\boldsymbol{\beta}^T), \tag{2}$$

$$\mathbf{0}^T = \boldsymbol{p}_1^T(\boldsymbol{\lambda}\boldsymbol{\alpha}^T\otimes\boldsymbol{I}) + \boldsymbol{p}_2^T(\boldsymbol{\Lambda}\oplus\boldsymbol{M}+\boldsymbol{\lambda}\boldsymbol{\alpha}^T\otimes\boldsymbol{I}), \tag{3}$$

with the condition of normalization

$$\sum_{k=0}^{2}\boldsymbol{p}_k^T\mathbf{1} = 1. \tag{4}$$

Hereafter $\boldsymbol{\lambda} = -\boldsymbol{\Lambda}\mathbf{1}$, $\boldsymbol{\mu} = -\boldsymbol{M}\mathbf{1}$, the sign $\otimes$ means the Kronecker product, and the sign $\oplus$ means the Kronecker sum of matrices.

Let's explain the conclusion of the SEE (1)–(3) using the MP transition scheme $X(t)$ on the interval $(t, t+\Delta)$, where Δ is a "small" time interval.

The subset of states $\mathscr{X}_0$ can be accessed from the subset $\mathscr{X}_1$ due to the end of the application service on the device with an intensity characterized by the vector $\boldsymbol{\mu}$ (Fig. 2). In addition to this transition, Fig. 2 shows two more situations in which the process does not go beyond the subset $\mathscr{X}_0$: during Δ there will be no change in the generation phases, or vice versa — the generation phases change.

The first situation is reflected by the elements of the main diagonal of the matrix $\boldsymbol{\Lambda}$, taken with the opposite sign, which we will denote $diag(\boldsymbol{\Lambda})$, and the second is the non-diagonal elements of this matrix, which we will denote $\boldsymbol{\Lambda} - diag(\boldsymbol{\Lambda})$. In a subset of states $\mathscr{X}_1$, during Δ, it is possible to get from the subset $\mathscr{X}_0$ due to the receipt of a new application, which occurs with an intensity characterized by the vector $\boldsymbol{\lambda}$. At the same time, we take into account that the generation of the next application immediately begins, and the choice of the initial phase occurs in accordance with the probability vector $\boldsymbol{\alpha}$. In addition, the subset $\mathscr{X}_1$ during Δ can be accessed from the subset $\mathscr{X}_2$ by the end of the service with an intensity determined by the vector $\boldsymbol{\mu}$. In this case, the initial phase of the service of the next application is selected in accordance with the probabilistic vector $\boldsymbol{\beta}$. The third possibility is to remain in this subset due to the fact that the passage of the current generation and maintenance phases will not be completed during Δ. This possibility is reflected by the intensities equal to the

elements of the main diagonal of the matrix $\boldsymbol{\Lambda} \oplus \boldsymbol{M}$, taken with the opposite sign. Either due to Δ there will be changes in the phases of generation or maintenance, which reflect intensities equal to the non-diagonal elements of the matrix $\boldsymbol{\Lambda} \oplus \boldsymbol{M}$ (Fig. 3).

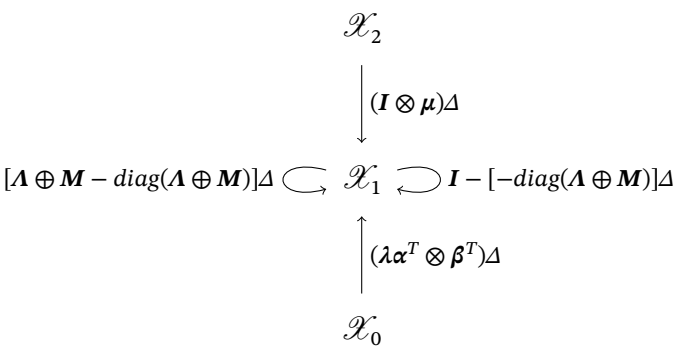

Figure 3. Diagram of transitions of MP $X(t)$ to states $\mathscr{X}_1$ for Δ

Now let's explain the derivation of equation (3). In the subset $\mathscr{X}_2$ for the time Δ you can get in by receiving a new application as from a subset $\mathscr{X}_1$, while remaining inside a subset of $\mathscr{X}_2$. In both cases, the given transition occurs with intensities equal to the corresponding coordinates of the vector $\boldsymbol{\lambda}$, and ends with the choice of a new generation phase in accordance with the vector $\boldsymbol{\alpha}$ (Fig. 4). In addition, as in the previous case, there are two possibilities to remain in the subset $\mathscr{X}_2$: due to the fact that there are no changes in Δ it will happen, or only a change of generation or maintenance phases will occur.

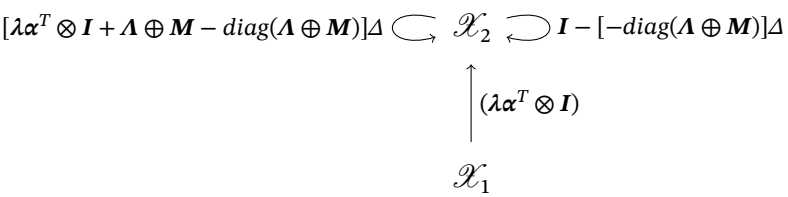

Figure 4. Diagram of transitions of MP $X(t)$ to states $\mathscr{X}_1$ for Δ

4. Solution of the SEE

Let's move on to solving the SEE (1)–(4), noting that we are interested not in numerical, but in the analytical solution of a system of equations, that is, analytical expressions in explicit form both to determine the performance indicators of the system itself and to obtain similar results for numerous special cases of the system under consideration. Before proceeding to the direct solution of the SEE, we introduce notation and prove the validity of a number of auxiliary correlations.

Let's introduce the matrices

$$\boldsymbol{V}_0 = \boldsymbol{\Lambda} \otimes (\boldsymbol{1}\boldsymbol{\beta}^T - \boldsymbol{I}) - \boldsymbol{I} \otimes \boldsymbol{M}, \tag{5}$$

$$\boldsymbol{V}_1 = \boldsymbol{\Lambda} \oplus \boldsymbol{M} + \boldsymbol{\lambda}\boldsymbol{\alpha}^T \otimes \boldsymbol{I}, \tag{6}$$

$$\boldsymbol{W}_0 = -(\boldsymbol{\Lambda} \otimes \boldsymbol{\beta}^T)\boldsymbol{V}_0^{-1}, \tag{7}$$

$$\boldsymbol{W}_1 = -(\boldsymbol{\lambda}\boldsymbol{\alpha}^T \otimes \boldsymbol{I})\boldsymbol{V}_1^{-1}, \tag{8}$$

$$\boldsymbol{W} = [(\boldsymbol{I} + \boldsymbol{W}_0 + \boldsymbol{W}_0\boldsymbol{W}_1)(\boldsymbol{I} \otimes \boldsymbol{1})]. \tag{9}$$

Let's prove that the following lemmas are valid.

Lemma 1.

$$\boldsymbol{V}_0(\boldsymbol{I} \otimes \boldsymbol{1}\boldsymbol{\beta}^T) = \boldsymbol{I} \otimes \boldsymbol{\mu}\boldsymbol{\beta}^T. \tag{10}$$

Proof. Given (5), we get:

$$\boldsymbol{V}_0(\boldsymbol{I} \otimes \boldsymbol{1}\boldsymbol{\beta}^T) = \big(\boldsymbol{\Lambda} \otimes (\boldsymbol{1}\boldsymbol{\beta}^T - \boldsymbol{I}) - \boldsymbol{I} \otimes \boldsymbol{M}\big)(\boldsymbol{I} \otimes \boldsymbol{1}\boldsymbol{\beta}^T) = \big(\boldsymbol{\Lambda} \otimes \boldsymbol{1}\boldsymbol{\beta}^T\boldsymbol{1}\boldsymbol{\beta}^T\big) - (\boldsymbol{\Lambda} \otimes \boldsymbol{1}\boldsymbol{\beta}) - (\boldsymbol{I} \otimes \boldsymbol{M}\boldsymbol{1}\boldsymbol{\beta}^T).$$

Further, noting that $\boldsymbol{\beta}^T\boldsymbol{1} = 1$ and $\boldsymbol{M}\boldsymbol{1} = -\boldsymbol{\mu}$, we obtain the right part (10). □

Lemma 2.

$$-(\boldsymbol{\Lambda} \otimes \boldsymbol{\beta}^T)(\boldsymbol{1}\boldsymbol{\alpha}^T \otimes \boldsymbol{\beta}^T). \tag{11}$$

Proof. Obviously, if we consider that $\boldsymbol{\Lambda}\boldsymbol{1} = -\boldsymbol{\lambda}$. □

Lemma 3.

$$-\boldsymbol{p}_0^T(\boldsymbol{\Lambda} \otimes \boldsymbol{\beta}^T) = \boldsymbol{p}_1^T\boldsymbol{V}_0. \tag{12}$$

Proof. Multiplying equation (2) on the right by the matrix $\boldsymbol{I} \otimes (\boldsymbol{1}\boldsymbol{\beta}^T - \boldsymbol{I})$, we get:

$$\boldsymbol{0}^T = \boldsymbol{p}_0^T\big(\boldsymbol{\lambda}\boldsymbol{\alpha}^T \otimes \boldsymbol{\beta}^T\big)\big(\boldsymbol{I} \otimes (\boldsymbol{1}\boldsymbol{\beta}^T - \boldsymbol{I})\big) + \boldsymbol{p}_1^T(\boldsymbol{\Lambda} \oplus \boldsymbol{M})\big(\boldsymbol{I} \otimes (\boldsymbol{1}\boldsymbol{\beta}^T - \boldsymbol{I})\big) + \boldsymbol{p}_2^T(\boldsymbol{I} \otimes \boldsymbol{\mu}\boldsymbol{\beta}^T)\big(\boldsymbol{I} \otimes (\boldsymbol{1}\boldsymbol{\beta}^T - \boldsymbol{I})\big).$$

Let's consider each term of the right part separately:

$$\boldsymbol{p}_0^T(\boldsymbol{\lambda}\boldsymbol{\alpha}^T \otimes \boldsymbol{\beta}^T)\big(\boldsymbol{I} \otimes (\boldsymbol{1}\boldsymbol{\beta}^T - \boldsymbol{I})\big) = \boldsymbol{p}_0^T(\boldsymbol{\lambda}\boldsymbol{\alpha}^T \otimes \boldsymbol{\beta}^T\boldsymbol{1}\boldsymbol{\beta}^T) - \boldsymbol{p}_0^T(\boldsymbol{\lambda}\boldsymbol{\alpha}^T \otimes \boldsymbol{\beta}^T) =$$
$$= \boldsymbol{p}_0^T(\boldsymbol{\lambda}\boldsymbol{\alpha}^T \otimes \boldsymbol{\beta}^T) - \boldsymbol{p}_0^T(\boldsymbol{\lambda}\boldsymbol{\alpha}^T \otimes \boldsymbol{\beta}^T) = \boldsymbol{0}^T,$$

$$\boldsymbol{p}_1^T(\boldsymbol{\Lambda} \oplus \boldsymbol{M})\big(\boldsymbol{I} \otimes (\boldsymbol{1}\boldsymbol{\beta}^T - \boldsymbol{I})\big) = \boldsymbol{p}_1^T(\boldsymbol{\Lambda} \otimes \boldsymbol{I} + \boldsymbol{I} \otimes \boldsymbol{M})\big(\boldsymbol{I} \otimes (\boldsymbol{1}\boldsymbol{\beta}^T - \boldsymbol{I})\big) =$$
$$= \boldsymbol{p}_1^T\big(\boldsymbol{\Lambda} \otimes (\boldsymbol{1}\boldsymbol{\beta}^T - \boldsymbol{I}) + \boldsymbol{I} \otimes \boldsymbol{M}(\boldsymbol{1}\boldsymbol{\beta}^T - \boldsymbol{I})\big) =$$
$$= \boldsymbol{p}_1^T\big(\big(\boldsymbol{\Lambda} \otimes (\boldsymbol{1}\boldsymbol{\beta}^T - \boldsymbol{I})\big) - \boldsymbol{I} \otimes \boldsymbol{\mu}\boldsymbol{\beta}^T - \boldsymbol{I} \otimes \boldsymbol{M}\big) = \boldsymbol{p}_1^T\big(\boldsymbol{V}_0 - \boldsymbol{I} \otimes \boldsymbol{\mu}\boldsymbol{\beta}^T\big),$$

$$\boldsymbol{p}_2^T(\boldsymbol{I} \otimes \boldsymbol{\mu}\boldsymbol{\beta}^T)\big(\boldsymbol{1} \otimes (\boldsymbol{1}\boldsymbol{\beta}^T - \boldsymbol{I})\big) = \boldsymbol{p}_2^T(\boldsymbol{I} \otimes \boldsymbol{\mu}\boldsymbol{\beta}^T\boldsymbol{1}\boldsymbol{\beta}^T) - \boldsymbol{p}_2^T(\boldsymbol{I} \otimes \boldsymbol{\mu}\boldsymbol{\beta}^T) = \boldsymbol{0}^T.$$

So,

$$\boldsymbol{0}^T = \boldsymbol{p}_1^T(\boldsymbol{V}_0 - \boldsymbol{I} \otimes \boldsymbol{\mu}\boldsymbol{\beta}^T).$$

Therefore,

$$\boldsymbol{p}_1^T\boldsymbol{V}_0 = \boldsymbol{p}_1^T(\boldsymbol{I} \otimes \boldsymbol{\mu}\boldsymbol{\beta}).$$

Next, multiplying equation (1) on the right by $(\boldsymbol{I} \otimes \boldsymbol{\beta}^T)$, we obtain

$$\boldsymbol{0}^T = \boldsymbol{p}_0^T(\boldsymbol{\Lambda} \otimes \boldsymbol{\beta}^T) + \boldsymbol{p}_1^T(\boldsymbol{I} \otimes \boldsymbol{\mu}\boldsymbol{\beta}^T).$$

Given the obtained equality in the expression for $\boldsymbol{p}_1^T\boldsymbol{V}_0$, we come to the formula (12). Thus, the lemma is proved. □

The main result of this section is formulated in the form of a theorem.

Theorem 5. *The stationary distribution* $\{p_x, x \in \mathscr{X}\}$ *is determined by the formulas:*

$$\boldsymbol{p}_0^T = \boldsymbol{q}^T \boldsymbol{W}^{-1}, \tag{13}$$

$$\boldsymbol{p}_1 = \boldsymbol{p}_0^T \boldsymbol{W}_0, \tag{14}$$

$$\boldsymbol{p}_2^T = \boldsymbol{p}_1^T \boldsymbol{W}_1, \tag{15}$$

where $\boldsymbol{q}$ *is the only solution to the system of equations*

$$\boldsymbol{q}^T(\boldsymbol{\Lambda} + \lambda\boldsymbol{\alpha}^T) = \boldsymbol{0}^T, \tag{16}$$

$$\boldsymbol{q}^T\boldsymbol{1} = 1. \tag{17}$$

Proof. First, we substitute (14) and (15) into the equations (1)–(3) and let's make sure that after substitution they turn into identities.

Let's start with equation (1). Multiply it on the right by the matrix $(\boldsymbol{I} \otimes \boldsymbol{\beta}^T)$. As a result, we get:

$$\boldsymbol{0}^T = \boldsymbol{p}_0^T(\boldsymbol{\Lambda} \otimes \boldsymbol{\beta}^T) + \boldsymbol{p}_0^T \boldsymbol{W}_0(\boldsymbol{I} \otimes \mu\boldsymbol{\beta}^T).$$

Considering (7) and (10), we arrive at the identity for equation (1):

$$\boldsymbol{0}^T = \boldsymbol{p}_0^T(\boldsymbol{\Lambda} \otimes \boldsymbol{\beta}^T) - \boldsymbol{p}_0^T(\boldsymbol{\Lambda} \otimes \boldsymbol{\beta}^T)\boldsymbol{V}_0^{-1}\boldsymbol{V}_0(\boldsymbol{I} \otimes \boldsymbol{1}\boldsymbol{\beta}^T).$$

Next, consider equation (2), which, taking into account (11) and (15), is written as:

$$\boldsymbol{0}^T = -\boldsymbol{p}_0^T(\boldsymbol{\Lambda} \otimes \boldsymbol{\beta}^T)(\boldsymbol{1}\boldsymbol{\alpha}^T \otimes \boldsymbol{I}) + \boldsymbol{p}_1^T(\boldsymbol{\Lambda} \oplus \boldsymbol{M}) + \boldsymbol{p}_1^T \boldsymbol{W}_1(\boldsymbol{I} \otimes \mu\boldsymbol{\beta}^T).$$

Next, taking into account (10) and (12), we get:

$$\boldsymbol{0}^T = \boldsymbol{p}_1^T \boldsymbol{V}_0(\boldsymbol{1}\boldsymbol{\alpha}^T \otimes \boldsymbol{I}) + \boldsymbol{p}_1^T(\boldsymbol{\Lambda} \oplus \boldsymbol{M}) + \boldsymbol{p}_1^T \boldsymbol{W}_1 \boldsymbol{V}_0(\boldsymbol{I} \otimes \boldsymbol{1}\boldsymbol{\beta}^T).$$

And finally, multiplying both parts of the obtained ratio on the right by $(\boldsymbol{I} \otimes \boldsymbol{1}\boldsymbol{\beta}^T)$, we arrive at the identity for equation (2).

Substitute (14) and (15) in equation (3):

$$\boldsymbol{0}^T = \boldsymbol{p}_0^T \boldsymbol{W}_0(\lambda\boldsymbol{\alpha}^T \otimes \boldsymbol{I}) + \boldsymbol{p}_0^T \boldsymbol{W}_0 \boldsymbol{W}_1(\boldsymbol{\Lambda} \oplus \boldsymbol{M} + \lambda\boldsymbol{\alpha}^T \otimes \boldsymbol{I}).$$

Considering (6) and (8), we obtain the identity for equation (3).

Next, we multiply the equations (1)–(3) to the right, look at the matrix $(\boldsymbol{I} \otimes \boldsymbol{1})$ of the corresponding dimension and sum up the obtained equalities. As a result, we get

$$\begin{aligned}\boldsymbol{0}^T = \boldsymbol{p}_0^T\boldsymbol{\Lambda} + \boldsymbol{p}_1^T(\boldsymbol{I} \otimes \boldsymbol{\mu}) + \boldsymbol{p}_0^T\lambda\boldsymbol{\alpha}^T + \boldsymbol{p}_1^T(\boldsymbol{\Lambda} \otimes \boldsymbol{1}) - \\ - \boldsymbol{p}_1^T(\boldsymbol{I} \otimes \boldsymbol{\mu}) + \boldsymbol{p}_2^T(\boldsymbol{I} \otimes \boldsymbol{\mu}) + \boldsymbol{p}_1^T(\lambda\boldsymbol{\alpha}^T \otimes \boldsymbol{1}) + \boldsymbol{p}_2^T(\boldsymbol{\Lambda} \otimes \boldsymbol{1}) - \\ - \boldsymbol{p}_2^T(\boldsymbol{I} \otimes \boldsymbol{\mu}) + \boldsymbol{p}_2^T(\lambda\boldsymbol{\alpha}^T \otimes \boldsymbol{I}) = \boldsymbol{p}_0^T(\boldsymbol{\Lambda} + \lambda\boldsymbol{\alpha}^T) + \\ + \boldsymbol{p}_1^T(\boldsymbol{\Lambda} + \lambda\boldsymbol{\alpha}^T)(\boldsymbol{I} \otimes \boldsymbol{1}) + \boldsymbol{p}_2^T(\boldsymbol{\Lambda} + \lambda\boldsymbol{\alpha}^T)(\boldsymbol{I} \otimes \boldsymbol{1}) = \\ = \left[\boldsymbol{p}_0^T + (\boldsymbol{p}_1^T + \boldsymbol{p}_2^T)(\boldsymbol{I} \otimes \boldsymbol{1})\right](\boldsymbol{\Lambda} + \lambda\boldsymbol{\alpha}^T).\end{aligned} \tag{18}$$

Substitute (14) and (15) into equations (18). As a result, we get:

$$\mathbf{0}^T = \left[\boldsymbol{p}_0^T + (\boldsymbol{p}_0^T\boldsymbol{W}_0 + \boldsymbol{p}_0^T\boldsymbol{W}_0\boldsymbol{W}_1)(\boldsymbol{I}\otimes\mathbf{1})\right](\boldsymbol{\Lambda} + \boldsymbol{\lambda}\boldsymbol{\alpha}^T). \tag{19}$$

Considering (9), we write the system (19) in the form (16):

$$\mathbf{0}^T = \boldsymbol{p}_0^T\boldsymbol{W}(\boldsymbol{\Lambda} + \boldsymbol{\lambda}\boldsymbol{\alpha}^T),$$

i.e. the vector $\boldsymbol{p}_0^T\boldsymbol{W}$ is satisfies the system (16) and, therefore,

$$\boldsymbol{p}_0^T\boldsymbol{W} = C\boldsymbol{q}^T. \tag{20}$$

Since the matrix of coefficients $\boldsymbol{\Lambda} + \boldsymbol{\lambda}\boldsymbol{\alpha}^T$ of the system (16) is indissoluble due to the irreducibility of the PH representation $(\boldsymbol{\alpha}, \boldsymbol{\Lambda})$, this system, taking into account (17) has a unique solution [18].

It remains for us to define C. According to (4), we have

$$\boldsymbol{p}_0^T\mathbf{1} + \boldsymbol{p}_1\mathbf{1} + \boldsymbol{p}_2^T\mathbf{1} = 1,$$

which, taking into account (14) and (15), we write in the form:

$$\boldsymbol{p}_0^T\mathbf{1} + \boldsymbol{p}_0^T\boldsymbol{W}_0\mathbf{1} + \boldsymbol{p}_0^T\boldsymbol{W}_0\boldsymbol{W}_1\mathbf{1} = 1,$$

The resulting equality, taking into account (9) and (20), has the form

$$\boldsymbol{p}_0^T\boldsymbol{W}\mathbf{1} = C\boldsymbol{q}^T\mathbf{1} = 1.$$

However, according to (17) $\boldsymbol{q}^T\mathbf{1} = 1$. Therefore, $C = 1$. So, we have shown that $\boldsymbol{p}_0^T\boldsymbol{W}$ coincides with the vector $\boldsymbol{q}$, i.e. to determine $\boldsymbol{p}_0$, we can use formula (13), having previously determined $\boldsymbol{q}$ from the system (16)–(17). We were convinced of the validity of formulas (14) and (15) earlier by substituting them in SEE (1)–(3) and turning the equations of the system into identities.

Thus, the theorem is proved. □

5. Markov chains nested at the moment of entry of applications into the system

Let's build a Markov chain (MC) embedded in MP $X(t)$ at the moments $t + 0$ of receipt of applications to the system over a set of states:

$$\mathscr{X}_A^+ = \bigcup_{k=1}^{2} \mathscr{X}_{A,k}^+,$$

where $\mathscr{X}_{A,k}^+ = \{(k, j), j = \overline{1,m}\}$, $k = 1, 2$.

The state (k, j) means that immediately after the application is received in the system, there are k applications in it and at the same time the service process is in phase j, $j = \overline{1,m}$, $k = 1, 2$. It was received immediately for maintenance and the selection of the maintenance phase was carried out instantly at the time of its receipt in accordance with the initial distribution set by the vector $\boldsymbol{\beta}$.

To determine the stationary probabilities $p_{A,x}^+$, $x \in \mathscr{X}_A^+$, the states of the nested MC will use the results of the work [19]. In accordance with the recommendations of this work, we will differentiate the jumps of MP $X(t)$, considering the jumps associated with the receipt of applications into the system to be "correct". At the same time, it should be noted that in our system, the incoming application cannot be lost, but when the drive is busy, it "kills" the application located in it. Considering the above, and also applying the formulas of [19] to calculate the stationary distribution of states of the nested MC, we come to the following result:

$$\boldsymbol{p}_{A,1}^{+\,T} = \frac{1}{\lambda}\boldsymbol{p}_0^T(\boldsymbol{\lambda}\otimes\boldsymbol{\beta}^T),$$

$$\boldsymbol{p}_{A,2}^{+}{}^{T} = \frac{1}{\lambda}\left[\boldsymbol{p}_1^T + \boldsymbol{p}_2^T\right](\boldsymbol{\lambda} \otimes \boldsymbol{I}) = \frac{1}{\lambda}\left[\mathbf{1}^T - \boldsymbol{p}_0^T\right]\boldsymbol{\lambda}, \tag{21}$$

where $\boldsymbol{p}_{A,k}^{+}{}^{T} = \left(p_{A,(k,1)}^{+}; ...; p_{A,(k,m)}^{+}\right)$, $\lambda = (-\boldsymbol{\alpha}^T\boldsymbol{\Lambda}^{-1}\mathbf{1})^{-1}$.

Next, we will build a MC embedded in the MP $X(t)$ at the moments $t-0$ of the receipt of applications into the system. The set of states of a given MC will be determined:

$$\mathscr{X}_A^- = \bigcup_{k=0}^{2} \mathscr{X}_{A,k}^-,$$

where $\mathscr{X}_{A,0}^- = \{(0)\}$, $\mathscr{X}_{A,k}^- = \left\{(k,j), j = \overline{1,m}\right\}$, $k = 1,2$.

The state (0) means that immediately before the first application was received into the system, the system was empty, and the state (k,j) means that immediately before the next application was received into the system, there were k applications in it, and at the same time the application on the device was serviced in phase j.

In accordance with the result of [19], we obtain:

$$p_{A,0}^- = \frac{1}{\lambda}\boldsymbol{p}_0^T\boldsymbol{\lambda},$$

$$\boldsymbol{p}_{A,k}^{-}{}^{T} = \frac{1}{\lambda}\boldsymbol{p}_k^T(\boldsymbol{\lambda} \otimes \boldsymbol{I}), \quad k = 1,2. \tag{22}$$

where $\boldsymbol{p}_{A,k}^{-}{}^{T} = \left(p_{A,(k,1)}^{-}; ...; p_{A,(k,m)}^{-}\right)$.

From (21) and (22) it follows that

$$\boldsymbol{p}_{A,2}^{+}{}^{T} = \boldsymbol{p}_{A,1}^{-}{}^{T} + \boldsymbol{p}_{A,2}^{-}{}^{T}. \tag{23}$$

Formula (23) means that in order for there to be two applications in the system immediately after the receipt of the next application, it is necessary and sufficient that there should be one or two applications immediately before the next application is received. In the first case, the incoming application will take up free space in the drive. In the second case, it will displace ("kill") the application in the drive and take its place.

6. Main indicators of system performance

As noted earlier, there is no loss of applications in the system we are considering due to lack of storage space. However, some applications leave the system without waiting for service. Let's call them "unsuccessful". An application becomes "unsuccessful" if there are two conditions: firstly, at the moment $t-0$ it enters the system, there must be one or two applications in the system, which will automatically turn it into an application waiting for the device to be released, and secondly, its waiting time should be longer before generation of the next application, which will force our application to leave the system by implementing a queue update mechanism. Thus, the time spent by the "unsuccessful" application in the system is equal to the time until the next application is generated. To calculate the probability that the application will be "unsuccessful", consider the following probabilities [20]:

$$\alpha_1 = \boldsymbol{\beta}^T \int_0^\infty e^{\boldsymbol{M}t} dA(t)\mathbf{1},$$

$$\alpha_2 = \boldsymbol{\alpha}^T \int_0^\infty e^{\boldsymbol{\Lambda}t} dB(t)\mathbf{1}.$$

Note that α_1 is a possibility that the service of the application on the device will not be completed during the time before the next request is generated.

Let's denote by $\bar{\gamma}$ the probability that the next application received by the system will be "unsuccessful". Considering the above, it can be argued that

$$\bar{\gamma} = \left(\boldsymbol{p}_{A,1}^{-}{}^{T} + \boldsymbol{p}_{A,2}^{-}{}^{T}\right)\mathbf{1} \cdot \alpha_1 = (1 - p_{A,0}^{-})\alpha_1, \tag{24}$$

$$\gamma = 1 - \bar{\gamma} = p_{A,0}^{-} + (1 - p_{A,0}^{-})\alpha_2.$$

Let's denote by λ_A the intensity of the flow of "successful" applications, and by λ_D the intensity of the outgoing flow. It is obvious that

$$\lambda_A = \lambda\gamma,$$

$$\lambda_D = \mu(1 - \boldsymbol{p}_0^{+}\mathbf{1}), \quad \text{where } \mu = (-\boldsymbol{\beta}^T\boldsymbol{M}^{-1}\mathbf{1})^{-1}.$$

In stationary mode, $\lambda_A = \lambda_D$, from where we get another not so obvious formula for γ:

$$\gamma = \frac{\mu}{\lambda}(1 - \boldsymbol{p}_0^T\mathbf{1}). \tag{25}$$

Obviously, if you enter the notation for the system load $\rho = \frac{\lambda}{\mu}$ and for the device utilization factor $u = 1 - \boldsymbol{p}_0^T\mathbf{1}$, then the formula (25) can be written as

$$\rho\gamma = u. \tag{26}$$

Considering that only "successful" applications are serviced in the system, formula (26) acquires a quite obvious probabilistic meaning: the utilization factor of the device is equal to the loading of the system with "successful" applications.

7. Conclusion

The paper investigates a single-line service system with queue updates and phase-type distributions. As a result, a recurrent matrix algorithm has been developed to calculate the stationary distribution of states of the Markov process describing the stochastic behavior of the system, and expressions for the main indicators of its performance have been obtained. The considered system is planned to be used as a mathematical model in the tasks of analyzing and managing the age of information. The authors are confident that this study will allow them to obtain sufficiently accurate estimates of the age of information for real technical systems in the future, due to the universality of phase-type distributions.

Author Contributions: Conceptualization, S.I. Matyushenko and K.E. Samouylov; methodology, S.I. Matyushenko; validation, S.I. Matyushenko and N.Yu. Gritsenko; writing—original draft preparation, S.I. Matyushenko; writing—review and editing, S.I. Matyushenko and N.Yu. Gritsenko; visualization, N.Yu. Gritsenko; supervision, K.E. Samouylov; project administration, K.E. Samouylov. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Data sharing is not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Kaul, S., Gruteser, M., Rai, V. & Kenney, J. *Minimizing age of information in vehicular networks* in *2011 8th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks* (2011), 350–358. doi:10.1109/SAHCN.2011.5984917.
2. Kaul, S., Yates, R. & Gruteser, M. *Real-time status: How often should one update?* in *2012 Proceedings IEEE INFOCOM* (2012), 2731–2735. doi:10.1109/INFCOM.2012.6195689.
3. Bedewy, A. M., Sun, Y. & Shroff, N. B. *Age-optimal information updates in multihop networks* in *2017 IEEE International Symposium on Information Theory (ISIT)* (2017), 576–580. doi:10.1109/ISIT.2017.8006593.
4. Bojan, T. M., Kumar, U. R. & Bojan, V. M. *An internet of things based intelligent transportation system* in *2014 IEEE International Conference on Vehicular Electronics and Safety* (2014), 174–179. doi:10.1109/ICVES.2014.7063743.
5. Hu, C. & Dong, Y. Age of information of two-way data exchanging systems with power-splitting. *Journal of Communications and Networks* **21,** 295–306. doi:10.1109/JCN.2019.000037 (2019).
6. Costa, M., Codreanu, M. & Ephremides, A. *Age of information with packet management* in *IEEE International Symposium on Information Theory (ISIT)* (2014), 1583–1587.
7. Costa, M., Codreanu, M. & Ephremides, A. On the Age of Information in Status Update Systems With Packet Management. *IEEE Transactions on Information Theory* **62,** 1897–1910. doi:10.1109/TIT.2016.2533395 (2016).
8. Chiariotti, F., Vikhrova, O., Soret, B. & Popovski, P. Peak Age of Information Distribution for Edge Computing With Wireless Links. *IEEE Transactions on Communications* **69,** 3176–3191. doi:10.1109/TCOMM.2021.3053038 (2021).
9. Kadota, I., Uysal-Biyikoglu, E., Singh, R. & Modiano, E. *Minimizing the Age of Information in broadcast wireless networks* in *2016 54th Annual Allerton Conference on Communication, Control, and Computing (Allerton)* (2016), 844–851. doi:10.1109/ALLERTON.2016.7852321.
10. Kosta, A., Pappas, N., Ephremides, A. & Angelakis, V. *Non-linear Age of Information in a Discrete Time Queue: Stationary Distribution and Average Performance Analysis* in *ICC 2020 — 2020 IEEE International Conference on Communications (ICC)* (2020), 1–6. doi:10.1109/ICC40277.2020.9148775.
11. Talak, R., Karaman, S. & Modiano, E. Improving Age of Information in Wireless Networks With Perfect Channel State Information. *IEEE/ACM Transactions on Networking* **28,** 1765–1778. doi:10.1109/TNET.2020.2996237 (2020).
12. Tripathi, V., Talak, R. & Modiano, E. *Age of information for discrete time queues* 2019.
13. Wijerathna Basnayaka, C. M., Jayakody, D. N. K., Ponnimbaduge Perera, T. D. & Vidal Ribeiro, M. *Age of Information in an URLLC-enabled Decode-and-Forward Wireless Communication System* in *2021 IEEE 93rd Vehicular Technology Conference (VTC2021-Spring)* (2021), 1–6. doi:10.1109/VTC2021-Spring51267.2021.9449007.
14. Bocharov, P. P. & Zaryadov, I. S. Stationary probability distribution in a queuing system with an update [Stacionarnoe raspredelenie veroyatnostey v sisteme massovogo obslugivaniya s obnovleniem]. *Vestnik RUDN Seriya Matematika. Informatika. Phizika.* in Russian, 14–23 (2007).
15. Kaul, S. K., Yates, R. D. & Gruteser, M. *Status updates through queues* in *2012 46th Annual Conference on Information Sciences and Systems (CISS)* (2012), 1–6. doi:10.1109/CISS.2012.6310931.
16. Inoue, Y., Masuyama, H., Takine, T. & Tanaka, T. A General Formula for the Stationary Distribution of the Age of Information and Its Application to Single-Server Queues. *IEEE Transactions on Information Theory* **65,** 8305–8324. doi:10.1109/TIT.2019.2938171 (2019).
17. Bocharov, P. P. & Pechinkin, A. V. *Queueing Theory [Teoriya massovogo obsluzhivaniya]* in Russian. 529 pp. (Izd-vo RUDN, Moscow, 1995).

18. Basharin, G. P., Bocharov, P. P. & Kogan, Y. A. *Queue analysis in computer networks [Analiz ocheredey v vichislitelnih setyah]* in Russian (Nauka, Moscow, 1989).
19. Naumov, V. A. *O predelnih veroyatnostyah polumarkovskogo processa [On the limiting probabilities of a semi-Markov process]* in Russian. 35–39 (Universitet drugbi narodov, Moscow, 1975).
20. Gelenbe, E. & Pujolle, G. *Introduction to queueing networks* (John Wiley, New York, 1987).

Information about the authors

Sergey I. Matyushenko—Candidate of Physical and Mathematical Sciences, Assistant professor of Department of Probability Theory and Cyber Security of RUDN University (e-mail: matyushenko-si@rudn.ru, ORCID: 0000-0001-8247-8988)

Konstantin E. Samouylov—Professor, Doctor of Technical Sciences, Head of the Department of Probability Theory and Cyber Security of RUDN University (e-mail: samuylov-ke@rudn.ru, ORCID: 0000-0002-6368-9680)

Nikolai Yu. Gritsenko—PhD student of Department of Probability Theory and Cyber Security of RUDN University (e-mail: 1142221032@rudn.ru)

УДК 519.21
DOI: 10.22363/2658-4670-2024-32-3-271–282 EDN: BAUGIT

Анализ системы обслуживания единичной ёмкости с распределениями фазового типа и обновлением очереди

С. И. Матюшенко, К. Е. Самуйлов, Н. Ю. Гриценко

Российский университет дружбы народов, ул. Миклухо-Маклая, д. 6, Москва, 117198, Российская Федерация

Аннотация. В данной работе исследуется однолинейная система массового обслуживания с накопителем единичной ёмкости и обновлением очереди. Под обновлением понимается следующий механизм: заявка, поступающая в систему и застающая в накопителе другую заявку, уничтожает её, занимая её место в накопителе. Следует заметить, что системы с тем или иным механизмом обновления давно привлекают внимание исследователей, поскольку имеют важное прикладное значение. В последнее время интерес к системам подобного рода вырос в связи с задачами оценки и управления возрастом информации. Система с механизмом обновления очереди, подобная рассматриваемой нами, уже исследовалась ранее в работах других авторов. Однако в этих работах речь шла о простейшем варианте системы с пуассоновским потоком и экспоненциальным обслуживанием. В данной работе мы рассматриваем систему с потоком и обслуживанием фазового типа. В результате проведённого исследования нами был разработан рекуррентный матричный алгоритм для расчёта стационарного распределения состояний марковского процесса, описывающего стохастическое поведение рассматриваемой системы, и получены выражения для основных показателей её производительности.

Ключевые слова: система массового обслуживания, распределение фазового типа, механизм обновления очереди

Discrete & Continuous Models & Applied Computational Science
ISSN 2658-7149 (Online), 2658-4670 (Print)

2024, 32 (3) 283–293
http://journals.rudn.ru/miph

Computer science

Research article

UDC 004.891.2
PACS 03B70, 68T27, 68W01
DOI: 10.22363/2658-4670-2024-32-3-283–293 EDN: EUNYIE

Stabilization and recovery assistant of people with disabilities based on artificial intelligence methods

Gleb A. Kiselev[1, 2], Nikolay A. Blagosklonov[2], Artem A. Nikolaev[2]

[1] RUDN University, 6 Miklukho-Maklaya St, Moscow, 117198, Russian Federation

[2] Federal Research Center "Computer Science and Control" of the Russian Academy of Sciences, 44 Vavilova St, bldg 2, Moscow, 119333, Russian Federation

(received: April 5, 2024; revised: April 29, 2024; accepted: May 12, 2024)

Abstract. Chronic non-communicable diseases account for more than 70% of global mortality statistics. The main share is made up of diseases of the cardiovascular system. Adequate preventive measures—impact on controllable and conditionally controllable risk factors—can reduce the contribution of these diseases to the structure of mortality. A significant effect can be achieved with an adequately selected level of physical activity, but doctors do not always recommend specific actions to patients. This article describes a prototype of a cognitive assistant for constructing personalized plans for therapeutic physical exercises for relatively healthy people and people suffering from cardiovascular diseases. The developed system consists of two main components: a cardiovascular risk assessment module and an exercise planning module. The risk assessment module consists of a knowledge base and an argumentative reasoning algorithm. Its task is to identify risk factors and levels, which is dual in nature: in the case of monitoring a relatively healthy user, the risk of developing cardiovascular disease is assessed, while in the case of interaction of the system with a user with cardiovascular disease, the risk of complications of a chronic form is assessed—development of a cardiovascular event. The exercise planning module includes an exercise database and a scheduler algorithm. The planning algorithm selects optimal therapeutic physical exercises according to optimal criteria, in order to form a plan that will not harm the patient and will increase his physical performance. The developed mechanism allows you to create training scenarios for users with any level of initial training, taking into account the available sports equipment, the preferred location for training (home, street, gym) and at any level of the cardiovascular continuum.

Key words and phrases: cognitive assistant, prevention, planning, risk analysis, semiotic network, knowledge base

For citation: Kiselev, G. A., Blagosklonov, N. A., Nikolaev, A. A. Stabilization and recovery assistant of people with disabilities based on artificial intelligence methods. *Discrete and Continuous Models and Applied Computational Science* **32** (3), 283–293. doi: 10.22363/2658-4670-2024-32-3-283–293. edn: EUNYIE (2024).

1. Introduction

According to the World Health Organization (WHO), there were 55.4 million deaths worldwide in 2019. The WHO Fact Sheet notes that cardiovascular diseases have been the leading cause of death worldwide for more than 20 years. The most common cause of death is coronary heart disease, accounting for 16% of total deaths worldwide. The greatest increase in mortality since 2000 was due to this disease: by 2019, mortality from it increased by more than 2 million cases and reached 8.9 million cases. Cerebral stroke is the second leading cause of death, accounting for approximately 11% of total deaths [1].

Together, coronary heart disease, stroke, diabetes, lung cancer, and chronic obstructive pulmonary disease accounted for nearly 100 million additional healthy life years lost in 2019 compared to 2000 [2].

In the Russian Federation, chronic non-communicable diseases (CNCDs) are found in a large number of the adult working population and are the leading causes of mortality. For example, arterial hypertension is found in 45.7% of the adult population of the Russian Federation [3]. As a result of mortality from CVDs in Russia in 2016, economic losses amounted to 2.7 trillion rubles (3.2% of gross domestic product) [4]. More than 90% of these losses are due to the mortality of people of working age.

The Russian Federation is implementing a comprehensive strategy for the prevention of non-communicable diseases, in which two directions can be distinguished—identifying people at high risk of chronic diseases or with an unidentified diagnosis [5]. Preventative measures include influencing lifestyle factors and other preventive measures in an identified group of people with increased risk factors [6]. Early identification of risk factors and prevention of the development of chronic non-diseases with personalization for a specific patient and his involvement in the process (4P medicine), according to some estimates, can increase the quality of life by 9.8%, reduce the number of years of life potentially lost due to disability by 27.3% [7]. The main goal of timely identification of risk factors and early prevention is to orient the patient towards a healthy lifestyle, and this is the main difference between preventive medicine and the traditional approach of palliative medicine [8].

Risk factors for the development of CNCDs (as well as other diseases) are usually divided into 2 groups: uncontrollable and controllable (unchangeable and modifiable or non-modifiable and modifiable) [9]. Uncontrollable factors include those factors that the patient and the doctor cannot influence—heredity, family history, trauma, past illnesses, etc. Controllable (changeable/modifiable) factors are those the degree of influence of which can be reduced, and ideally eliminated, due to medical and non-medical influence [9]. Among the controllable ones, we can distinguish a group of conditionally controllable risk factors, which include chronic diseases, which, with proper drug control, remain in the compensation phase (remission) and do not have a direct or indirect negative impact on the human body. Thus, the main task is the timely identification of controllable and conditionally controllable risk factors in the patient and adequate influence on them.

The modern approach to preventing exposure to risk factors is not only about changing behavioral habits and lifestyles and medicinal control of conditions. WHO currently gives a major role to the adequate prescription of physical exercise as a component of an integrated approach to the prevention and treatment of cardiovascular diseases [10]. In addition, it was noted that 5 million deaths per year can be prevented by increasing the level of physical activity of the population, thus reducing the risk of mortality from chronic NCDs by 20%–30%. However, in modern prevention, recommendations for optimizing physical/motor activity are too superficial. Although, when recommending certain types of physical activity to patients, even such as regular physical activity (walking), it is possible to reduce systolic blood pressure in people suffering from arterial hypertension to the target level of 140 mm Hg. Art. and below [11]. Doctors need to explain in detail to patients which exercises they can use and which they cannot. It is also advisable to help patients create individual training

plans. A number of multicenter large studies have demonstrated that regular dosed aerobic dynamic exercise lasting 150–300 minutes can significantly reduce the development of cardiovascular diseases in healthy individuals and reduce the risk of developing cardiovascular events in individuals who already have a number of cardiac nosologies [12]. Dosed anaerobic with static loads did not have a significant effect on these points, but they improved the quality of life and overall tolerance to physical activity.

In this regard, the purpose of this study is to develop an approach to developing personalized recommendations to the patient for certain physical exercises that he should perform independently. To achieve this goal, the patient is provided with an assistant who creates a personalized physical activity plan, approved by the doctor, with tips for implementation and gradually increasing intensity.

2. Materials and methods

The assessment of risk factors is based on a combined approach based on domestic and foreign clinical models and recommendations [4, 13, 14]. The assessment of risk levels is based on the cardiovascular continuum model [13], from which patients were divided into two main groups: patients without cardiovascular diseases and patients already having one or more cardiovascular diseases. Thus, for patients, the identification of risk factors and levels is of a different nature: in the first case, we are talking about assessing the risk of developing cardiovascular disease, while in the second case, the risk of complications of chronic non-diseases is assessed—the development of a cardiovascular event [15].

A formalized representation of knowledge about assessing the risks of developing a cardiovascular disease or a cardiovascular event in a person is carried out on the basis of a knowledge base implemented in the form of a heterogeneous semantic network (HSN) [16].

The construction of a set of hypotheses and the final solution is carried out on the basis of an argumentative algorithm proposed by G. S. Osipov [17]. Based on the information about the patient available in the system, a primary set of nodes is activated, then the algorithm sequentially performs the operations of expanding and narrowing the set of arguments, activating and eliminating hypotheses until the sets of hypotheses and arguments are stabilized. Stabilized sets are the result of the algorithm—risk levels (hypotheses) are explained by risk factors (arguments).

3. Results

A prototype of an assistant for the stabilization and improvement of people with disabilities c.Live using therapeutic physical education methods has been developed. The created intelligent recommendation system is focused on preventing the development of CNCDs in healthy patients and preventing the deterioration of development in people suffering from CNCDs, with the help of recommendations for lifestyle correction and a personalized plan for physical therapy exercises [18]. The developed prototype, in the first version, is focused on personifying risk assessment and issuing recommendations for the prevention of coronary heart disease using exercise therapy methods [19]. For this purpose, expert knowledge was collected not only from cardiologists, but also from rehabilitation doctors [20]. It is possible to assess risk factors based on the characteristics (parameters) of a person. In implementing the system, two main sources of obtaining information about the patient were identified: questionnaires in the application and electronic medical records (EMR). These sources are complementary, but a situation is envisaged when there is no access to the patient's EMR.

Knowledge about the names of risk factors, risk levels, physical exercise and general preventive recommendations was obtained from international and domestic literature, including clinical guidelines and individual scientific publications. The knowledge was adapted to the characteristics of the Russian population, for which expert work was carried out by cardiologists together with rehabilitation doctors. Thus, a structured representation of knowledge was formed.

Before determining and prescribing treatment or prevention tactics for CNCDs, the doctor needs to comprehensively assess the patient's health status in order to prescribe adequate (in a particular case) recommendations. To do this, the first stage of the assistant's work is to assess the risk levels of CNCDs.

After the risk levels are assessed, the physical exercises acceptable for the patient to perform are selected. For this purpose, a database was created in which attributes were added to all types of physical activity: indications and contraindications for use. The risk levels calculated by the system are indications and contraindications depending on their nature. If the patient does not have a disease of the cardiovascular system, and the maximum calculated risk level is not higher than high, then this situation is considered a partial limiter—the variety of possible exercises is not reduced, but the duration of training and the number of approaches for one exercise are reduced. For situations of patients with cardiovascular diseases with a calculated level of risk of a cardiovascular event that is very high or extreme, then in such situations all strength exercises are excluded and restrictions are placed on high-intensity exercises that such exercises should be performed by the patient no more than twice per week at approximately equal intervals.

To personalize the exercises, the following parameters were included in the exercise database:

- Exercise ID. The parameter is necessary for linking with video files demonstrating the correct execution of exercises;
- Brief description of the exercise;
- The number of MET units (universal endurance units) expended during the exercise. This parameter allows you to create a training plan that is optimal for the user;
- Locations where the exercise can be performed. The parameter allows the user to choose one of three locations where he will perform exercises: at home, on the street, in the gym. Each of the locations includes a list of exercise machines and other aids available for use, for example, a fitness expander, which expands the selection of available exercises;
- A list of diseases that this exercise has a positive effect on. This parameter plays an important role in planning exercises that help the user improve their condition. Moreover, for each disease from the list, a numerical representation of the strength of influence is stored, where 1 means weak influence and 5 means strong. This view is used to plan exercises that will bring maximum benefit to the user;
- List of diseases that this exercise negatively affects. This parameter is necessary when planning exercises that will not bring negative consequences to the user. For each disease from this list, a numerical representation of the strength of influence is also stored, where 1 means weak influence and 5 means strong. This representation is used when planning exercises to minimize the possible negative impact of the exercise on the user.

The developed algorithm is presented in Figure 1.

To create a personalized training plan, a planner algorithm has been developed, which is based on the following statements:

- In sports and general physical training, a training plan is drawn up for three months.
- When drawing up a plan, the user's current illnesses are taken into account. Based on knowledge about the user's diseases, the set of exercises acceptable for the user is narrowed.

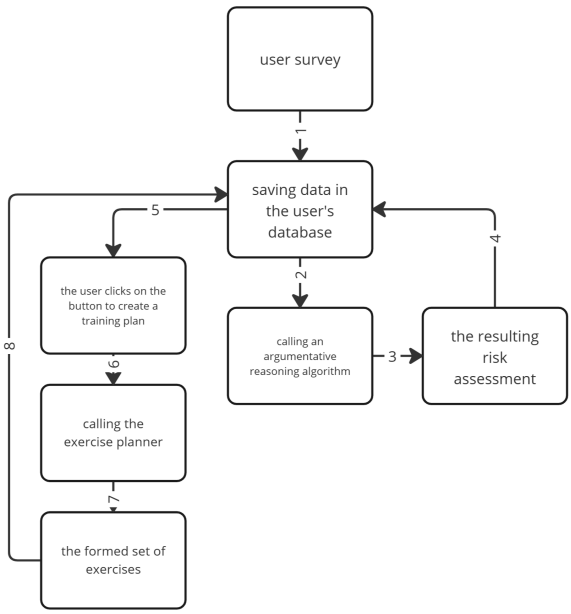

Figure 1. Scheme of synthesis of a personalized exercise plan

- When drawing up a plan, the current health status of the user is taken into account. Before drawing up a plan, potential risks are assessed based on existing health data.
- As the user's training level increases, the load on the user's muscles increases.
- As the user becomes more trained, the intensity of the exercises increases.
- The user can choose a location for training: home/street/gym.

Generating a personalized training plan consists of the following steps:

1. The user takes a survey.
2. The survey results are sent to the risk assessment module. The results of the module's operation are added to the database, which stores information about the user's health status.
3. In the user interface, the "Create a training plan" button becomes available for clicking. When you click on the button, the scheduler algorithm is launched, to which all available information about the user is transmitted as input.
   3.1. The scheduling algorithm receives user data.
   3.2. The planner algorithm receives a set of exercises, each exercise is described by attributes (number of METs per 1 repetition, indications, contraindications, muscle groups involved, location).
   3.3. The planner algorithm excludes from the set of exercises those that have a negative impact on the user's concomitant diseases and those that have a negative impact on diseases whose risk level for the user is quite high.
   3.4. The scheduler algorithm divides the remaining exercises into three sets, each set corresponding to one of the possible locations: home, street, gym.
   3.5. For each concomitant disease of the user, the planner algorithm ranks sets of exercises in descending order of positive effect on this concomitant disease.

3.6. For each muscle group (a set of muscles involved during a workout; a total of 4 muscle groups are considered based on the statement of 4 workouts per week), the planner algorithm filters the resulting sets by the attribute "muscle groups involved" and from the remaining exercises based on the principle of maximum benefit for the user compiles sets of exercises for each location. The limit is the amount of METs allowed per workout.
3.7. The planner algorithm repeats the previous steps for each week, taking into account that over time the allowable amount of METs per workout increases.
3.8. The result of the scheduler algorithm is stored in the user database.

4. The interface prompts the user to select the location where he plans to conduct the next training session. Data is loaded from the database and displayed to the user.

The scheduling algorithm itself works as follows:

Let given:

$E = \{e_1, \dots .e_n\}$—set of all exercises,
$D = \{d_1, \dots, d_m\}$—the set of all diseases that are taken into account when drawing up a training plan,
$L = \{l_1, \dots, l_k\}$—the set of all locations in which the user can perform exercises.
$M = \{m_1, \dots, m_r\}$—set of muscle groups.
Each exercise e from the set E is characterized by:
l_e—location where exercise should be performed $l_e \in L$,
m_e—the amount of MET, given by some natural number,
t_e—time that will be spent on performing the exercise,
M_e—muscle groups that are affected by the exercise,
P_e—a list of diseases in which this exercise has a positive effect on health,
N_e—a list of diseases in which this exercise has a negative effect on health,
$fp_e : P_e \to \{1, 2, 3, 4, 5\}$—a function that evaluates the positive effect of performing an exercise on human health for a disease selected from P_e, the higher the value, the more positive the effect,
$fn_e : N_e \to \{1, 2, 3, 4, 5\}$—a function that evaluates the negative effect of performing an exercise on human health for a disease selected from N_e; the higher the value, the more dangerous the exercise is for health.
Each user is described by the parameter:
UD—diseases that the user has or the likelihood of having them is high.

The goal of the algorithm is to create sets of exercises for 12 weeks, with four workouts planned each week. The duration of training in the first four weeks is t_1 minutes, in the next four weeks t_2 and in the final four weeks t_3 minutes.
Step 1. We remove all exercises from the list that can cause harm to the user, i.e. exercises that can have a negative effect on his health; we use AE to denote the set of exercises available to the user.

$$AE = \{e \in E : N_e \cap UD = \varnothing\}.$$

Step 2. Organize the many exercises available to the user by location. Let us denote such sets by AE_l, where l indicates the location.

$$AE_l = \{e \in AE : l_e = l\}.$$

Step 3. Let's create sets of exercises for each workout in each location for each muscle group, sets of exercises will be stored in $w_{month,week,muscle\_group}$.

A library of video materials has been created for users, in which an experienced trainer clearly shows how to perform certain exercises. A total of 118 explanatory videos were recorded on the correct execution of exercises in three locations: 41 videos for practicing at home, 36 for practicing outdoors, 41 for practicing in the gym.

Algorithm 1: Create a set of exercises for training

```
months = {1, 2, 3}
weeks = {1, 2, 3, 4}
for l ∈ L
 do
    for muscle_group ∈ M
     do
        exercise_list = create_exercise_list(l, m, UD, user_met_limit)
        n=0
        for month ∈ months
         do
            for week ∈ weeks
             do
                total_time := 0
                while total_time < t_month
                 do
                    w_{month,week,muscle_group}.add(exercise_list[n])
                    total_time := total_time + t_{exercise_list[n]}
                    n=n+1
                    if n >= len(exercise_list) then
                        n=0
```

Algorithm 2: Create exercise list

```
Function create_exercise_list(location, muscle_group, user_diseases, user_met_limit):
    n := 0
    answer_list := []
    current_size := −1
    exercises_after_filtration =
      filter_exercise_list(location, muscle_group, user_diseases, user_met_lmit)
    while current_size! = len(answer_list)
     do
        current_size := len(answer_list)
        for e ∈ exercises_after_filtration
         do
            if max_{a∈user_diseases} fp_e(a) > n then
                answer_list.add(e)
                n := n + 1
    return answer_list
```

Algorithm 3: Filter exercise list

```
Function filter_exercise_list(location, muscle_group,user_diseases, user_met_lmit):
    answer_list := []
    for e ∈ AE_l
     do
        if muscle_group ∈ M_e & user_met_limit <= m_e then
            answer_list.add(e)
    sort(answer_list) #Sorting exercise in descending order of total utility
    return answer_list
```

4. Conclusions

A feature of the developed architecture of the c.Live cognitive assistant is its easy scalability and extensibility. A formalized representation of knowledge in the form of a heterogeneous semantic network allows you to both reuse existing feature nodes and logical operation nodes, and create new nodes of various types. The argumentative reasoning algorithm is capable of working with a large number of hypotheses and their arguments, while the computation time will increase slightly.

Expanding the list of exercises available to users will allow for more fine-tuning of individual training plans. To expand functionality, close cooperation with practicing doctors of various specialties, including rehabilitation doctors, is necessary.

Due to the fact that a relational structure was chosen to store the knowledge base in a form suitable for machine processing, the current version of the knowledge base can be quite easily expanded by adding new diseases with their risk factors and recommendations. This makes it possible to add treatment scenarios for new diseases based on unified system mechanisms.

As a consistent development of this development, it is advisable at the following stages to pay attention to other diseases: the cardiovascular system (risks of developing arrhythmias, arterial hypertension, heart failure, etc.), the endocrine system (diabetes mellitus, primarily type 2), the digestive system (gastritis, stomach ulcer), pulmonary system (chronic bronchitis, chronic obstructive pulmonary disease, bronchial asthma), etc. Expanding the list of diseases in the application will allow us to cover as many segments of the population as possible from various professions and social status.

Author Contributions: Conceptualization, Gleb A. Kiselev and Nikolay A. Blagosklonov; methodology, Gleb A. Kiselev and Artem A. Nikolaev; software, Artem A. Nikolaev; investigation, Nikolay A. Blagosklonov; resources, Gleb A. Kiselev; data curation, Gleb A. Kiselev and Nikolay A. Blagosklonov; writing—original draft preparation, Nikolay A. Blagosklonov and Artem A. Nikolaev; writing—review and editing, Gleb A. Kiselev; visualization, Artem A. Nikolaev; supervision, Gleb A. Kiselev; project administration, Gleb A. Kiselev; funding acquisition, Gleb A. Kiselev. All authors have read and agreed to the published version of the manuscript.

Funding: This paper has been supported by the RUDN University Strategic Academic Leadership Program.

Data Availability Statement: Data sharing is not applicable.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; in the decision to publish the results.

References

1. *Top 10 leading causes of death in the world* https://www.who.int/ru/news-room/fact-sheets/detail/the-top-10-causes-of-death/. 2020.
2. *WHO publishes statistics on the leading causes of death and disability worldwide for the period 2000–2019* https://www.who.int/ru/news/item/09-12-2020-who-reveals-leading-causes-of-death-and-disability-worldwide-2000-2019/. 2019.
3. Balanova, Y. *Arterial hypertension in the Russian population: prevalence, contribution to survival and mortality, possibilities for reducing socio-economic damage* (Abstract for the scientific degree of Doctor of Medical Sciences, 2021).
4. Kontsevaya, A., Drapkina, O. & Balanova, Y. Economic damage from cardiovascular diseases in the Russian Federation in 2016. *Rational pharmacotherapy in cardiology* **14,** 156–166 (2018).
5. Boytsov, S., Chuchalin, A. & Arutyunov, G. Prevention of chronic non-infectious diseases: Recommendations. *Profmedforum* (2013).
6. *Cardiovascular prevention 2017. Russian national recommendations* 122 pp. (2018).
7. *P4-medicine - a new direction in the development of healthcare* https://www.dirklinik.ru/article/500-4p-meditsina-novoe-napravlenie-razvitiya-zdravoohraneniya/. 2023.
8. Suchkov, S., H., A. & Antonova, E. Personalized medicine as an updated model of the national healthcare system.Part 1. Strategic aspects of infrastructure. *Russian Bulletin of Perinatology and Pediatrics* **62,** 7–14 (2017).
9. Savilov, E. & Shugaeva, S. Risk factor: theory and practice of application in epidemiological studies. *Epidemiology and infectious diseases* **22,** 306–310 (2017).
10. *World Health Organization. News bulletin. Physical activity.* https://www.who.int/ru/news-room/fact-sheets/detail/physical-activity/. 2023.
11. Shebeko, L., Vlasova, S., Germanovich, L. & Belyakovskaya, N. Physical rehabilitation of patients with arterial hypertension. *Bulletin of the Transbaikal State University* **93,** 80–87 (2013).
12. Bubnova, M. & Aronov, D. Methodic recommendations. Maintaining physical activity of those with limitations in health. In Russ. *CardioSomatics* **7** (ed Boytsov, S.) 5–50 (2016).
13. Trenkwalder, P. Preventing thecardiovascular complications of hypertension. *European Heart Journal Supplements* **6,** H37–H42 (2004).
14. ESC recommendations for the prevention of cardiovascular diseases in clinical practice. *Russian Journal of Cardiology. Clinical recommendations. Arterial hypertension in adults* **27,** 191–288 (2022).
15. Tolpygina, S. & Martsevich, S. Cardiac risk stratification in stable coronary artery disease. In Russ. *The Clinician* **14,** 24–33. doi:10.17650/1818-8338-2020-14-1-2-24-33 (2020).
16. Kobrinsky, B., Kadykov, A., Poltavskaya, M., Blagoslonov, N. & Kovelkova, M. Principles of functioning of an intelligent system for dynamic control of risk factors and the formation of recommendations for health care. *Preventive Medicine* **22,** 78–84 (2019).
17. Osipov, G. *Acquisition of knowledge by intelligent systems: Fundamentals of theory and technology* Russian (Fizmatlit, 1998).
18. Drapkina, O., Novikova, N. & Dzhioyeva, O. Methodological recommendations: "Current opportunities and prospects of complex physical activity of patients with cardiovascular pathology". In Russ. *Russian Journal of Preventive Medicine* **23,** 61–119. doi:10 . 17116 / profmed20202303261 (2020).
19. Dibben, G., Faulkner, J., Oldridge, N., Rees, K., Thompson, D., Zwisler, A.-D. & Taylor, R. Exercise-based cardiac rehabilitation for coronary heart disease. *Cochrane Database of Systematic Reviews* **11.** doi:10.1002/14651858.CD001800.pub4 (2021).
20. Lyamina, N., Karpova, E., Kotelnikova, E. & Bizyaeva, E. Physical training in the rehabilitation and prevention in patients with ischemic heart disease after percutaneous coronary

interventions: the borders of efficiency and safety. In Russ. *Russian Journal of Cardiology,* 93–98. doi:10.15829/1560-4071-2014-6-93-98 (2014).

Information about the authors

Gleb A. Kiselev—Candidate of Technical Sciences, Senior Lecturer at the Department of Mathematical Modeling and Artificial Intelligence of RUDN University; Researcher of Federal Research Center "Computer Science and Control" of the Russian Academy of Sciences (e-mail: kiselev@isa.ru, phone: +7(906)7993329, ORCID: 00000-0001-9231-8662, ResearcherID: Y-6971-2018, Scopus Author ID: 57195683637)

Nikolay A. Blagosklonov—Researcher of Federal Research Center "Computer Science and Control" of the Russian Academy of Sciences (e-mail: nblagosklonov@frccsc.ru, phone: +7(499)1354246, ORCID: 0000-0002-5293-8469, ResearcherID: ABG-2002-2021, Scopus Author ID: 57206274545)

Artem A. Nikolaev—Senior developer of Federal Research Center "Computer Science and Control" of the Russian Academy of Sciences (e-mail: nicepeopleproject@gmail.com, phone: +7(977)2790346, ORCID: 0000-0003-4561-8990, ResearcherID: G-9622-2018)

УДК 004.891.2
PACS 03B70, 68T27, 68W01
DOI: 10.22363/2658-4670-2024-32-3-283–293 EDN: EUNYIE

Ассистент стабилизации и восстановления людей с ограниченными возможностями на основе методов искусственного интеллекта

Г. А. Киселёв[1, 2], Н. А. Благосклонов[2], А. А. Николаев[2]

[1] Российский университет дружбы народов, ул. Миклухо-Маклая, д. 6, Москва, 117198, Российская Федерация

[2] Федеральный исследовательский центр «Информатика и управление» Российской академии наук, ул. Вавилова, д. 44, корп. 2, Москва, 119333, Российская Федерация

Аннотация. Хронические неинфекционные заболевания составляют более 70% в статистике общемировой смертности. Основную долю составляют заболевания сердечно-сосудистой системы. Снизить вклад данных заболеваний в структуру смертности могут адекватные меры профилактики — воздействие на управляемые и условно управляемые факторы риска. Значительного эффекта можно добиться с помощью адекватно подобранного уровня физической активности, однако врачи не всегда рекомендуют пациентам конкретные действия. В настоящей статье описан прототип когнитивного ассистента построения персонифицированных планов лечебных физических упражнений для условно здоровых людей и лиц, страдающих сердечно-сосудистыми заболеваниями. Разработанная система состоит из двух основных компонентов: модуль оценки рисков сердечно-сосудистых заболеваний и модуль планирования упражнений. Модуль оценки рисков состоит из базы знаний и алгоритма аргументационных рассуждений. Его задача — выявление факторов и уровней риска, которое носит двойственный характер: в случае мониторинга условно здорового пользователя происходит оценка риска развития сердечно-сосудистого заболевания, в то время как в случае взаимодействия системы с пользователем с сердечно-сосудистым заболеванием, оценивается риск осложнения хронической формы — развитие сердечно-сосудистого события. Модуль планирования упражнений включает базу данных упражнений и алгоритм-планировщик. Алгоритм планирования осуществляет подбор оптимальных лечебных физических упражнений по оптимальным критериям, с целью формирования такого плана, который не навредит пациенту и увеличит его физические показатели. Разработанный механизм позволяет составлять сценарии тренировок для пользователей с любым уровнем исходной подготовки, с учётом имеющегося спортивного инвентаря, предпочитаемой локации для выполнения тренировок (дом, улица, зал) и на любом уровне сердечно-сосудистого континуума.

Ключевые слова: когнитивный ассистент, профилактика, планирование, анализ рисков, семиотическая сеть, база знаний

Discrete & Continuous Models
& Applied Computational Science
ISSN 2658-7149 (Online), 2658-4670 (Print)

2024, 32 (3) 294–305
http://journals.rudn.ru/miph

Computer science

Research article

UDC 535.8:004.725
PACS 42.81.–i, 42.60.Lh,
DOI: 10.22363/2658-4670-2024-32-3-294–305 EDN: BGKJTK

Maintaining the reliability of communication networks while continuing operation of optical cables beyond their warranty period

Denis A. Paltsin, Alexander Yu. Tsym

M.I. Krivosheev National Research Centre for Telecommunication, 16 Kazakova St, Moscow, 105064, Russian Federation

(received: July 15, 2024; revised: July 30, 2024; accepted: August 2, 2024)

Abstract. Currently, an increasing number of fiber-optic communication lines are reaching the end of their predetermined service life, yet the quality indicators of these lines still allow for continued operation. To extend the actual operational life of these lines, it is necessary to conduct high-quality monitoring of both the current status of all components and the dynamics of key indicators. This article proposes a method for addressing the challenge of maintaining communication network reliability while continuing to use optical cables after their warranty period has expired. A study of random values of the attenuation coefficient and polarization mode dispersion of an optical fiber, supported by actual operational data from a network segment, shows high temporal stability in the attenuation coefficient and polarization mode dispersion of optical fiber type G.652. This conclusion allows us to discuss the continued operation of optical cables after the warranty period. To analyze the key aging metric, mathematical models are used that take into account the physical and chemical properties of cables as well as the conditions of their proof-tests. Using an example related to the current state of Russian fiber optic networks, we calculate the number of emergency reserve elements necessary to maintain the reliability of their operation. Practical recommendations for the placement of emergency reserve are also provided.

Key words and phrases: communication networks, fiber-optic cables, service life, operational reliability index

For citation: Paltsin, D. A., Tsym, A. Y. Maintaining the reliability of communication networks while continuing operation of optical cables beyond their warranty period. *Discrete and Continuous Models and Applied Computational Science* **32** (3), 294–305. doi: 10.22363/2658-4670-2024-32-3-294–305. edn: BGKJTK (2024).

1. Introduction

By definition, the service life of a cable is the period of time from the start of its operation until it reaches its limit state, where further operation becomes unacceptable, impractical or impossible, or restoration of its operational state is not feasible. This limit state is determined in advance in regulatory, technical, or project documentation. Thus, on the one hand, it may seem contradictory to talk about the reliability of a communication transport network when its basis—the optical cable [1]—has reached its limit state.

However, it is difficult to determine the exact limit state of an optical cable. In practice, the term "service life" refers to the duration of time determined by the manufacturer, during which the failure rate caused by aging (fatigue) is not expected to increase exponentially. This usually occurs between 25 and 30 years. So, the service life can be seen as an extrapolated indicator of reliability, which can be estimated (in point or on interval) by extrapolating calculations, tests, and/or operational data to a different duration of operation and different operating conditions.

It is important to consider two factors. Firstly, the expected service life of optical cable in general and optical fibers in particular is determined with a high degree of uncertainty, and therefore, cable manufacturers specify this in the passport-certificate for the construction length with a considerable margin. Secondly, operating companies have technical and organizational resources to influence the operational reliability indicator, which can be assessed by point or interval analysis based on operational data. In particular, the conditions of laying and operation with a limitation of the permissible bending radii of optical fibers can significantly reduce the mechanical loads on them and increase their service life [2].

Based on the above, it is clear that the problem formulation is correct and relevant. To find an informed solution, we will separately consider the issues of the temporal stability of optical fiber parameters.

2. Study of the temporal stability of the optical fiber attenuation coefficient

The main distinguishing feature of optical fiber as a signal transmission medium is the random nature of the attenuation coefficient. This coefficient depends on a number of design parameters, including the longitudinal uniformity of the core, the eccentricity and non-roundness of the core, and the longitudinal uniformity of the material. Other factors that can affect the attenuation include extraneous inclusions in the quartz, microcracks, and microbends.

The culture of producing optical fiber and optical cable significantly impacts the attenuation coefficient value. Studies [3–5] indicate that the range of the attenuation coefficient spans from the minimum values determined by the so-called Rayleigh (or dipole) scattering [3] to the maximum values defined in manufacturer technical specifications.

The same can be said about spliced optical fiber connections. Almost everywhere, they are made by welding. In this case, optical power losses depend on several factors, including the fiber chopping angle, the tolerances on the outer diameter of the optical fiber, the eccentricity and non-roundness of the core, the welding mode, and the compressive force. These factors all have an inherent random nature, which means that the optical loss value is also random. The minimum value of the loss is zero, while the maximum value is determined by the technical specifications for connector couplings.

The resulting attenuation of optical fiber in the amplifying (regeneration) span is thus a sum of random variables. In general, finding the distribution law for the sum of random variables is a difficult task. However, in this case, where there are a large number of terms due to the long length of the span, there is a simple and fairly accurate solution to the problem. Based on the central limit theorem [6], it can be argued that the distribution of the total attenuation in the amplification span will follow the normal distribution. The numerical values of the mean and standard deviation of this distribution, $\overline{a}_L$ and $\sigma(a_L)$, respectively, can be defined as follows

$$\overline{a}_L = \overline{\alpha} L + \overline{a}_{sc} \left(\frac{L}{l_{cs}} - 1 \right),$$

$$\sigma(a_L) = \sqrt{\sigma^2(\alpha)L + \sigma^2(a_{sc})\left(\frac{L}{l_{cs}} - 1\right)},$$

where $\overline{\alpha}$ is the average of the attenuation coefficient; $\overline{a}_{sc}$, $\sigma(a_{sc})$ are the average and the standard deviation of losses at the optical fiber spliced connection; L is the length of the regeneration span, l_{cs} is the length of optical cable construction span.

The right bound of the confidence interval with probability 0.996 is usually used as a norm in practice. This means that, on average, one fiber out of every 1000 would have an attenuation slightly greater than that calculated by the formula

$$a_L(0.9986) = \overline{\alpha}L + \overline{a}_{sc}n + 3\sqrt{\sigma^2(\alpha)L + \sigma^2(a_{sc})n},$$

where $n = \frac{L}{l_{cs}} - 1$.

Let us consider the initial laws of distribution of the attenuation coefficient and losses in spliced optical fiber connections. A number of studies and ITU-T recommendations [7–9] indicate that the attenuation coefficient follow a normal probability distribution. The range of its value, denoted by $\Delta\alpha$, extends from the lowest value of Rayleigh losses, ($\alpha \approx 0.8/\lambda^4$), to the highest value specified by the manufacturer or in ITU-T recommendations [7–9]. The mathematical expectation of this distribution corresponds to the midpoint of the range, $\Delta\alpha/2$, while the standard deviation is one-sixth of the range, $\Delta\alpha/6$.

The results of the study [10] indicate, that the attenuation coefficient of optical fiber

- is random;
- with high accuracy obeys the normal Laplace–Gauss law with a mathematical expectation of 0.193 dB/km when the wavelength is 1.55 μm;
- in typical specifications (0.35 and 0.22 dB/km) takes values close to the maximum;
- in the range from 0.14 dB/km (Rayleigh scattering) to 0.22 dB/km has a standard deviation equal to 0.013 dB/km if determined by the three-sigma rule.

To assess temporary changes, we determine the confidence interval of the mathematical expectation Δ with a confidence probability of 0.95 (t = 1.96 is a parameter from Laplace tables).

$$\Delta = \sigma \times t/\sqrt{m} = 1.96 \times 0.013/\sqrt{8} = 0.009,$$

where m is the number of optical fibers measured. Thus, the confidence interval for the mathematical expectation of the attenuation coefficient lies between the values of 0.184 and 0.202 dB/km.

After many years of research on optical fibers under real operating conditions, we have obtained the following results. To illustrate, let's consider the optical losses measured on eight optical fibers of a separate amplifying section within the Kadala–Skovorodino regeneration span in Siberia. Table 1 shows the calculation of average attenuation coefficient based on optical loss measurements.

As can be seen from Table 1, all average sample values of the attenuation coefficient (0.190–0.197) are within the confidence interval of the mathematical expectation (0.184–0.202). These results coincide with a number of studies conducted on the parameters of optical fibers over several years [11]. This indicates that the attenuation coefficient of G.652 optical fiber is highly stable over time.

3. Study of the temporal stability of optical fiber chromatic and polarization-mode dispersion

The chromatic dispersion of an optical fiber over a given distance is calculated by multiplying the chromatic dispersion coefficient by the length of that distance [12]. Thus, chromatic dispersion accumulates in proportion to the length of an optical fiber.

Table 1

Optical loss statistics

Optical fiber #	Average losses in weldings, dB	Welding losses, normalized to a length of 1 km, dB/km	Optical losses normalized to a length of 1 km, dB/km	Attenuation coefficient, dB/km
05	0.022	0.004	0.200	0.196
06	0.044	0.009	0.201	0.192
09	0.063	0.013	0.203	0.190
10	0.027	0.005	0.200	0.195
13	0.043	0.009	0.204	0.195
14	0.021	0.004	0.201	0.197
15	0.035	0.007	0.202	0.195
16	0.019	0.004	0.198	0.194

The term "single-mode fiber" is relative, as there can be no truly single-mode fiber. Strictly, a "mode" is a solution to Maxwell's equation for a dielectric waveguide. These equation for a cylindrical waveguide with a core-to-cladding diameter ratio (μm) of 8-10/125 always have two solutions, producing two modes with the same attenuation coefficient but with two polarization planes shifted by 90 degrees from each other. Due to various factors such as structural inhomogeneities in the optical fiber and local mechanical stresses, polarization mode dispersion (PMD) has a pronounced random character and accumulates in proportion to the square root of the amplifying (or regeneration) span length. The maximum value of PMD is typically characterized by its 99.99th percentile.

For G.652 optical fiber, the chromatic dispersion coefficient at the 1.55 μm wavelength is 18 ps/(nm×km), the minimum and maximum values of the coefficient in the 3rd optical transparent window (1525–1625 nm) are 14.5 and 20.5 ps/(nm×km) respectively, and the maximum value of PMD, normalized to a length of 1 km, is equal to 0.5 ps/km$^{1/2}$. Chromatic dispersion coefficients and PMD values, normalized to 1 km, measured on 8 optical fibers of a separate amplifying span inside the Kadala–Skovorodino regeneration span in Siberia, are given in Table 2.

The data in Table 2 shows: a) the dispersion parameters of the optical fiber are stable over time; b) the average of the maximum PMD values normalized to a length of 1 km (0.235 ps/km$^{1/2}$) is significantly less than the normalized quantile (0.5 ps/km$^{1/2}$). With this value of the normalized PMD, the total value of PMD on the entire regeneration span Kadala–Skovorodino (1027.3 km) will be no more than 7.5 ps.

Thus, the research findings indicate a high degree of temporal stability in the attenuation coefficient and PMD of G.652 optical fiber. These results allow us to proceed with solving the main problem.

4. Russian transport networks features

Within the context of this task, it is important to note the following features of the transport networks that form the Russian information infrastructure.

Table 2

Chromatic and polarization-mode dispersion

Characteristic	**Optical fiber #**							
	05	06	09	10	13	14	15	16
Chromatic dispersion coefficient, minimum, ps/(nm×km)	14.94	14.93	14.86	14.88	14.88	15.03	14.83	14.91
Chromatic dispersion coefficient, maximum, ps/(nm×km)	20.19	20.18	20.10	20.16	20.14	20.21	20.19	20.18
Maximum PMD value, normalized to a 1 km, ps/km$^{1/2}$	0.216	0.238	0.256	0.283	0.227	0.205	0.224	0.231

1. The main backbone fiber-optic communication lines (FOCL) of national transport networks were built in a relatively short period of time in the late 1990s and early 2000s, and they are currently approaching the service life limits of 25–30 years [13, 14]. At the same time, the cable infrastructure is not only an expensive component of the fiber optic network, but also an element requiring special operational support [15].
2. In accordance with the requirements for the construction of FOCL, the nominal length of elementary cable sections (ECS) has been calculated based on the maximum values of the attenuation coefficient and losses in permanent connection of optical fibers [16]. The maximum length has also been calculated taking into account three-sigma deviations, which corresponds to a 99.86% probability of attenuation point. Statistical normalization allows for an increase in the length of individual ECS of 20–35% relative to the specified value. However, this reduces the reserve in optical transmission system power budget, which is so necessary when the failure rate of the optical fiber increases after the end of the cable's service life.
3. The design and technical features of optical cables have led to the development of a fundamentally new technology for constructing long-distance communication networks—transport multichannel communication (TMC). In all-dielectric optical cables, it has been possible to "peel off" the massive outer coatings from the lightweight optical core, turning the outer covers into an independent protective polyethylene pipe (PPP), which is an element of the cable sewer, and the optical core into a small-sized optical cable that is blown inside the PPP.

When constructing the engineering infrastructure of transport networks using TMC technology, first a PPP package is laid along the route, then viewing devices are installed and, finally, optical cables are blown into the formed channels. TMC technology can be implemented along highways, and/or along the routes of decommissioned copper communication cables.

The high competitiveness of TMC's technology is based on a range of features, including legal, financial, technical, and organizational aspects. These features include universal right of passage, an extended construction season, lower final capital and operating costs, long service life (TMC

lasts 50 years, optical fiber 25–30), localization of accidents using a geographic information system, low damage density (less than 0.1 per 100 km of TMC per year), short cable recovery times (within 8 hours), comfortable ambient temperature from optical fiber aging point of view (−2°C – +18°C) and high protection against vandalism, and ammunition explosions.

5. Aging of optical fiber

In accordance with the classification adopted in the reliability theory, optical fiber is an object of a specific purpose, continuous long-term use, renewable during its service life and susceptible to aging. The strength theory reveals the process of aging (fatigue) of solids through the thermal fluctuation mechanism of silicate-oxygen bonds. Brownian motion of atoms constantly breaks and recreates these bonds. At the moment when the number of broken bonds significantly exceeds the number of restored ones, a microcrack forms. The presence of even a small tensile force provokes crack growth.

The basic reliability model of a fiber that has passed proof tests is presented as follows

$$t_f = \left\{ \left[B^{\frac{m}{n-2}} S_0^m \frac{L_0}{L} \ln \frac{1}{1-F} + \left(\sigma_p^n t_p\right)^{\frac{m}{n-2}} (1+C)^{\frac{m}{n-2}} \right]^{\frac{n-2}{m}} - \sigma_p^n t_p \right\} - \frac{B}{\sigma_a^2},$$

where $C = \frac{\frac{B}{\sigma_p^2} - \frac{t_u}{n+1}}{t_p}$, when the loading time of the proof test is less than or equal to the unloading rate $t_u \le (n-2)\frac{B}{\sigma_p^2}$; B and n—fiber fatigue parameters; m—Weibull modulus of the unimodal Weibull distribution; S_0^m—Weibull strength measure; F—tensile force; σ_p—load during proof test; $t_p = t_d + \frac{t_l + t_u}{n+1}$; t_d—holding time; t_l and t_u are the loading and unloading times of the proof test, respectively; L_0/L is the ratio of the test length L_0 to the simulated operational length L.

The intrinsic inert strength of quartz in liquid nitrogen is about 14 GPa [17, 18]. The typical fiber load during the proof test is 0.7 GPa, and the recommended permanent operating load is no more than 80% of the load during factory technical control. In this case, with a probability close to 1, the service life of optical fiber is 25 years. After that period, the rate of damage to the optical fiber will increase.

6. Aging of optical cable

The basis of the design of optical cables is plastics that are susceptible to aging. The aging process occurs when plastics gradually lose substances over time, such as plasticizers (polyvinylchloride) and antioxidants (polyethylene et al.), which are slowly evaporated.

Evaporation (diffusion) of plasticizers plasticizer is described by Fick's law

$$\frac{d}{dx}\left(D\frac{\partial c}{\partial x}\right) = \frac{dc}{d\tau},$$

where c is the plasticizer concentration per unit volume, kg/m^3; x—diffusion depth, m; τ—time, s; D—diffusion coefficient, m^2/s.

Chemical and physical parameters of plasticizers are given in Table 3.

The partial vapor pressure of the plasticizer P is expressed by the approximate ratio

$$P \approx P_0 e^{-\frac{E}{RT}},$$

Table 3

Chemical and physical parameters of plasticizers

Parameter name	**Parameter value for plasticizer**		
	Dibutyl phthalate	Dioctyl phthalate	Tricresyl phthalate
Chemical formula	C11H22O4	C24H38O4	C21H21O4P
Molecular mass μ, g/mol	278	390	368
Density, g/cm^3	1.046	0.986	1.162
Partial pressure constant P_0, Pa	2.1×10^{11}	8.2×10^{11}	9.4×10^{12}
Evaporation energy E, J/mol	7.40×10^4	8.73×10^4	10.44×10^4

where P_0—partial pressure constant given in Table 3; $R = 8.314$ J/(mol×K)—universal gas constant; T—temperature, K.

The plasticizer evaporation rate constant, reduced to the effective operating temperature k_{re}, is expressed by the following equation:

$$k_{re} = k_i e^{-\frac{E}{R}\left(\frac{1}{T_p} - \frac{1}{T_i}\right)},$$

where k_i is the evaporation (desorption) rate constant of the plasticizer at the i-th temperature, s^{-1}; T_i—test temperature, K; T_p—effective operating temperature of the product, K. The dependence of the rate of antioxidant consumption on temperature is determined by the Arrhenius equation

$$k_2 = k_b e^{a\vartheta}, \quad (1)$$

where k_2 is the constant of the rate of antioxidant consumption during polyethylene oxidation; $k_b = k_{02} e^{-\frac{U_A}{RT_b^2}}$; $a = \frac{U_A}{RT_b^2}$; $\vartheta = T - T_b$; U_A—activation energy of the antioxidant consumption process; k_{02}—constant factor; T_b is the base temperature above which the oxidation reaction proceeds at a rate recorded by a microcalorimeter.

According to (1), the deviation in antioxidant concentration has the form

$$W_a^0 - W_a = \frac{k_b}{a\vartheta_T}(e^{a\vartheta} - 1),$$

where W_a^0 is the initial concentration of the antioxidant; W_a—the final concentration of antioxidant; ϑ_T—rate of temperature change.

As a result of aging, the mass of plastic changes (decreases) over time. So, a diagnostic parameter for cable durability is

$$\Delta G = 100\frac{G_i - G_f}{G_i}\%,$$

where ΔG—relative mass loss, %; G_i—initial sample weight, mg; G_f—final sample weight after heating, mg.

Along with the change in mass due to evaporation of the plasticizer or antioxidant the electrical conductivity, dielectric loss tangent, electrical strength, elongation at break, deformation and deformation rate of plastics, glass transition temperature and cold resistance of plastics change. These physical characteristics can also be used to determine the durability ratings.

7. Emergency cable reserve

When the risk of failures of a fiber-optic cable increases, it is advisable to optimize the emergency reserve (ER), which includes the estimated cable length, couplings, and fittings.

For optimization, the initial data for calculation are determined:

- the number of identical elements with the same time-to-failure value in a component of the i-th type (cable, coupling)—m_i;
- failure rate of elements of the i-th type—λ_i, 1/hour;
- time periods $t_{d1} = 10$ thousand hours—the period of ER revision, comparable to one year of operation of the FOCL, and $t_{d2} = 100$ thousand hours—period comparable to the service life of storage of the optical cable;
- number of components—S;
- FOCL operation intensity coefficient—K_I;
- indicator of the availability of the product with the i-th component in the ER set—P_A;
- availability factor—K_A.

The calculation algorithm is the following:

- b_i is calculated as the average number of failures of components of the i-th type for one product during the estimated time t_d: $b_i = m_i \lambda_i t_d K_I$;
- the indicator of spare parts availability for a product P_A is set;
- the average indicator of the lack of ER components is calculated: $q_{iA} = (1 - P_A)/S$;
- using the known values of q_{iA} and b_i, the values of m_i (the number of elements or components in the ER) are determined.

Consider, as an example, the calculation of the ER of an optical cable in a lightning-protection cable (OCLC) for a communication line built using FOCL-OHL (overhead power line) technology. A FOCL, constructed using FOCL-OHL technology, consisting of OCLC, connecting cable joints, support and tension fittings, is considered as a whole, since data on the damageability of lightning protection cable on OHL with a rated voltage of 110, 220, 330 and 500 kV are known. From the point of view of reliability theory, one of the key components of a fiber-optic transmission system is the average length of the OCLC insert (1.5 km), which is used in damage repair technology. Initial data are given in Table 4. The calculation results for b_i, $m(t_{d1})$, $m(t_{d2})$, $l(t_{d1})$, $l(t_{d2})$ are given in Table 5, where b is the average number of failures during the estimated time; m—number of elements; t_{d1}—the period of revision of the spare set (comparable to one year of operation); t_{d2}—period comparable to service life; λ—failure rate of elements of the i-th type λ_i; l—optical cable insertion length (multiple of the average insertion length of 1.5 km). The average indicator of the product's lack of components in the spare parts kit according to the formula $q_{iA} = (1 - P_A)/S$, where S is the number of components.

Let us consider the time it takes to deliver an ER to the accident site. According to current regulations for the design, construction, and operation of FOCL, the standard average time for restoring the working condition of a FOCL is 10 hours. Based on expert estimates, restoring the working state of an FOCL using a temporary optical cable insert takes no more than 6 hours, so the standard time for delivering an ER is 4 hours. Assuming an average speed of 50 kilometers per hour on Russian roads, the distance between the accident site and the nearest ER storage base should not exceed 200 kilometers, and the distance between ER bases should not exceed 400 kilometers. As the service life of the equipment approaches its end (and, consequently, the failure rate increases), operational reliability can be ensured by locating ERs at amplification points, where the distance between storage bases is less than 150 km.

Table 4

Initial data for optimizing ER

Indicator	**Rated voltage of overhead power line (OHL), kV**			
	110	220	330	500
Accident density per 100 km per year	0.08	0.05	0.04	0.03
Normative time to repair t_r, hour	10	10	10	10
Coefficient of availability K_A (100 km)	0.99990	0.99994	0.99995	0.99996
Coefficient of availability K_A (1.5 km)	0.9999985-074	0.9999991-045	0.9999992-537	0.9999994-030
Time to failure T_0, hour	6699709	11166936	13399427	16750408
λ, $1/h \times 10^7$	1.4926	0.8955	0.7463	0.5970
t_{d1}, h	10 000			
t_{d2}, h	100 000			
S	1			
K_I	1 (24 hours per day)			
P_A	0.995			
q_A	0.005			

Table 5

Results of calculation of ER

Indicator	**Nominal voltage of overhead lines, kV**			
	110	220	330	500
λ, $1/h \times 10^7$	1.4926	0.8955	0.7463	0.5970
$b(t_{d1})$	0.1	0.06	0.05	0.04
$b(t_{d2})$	1.0	0.6	0.5	0.4
$m(t_{d1})$	1	1	1	1
$l(t_{d1})$, km	1.5	1.5	1.5	1.5
$m(t_{d2})$	4	3	3	2
$l(t_{d2})$, km	6	4.5	4.5	3.0

8. Conclusion

Fiber optic cables are an essential part of the national information infrastructure, widely used in fixed communication networks, mobile networks, including the upcoming 5G, 6G, and 7G technologies [19, 20]. Optical fibers and optical cable are considered aging objects according to reliability theory. Scientists from both domestic and abroad have developed advanced models to predict their aging. Due to various factors, the accuracy of predicting the maximum service life for optical fiber and optical cable can be quite inaccurate. Cable manufacturers, however, typically provide a lower estimate for the service life.

With sufficient evidence, it has been shown that it is possible to maintain cable reliability beyond the specified service life by optimizing ER locations and, accordingly, reducing the time to repair in case of failures.

9. Results

Long-term observations of the parameters of the G.652 optical fiber in the backbone cable have shown that all average sample values of the attenuation coefficient (0.190–0.197) fall within the confidence interval of the mathematical expectation (0.184–0.202). It has also been shown that the average maximum value of polarization-mode dispersion reduced to a length of 1 km (0.235 ps/$\sqrt{km}$) is significantly lower than the normalized quantile (0.5 ps/$\sqrt{km}$). With this reduced polarization-mode dispersion value, the total polarization-mode dispersion over the entire observation area between Kadala and Skovorodino (1027.3 km) will not exceed 7.5 ps.

10. Discussion

When calculating the required emergency cable reserve in case of continued operation after the warranty period, it is necessary to take into account several factors. These include the number of identical components with the same time to failure, the intensity of failure flow, the audit periods compared to the warranty period, security indicators, and the intensity factor of fiber-optic communication line operation. The optimization of the calculation using a proposed algorithm allows us to ensure the specified availability coefficient of the communication line.

Author Contributions: Conceptualization, Alexander Tsym; methodology, Denis Paltsin; validation, Alexander Tsym; investigation, Denis Paltsin; writing—original draft preparation, Denis Paltsin; writing—review and editing, Denis Paltsin. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Data supporting this study cannot be made available due to commercial restrictions.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Mahlke, G. & Gössing, P. *Fiber Optic Cables: Fundamentals, Cable Planning, Systems Planning* (Siemens Aktiengesellschaft, 1993).
2. Tarasov, D., Ovchinnikova, I., Meschanov, G., Gordienko, V. & Tsym, A. *Quartz-glass Optical Fibre Time to Fracture at Small Bending Radiuses* in (Mar. 2020), 1–5. doi:10.1109/IEEECONF48371.2020.9078607.
3. Snyder, A. & Love, J. *Optical Waveguide Theory* (Springer US, 2012).
4. *Stress-strein characteristics of selfsupporting aerial optical fibre cables* IWCS (1991), 178–185.
5. *A High-speed coating process for optical fibre ribbon* IWCS (1991), 550–555.
6. Korn, G. & Korn, T. *Mathematical Handbook for Scientists and Engineers: Definitions, Theorems, and Formulas for Reference and Review* (Dover Publications, 2013).

7. *Definitions and test methods for linear, deterministic attributes of single-mode fibre and cable* tech. rep. G.650.1 (Recommendation ITU-T, Jan. 2024).
8. *Characteristics of a single-mode optical fiber and cable* tech. rep. G.652 (Recommendation ITU-T, Nov. 2009).
9. *Characteristics of a non-zero dispersion-shifted single-mode optical fibre and cable* tech. rep. G.655 (Recommendation ITU-T, Nov. 2009).
10. *Characteristics of a dispersion-shifted, single-mode optical fibre and cable* tech. rep. G.653 (Recommendation ITU-T, July 2010).
11. Maslo, A., Hodzic, M., Skaljo, E. & Mujcic, A. Aging and Degradation of Optical Fiber Parameters in a 16-Year-Long Period of Usage. *Fiber and Integrated Optics* **39,** 1–14. doi:10.1080/01468030.2020.1725185 (Feb. 2020).
12. *Characteristics of a fibre and cable with non-zero dispersion for wideband optical transport* tech. rep. G.656 (Recommendation ITU-T, July 2010).
13. Sultanov, A. & Vinogradova, I. *Optical Fiber for Telecommunication in Russia* in *Proceedings of SPIE* (Oct. 2001), 78–88. doi:10.1117/12.445695.
14. I., B. V. The effect of cross-border fibre-optic transitions on the information and communication connectivity of the Russian cities. *Baltic Region* (2018).
15. Mane, S. Fiber Optics in Communication Networks: Trends, Challenges, and Future Directions. *International Journal of All Research Education and Scientific Methods (IJARESM),* 607–612 (July 2023).
16. *Characteristics of a bending-loss insensitive single-mode optical fibre and cable* tech. rep. G.657 (Recommendation ITU-T, Nov. 2016).
17. Glaesemann, G. *Optical Fiber Mechanical Reliability. Review of Research at Corning's Optical Fiber Strength Laboratory. White Paper.* tech. rep. WP8002 (Corning Incorporated, Corning, New York, USA, July 2017), 62.
18. *Characteristics of a cut-off shifted single-mode optical fibre and cable* tech. rep. G.654 (Recommendation ITU-T, Mar. 2020).
19. Tonkih, E. Analysis of ITU-T and ITU-R recomendations on fith generation communication networks. Part II. *Work of NIIR.* doi:10.34832/NIIR.2021.7.4.001 (Dec. 2021).
20. Tonkih, E. Analysis of ITU-T and ITU-R recomendations on fith generation communication networks. Part I. *Work of NIIR.* doi:10.34832/NIIR.2021.6.3.001 (Sept. 2021).

Information about the authors

Denis A. Paltsin—Director at the Access Network Research Center, M.I. Krivosheev National Research Centre for Telecommunication (e-mail: palcinda@niir.ru, phone: +7(495) 647-17-77, ORCID: 0009-0005-6394-5022)

Alexander Yu. Tsym—Doctor of Technical Sciences, Chief researcher at the Access Network Research Center, M.I. Krivosheev National Research Centre for Telecommunication (e-mail: tsymay@niir.ru, ORCID: 0009-0002-2708-7065)

УДК 535.8:004.725
PACS 42.81.–i, 42.60.Lh,
DOI: 10.22363/2658-4670-2024-32-3-294–305 EDN: BGKJTK

Сохранение надежности сетей связи при продолжении эксплуатации оптических кабелей за пределами их гарантийного срока службы

Д. А. Пальцин, А. Ю. Цым

Ордена Трудового Красного Знамени Российский научно-исследовательский институт радио имени М. И. Кривошеева, ул. Казакова, д. 16, Москва, 105064, Российская Федерация

Аннотация. В настоящее время, все большее количество волоконно-оптических линий связи оказываются в ситуации, когда гарантированная производителем продолжительность эксплуатации кабеля достигает заданного срока службы, однако качественные показатели линии допускают продолжение ее работы. Продление срока реальной эксплуатации требует качественного учета, как текущего состояния всех составляющих, так и динамики основных показателей. В статье предлагается метод решения проблемы сохранения надежности сетей связи при продолжении эксплуатации оптических кабелей за пределами их гарантийного срока службы. Исследование случайных показателей коэффициента затухания и поляризационно-модовой дисперсии оптического волокна, подкрепленное реальными эксплуатационными данными фрагмента сети, показывает высокую временну́ю стабильность коэффициента затухания и поляризационно-модовой дисперсии оптического волокна G.652. Данный вывод позволяет говорить о продолжении эксплуатации оптических кабелей за пределом гарантийного периода. Для анализа ключевой метрики — старения — используются математические модели, учитывающие физико-химические свойства кабелей и условия проводимых для них контрольных испытаний. На примере, актуальном для текущего состояния российских оптоволоконных сетей, расчитывается количество элементов аварийного запаса, необходимых для поддержания уровня надежности их эксплуатации. Также приводятся практические рекомендации по размещению аварийного запаса.

Ключевые слова: сети связи, волоконно-оптические кабели, срок службы, эксплуатационный показатель надежности

Discrete & Continuous Models & Applied Computational Science
ISSN 2658-7149 (Online), 2658-4670 (Print)

2024, 32 (3) 306–318
http://journals.rudn.ru/miph

Modeling and simulation

Research article

UDC 004.021:519.2:519.6
DOI: 10.22363/2658-4670-2024-32-3-306–318
EDN: FEMNAB

Symbolic-numeric approach for the investigation of kinetic models

Ekaterina A. Demidova[1], Daria M. Belicheva[1], Victoria M. Shutenko[1], Anna V. Korolkova[1], Dmitry S. Kulyabov[1,2]

[1] RUDN University, 6 Miklukho-Maklaya St, Moscow, 117198, Russian Federation

[2] Joint Institute for Nuclear Research, 6 Joliot-Curie St, Dubna, 141980, Russian Federation

(received: June 4, 2024; revised: June 20, 2024; accepted: June 25, 2024)

Abstract. Our group has been investigating kinetic models for quite a long time. The structure of classical kinetic models is described by rather simple assumptions about the interaction of the entities under study. Also, the construction of kinetic equations (both stochastic and deterministic) is based on simple sequential steps. However, in each step, the researcher must manipulate a large number of elements. And once the differential equations are obtained, the problem of solving or investigating them arises. The use of symbolic-numeric approach methodology is naturally directed. When the input is an information model of the system under study, represented in some diagrammatic form. And as a result, we obtain systems of differential equations (preferably, in all possible variants). Then, as part of this process, we can investigate the resulting equations (by a variety of methods). We have previously taken several steps in this direction, but we found the results somewhat unsatisfactory. At the moment we have settled on the package Catalyst.jl, which belongs to the Julia language ecosystem. The authors of the package declare its relevance to the field of chemical kinetics. Whether it is possible to study more complex systems with this package, we cannot say. Therefore, we decided to investigate the possibility of using this package for our models to begin with standard problems of chemical kinetics. As a result, we can summarize that this package seems to us to be the best solution for the symbolic-numerical study of chemical kinetics problems.

Key words and phrases: chemical kinetics equations, stochastic differential equations, population models, one-step processes

For citation: Demidova, E. A., Belicheva, D. M., Shutenko, V. M., Korolkova, A. V., Kulyabov, D. S. Symbolic-numeric approach for the investigation of kinetic models. *Discrete and Continuous Models and Applied Computational Science* **32** (3), 306–318. doi: 10.22363/2658-4670-2024-32-3-306–318. edn: FEMNAB (2024).

1. Introduction

The chemical kinetics equations (chemical reaction networks, CRN) are a simplified version of the stochastic kinetic equations. In the works of C. W. Gardiner [1] and N. G. Van Kampen [2] the chemical kinetics equations are derived from the more general stochastic kinetic equations.

Model representation in the form of CRN is mostly used in biochemistry, theoretical chemistry and biology. Such models are based on a combination of a set of substances, which defines the state of the system, and a set of reaction events, which describe reaction rates and rules for changing the state of the system as they occur. This structure makes it easier to understand and analyze the model. Each reaction includes reagents (initial substances) and products (substances that are formed as a result of a reaction), which are written as concentrations of objects in the extended sense.

For instance, in the work [3] A. Lotka derived a system of stochastic equations from hypothetical chemical reaction. The same model was independently arrived at by V. Volterra [4]. This model describes predator–prey species interaction.

The equations of chemical kinetics are constructed according to rather primitive (but also somewhat cumbersome) rules. It seems justified to use the analytic–numerical approach for these tasks. This paper provides an overview of the main functions of the Catalyst.jl library [5, 6] for the programming language Julia [7, 8], which provides a toolkit for symbolic–numerical exploration of chemical kinetics models [9]. The Julia programming language is widely used in biology, providing a large number of tools [10, 11]. Due to Julia's support for the metaprogramming paradigm, this language is well-suited for implementing custom domain–specific languages. For example, it is possible to implement a computer algebra subsystem integrated into the language [12].

1.1. Paper structure

The section 2 briefly provides an overview of the basic principles of chemical kinetics. The section 3 provides brief information about the Catalyst.jl package. The sections 4 and 5 consider examples of constructing and solving one-dimensional and multidimensional models, respectively. The models are presented in both deterministic and stochastic forms.

2. Chemical kinetics equations

Representation of kinetics equations in the form of chemical reactions is widely used in modeling. We will briefly examine the structure of equations of this type. Consider mixtures of chemical substances X_a, $a = \overline{1,n}$. Each substance has a concentration x_a. A reaction occurs under the influence of a catalyst, and the mixture changes. This is expressed as changes in the reaction orders with respect to the substances, i.e., changes in the amounts of substances, which are represented by the sets of coefficients N_a^A, M_a^A. The total amount of substances involved in a reaction is called the reaction order. In this case, the reaction is characterized by constant proportionality coefficients k^+ и k^-, which characterize the intensity of the processes—the reaction rate.

The equation of a reversible chemical reaction in a general form is considered [13–16]:

$$\sum_a N_a^A X_a \underset{k_A^-}{\overset{k_A^+}{\rightleftarrows}} \sum_a M_a^A X_a, \qquad A = \overline{1,m}, \quad a = \overline{1,n}. \tag{1}$$

In the equation (1), M_a^A and N_a^A represent the number of components of type X_a on the left and right sides, respectively. In the interaction of type A, a set N_a^A of components of type X_a enters, and a set M_a^A of components of type X_a, or M_b^A of type X_b ($b \neq a$) is informed. The interaction is similar in the reverse direction.

Velocity is considered proportional to the concentration of substances [2]:

$$s_A^+ = k^+ \prod_a {x_a}^{N_a},$$

$$s_A^- = k^- \prod_a {x_a}^{M_a}.$$

The state of the system $\mathbf{x} = (x_1, x_2, \dots, x_n)$ is introduced, where x_a is the substance concentration, i.e., the number of elements of type X_a. The change in the state of the system is described by the vector $\mathbf{r}^A$:

$$\mathbf{r}^A = \mathbf{M}^A - \mathbf{N}^A.$$

To construct a system of differential equations corresponding to the interaction scheme (1), we will consider the Fokker–Planck equation [1]:

$$\partial_t p(X,t) = -\sum_a \partial_a [A_a(x)P(X,t)] + \frac{1}{2}\sum_{a,b} \partial_a \partial_b [B_{ab}(X)P(X,t)], a = \overline{1,n}, b = \overline{1,n},$$

where

$$\begin{aligned} A_a(X) &= \sum_A r_a^A [s_A^+(x) - s_A^-(x)], \\ B_{a,b}(X) &= \sum_A r_a^A r_b^A [s_A^+(x) + s_A^-(x)]. \end{aligned} \tag{2}$$

The Fokker–Planck equation is mathematically equivalent to the Langevin equation. In the Langevin equation

$$\mathrm{d}\mathbf{x} = \mathbf{a}(\mathbf{x})\,\mathrm{d}t + \mathbf{b}(\mathbf{x})\,\mathrm{d}\mathbf{W},$$

where W—n-dimensional Wiener process, the coefficient $a(x)$ ccorresponds to the coefficient $A(X)$, $b(x) = B(x)B(x)^T$ [17, 18].

We discard the stochastic term and use the coefficients of A_a from equation (2) to obtain a system of differential equations:

$$\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = \sum_A r_a^A [s_A^+(x) - s_A^-(x)].$$

3. The Catalyst library

Numerical modeling of chemical reactions written with the `Catalyst.jl` library is usually performed with the `DifferentialEquations.jl` [19] package. To work with it, the reaction system is converted into the desired problem type from this package. It contains a large number of numerical solution methods and additional functions. The obtained solutions can be visualized with the `Plots.jl` package by specifying the necessary parameters and time interval. Using `Catalyst.jl`, reaction systems can be represented as deterministic and stochastic models.

In terms of performance, the `Catalyst.jl` package significantly outperforms other chemical reaction network (CRN) modeling tools such as `BioNetGen`, `COPASI`, `GillesPy2`, `Matlab` and `SimBiology` [6]. The performance of models created with Catalyst depends on a variety of factors. For instance, Catalyst builds in all reaction conditions in the ordinary differential equations (ODE), which allows the compiler to optimize computation and reduces function call overhead.

Consider the simplest reversible reaction:

$$X + Y \underset{k^-}{\overset{k^+}{\rightleftarrows}} XY,$$

where X, Y and XY are the types of some objects, and k^+, k^- are reaction parameters. Parameters can be either constants and functions of time or other components of the system.

The `@reaction_network` macro is used to symbolize a chemical reaction.

First, we load the `Catalyst.jl` package using the command `using Catalyst`.

Then we assign the reaction to the variable `rn`, where the reaction is described using the `@reaction_network` macro:

```
rn = @reaction_network begin
1.0, X + Y --> XY
1.5, XY --> X + Y
end
```

Here, the proportionality constants k^+ and k^- are specified first, followed by the corresponding reaction, separated by comma. One reversible reaction is written as two irreversible reactions. If it is necessary to specify a reaction in which reactants are produced/destroyed from nothing, we write the necessary part as 0 (it is perceived as an empty set).

The `Catalyst.jl` package allows to use different ways of writing arrows. Unidirectional arrows can be written in both directions and with Unicode characters. Accordingly, one reaction can be represented as four equivalent variants:

```
rn = @reaction_network begin
  1.0, X + Y --> XY
  1.0, X + Y → XY
  1.0, XY ← X + Y
  1.0, XY <-- X + Y
end
```

As a result of executing the code, we will get:

$$X + Y \xrightarrow{1.0} XY,$$
$$X + Y \xrightarrow{1.0} XY,$$
$$X + Y \xrightarrow{1.0} XY,$$
$$X + Y \xrightarrow{1.0} XY.$$

Speaking of bidirectional arrows, both two-line and one-line entries are possible. The following variants of writing, which are also equivalent, are obtained. The first variant consisting of two reactions (forward and reverse):

```
rn = @reaction_network begin
  2.0, X + Y --> XY
  2.0, XY <-- X + Y
end
```

The second option represents a bidirectional reaction:

```
rn = @reaction_network begin
  (2.0, 2.0), X + Y <--> XY
end
```

The output will be the same in both cases:

$$X + Y \underset{2.0}{\overset{2.0}{\rightleftarrows}} XY.$$

The `Catalyst.jl` package also allows to use different ways to combine reactions. The combining of two or more reactions are considered:

– write both reactions on the same line – `1.0, S --> (P1, P2);`
– write each reaction on a separate line –

```
1.0, S --> P1
1.0, S --> P2
```

– combine reactions with different parameters – `(1.0, 2.0), (S1, S2) --> (P1,P2).`

In the case where reversible reactions are to be combined, the following notation may occur:

`(1.0,(1.0,2.0)), S <--> (P1,P2);`

As a result, we get:

$$S \xrightarrow{1.0} P1,$$
$$S \xrightarrow{1.0} P2,$$
$$P1 \xrightarrow{1.0} S,$$
$$P2 \xrightarrow{2.0} S.$$

4. Example of one-dimensional model

Consider the birth–death model of kinetic reactions:

$$\begin{gathered} X \xrightarrow{1} 2X, \\ X \underset{100}{\overset{2}{\rightleftarrows}} 0. \end{gathered} \tag{3}$$

We set the differential equations in stochastic and deterministic forms that correspond to this chemical reaction (3). We write down the vectors describing the state of the system. In our case, they will be one-dimensional:

$$r^1 = 2 - 1 = 1,$$
$$r^2 = 0 - 1 = -1.$$

We consider the probabilities of the transitions. The first reaction is irreversible, so $s_1^- = 0$.

$$s_1^+ = 1x^1 = x,$$
$$s_2^+ = 2x^1 = 2x,$$
$$s_2^- = 100x^0 = 100.$$

Using the transition probabilities, we can write the Fokker–Planck equation:

$$\partial_t p(X,t) = -\sum_a \partial_a [A_a(x)P(X,t)] + \frac{1}{2}\sum_{a,b}[B_{ab}(X)P(X,t)],$$

where

$$A(X) = r^1 s_1^+ + r^2[s_2^+ - s_2^-] = x - 2x + 100 = 100 - x,$$
$$B(X) = r^1(r^1)^T s_1^+ + r^2(r^2)^T[s_2^+ - s_2^-] = x + 2x - 100 = 3x - 100.$$

We proceed to write a stochastic differential equation in Langevin form. To do this, we need to extract the square root of $B(x)$.

$$\mathrm{d}x = a(x)\,\mathrm{d}t + \sqrt{B(x)}\,\mathrm{d}W = (100 - x)\,\mathrm{d}t + \sqrt{3x - 100}\,\mathrm{d}W.$$

[24]:

$$\mathrm{X} \xrightarrow{1.0} 2\mathrm{X}$$
$$\mathrm{X} \underset{100.0}{\overset{2.0}{\rightleftharpoons}} \varnothing$$

Figure 1. Result of the chemical reaction initialization code

By removing the stochastic term, we get a deterministic differential equation.

Now we will consider an example of writing and finding a solution to a birth–death model using `Catalyst.jl`. The system (3) is set by the variable `rn`:

```
rn = @reaction_network begin
  1.0, X --> 2X
  2.0, X --> 0
  100.0, 0 --> X
end
```

We get a system of chemical reactions (Fig. 1).

Here, the first reaction means reproduction of species, the second one means extinction, and the third one sets a constraint on the population size.

For numerical modeling we use the package `DifferentialEquations.jl`, which allows solving a wide range of DEs. The resulting reaction system is converted into a differential equation using the ODESystem method, and then the initial condition and the time interval on which the solution is sought are set:

```
rnsys = convert(ODESystem, rn)

@variables t
@species X(t)
u0 = [X => 10]
tspan = (0, 10.0)
symsys = structural_simplify(rnsys)
rnprob = ODEProblem(symsys, u0, tspan)

sol = solve(rnprob, Tsit5())
```

The macros `@variables` and `@species` are used here, where t is time and $X(t)$ is the population concentration (number of individuals) dependent on t.

We can notice that the system of differential equations (Fig. 2) coincides with the theoretically derived deterministic part of the equation (4).

The results of modeling can be visualized using the Plots.jl package (Fig. 3).

It is also possible to convert the system of chemical reactions into a stochastic differential equation in Langevin form, which will have the form:

$$\frac{\mathrm{d}X}{\mathrm{d}t} = \big(100 - X(t)\big)\,\mathrm{d}t + 100\,\mathrm{d}W_1 - X(t)\,\mathrm{d}W_2.$$

In this case, the stochastics are different from those derived in (4).

[30]:

$$\frac{\mathrm{d}X(t)}{\mathrm{d}t} = 100 - X(t)$$

Figure 2. The result of code converting a death-birth chemical reaction systems into deterministic differential equations

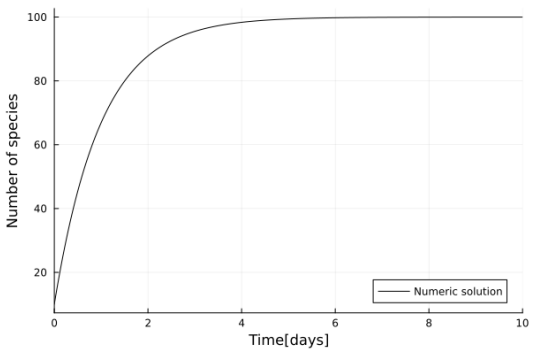

Figure 3. The birth–death model

```
sprob = SDEProblem(rn, u0, tspan)
sol_stoch = solve(sprob, EM(), dt = 0.001)
```

As a result, we get a graph of the solution (Fig. 4).

In addition, we can represent the model as a jump process. To define a jump process, we use the `JumpProblems` method. First, from the network of chemical reactions, we construct another type of problem with which transitions will be associated, namely we create a `DiscreteProblem` using Gillespie's direct method (`Direct()`), which sets constant transition rates. For the solution, a high-performance integrator `SSAStepper()` for pure jump problems (with constant transition rates) is passed to the `solve` function:

```
dprob = DiscreteProblem(rn, u0, tspan)
jprob = JumpProblem(rn, dprob, Direct())
jsol = solve(jprob, SSAStepper())
```

As a result, we get the solution graph (Fig. 5).

5. Example of a multidimensional model

Also, using this library we can set n-dimensional reactions. In general, the n-dimensional predator prey system can be written as:

$$\frac{\mathrm{d}x^i}{\mathrm{d}t} = x^i\left(b^i + \sum_{j=1}^{n} a_j^i x^j\right),$$

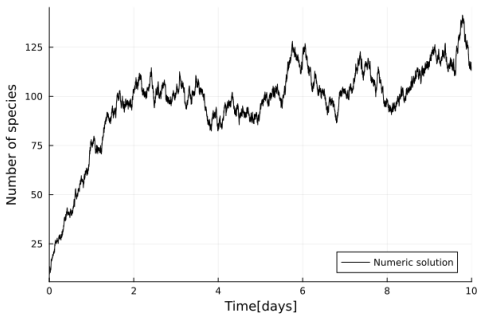

Figure 4. Stochastic birth–death model. Solution using the Euler–Maruyama method

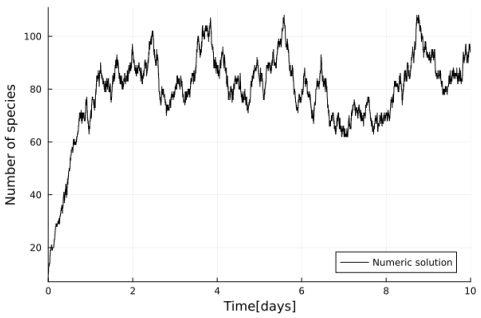

Figure 5. A stochastic birth–death model. Solution using the direct Gillespie method

where $i \neq j$, $\vec{x} = (x^1, x^2, \dots, x^n)$—$n$ species, $\vec{b} = (b^1, b^2, \dots, b^n)$—natural death or birth parameters (greater than zero for autotrophic species, less than zero for heterotrophic species); $\hat{A} := a^i_j$—parameters describing the interaction between species (if greater than zero, the individual is eaten, if less, the individual is born):

$$\hat{A} := a^i_j = \begin{pmatrix} 0 & a^1_2 & a^1_3 \\ a^2_1 & 0 & a^2_3 \\ a^3_1 & a^3_2 & 0 \end{pmatrix}.$$

Bazykin's work [20] examines various types of interactions between three populations, consider the producer-consultant-predator system (4).

$$\begin{aligned} X_i &\overset{p_{i0}}{\to} 2X_i, \\ X_i + X_j &\overset{p_{ij}}{\to} X_j, \end{aligned} \tag{4}$$

where $i, j = 1, \dots, n$.

We set it with specific parameters using `Catalyst` and find the solution using the numerical method `Tsit5()`:

```
lv = @reaction_network begin
  1.5, X + Y --> 2*Y
  3, Y --> 0
  1.5, X + Z --> 2*X
  3, Z --> 2*Z
  end

  @variables t
  @species X(t) Y(t) Z(t)
  u0 = [X => 1.0, Y => 1.0, Z => 1.0]
  tspan = (0.0, 16.0)
  symsys = structural_simplify(lvsys)
  lvoprob = ODEProblem(symsys, u0, tspan)
  sol_lv_ode = solve(lvoprob, Tsit5())
```

[5]:

$$\begin{aligned}\frac{\mathrm{d}X(t)}{\mathrm{d}t} &= -X(t) - 1.5Y(t)X(t) + 1.5X(t)Z(t)\\ \frac{\mathrm{d}Y(t)}{\mathrm{d}t} &= -3Y(t) + 1.5Y(t)X(t)\\ \frac{\mathrm{d}Z(t)}{\mathrm{d}t} &= 3Z(t) - 1.5X(t)Z(t)\end{aligned}$$

Figure 6. The result of code converting a three-dimensional predator–prey chemical reaction systems into deterministic differential equations

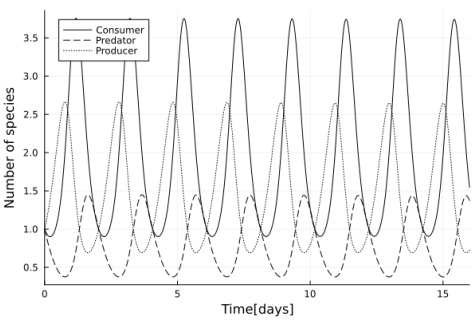

Figure 7. Solutions of the predator–prey three-dimensional system

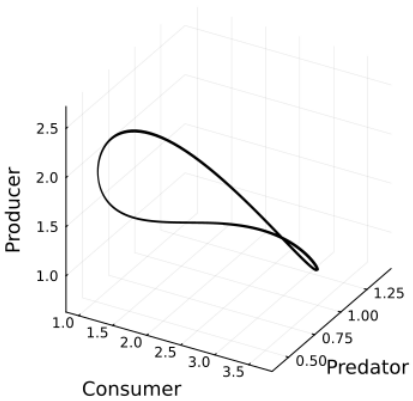

Figure 8. Phase portrait of the predator–prey three-dimensional system

In the form of differential equations, this system of reactions, transformed with `Catalyst`, looks as follows (Fig. 6).

As a result, using the `solve` method, we get a graph of the solution (Fig. 7), and we can display a phase portrait (Fig. 8).

This system can also be solved by set a stochastic problem using the function `SDEProblem` (see Fig. 9, 10):

```
lvoprob = SDEProblem(lvsys, u0, tspan)
```

We also display phase portraits (see Figs. 11, 12). Here additionally, a constraint was set to ensure that the number of individuals does not fall below zero:

```
function condition(u, t, integrator)
    any(u .< 0)
end

function affect!(integrator)
    integrator.u .= max.(integrator.u, 0)
end
```

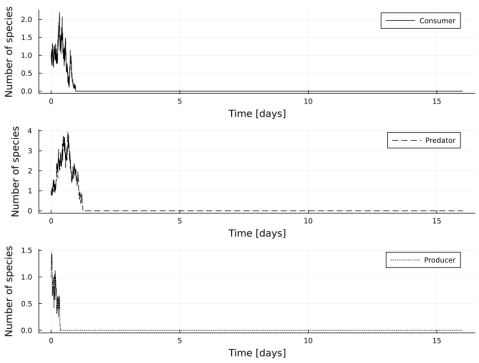

Figure 9. Solution of a three-dimensional stochastic predator–prey system. The case of extinction of all species

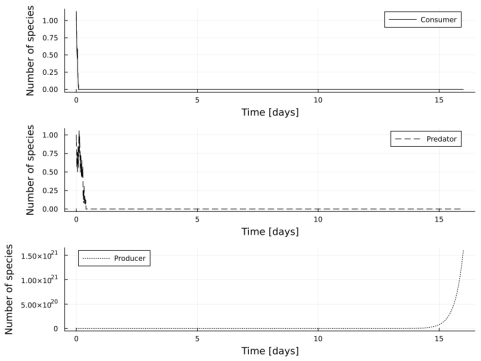

Figure 10. Solution of a three-dimensional stochastic predator–prey system. The case of unrestricted growth of the producer population

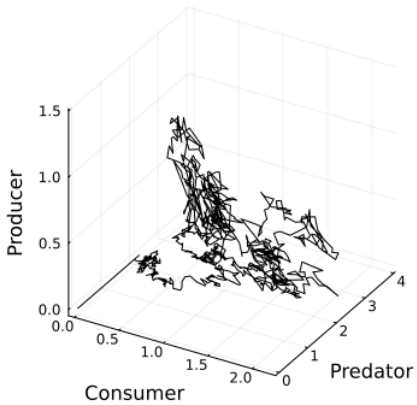

Figure 11. Phase portrait of a three-dimensional stochastic predator–prey system. Case of unrestricted growth of the producer population

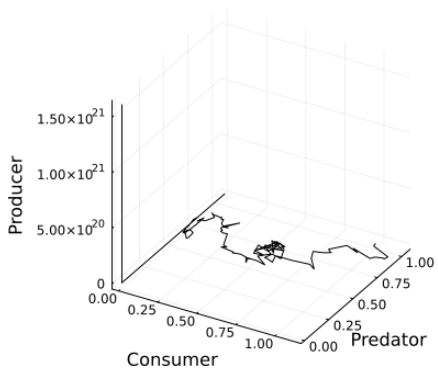

Figure 12. Phase portrait of a three-dimensional stochastic predator–prey system. The case of unrestricted growth of the producer population

```
cb = DiscreteCallback(condition, affect!)

sol_lv_sde = solve(lvsdeprob, EM(), dt = 0.001, callback = cb)
```

We can notice that in the stochastic case, the system can behave in two ways: all species go extinct or the consumers and predators go extinct and the producers multiply unrestrictedly.

6. Conclusion

The paper reviews the tools of the Catalyst.jl library for working with symbolic–numerical notation of kinetic equations, which allows to describe and analyze kinetic processes in a convenient and flexible form. Examples of constructing and solving one-dimensional and multidimensional models in the form of ODEs and SDEs are considered, specifically, birth–death and three-dimensional predator–prey models are constructed and their numerical solutions are found.

Future research plan include exploring the extended functionality of Catalyst.jl with the application of neural networks. For example, in this library it is possible to set reaction rate parameters not only as constants but also as neural networks. Also, complex CRN structures can be approximated using deep learning methods.

Author Contributions: Conceptualization, methodology, Dmitry S. Kulyabov; writing—original draft preparation, Ekaterina A. Demidova, Daria M. Belicheva; writing—review and editing, Victoria M. Shutenko, Anna V. Korolkova. All authors have read and agreed to the published version of the manuscript.

Funding: This work was carried out in the framework of grant support of the RUDN University, project 021934-0-000 (Anna V. Korolkova) and was supported by the program of strategic academic leadership of the RUDN University.

Data Availability Statement: Data sharing is not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Gardiner, C. W. *Handbook of Stochastic Methods: for Physics, Chemistry and the Natural Sciences* (Springer Series in Synergetics, 1985).
2. Van Kampen, N. G. *Stochastic Processes in Physics and Chemistry* (Elsevier Science, 2011).
3. Lotka, A. J. *Elements of Physical Biology* 435 pp. (Williams and Wilkins Company, Baltimore, 1925).
4. Volterra, V. *Leçons sur la Théorie mathématique de la lutte pour la vie* French (Gauthiers-Villars, Paris, 1931).
5. Loman, T. E., Ma, Y., Ilin, V., Gowda, S., Korsbo, N., Yewale, N., Rackauckas, C. & Isaacson, S. A. *Catalyst: Fast Biochemical Modeling with Julia* Aug. 2022. doi:10.1101/2022.07.30.502135. bioRxiv: 2022.07.30.502135.
6. Loman, T. E., Ma, Y., Ilin, V., Gowda, S., Korsbo, N., Yewale, N., Rackauckas, C. & Isaacson, S. A. Catalyst: Fast and flexible modeling of reaction networks. *PLOS Computational Biology* **19** (ed Ouzounis, C. A.) e1011530.1–19. doi:10.1371/journal.pcbi.1011530 (Oct. 2023).
7. Bezanson, J., Karpinski, S., Shah, V. B. & Edelman, A. Julia: A Fast Dynamic Language for Technical Computing, 1–27. arXiv: 1209.5145 (2012).
8. Bezanson, J., Edelman, A., Karpinski, S. & Shah, V. B. Julia: A fresh approach to numerical computing. *SIAM Review* **59,** 65–98. doi:10.1137/141000671. arXiv: 1411.1607 (Jan. 2017).
9. Fedorov, A. V., Masolova, A. O., Korolkova, A. V. & Kulyabov, D. S. *Implementation of an analytical-numerical approach to stochastization of one-step processes in the Julia programming language* in *Workshop on information technology and scientific computing in the framework of the XI International Conference Information and Telecommunication Technologies and Mathematical Modeling of High-Tech Systems (ITTMM-2021)* (eds Kulyabov, D. S., Samouylov, K. E. & Sevastianov, L. A.) **2946** (Moscow, Apr. 2021), 92–104.
10. Roesch, E., Greener, J. G., MacLean, A. L., Nassar, H., Rackauckas, C., Holy, T. E. & Stumpf, M. P. H. *Julia for Biologists* 2021. doi:10.48550/ARXIV.2109.09973. arXiv: 2109.09973.
11. Pal, S., Bhattacharya, M., Dash, S., Lee, S.-S. & Chakraborty, C. A next-generation dynamic programming language Julia: Its features and applications in biological science. *Journal of Advanced Research,* 1–12. doi:10.1016/j.jare.2023.11.015 (Nov. 21, 2023).

12. Kulyabov, D. S. & Korol'kova, A. V. Computer Algebra in JULIA. *Programming and Computer Software* **47,** 133–138. doi:10.1134/S0361768821020079. arXiv: 2108.12301 (Mar. 2021).
13. Laidler, K. J. *Chemical Kinetics* 3rd ed. 531 pp. (Prentice Hall, Inc., Jan. 17, 1987).
14. Korolkova, A. V. & Kulyabov, D. S. One-step Stochastization Methods for Open Systems. *EPJ Web of Conferences* **226** (eds Adam, G., Buša, J. & Hnatič, M.) 02014.1–4. doi:10.1051/epjconf/202022602014 (Jan. 2020).
15. Doi, M. Stochastic theory of diffusion-controlled reaction. *Journal of Physics A: Mathematical and General* **9,** 1479–1495. doi:10.1088/0305-4470/9/9/009 (1976).
16. Schlögl, F. Chemical reaction models for non-equilibrium phase transitions. *Zeitschrift für Physik* **253,** 147–161. doi:10.1007/BF01379769 (1972).
17. Hnatič, M., Eferina, E. G., Korolkova, A. V., Kulyabov, D. S. & Sevastyanov, L. A. Operator Approach to the Master Equation for the One-Step Process. *EPJ Web of Conferences* **108** (eds Adam, G., Buša, J. & Hnatič, M.) 58–59. doi:10.1051/epjconf/201610802027. arXiv: 1603.02205 (2016).
18. Korolkova, A. V., Eferina, E. G., Laneev, E. B., Gudkova, I. A., Sevastianov, L. A. & Kulyabov, D. S. *Stochastization Of One-Step Processes In The Occupations Number Representation* in *Proceedings 30th European Conference on Modelling and Simulation* (ECMS, Regensburg, Germany, June 2016), 698–704. doi:10.7148/2016-0698.
19. Rackauckas, C. & Nie, Q. DifferentialEquations.jl - A Performant and Feature-Rich Ecosystem for Solving Differential Equations in Julia. *Journal of Open Research Software* **5.** doi:10.5334/jors.151 (2017).
20. Bazykin, A. D. *Nonlinear Dynamics of Interacting Populations* ed. by Khibnik, A. I. Ed. by Krauskopf, B. doi:10.1142/2284 (World Scientific, Singapore, May 1998).

Information about the authors

Ekaterina A. Demidova (Russian Federation)—Student of Department of Probability Theory and Cyber Security of RUDN University (e-mail: 1032216451v@rudn.ru, phone: +7 (495) 955-09-27, ORCID: 0009-0005-2255-4025)

Daria M. Belicheva (Russian Federation)—Student of Department of Probability Theory and Cyber Security of RUDN University (e-mail: 1032216453@rudn.ru, phone: +7 (495) 955-09-27, ORCID: 0009-0007-0072-0453)

Victoria M. Shutenko (Russian Federation)—Student of Department of Probability Theory and Cyber Security of RUDN University (e-mail: shutenkovika@yandex.ru, phone: +7 (495) 955-09-27, ORCID: 0000-0003-3922-4805)

Anna V. Korolkova (Russian Federation)—Docent, Candidate of Sciences in Physics and Mathematics, Associate Professor of Department of Probability Theory and Cyber Security of RUDN University (e-mail: korolkova-av@rudn.ru, phone: +7(495) 952-02-50, ORCID: 0000-0001-7141-7610, ResearcherID: I-3191-2013, Scopus Author ID: 36968057600)

Dmitry S. Kulyabov (Russian Federation)—Professor, Doctor of Sciences in Physics and Mathematics, Professor of Department of Probability Theory and Cyber Security of Peoples' Friendship University of Russia named after Patrice Lumumba (RUDN University); Senior Researcher of Laboratory of Information Technologies, Joint Institute for Nuclear Research (e-mail: kulyabov-ds@rudn.ru, phone: +7 (495) 952-02-50, ORCID: 0000-0002-0877-7063, ResearcherID: I-3183-2013, Scopus Author ID: 35194130800)

УДК 004.021:519.2:519.6
DOI: 10.22363/2658-4670-2024-32-3-306–318 EDN: FEMNAB

Символьно-численный подход для исследования кинетических моделей

Е. А. Демидова[1], Д. М. Беличева[1], В. М. Шутенко[1], А. В. Королькова[1], Д. С. Кулябов[1,2]

[1] Российский университет дружбы народов, ул. Миклухо-Маклая, д. 6, Москва, 117198, Российская Федерация

[2] Объединённый институт ядерных исследований, ул. Жолио-Кюри, д. 6, Дубна, 141980, Российская Федерация

Аннотация. Наша группа достаточно долго исследует кинетические модели. Структура классических кинетических моделей описывается достаточно простыми предположениями о взаимодействии исследуемых сущностей. Также построение кинетических уравнений (как стохастических, так и детерминистических) основывается на простых последовательных шагах. Однако на каждом шаге исследователь должен манипулировать большим количеством элементов. А после получения дифференциальных уравнений возникает проблема их решения или исследования. Естественным образом напрашивается использование методологии символьно–численного подхода. Когда на входе представляется информационная модель исследуемой системы, представленная в каком-либо диаграммном виде. А в результате мы получаем системы дифференциальных уравнений (желательно, во всех возможных вариантах). Далее, в рамках этого процесса мы можем исследовать полученные уравнения (разнообразными методами). Ранее нами было предпринято несколько шагов в этом направлении, однако результаты нам показались несколько неудовлетворительными. На данный момент мы остановились на пакете Catalyst.jl, принадлежащему экосистеме языка Julia. Авторы пакета декларируют соответствие пакета области химической кинетики. Возможно ли исследовать с помощью этого пакета более сложные системы, мы сказать не можем. Поэтому исследование возможности применения данного пакета для наших моделей мы решили начать со стандартных задач химической кинетики. В результате мы можем резюмировать, что данный пакет видится нам наилучшим решением для символьно–численного исследования задач химической кинетики.

Ключевые слова: уравнения химической кинетики, стохастические дифференциальные уравнения, популяционные модели, одношаговые процессы

Discrete & Continuous Models & Applied Computational Science
ISSN 2658-7149 (Online), 2658-4670 (Print)

2024, 32 (3) 319–324
http://journals.rudn.ru/miph

Physics and astronomy

Research article

UDC 538.91
DOI: 10.22363/2658-4670-2024-32-3-319–324 EDN: DLVHXG

Liquid radial flows with a vortex through porous media

Yuri P. Rybakov, Natalia V. Semenova

RUDN University, 6 Miklukho-Maklaya St, Moscow, 117198, Russian Federation

(received: April 10, 2024; revised: April 30, 2024; accepted: May 30, 2024)

Abstract. The filtration process is studied for a popular class of filters with radial cartridges that proved their high effectiveness in purification of water. The mass balance equation for radial flows in porous media is obtained by using the lattice approximation method, the transverse diffusion process being taken into account. The Euler dynamical equations are modified by including the Darcy force proportional to the velocity of the filtration flow. The system of equations is written for the stationary axially symmetric radial flow and solved by the perturbation method, if the vertical velocity is supposed to be small.

Key words and phrases: filtration, porous medium, Darcy force

For citation: Rybakov, Y. P., Semenova, N. V. Liquid radial flows with a vortex through porous media. *Discrete and Continuous Models and Applied Computational Science* **32** (3), 319–324. doi: 10.22363/2658-4670-2024-32-3-319–324. edn: DLVHXG (2024).

1. Introduction. The mass balance equation in porous media

The hydrodynamics of liquid flow in a porous medium modeling the grain filling in filters is studied [1–11]. The main concept behind this research appears to be the necessity to modify the fundamental equations of hydrodynamics to meet the requirements of mass and momentum balance under specific conditions of liquid flows through porous media. As can be shown later, bearing on the lattice approximation, the structure of the fluid current and the transverse diffusion coefficient D are derived, the latter proving to be proportional to the diameter d of the grains as constituents of the medium [12–16].

Our study concerns radial filtration process, where the purification proves to be more effective than that for cylindrical geometry. First, let us apply the lattice approximation to the mass balance equation and use the cylindrical coordinates ρ, ϕ, z, with ϕ being the azimuth angle. Let us number the lattice vertices by the indices i, j (transverse to the flow) and k (along the flow), the corresponding cylindrical coordinates being ϕ, z and ρ, respectively. Let us denote the local radial stream of the fluid by

$$G_{ijk} = \Delta S_k \, u_{ijk}, \tag{1}$$

where u_{ijk} stands for the radial velocity of the flow and ΔS_k is the area of the gap between the grains. It means that

$$\Delta S_k = \rho_k \, \Delta\phi \, \Delta z \, S_k, \tag{2}$$

where S_k denotes the porosity of the medium, with the fluid density being taken unity. Therefore, the local mass conservation law reads

$$G_{ijk} = r_{k-1}G_{ijk-1} + p_{k-1}(G_{i-1jk-1} + G_{i+1jk-1}) + q_{k-1}(G_{ij-1k-1} + G_{ij+1k-1}), \tag{3}$$

where the branching coefficients $r,\ p,\ q$ are introduced. Thus, the mass conservation equation reads

$$\sum_{ij} G_{ijk} = \sum_{ij} G_{ijk-1},$$

and implies the constraint on the branching coefficients:

$$r_k + 2(p_k + q_k) = 1. \tag{4}$$

In view of (1), (2) and the constraint (4) one can represent the equation (3) in the form:

$$r_k\rho_k S_k\, u^{(k)}_{\rho ij} - r_{k-1}\rho_{k-1} S_{k-1}\, u^{(k-1)}_{\rho ij} = p_{k-1}\rho_{k-1}S_{k-1}\left(u^{(k-1)}_{\rho i-1j} + u^{(k-1)}_{\rho i+1j}\right) + \\ + q_{k-1}\rho_{k-1}S_{k-1}\left(u^{(k-1)}_{\rho ij-1} + u^{(k-1)}_{\rho ij+1}\right) - 2\rho_k S_k(p_k + q_k)\, u^{(k)}_{\rho ij}.$$

Identifying now the lattice spacing with the diameter d of the grain, one can prove through the latter relation that in the continuous limit the following differential equation is valid:

$$\partial_z\big[\rho S(u - D_z\,\partial_z w)\big] + \partial_\rho(r\,S\rho w) + \partial_\phi(S\,v) - \partial^2_\phi[D_\phi S\,w/\rho] = 0, \tag{5}$$

where the transverse diffusion coefficients are introduced: $D_z = q\,d$, $D_\phi = p\,d$ and the following denotations for the components of the fluid velocity are used: $u_\rho = w$, $u_\phi = v$, $u_z = u$. However, in virtue of the axial symmetry of the flow it is necessary to put $p_k = 0$. Therefore, the equation (5) takes the form of the stationary mass conservation law:

$$div\,\mathbf{j} = 0, \tag{6}$$

with the components of the current $\mathbf{j}$ reading:

$$j_\rho = r(\rho)S(\rho)w, \quad j_z = S(\rho)\big[u - D(\rho)\,\partial_z w\big], \tag{7}$$

where the local transverse diffusion coefficient is introduced:

$$D(\rho) = q(\rho)d(\rho). \tag{8}$$

It is worth while to stress that the effect of the transverse diffusion in porous media is widely discussed in literature [13, 14].

2. The hydrodynamics of the radial flow in porous media: Darcy's law

To find the profiles of the velocity $\mathbf{u}$ and the pressure P, it is necessary to solve the Euler equation, with the force density $\mathbf{f}$ including the gravity acceleration $\mathbf{g}$ and the Darcy force $\mathbf{f}_D = -k_D\mathbf{u}$. In the first approximation, the Darcy coefficient k_D appears to be constant: $k_D = k_0 = const$, but in general it should be some function of the velocity and the pressure [7, 15–21]. In particular, recently some deviations from the standard Darcy's law appear to be evident [22, 23]. Let us now add to the mass balance equation also the stationary Euler equation:

$$(\mathbf{u}\nabla)\mathbf{u} + \nabla P = \mathbf{g} - k_0\mathbf{u}. \tag{9}$$

Let us now rewrite the equations (6) and (9) in cylindrical coordinates:

$$\partial_\rho(r\rho S\, w) + \partial_z\big[\rho S(u - D\,\partial_z\, w)\big] = 0,$$
$$(w\,\partial_\rho + u\,\partial_z)u + \partial_z\, P + g + k_0\, u = 0,$$
$$(w\,\partial_\rho + u\,\partial_z)w - \frac{v^2}{\rho} + \partial_\rho\, P + k_0\, w = 0,$$
$$(w\,\partial_\rho + u\,\partial_z)v + \frac{v\, w}{\rho} + k_0\, v = 0.$$

3. Perturbation method

Let us suppose that the radial part of our filter has the external diameter $2b$, the internal one $2a$ and several plates (layers or cartridges) of the height $2l \ll a$. Therefore, $a \le \rho \le b$, $-l \le z \le l \ll a$ and the boundary condition for the fluid vertical velocity reads:

$$u(\rho, z = \pm l) = 0.$$

In view of the condition $l \ll a$, the vertical velocity u is supposed to be small: $u \ll v,\ w$. Therefore, in the first approximation one can put in equations (2), (2), (2), (2) $u = 0$, $\partial_z\, w = \partial_z\, v = 0$ and obtain the following structure of the velocities $w_0(\rho)$, $v_0(\rho)$ and the pressure $P = -gz + p_0(\rho)$:

$$v_0(\rho) = \frac{C_1\, a}{\rho}\ \exp\!\left(-\int\limits_a^\rho \frac{d\rho}{w_0}\right),$$

$$p_0(\rho) = C_2 - \frac{w_0^2}{2} + \int\limits_a^\rho \left(\frac{v_0^2}{\rho} - k_0\, w_0\right) d\rho,$$

where $w_0(\rho) = C_0(r\rho S)^{-1}$ and C_0, C_1, C_2 denote some constants.

In the second approximation, in view of the boundary condition (3), one can put:

$$u = \alpha z(l^2 - z^2), \quad w = w_0 + \beta z^2, \quad v = v_0 + \gamma z^2;$$

where the functions $\alpha(\rho)$, $\beta(\rho)$, $\gamma(\rho)$ take the form:

$$\alpha(\rho) = \frac{2C_3\, D}{r\rho S\, l^2}\ \exp\!\left(\int\limits_a^\rho \frac{6D}{r\, l^2}\, d\rho\right);$$

$$\beta(\rho) = \frac{C_3}{r\rho S}\ \exp\!\left(\int\limits_a^\rho \frac{6D}{r\, l^2}\, d\rho\right);$$

$$\gamma(\rho) = C(\rho)\frac{a}{\rho}\ \exp\!\left(-\int\limits_a^\rho \frac{k_0 + 4\beta D}{w_0}\, d\rho\right).$$

Finally, inserting functions (3), (3), (3) into the equations (2) and (2), one can obtain the modified pressure:

$$P = -g\, z + p_0 - (k_0\,\alpha + w_0\,\partial_\rho\,\alpha)\, z^2(l^2 - z^2/2)/2 -$$
$$- z^2\left[w_0\beta - \int\limits_a^\rho \left(\frac{2v_0\gamma}{\rho} - k_0\,\beta\right) d\rho\right],$$

where

$$C(\rho) = C_4 + \int\limits_a^\rho \frac{k_0 v_0 \beta \rho}{w_0^2 a} \exp\left(\int\limits_a^\rho \frac{k_0 + 4\beta D}{w_0}\, d\rho\right)$$

and C_3, C_4 denote some constants.

4. Conclusion

Several important effects were revealed in our study of radial flows with a vortex through porous media. First, the unusual structure of the mass balance equation (2), having the form of the transverse diffusion law, was found. In this equation the transverse diffusion coefficient $D(\rho)$ appears to be proportional to the diameter $d(\rho)$ of the grain filling modeling a porous medium.

Second, the simplest Darcy's force with constant Darcy coefficient $k_D = k_0$ was used, the important dependence of the vortex velocity (3) on k_0 being established. This fact supports the necessity of generalizing the Darcy's law, in accordance with the effects mentioned in [22, 23].

Third, the important influence of the vortex velocity $v(\rho, z)$ on the purification efficiency becomes evident from the structure of the pressure (3) and (3). It is worth while to stress the connection of this effect with the fluidization process discussed in [24, 25].

Author Contributions: The authors' contributions are equal. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: No new data were created or analysed during this study. Data sharing is not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Bear, J. *Dynamics of Fluids in Porous Media* (Dover Publications, Mineola, 1988).
2. Cheremisinof, N. P. & Azbel, D. S. *Liquid Filtration* (Butterworth-Heinemann, Boston, 1998).
3. Pinder, G. F. & Gray, W. G. *Essentials of Multiphase Flow and Transport in Porous Media* (John Wiley & Sons, New York, 2008).
4. Dullien, F. A. L. *Porous Media: Fluid Transport and Pore Structure* (Academic Press, San Diego, 2012).
5. Kim, S. & Karila, S. J. *Microhydrodynamics: Principles and Selected Applications* (John Wiley & Sons, Boston, York, 1991).
6. Sahimi, M. *Flow and Transport in Porous Media and Fractional Rock* (John Wiley & Sons, New York, 2011).
7. Sheidegger, A. E. *The Physics of Flow through Porous Media* (MacMillan, New York, 1960).
8. Sheidegger, A. E. Statistical hydrodynamics in porous media. *Journal of Applied Physics* **25,** 997 (1964).
9. Vafai, K. *Handbook of Porous Media* (CRC Press, Taylor & Francis Group, Boca Raton, 2015).
10. Polubarinova-Kochina, P. Y. & Falcovich, S. V. Theory of fluid filtration in porous media. *Applied Mathematics and Mechanics* **11,** 629 (1947).
11. Frog, B. N. & Levchenko, A. P. *Water Purification* (Moscow State University Publishing, Moscow, 1996).
12. Fara, H. D. & Sheidegger, A. E. Statistical geometry of porous media. *Journal of Geophysical Research* **66,** 3279 (1961).

13. Harlemaii, D. R. F. & Rumer, R. R. Longitudinal and lateral dispersion in an isotropic porous medium. *Journal of Fluid Mechanics* **16,** 385 (1963).
14. Josselin de Jong, G. Longitudinal and transverse diffusion in granular deposits. *Transactions of American Geophysical Union* **39,** 67 (1958).
15. Saffman, P. G. A theory of dispersion in a porous media. *Journal of Fluid Mechanics* **6,** 321 (1959).
16. Sheidegger, A. E. General theory of dispersion in porous media. *Journal of Geophysical Research* **66,** 3273 (1961).
17. Darcy, H. *Les Fontaines Publiques de la Ville de Dijon* French (Dalmont, Paris, 1856).
18. Polubarinova-Kochina, P. Y. *Theory of Ground Water Movement* (Princeton University Press, Princeton, 1960).
19. Rybakov, Y. P., Semenova, N. V. & Safarov, J. S. Generalizing Darcy's law for filtration radial flows. *IOP Conference Series: Materials Science and Engineering* **675,** 012064 (2019).
20. Whitaker, S. The equations of motion in porous media. *Chemical Engineering Science* **21,** 291 (1966).
21. Whitaker, S. Flow in porous media I: A theoretical derivation of Darcy's law. *Transport in porous media* **1,** 3 (1986).
22. Olsen, H. W. *Deviations from Darcy's law in saturated clays* in *Proceedings of Soil Scientific Society of America* **29** (1965), 135.
23. Firdaouss, M. *et al.* Nonlinear corrections to Darcy's law at low Reynolds numbers. *Journal of Fluid Mechanics* **343,** 331 (1997).
24. Leva, M. *Fluidization* (McGraw-Hill, New York, Toronto, London, 1959).
25. Xu, C. C. & Zhu, J. Prediction of the minimal fluidization velocity for fine particles of various degrees of cohesiveness. *Chemistry Engineering Communications* **196,** 499 (2008).

Information about the authors

Yuri P. Rybakov (Russian Federation)—Professor, Doctor of Sciences in Physics and Mathematics, Professor at the Institute of Physical Research and Technologies of RUDN University (e-mail: rybakov-yup@rudn.ru, phone: +7(916)262-55-36, ORCID: 0000-0002-7744-9725, ResearcherID: S-4813-2018, Scopus Author ID: 16454766600)

Natalia V. Semenova (Russian Federation)—Junior member of teaching at the Institute of Physical Research and Technologies of RUDN University (e-mail: semenova-nv@rudn.ru, phone: +7(915)478-10-23, ORCID: 0000-0001-6894-6255, ResearcherID: AAC-8298-2020, Scopus Author ID: 57200754585)

УДК 538.91
DOI: 10.22363/2658-4670-2024-32-3-319–324 EDN: DLVHXG

Радиальные потоки жидкости с вихрем через пористые среды

Ю. П. Рыбаков, Н. В. Семёнова

Российский университет дружбы народов, ул. Миклухо-Маклая, д. 6, Москва, 117198, Российская Федерация

Аннотация. Изучается процесс фильтрации для популярного класса фильтров с радиальными картриджами, доказавших свою высокую эффективность при очистке воды. Уравнение баланса массы для радиальных потоков в пористых средах получено с использованием метода решёточного приближения с учётом процесса поперечной диффузии. Динамические уравнения Эйлера модифицированы путём включения силы Дарси, пропорциональной скорости фильтрационного потока. Система уравнений записана для стационарного осесимметричного радиального потока и решена методом возмущений, если вертикальная скорость предполагается малой.

Ключевые слова: фильтрация, пористая среда, сила Дарси

Discrete & Continuous Models & Applied Computational Science
ISSN 2658-7149 (Online), 2658-4670 (Print)

2024, 32 (3) 325–336
http://journals.rudn.ru/miph

Letters

Research article

UDC 93/94+811.58+75.04+003.324+004.9
PACS 01, 02.60.Ed, 68.37.-d, 93.85.Bc, 95.75.De
DOI: 10.22363/2658-4670-2024-32-3-325–336 EDN: FHPKYK

New method for correct identification of structural elements of ancient hieroglyphs

Maia A. Egorova, Alexander A. Egorov

RUDN University, 6 Miklukho-Maklaya St, Moscow, 117198, Russian Federation

(received: October 14, 2023; revised: December 20, 2023; accepted: January 20, 2024)

Abstract. A new method for the correct identification of various complex structural elements of ancient hieroglyphs is described. The method is based on photometry of the studied surface of the ancient artifacts. The obtained data are converted into digital form in order to determine the characteristics and parameters characterizing the properties of the investigated artifact surface. Digitized data is processed in various graphic applications, including those working with vector images. Several control experiments were also carried out. In particular, the corresponding statistical characteristics and parameters of the studied artifact surface profiles were determined. The data obtained made it possible to unambiguously detect the ancient hieroglyphs on the artifact surface and determine their number. Described method of studying ancient artifacts makes it possible to obtain sufficiently reliable results that will undoubtedly be useful and promising in the study of ancient hieroglyphic signs. Our research method is characterized as: non-contact, informative, and sensitive. This testifies to its importance and prospects in the study of similar ancient artifacts.

Key words and phrases: hieroglyphic inscription, structural elements, surface, photometry, statistical characteristics and parameters, Jiǎgǔwén, Jīnwén, Chinese radicals (bùshǒu), computer data processing

For citation: Egorova, M. A., Egorov, A. A. New method for correct identification of structural elements of ancient hieroglyphs. *Discrete and Continuous Models and Applied Computational Science* **32** (3), 325–336. doi: 10.22363/2658-4670-2024-32-3-325–336. edn: FHPKYK (2024).

1. Introduction

Writing on bronze vessels (Jinwen) and tortoise shells and fortunetelling bones Jiaguwen relate to the oldest examples of art and culture [1–10]. They are unique in their own way. Jiaguwen (甲骨文 / Jiǎgǔwén, XIV–XI c. BC) are hieroglyphic inscriptions fixing the results of fortune telling or predictions [3, 6–9].

Bronze vessels Jinwen (jīnwén / 金文, the earliest of which date back to the end of 2nd millennium BC) belong also to the most ancient samples of Chinese history, linguistics, archeology, and science [4, 7, 8]. These ancient objects are often poorly preserved, but these objects are of great historical, cultural and scientific value [11]. Basically, their research is devoted to the study of hieroglyphic inscriptions written in the surface in order to interpret their contents, as well as the identification of ancient keys (hieroglyphs) or Chinese radicals (bùshǒu / 部首).

Figure 1. An example of Jiaguwen (甲骨文, "Turtle shell inscriptions")—hieroglyphic inscriptions fixing the results of fortune telling or predictions. Vertical lines in the center: "1" corresponds to the central part of the investigated surface of the shell, where there are no hieroglyphs; "2" corresponds to the investigated part of the surface of the shell, where there is a column of ancient Chinese hieroglyphs

The present paper is devoted to the development of a new method for solving some problem for recognition of ancient artifacts [7, 12]. Using digital photometric data of the artifact relief, this method allows determining various characteristics and parameters, including statistical ones, characterizing the properties of the studied ancient surface, e.g. Jiǎgǔwén, jīnwén [7, 11–20]. The main task of the work is to reveal the possibility of correct identification of groove-type elements that are an integral part of various structural surface elements. In technical applications, these elements can be various single grooves, which are the constituent elements of such widely used structures as diffraction gratings and grating's type structures. In the study of the oldest artifacts, for example, Jiaguwen or Jinwen, there is a need to study ancient hieroglyphs written on the surface of a tortoise shell, copper vessel or stone [3, 4, 6–9].

2. Materials and methods

This article uses a new statistical method of research, which can be a good addition to the currently used traditional methods (mainly visual) of research [8, 9, 11–18]. This method allows using various digital photometric data on the surface of the shell (see Figure 1) to determine various characteristics and parameters characterizing different properties (including statistical) of the investigated surface [8, 12].

The main advantages of our research method are: contactless, informativeness, non-destructivity, and potentially high enough resolution (for example, according to the statistical parameters of the surface profile). This allows us to talk about its promise in the study of such ancient art and culture as Jiaguwen. The main goal at this stage of the study is to demonstrate the possibility of a certain identification of ancient hieroglyphs written on the surface of the tortoise shell using the obtained digital surface profiles of the test sample.

It is important to emphasize that the results obtained allowed us to solve the formulated problem of identifying the studied surface images, namely, to determine approximately the number of ancient Chinese radicals placed on the scan line of the surface profile with hieroglyphs. It is shown that the discrepancy with predetermined number of hieroglyphs does not exceed 10%. At this stage of the study, the results obtained can be considered good.

Our method is based on photometry of artifact surface profiles, in which a sample of brightness levels in a discrete set of points is carried out with subsequent conversion to digital form [7, 11, 12]. Photometric data of the surface profile can be obtained using for example "Fiji ImageJ" program for Windows [7, 11–16]. The obtained data are digitally processed in order to determine the characteristics and parameters characterizing the properties of the researched artifact (see Appendix) [7, 11, 12].

We utilize different research methods including a statistical approach, which allows us to find the characteristics and profile parameters of the test sample according to photometry of its surface [7]. This research method will finally solve an important inverse problem of pattern recognition: find the number of characters in the studied sections of the surface profile, as well as identify a solitary micro-object of the groove type. Properties of these structures will be useful in further research of different type ancient hieroglyphs, for example Chinese hieroglyphs (see Appendix).

The use of digital methods for processing large amounts of experimental data, for example, using such well-known signal processing methods as the fast Fourier transforms or wavelet transform, undoubtedly allows us to obtain new interesting results [7, 12]. However, we should note that these studies are quite expensive and time-consuming.

Figure 1 shows a processed photo of a shell, which is one of the objects under study. The images used in this paper are taken from open sources and are used as illustrations [4, 6, 7, 9, 11].

The subject of the study is the ancient shell with or without hieroglyphs written on it surface by the antique scribe. Major attention is paid to the improvement of innovative methods for the correct identification of the same type of hieroglyphic signs in the ancient hieroglyphic inscriptions according to the typical digitized data. The solution of this key problem makes it possible to avoid mistakes at the initial stage of the study. As a result, at all subsequent stages of research it becomes possible to examine the accurate data. Finally, when the correct recognition of hieroglyphic signs is carried out, one can obtain the quantity of ancient hieroglyphs.

3. Results

Our main goal is to find the quantity of ancient Chinese hieroglyphs (radicals / bùshǒu / 部首) $\mathfrak{I}$. For this purpose we propose several procedures (see Appendix) [7, 12, 14, 15]: 1) as an initial estimate – an integral method that permits to define the number of grooves (structural elements of hieroglyphic inscriptions) on the scanning line; 2) division the scan interval length L of the artifact profile to the correlation radii r, characterizing the statistics of these profiles; 3) numerical Fourier transform of the photometric data of the surface profile.

Indeed, the correlation radius determines the characteristic size of the "particles" of scatterers on the surface of the object under study [7, 12]. These "particles" should not be taken literally. This term is used to indicate the area within which correlation takes place, i.e. areas of "distortion" of the surface. In our case, this distortion is created by the ancient hieroglyphs; therefore, our estimate $\mathfrak{I}$ is quite fair. The approximate average length of the scan line of the surface profile of the shell: $L = 18$ cm. In this case, both shorter and longer sections were used in the calculations.

Let us make a remark about which correlation radius should be taken in the calculations. Formally, one need to take the correlation radius defined for the surface profiles of the shell with hieroglyphs. However, it is necessary to take into account the certain contribution of a surface free of hieroglyphs, since its role in the formation of the final profile is undoubtedly also present.

Note that to solve this problem, it is not necessary to determine the rms height with high resolution, it is sufficient that the signal-to-noise ratio level allows for the reliable identification of the presence of the relevant objects (hieroglyphic signs) on the surface of the test sample.

As an initial step, the following approach can be proposed. Where a sufficiently dense arrangement of ancient Chinese radicals is visually observed, the contribution of a surface free of hieroglyphs will be taken into account to a lesser extent. Given the statistical approach in our analysis, for simplicity, we will find the average of two estimates of the correlation radii: for surface profiles without hieroglyphs and with hieroglyphs. The data obtained above for a comparative analysis of the graphs in Figure 2 and Figure 3 confirm the validity of this approach (see Appendix).

Therefore, we will determine the number of ancient Chinese hieroglyphs by the following simple formula [12]:

$$\mathfrak{I} = L/\bar{r}. \tag{1}$$

Let's find the average value of the correlation radius $\bar{r}$, taking into account the above: $\bar{r} = (r_1 + r_2)/2$, where r_1 is the correlation radius of the shell surface, where there are no hieroglyphs, and r_2 is the correlation radius of the shell surface, where there is a column of the ancient hieroglyphs (see Figure 1). We take into account a Gaussian approximating correlation function. We can find currently $\bar{r}$: $\bar{r} = [(25.7 + 2.9)/2]\,\text{mm} = 14\,\text{mm}$.

Now we can finally find the approximate number of Chinese radicals in accordance with the given above formula (1):

$$\mathfrak{I} = 180/14 = 13. \tag{2}$$

So, we found from equation (2) that the number of ancient hieroglyphs (or ancient Chinese radicals (部首 / bùshǒu)) is equal to 13. According to our preliminary data, this vertical scan line (marked with the number "2" in Figure 1) contains 12 hieroglyphs.

Thus, the value close to the exact quantity of Chinese radicals was determined: error is 8.3%. If the length of the scan line changes, the number of ancient hieroglyphs changes a little too. Indeed, if in calculations according to formula we take not the average value L, and the dimensions of the used scanning lines of the shell vertically in different sections (with a horizontal shift), and take others r, we obtain that $\mathfrak{I}$ may vary from about 11 to 15. Consequently, the greatest error in finding the number of radicals $\mathfrak{I}$ will be approximately 25%. At this stage of research, such an error can be considered quite acceptable.

4. Conclusion

It is important to emphasize that the results of the work made it possible to finally solve the recognition problem posed, namely, to determine the number of ancient hieroglyphs on the artifact under study. It is shown that the discrepancy with the specified number of hieroglyphs located on the studied scan lines is less than 10%. At this stage of the study, the results obtained can be considered good.

The advantages of the research method implemented in this work are its non-contact, non-destructive nature, informative value and high resolution in terms of statistical parameters of the surface profile. This indicates its promise in studying such ancient art and culture samples as Jiaguwen or Jingwen—ancient hieroglyphic inscriptions that record the results of fortune-telling or predictions. Moreover, a certain simplicity and clarity of implementation allows it to be used in interdisciplinary research involving specialists from different subject areas of knowledge. The results obtained in the article will undoubtedly be useful in areas such as graphic document processing and applied linguistics, especially in the statistical analysis of ancient hieroglyphic inscriptions. Considering that

the most ancient studied samples often have poor preservation, as well as of high historical, cultural and scientific value, our method can be a good addition to the traditional (mainly visual) research methods currently used.

The methods developed in this paper can be used in the future, for example in computer intelligent systems for recognition of text containing large arrays of various hieroglyphic signs. A particularly promising advantage, in our opinion, is the ability to identify the statistical properties of large arrays of studied data. Knowledge of statistics allows, for example, to use the intelligent computer classification of ancient objects under study in a confident way. The results obtained can be useful in linguistics, lexicostatistics and sociolinguistics, especially in the statistical analysis of such old hieroglyphic inscriptions, as well as the information contained in them. Of particular interest may be the study of the relationship between the information contained, for example, in the Jiaguwen or Jinwen and the socio-cultural conditions for the development of the language and society in the corresponding eras.

Appendix

A new research method used in this work allows determining different characteristics and parameters describing the properties of the studied artifact surface [7, 12]. As a result, the form of approximating autocorrelation functions (ACF) can be found, both without hieroglyphs and with hieroglyphs. And the corresponding statistical parameters are accordingly determined: the standard deviation and the radius of correlation of the surface profile irregularities.

The most important statistical characteristic of an uneven surface is *autocorrelation functions*, which characterizes the relationship between the analyzed function and its shifted copy of the shift value of the argument (the process under consideration is stationary).

The most important statistical parameters are the root-mean-square (rms or standard deviation) σ and correlation radius r of surface irregularities. *Root mean square* characterizes the amount of dispersion of the values of a random variable relative to its mathematical expectation, i.e. average value. A larger value of the standard deviation indicates a larger scatter of values in the presented set of values, and a smaller value of it, respectively, shows that the values are grouped around the average value. *Correlation radius* determines the characteristic size of the "particles" of scatterers on the surface of the investigated object. These "particles" should not be taken literally. This term is used to indicate the area within which correlation takes place, i.e. areas of "distortion" of the surface [7, 12].

It is possible to determine the statistical characteristics and parameters describing the properties of investigated surface without involving complex scattering theories, requiring more laborious and subtle experiments and complex calculations than the use of more simple scalar equation (3) [11–15]. However, if necessary, more complex, for example, vector scattering theories and appropriate calculation algorithms can be used [7, 11–15].

Incident and reflected radiation intensities I_i and I_r are related by the next known expression:

$$I_r = I_i \exp\left\{-\left[(\pi h/\lambda)\cos\theta\right]^2\right\}, \tag{3}$$

where h is the height of the surface profile irregularities (e.g. ordinary roughness and hieroglyphs), $h \approx 4\sigma$; $\pi \approx 3.14$; *theta* is the angle of incidence of light on the artifact surface, λ is the radiation wavelength (in the case of non-monochromatic light, one can take, for example, the average value); light intensities incident on the surface and reflected (scattered) from it, satisfy the condition: $I_r/I_i \ll 1$.

From expression (3) we can find the height h of the studied artifact surface:

$$h = \left(\frac{\lambda}{2\pi \cos\theta}\right) \ln(I_i/I_r). \tag{4}$$

Consequently the height of the artifact (surface) profile irregularities is directly proportional to $\ln(I_i/I_r)$, obtained from the photometric data of a certain section of the profile (at the fixed values of θ and λ).

It follows from the expression (4) that to determine the profile heights of the surface under study, it is sufficient to know the values of λ, θ and I_i/I_r.

Since in our case the quantity I_i is also fixed, we assume that h depends only on the intensity I_r of the light reflected (scattered) from the artifact surface, i.e. in accordance with expression (4) the spatial distribution of h is a reflection of dependence I_r (or simply I_i) on the average surface profile: $h \propto h(I) \propto h(y)$. Therefore, by measuring the spatial distribution I, for example, along a certain direction, we can get an average surface profile. In the case of using a statistical research methods, this approach is justified and, moreover, necessary [7, 12].

Now, some approximating autocorrelation function can be found, that characterizes the relationship between the analyzed function and its shifted copy of the shift value of the argument $y = y_1 - y_2$:

$$R(y_1, y_2) = \langle h(y_1)h(y_2)\rangle, \tag{5}$$

where brackets $\langle ...\rangle$ mean (statistical) average.

An approximate estimate of the experimental (empirical) autocorrelation function (5) on a discrete set of points N (sample implementation of an encoded profile $h(y)$, see Figure 1) can be calculated using the following formula:

$$R(mu) = N^{-1} \sum_{i=1}^{N} H(iu)H(iu + mu), \tag{6}$$

where mu is the shift interval, $m = 0, 1, 2, \ldots, M - 1$, where M is number of measured ordinates (counts) of the function $h(y)$.

This procedure is carried out by recording an array of digitized values $h(y)$, followed by their digital processing in accordance with the expression (6). The resulting numerical data allows us to construct a graph of the approximating ACF and determine the corresponding statistical parameters of the irregularities of the artifact surface: standard deviation and correlation radius.

It should be noted that the most accurate (consistent and unbiased) estimates of the statistical characteristics of a random profile can be obtained as an average over the entire ensemble of realizations. It is clear that this will require a significant increase in processing time for all received data.

Now we can give the formulas of two well-known autocorrelation functions that are most frequently used in statistical analysis: exponential and Gaussian. They are often used by researchers to approximate experimental ACFs.

Exponential and Gaussian ACF are described by the following formulas accordingly [7]:

$$R(y) = \sigma^2 \exp\lfloor -|y/r|\rfloor, \quad R(y) = \sigma^2 \exp\lfloor -(y/r)^2\rfloor. \tag{7}$$

We should note that not only ACFs (7) were used in the research, but also other approximating functions, for example: decaying cosine, logistic (see Figure 2; 1 arb.u. = 100 m), polynomial, logarithmic, etc. In the calculations, we took sections of different lengths, from about 3 to 10 cm. Several results were obtained by averaging over these profiles realizations.

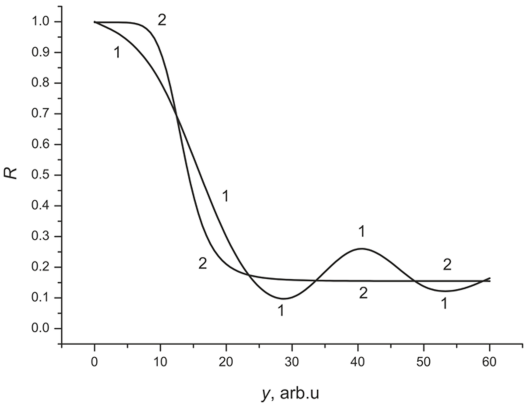

Figure 2. Approximate experimental (1) and logistic (2) normalized functions $R(y)$

As an example Figure 2 shows the logistic fitting function (indicated by the number "2"). To test the method, we used onyx, jasper, and jade specimens, the surface of which is similar to the surface of ancient specimens [9, 12]. These results confirmed the possibilities and prospects of the described method.

It should be noted that a solitary object of the groove type that is a part of the complex element similar to the old hieroglyph can be easy determined if its depth and width are not less than 0.1 mm. We should outline that the groove width must be greater than wavelength of radiation incident on the object under study. A more detailed analysis of this problem, especially the issues of the incorrect inverse problem of restoring the geometric profile of a groove, is beyond the scope of this article. We only note that this problem can be solved using fairly complex algorithms that implement a numerical reconstruction of such geometric profiles, in particular with super-resolution, i.e. exceeding the Rayleigh limit.

The mathematical model of the groove in the broad optical range of wavelengths is described by the next expression:

$$f(y) = \exp\left[-i2kh(y)\right]. \tag{8}$$

The notation in the formula (8) is as follows: $k = 2\pi/\lambda$ is wave number, $\pi = 3.14$; $h(y) = -h$, at $|y| \leq \Delta y/2$ and $h(y) = 0$ at $|y| > \Delta y/2$; h is a groove depth ($h \leq 2$ mm); Δy is a groove width ($\Delta y \leq 1$ mm). In one of the subsequent works, we plan to show the possibility of reconstructing a geometric profile $h(y)$ from its experimentally obtained optical relief $f(y)$ (solution of an ill-posed problem).

Figure 3 shows $I(y)$ (i.e. a profile $h(y) \propto I(y)$) for a line containing ancient Chinese hieroglyphs. As can be seen in the figure, the number of grooves is 39 considering some level of intensity $I \geq 0.15$. If we take into account smaller details, we get a little more: the number of grooves is about 45. In this case, only one groove with number "18" is uniquely identified with the element in the center of the scan line on the surface under study (see Figure 1). Obviously, it is quite difficult to identify all ancient Chinese hieroglyphs based on the data shown in Figure 3 (1 arb.u. = 100 m). One can make only some assumptions about the number of ancient hieroglyphs and/or Chinese radicals (bùshǒu).

To test the method, samples of onyx, jasper and jade were used, which surface is similar to the surface of ancient artifacts. Eventually, several control experiments were conducted, during which the following were investigated: 1) systems of structural elements, similar to the system of elements in hieroglyphic writing on the surface of antique samples, each of which contains elements similar to

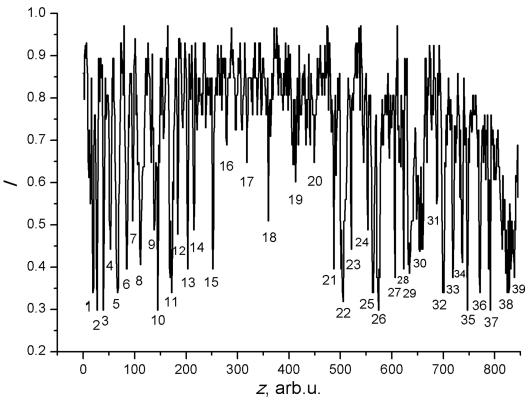

Figure 3. Experimental normalized intensity $I(z)$ along a vertical line close to the line "2" on the Figure 1, where there is a column of ancient Chinese hieroglyphs

Figure 4. Two Chinese hieroglyphs "yué" carved into the onyx surface

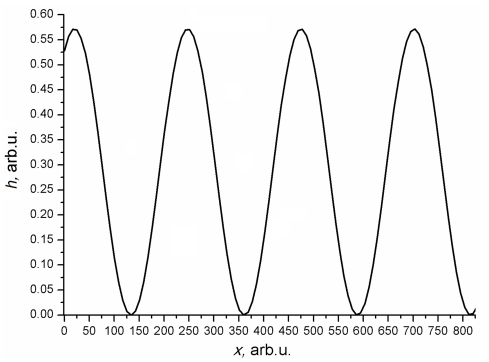

Figure 5. Experimental smoothed approximate profile $h(x)$ along a horizontal line close to the centers of two ancient Chinese hieroglyphs "yué" from the Fig. 4.

a groove; 2) hieroglyphs "yué" – month; meat (depicted on the surface of the sample by the co-authors of this article).

As an example, data are given on the study of two hieroglyphs "yué", carved with a fine chisel on the surface of onyx (written on the surface by co-authors of this paper). Figure 4 shows a processed photograph of the surface of the onyx sample, which is one of the objects under study (photo of co-authors of this paper).

Figure 5 shows a smoothed approximated curve for the horizontal profile of two ancient Chinese hieroglyphs "yué" (month; meat) depicted on the Figure 4 (photometry was carried out approximately in the middle of hieroglyphs; 1 arb.u. is about 10 m). Digital profiles were obtained using the "Fiji ImageJ" program for Windows. The length of the horizontal scan profile was approximately 1–5 cm (or 200–600 pixels).

We can note that the curve from the Figure 5 can be approximated by the function like $\sin(x)$:

$$I(x) = I_{\max}\left[\sin(2\pi x_w + \varphi_0)\right], \tag{9}$$

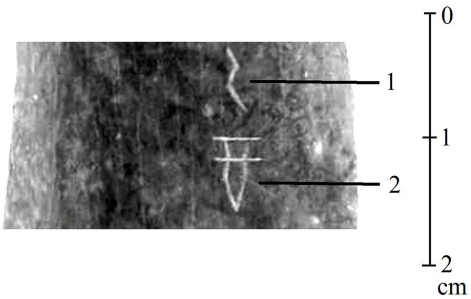

Figure 6. Jiaguwen with two numbered hieroglyphs (Chinese radicals (keys)): 1 is 乙 (2nd cyclic sign); 2 is 酒 (酉) (wine (key "wine", "vessel for wine"))

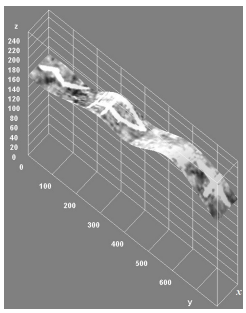

Figure 7. Digital 3D image of Jiaguwen (see Figure 6)

where $I_{\max}$ is a peak amplitude, $\pi \approx 3.14$, x_w is a dimensionless parameter, φ_0 is an initial phase.

We use different types of curves as fitting functions and not only the type given in equation (9). The research showed that the final profile under study is well characterized by adjustable functions of the sinusoidal (periodic) types, i.e. the functions that describe harmonic vibrations. This approach demonstrates additional possibilities for processing the received digital data. We plan to give a more detailed description in subsequent works. It can only be noted that such an approach can be promising in solving the problems of ancient hieroglyphs identification and certain automation of the research process.

The results obtained allow, in particular, to express the idea of using the series expansion of the obtained profiles in some basis functions, for example, like $\sin(x)$ and $\cos(x)$. So, we can use functions of periodic type not only as approximating functions in similar study, but also as the basis functions, i.e. as some orthogonal basis.

A comparison of Figure 1, Figure 3 and Figure 5 shows that profiles similar to those shown in the Figure 5, can be found in the Figure 3, for example, at the beginning, middle and the end of the profile. It is in those places that there are hieroglyphs containing elements similar to the hieroglyph "yué" (see Figure 4).

In the Figure 6 a part of the bone (scapula) presumably of a large horned animal is presented, on which written signs are depicted, recording the interpretation of fortune-telling or predictions [6].

The Figure 7 shows a digital 3D image of this Jiaguwen from Figure 6, obtained with the computer program "Fiji ImageJ" for Windows. The hieroglyphs are shown in Figure 7 as in Figure 6, from top to bottom and from left to right (along the vertical, i.e. y-axis in Figure 6, but slightly at an angle to the horizontal x-axis). Note that a section along the horizontal axis gives 2D profiles similar to those shown in the Figure 5 (after appropriate smoothing).

We emphasize that the computer program "Fiji ImageJ" for Windows allows you to export data in raster and vector formats. Then the digitized data can be processed in various graphic applications (for example, CorelDraw, etc.) that work, among others, with vector images. In conclusion, it is important to note that at a low noise level, it is possible to reliably identify a groove with a depth and width of at least 0.1 mm. It is unlikely that scribes used in antiquity (about 3–4 thousand years ago) finer tools for drawing hieroglyphs. Therefore, we can conclude that the described method of studying ancient artifacts makes it possible to obtain sufficiently reliable results that will undoubtedly be useful and promising in the study of ancient hieroglyphic signs. It should be emphasized that, in general, the error of the described research method at this stage does not exceed 20%.

For this study the following sources of illumination were used: natural (sunlight) light, household light sources ($\lambda = 0.5\ \mu$m), and LEDs (with illumination of 150–300 lux). Digital surface profiles were obtained using the "Fiji ImageJ" program for Windows [12, 16]. Computer processing of digital photometric data makes it possible to judge not only the different structural complexity of the test object and hieroglyphs under study, but also to determine their number and location [7, 12]. The processing of digital photometric data and the necessary calculations were carried out on the computer "Intel Pentium 4".

The methods proposed in our paper cannot completely replace the traditional research methods in linguistics, history and archeology; however, they undoubtedly allow us to look at current scientific problems from new, non-standard positions and obtain a number of innovative results. It should be noted that other approaches to the study of different kind of hieroglyphs have recently been developed (see e.g. [17, 18]). At the same time, it must be emphasized that the process of combining ancient Chinese hieroglyphic characters with modern digital systems is the significant milestone towards the popular and widespread use of this character system [19]. At the same time, it is obviously necessary to see the entire research process in broader historical, philosophical and cultural aspects [3–8, 19, 20]. In our opinion methods described in this article can be used for studying different symbol inscriptions like oldest hieroglyphs and cuneiform in various antique artifacts (bronze vessels, tortoise shells, stones, and clay tablets), for example: Chinese, Hittite, Akkadian, Sumerian, and Egyptian.

The possibilities of our innovative method can be expanded by using: a) ultraviolet and infrared light sources, including coherent radiation sources; b) various computer processing methods, both digitized images and the resulting digital surface profiles of the studied antique samples. The main advantages of the research method used are its non-contact, informative nature, and potentially sufficiently high resolution, which makes it possible to speak about its prospects in the study of such ancient artifacts.

Author Contributions: Conceptualization, A.E. and M.E.; methodology, A.E. and M.E.; software, A.E.; validation, M.E. and A.E.; formal analysis, M.E. and A.E.; investigation, A.E. and M.E.; resources, M.E. and A.E.; data curation, M.E. and A.E.; writing—original draft preparation, A.E.; writing—review and editing, A.E. and M.E.; visualization, M.E. and A.E.. All authors have read and agreed to the published version of the manuscript.

Funding: The publication was prepared with the partial support of the RUDN "Strategic Academic Leadership Program".

Data Availability Statement: Data sharing is not applicable.

Acknowledgments: The authors would like to thank their colleagues for their kind feedback, which helped in preparing the paper for publication.

Conflicts of Interest: There are no conflicts to declare.

References

1. Comrie, B., Matthews, S. & Polinsky, M. *The Atlas of Languages: The origin and development of languages throughout the world* (Rev. ed. NY: Facts on File, 2003).
2. Blench, R., Sagart, L. & Sanchez-Mazas, A. *The Peopling of East Asia: putting together archaeology, linguistics and genetics* (Routledge Curzon, London, 2005).
3. Keightley, D. N. *Sources of Shang history: the oracle-bone inscriptions of Bronze Age China* (Berkeley, London, 1985).
4. Shaughnessy, E. L. *Sources of Western Zhou history: Inscribed bronze vessels* (University of California Press, Los Angeles, 1991).
5. Kryukov, M. V. & Shu-In, K. *Ancient Chinese* (Vostochnaya kniga, Moscow, 2020).

6. Cheng, Z. 簡明的海龜和動物骨骼偵銘文語言詞典 *(A brief dictionary of the language of inscriptions on the scutes of turtles and animal bones. A systematized reader of fortune-telling inscriptions)* (Beijing, 1988).
7. Egorova, M. & Egorov, A. *The Role of Ancient Written Signs in the Preservation and Development of the Chinese Language* in *Proceedings of the 2020 International Conference on Language, Communication and Culture Studies (ICLCCS 2020)* **537** (Atlantis Press, 2021), 41–47. doi:10.2991/assehr.k.210313.008.
8. Egorova, M. A., Egorov, A. A. & Solovieva, T. M. Features of Archaic Writing of Ancient Chinese in Comparison with Modern: Historical Context. *Voprosy istorii,* 189–207. doi:10.31166/VoprosyIstorii202111Statyi17 (2021).
9. Egorova, M. A., Egorov, A. A., Orlova, T. G. & Trifonova, E. D. Methods of research of hieroglyphs on the oldest artifacts — introduction to problem: history, archeology, linguistics. *Voprosy istorii,* 17–25. doi:10.31166/VoprosyIstorii202203Statyi10 (2022).
10. Egorova, M. A., Egorov, A. A. & Solovieva, T. M. Modeling the distribution and modification of writing in proto-Chinese language communities. *Automatic Documentation and Mathematical Linguistics* **54,** 92–104. doi:10.3103/S0005105520020065 (2020).
11. Yin, X. *Convention concerning the protection of the World Cultural and Natural Heritage nomination of Cultural Property for Inscription on the World Heritage List* (China, 2006).
12. *High-resolution digital images of oracle bones, Cambridge Digital Library* 2024.
13. *Request for comment on encoding Oracle Bone Script, L2/15-280. Working Group Document, ISO/IEC JTC1/SC2/WG2 and UTC. 2015-10-21. Retrieved 2016-01-23* 2024.
14. Jahne, B. *Digital image processing* (Springer, NY, 2005).
15. Egorova, M. A., Egorov, A. A. & Solovieva, T. M. Identification of hieroglyphs in the ancient inscriptions according to typical digitized data: application in history, archeology, linguistics. *Voprosy istorii,* 124–131. doi:10.31166/VoprosyIstorii202301Statyi42 (2023).
16. Haslam, M., Robertson, G., Crowther, A., Nugent, S. & Kirkwood, L. *Archaeological science under a microscope* (ANU Press, 2009).
17. Born, M. & Wolf, E. *Principles of optics* (Pergamon Press, NY, 1986).
18. Beckmann, P. & Spizzichino, A. *The scattering of electromagnetic waves from rough surfaces* (Pergamon Press, NY, 1963).
19. Broeke, J., Perez, J. M. M. & Pascau, J. *Image processing with ImageJ, 2nd Ed.* (Packt Publishing, NY, 2015).
20. Santanam, K., Vaithiyanathan, R. & Tripati, S. *Digital image processing* (Harman Publishing House, London, 2004).

Information about the authors

Maia A. Egorova—Candidate of Political Sciences, Associate Professor at the Department of Foreign Languages of the Faculty of Humanities and Social Sciences of RUDN University (e-mail: Mey1@list.ru, ORCID: 0000-0003-2931-8330)

Alexander A. Egorov—Doctor of Physical and Mathematical Sciences, Consulting Professor of RUDN University (e-mail: alexandr_egorov@mail.ru, ORCID: 0000-0002-1999-3810)

УДК 93/94+811.58+75.04+003.324+004.9
PACS 01, 02.60.Ed, 68.37.-d, 93.85.Bc, 95.75.De
DOI: 10.22363/2658-4670-2024-32-3-325–336 EDN: FHPKYK

Новый метод корректной идентификации структурных элементов древних иероглифов

М. А. Егорова, А. А. Егоров

Российский университет дружбы народов, ул. Миклухо-Маклая, д. 6, Москва, 117198, Российская Федерация

Аннотация. В статье описан новый метод корректной идентификации различных структурных элементов древних иероглифов. Метод основан на фотометрии исследуемой поверхности древнего артефакта. Полученные данные преобразуются в цифровую форму с целью определения характеристик и параметров, характеризующих свойства исследуемой поверхности артефакта. Оцифрованные данные обрабатываются в различных графических приложениях, в том числе работающих с векторными изображениями. Проведено также несколько контрольных экспериментов. Полученные данные позволили однозначно обнаружить на поверхности артефакта древние иероглифы и определить их количество. Описанный метод изучения древних артефактов позволяет получить достаточно достоверные результаты, которые, несомненно, будут полезны и перспективны при изучении древних иероглифических знаков. Наш метод исследования характеризуется как: бесконтактный, информативный, чувствительный. Это свидетельствует о его важности и перспективности в исследовании подобных древних артефактов.

Ключевые слова: иероглифическая надпись, структурные элементы, поверхность, фотометрия, Цзягувэнь, Цзиньвэнь, китайские ключи, компьютерная обработка данных