

ISSN 1684-8853 (print); ISSN 2541-8610 (online)

ИНФОРМАЦИОННО- УПРАВЛЯЮЩИЕ СИСТЕМЫ

НАУЧНЫЙ ЖУРНАЛ

2(99)/2019

2(99)/2019

PEER REVIEWED JOURNAL

INFORMATSIONNO- UPRAVLIAIUSHCHIE SISTEMY (INFORMATION AND CONTROL SYSTEMS)

Founder

«Information and Control Systems», Ltd.

PublisherSaint-Petersburg State University
of Aerospace Instrumentation**Editor-in-Chief**

M. Sergeev

Dr. Sc., Professor, Saint-Petersburg, Russia

Deputy Editor-in-Chief

E. Krouk

Dr. Sc., Professor, Moscow, Russia

Executive secretary

O. Muravtsova

Editorial Board

S. Andreev

PhD, Tampere, Finland

V. Anisimov

Dr. Sc., Professor, Saint-Petersburg, Russia

B. Bezruchko

Dr. Sc., Professor, Saratov, Russia

N. Blaustein

Dr. Sc., Professor, Beer-Sheva, Israel

C. Christodoulou

PhD, Professor, Albuquerque, New Mexico, USA

A. Dudin

Dr. Sc., Professor, Minsk, Belarus

I. Dumer

PhD., Professor, Riverside, USA

M. Favorskaya

Dr. Sc., Professor, Krasnoyarsk, Russia

L. Fortuna

PhD, Professor, Catania, Italy

A. Fradkov

Dr. Sc., Professor, Saint-Petersburg, Russia

A. Hramov

Dr. Sc., Professor, Innopolis, Russia

L. Jain

PhD, Professor, Canberra, Australia

V. Khimenko

Dr. Sc., Professor, Saint-Petersburg, Russia

G. Maltsev

Dr. Sc., Professor, Saint-Petersburg, Russia

G. Matvienko

Dr. Sc., Professor, Tomsk, Russia

A. Myllari

PhD, Professor, Grenada, West Indies

Y. Podoplyokin

Dr. Sc., Professor, Saint-Petersburg, Russia

J. Seberry

PhD, Professor, Wollongong, Australia

A. Shalyto

Dr. Sc., Professor, Saint-Petersburg, Russia

A. Shepeta

Dr. Sc., Professor, Saint-Petersburg, Russia

Yu. Shokin

RAS Academician, Dr. Sc., Novosibirsk, Russia

A. Smirnov

Dr. Sc., Professor, Saint-Petersburg, Russia

T. Sutikno

PhD, Associate Professor, Yogyakarta, Indonesia

Z. Yuldashev

Dr. Sc., Professor, Saint-Petersburg, Russia

R. Yusupov

RAS Corr. Member, Dr. Sc., Professor, Saint-Petersburg, Russia

A. Zeifman

Dr. Sc., Professor, Volgda, Russia

Editor: A. Larionova**Proofreader:** T. Zvertanovskaia**Design:** M. Chernenko, Y. Umnitsina**Layout and composition:** Y. Umnitsina**Contact information**

The Editorial and Publishing Center, SUAI

67, B. Morskaia, 190000, St. Petersburg, Russia

Website: <http://i-us.ru/en>, e-mail: i.us.spb@gmail.com

Tel.: +7 - 812 494 70 02

THEORETICAL AND APPLIED MATHEMATICS**Vassiliev N. N., Parasidis I. N., Providas E.** Exact solution method for Fredholm integro-differential equations with multipoint and integral boundary conditions. Part 2. Decomposition-extension method for squared operators 2**Khimenko V. I.** Probabilistic analysis of two-dimensional vector random processes on the phase plane 10**INFORMATION PROCESSING AND CONTROL****Konovalev P. A., Kornilova A. V., Korolev D. M., Reneva G. V., Fradkov A. L., Shirokolobov I. Yu., Yarosh D. S.** A team of soccer robots for RoboCup competitions in SSL league: system and algorithms 19**Ivanko D. V., Ryumin D. A., Karpov A. A., Zelezny M.** Measuring the effect of high-speed video data on the audio-visual speech recognition accuracy 26**Rukavitsyn A. N.** Data clustering in distributed monitoring systems 35**SYSTEM AND PROCESS MODELING****Klucharev A. A., Fomenkova A. A.** Mathematical model of a fixed-biomass anaerobic bioreactor as a control object 44**HARDWARE AND SOFTWARE RESOURCES****Gordeyev A. V., Andreyeva A. V.** OpenLDAP directory service structure as a search filter 52**INFORMATION SECURITY****Novikova E. S., Kotenko I. V.** Open challenges in visual analytics for security information and event management 57**INFORMATION CODING AND TRANSMISSION****Sovetov B. Ya., Tatarnikova T. M., Poymanova E. D.** Organization of multi-level data storage 68**Belim S. V., Rakitskiy Yu. S.** Using families of orthogonal functions for coding messages in a chaotic masking scheme 76**INFORMATION INSTRUMENTATION AND EDUCATION****Komarov V. A., Sarafanov A. V., Tumkovskiy S. R.** Multi-user distributed information-control systems as an element of modern digital educational environment 83**CHRONICLES AND INFORMATION**

Computer Assisted Mathematics – CAM-2019 95

INFORMATION ABOUT THE AUTHORS

96

Submitted for publication 27.02.19. Passed for printing 17.04.19. Format 60x84_{1/8}.
Phototype SchoolBookC. Digital printing.Layout original is made at the Editorial and Publishing Center, SUAI.
67, B. Morskaia, 190000, St. Petersburg, Russia
Printed from slides at the Editorial and Publishing Center, SUAI.
67, B. Morskaia, 190000, St. Petersburg, RussiaThe journal is distributed by subscription. Subscription can be made in the Editorial and publishing center, SUAI as well as in any post office based on «Pressa Rossiya» catalogue.
№ 15385 — semiannual subscript.The Journal was registered in the Ministry of Press, Broadcasting and Mass Media of the Russian Federation. Registration Certificate JD № 77-12412 from April, 19, 2002. Re-registration in the Federal Service for Supervision in the Sphere of Telecom, Information Technologies and Mass Communications (ROSKOMNADZOR) due to change of the founder: «Information and Control Systems», Ltd., JD № FS77-49181 from March, 30, 2012.
© Corporate authors, 2019

2(99)/2019

РЕЦЕНЗИРУЕМОЕ ИЗДАНИЕ

ИНФОРМАЦИОННО-
УПРАВЛЯЮЩИЕ
СИСТЕМЫ

Учредитель
ООО «Информационно-управляющие системы»

Издатель
Санкт-Петербургский государственный университет
аэрокосмического приборостроения

Главный редактор
М. Б. Сергеев,
д-р техн. наук, проф., Санкт-Петербург, РФ

Зам. главного редактора
Е. А. Крук,
д-р техн. наук, проф., Москва, РФ

Ответственный секретарь
О. В. Муравцова

Редакционная коллегия:
С. Д. Андреев,
канд. техн. наук, Тампере, Финляндия
В. Г. Анисимов,
д-р техн. наук, проф., Санкт-Петербург, РФ
Б. П. Безручко,
д-р физ.-мат. наук, проф., Саратов, РФ
Н. Блаунштейн,
д-р физ.-мат. наук, проф., Беэр-Шева, Израиль
Л. С. Джайн,
д-р наук, проф., Канберра, Австралия
А. Н. Дудин,
д-р физ.-мат. наук, проф., Минск, Беларусь
И. И. Думер,
д-р наук, проф., Риверсайд, США
А. И. Зейфман,
д-р физ.-мат. наук, проф., Вологда, РФ
К. Кристоделу,
д-р наук, проф., Альбукерке, Нью-Мексико, США
Г. Н. Мальцев,
д-р техн. наук, проф., Санкт-Петербург, РФ
Г. Г. Матвиенко,
д-р физ.-мат. наук, проф., Томск, РФ
А. А. Мюллери,
д-р наук, профессор, Гренада, Вест-Индия
Ю. Ф. Подоплёкин,
д-р техн. наук, проф., Санкт-Петербург, РФ
Д. Себерри,
д-р наук, проф., Волонгонг, Австралия
А. В. Смирнов,
д-р техн. наук, проф., Санкт-Петербург, РФ
Т. Суткиноу,
д-р наук, доцент, Джокарта, Индонезия
М. Н. Фаворская,
д-р техн. наук, проф., Красноярск, РФ
Л. Фортуна,
д-р наук, проф., Катания, Италия
А. Л. Фрадков,
д-р техн. наук, проф., Санкт-Петербург, РФ
В. И. Хименко,
д-р техн. наук, проф., Санкт-Петербург, РФ
А. Е. Храмов,
д-р физ.-мат. наук, Иннополис, РФ
А. А. Шальто,
д-р техн. наук, проф., Санкт-Петербург, РФ
А. П. Шепета,
д-р техн. наук, проф., Санкт-Петербург, РФ
Ю. И. Шокин,
акад. РАН, д-р физ.-мат. наук, проф., Новосибирск, РФ
З. М. Юлдашев,
д-р техн. наук, проф., Санкт-Петербург, РФ
Р. М. Юсупов,
чл.-корр. РАН, д-р техн. наук, проф., Санкт-Петербург, РФ

Редактор: А. Г. Ларионова
Корректор: Т. В. Звертановская
Дизайн: М. Л. Черненко, Ю. В. Умницына
Компьютерная верстка: Ю. В. Умницына

Адрес редакции: 190000, Санкт-Петербург,
Б. Морская ул., д. 67, ГУАП, РИЦ
Тел.: (812) 494-70-02, эл. адрес: ius.spb@gmail.com,
сайт: http://i-us.ru

ТЕОРЕТИЧЕСКАЯ И ПРИКЛАДНАЯ МАТЕМАТИКА

Vassiliev N. N., Parasidis I. N., Providas E. Exact solution method for Fredholm integro-differential equations with multipoint and integral boundary conditions. Part 2. Decomposition-extension method for squared operators 2

Хименко В. И. Вероятностный анализ двумерных векторных случайных процессов на фазовой плоскости 10

ОБРАБОТКА ИНФОРМАЦИИ И УПРАВЛЕНИЕ

Коновалов П. А., Корнилова А. В., Королев Д. М., Ренева Г. В., Фрадков А. Л., Широколов И. Ю., Ярош Д. С. Команда роботов-футболистов для соревнований RoboCup в лиге SSL: система и алгоритмы 19

Иванько Д. В., Рюмин Д. А., Карпов А. А., Железны М. Исследование влияния высокоскоростных видеоданных на точность распознавания аудиовизуальной речи 26

Рукавицын А. Н. Кластеризация данных в распределенных системах мониторинга 35

МОДЕЛИРОВАНИЕ СИСТЕМ И ПРОЦЕССОВ

Ключарёв А. А., Фоменкова А. А. Математическая модель анаэробного биореактора с закрепленной биомассой как объекта управления 44

ПРОГРАММНЫЕ И АППАРАТНЫЕ СРЕДСТВА

Гордеев А. В., Андреев А. В. Структура каталога как фильтр поиска записей службы каталогов OpenLDAP 52

ЗАЩИТА ИНФОРМАЦИИ

Новикова Е. С., Котенко И. В. Открытые задачи визуального анализа в системах управления информационной безопасностью 57

КОДИРОВАНИЕ И ПЕРЕДАЧА ИНФОРМАЦИИ

Советов Б. Я., Татарникова Т. М., Пойманова Е. Д. Организация многоуровневого хранения данных 68

Белии С. В., Ракицкий Ю. С. Использование семейств ортогональных функций для кодирования данных в схеме хаотической маскировки 76

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ И ОБРАЗОВАНИЕ

Комаров В. А., Сарафанов А. В., Тумковский С. Р. Многопользовательские распределенные измерительно-управляющие системы как элемент современной цифровой образовательной среды 83

ХРОНИКА И ИНФОРМАЦИЯ

Международная конференция по компьютерной математике — SAM-2019 95

СВЕДЕНИЯ ОБ АВТОРАХ

96

Журнал входит в БД SCOPUS, в RSCI на платформе Web of Science и в Перечень рецензируемых научных изданий, в которых должны быть опубликованы основные научные результаты диссертаций на соискание ученой степени кандидата наук, на соискание ученой степени доктора наук.

Сдано в набор 27.02.19. Подписано в печать 17.04.19. Формат 60×84/8. Гарнитура SchoolBookC. Печать цифровая. Усл. печ. л. 11,7. Уч.-изд. л. 16,2. Тираж 1000 экз (1-й завод 84 экз.). Заказ № 148. Оригинал-макет изготовлен в редакционно-издательском центре ГУАП. 190000, Санкт-Петербург, Б. Морская ул., 67.

Отпечатано с готовых диапозитивов в редакционно-издательском центре ГУАП. 190000, Санкт-Петербург, Б. Морская ул., 67.

Журнал распространяется по подписке. Подписку можно оформить в любом отделении связи по каталогу «Пресса России»: № 15385 — полугодовой индекс.

UDC 338.984

doi:10.31799/1684-8853-2019-2-2-9

Exact solution method for Fredholm integro-differential equations with multipoint and integral boundary conditions. Part 2. Decomposition-extension method for squared operators

N. N. Vassiliev^{a,b}, PhD, Tech., Senior Researcher, orcid.org/0000-0002-0841-1168, vasiliev@pdm.ras.ru

I. N. Parasidis^c, PhD, Associate Professor, paras@teilar.gr

E. Providas^d, PhD, Associate Professor, providas@teilar.gr

^aSaint-Petersburg Department of V. A. Steklov Institute of Mathematics of the RAS, 27, Fontanka, 191023, Saint-Petersburg, Russian Federation

^bSaint-Petersburg Electrotechnical University «LETI», 5, Prof. Popov St., 197376, Saint-Petersburg, Russian Federation

^cDepartment of Electrical Engineering, Technological Educational Institute of Thessaly, 41110, Larissa, Greece

^dDepartment of Mechanical Engineering, Technological Educational Institute of Thessaly, 41110, Larissa, Greece

Introduction: In Part 1 of this article, a direct method was presented for examining the solvability and uniqueness problem, and for obtaining a closed-form solution of boundary value problems which incorporate an m^{th} order linear ordinary Fredholm integro-differential operator, or a differential operator, along with multipoint and integral boundary conditions. Here, we focus on a special class of boundary value problems including the composite square of an integro-differential operator and the corresponding non-local boundary conditions. **Purpose:** To investigate the construction of the unique solution of $2m^{\text{th}}$ order boundary value problems in the special case of an operator which can be presented as composite squares of lower m^{th} order ones, and to develop an algorithm for constructing an exact solution for this special case. **Results:** By decomposition and applying the extension method explicated in Part 1, we provide a formula for obtaining an exact solution of boundary value problems for squared integro-differential operators, or differential operators, with multipoint and integral boundary conditions. This method is simple to use and can be easily incorporated to any Computer Algebra System.

Keywords – differential and Fredholm integro-differential equations, multipoint and non-local integral boundary conditions, decomposition of operators, correct operators, exact solutions.

For citation: Vassiliev N. N., Parasidis I. N., Providas E. Exact solution method for Fredholm integro-differential equations with multipoint and integral boundary conditions. Part 2. Decomposition-extension method for squared operators. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2019, no. 2, pp. 2–9. doi:10.31799/1684-8853-2019-2-2-9

Introduction

In the article [1], we presented the development, the applications and the necessity for studying boundary value problems encompassing m^{th} order linear ordinary Fredholm integro-differential operators and general nonlocal boundary conditions such as multipoint and integral boundary conditions. We proposed a direct constructive method for examining the existence and uniqueness of the solution and obtaining it in closed-form. The method was based on the extension theory of linear operators in Banach spaces, in particular on the technique developed in [2] and [3] for solving exactly linear and nonlinear, respectively, integro-differential equations subject to initial and classical boundary conditions.

In this paper, which is a sequel to [1], we study separately a specific type of boundary value problems involving the composite square of an m^{th} order linear ordinary Fredholm integro-differential,

or differential, operator, and analogous multipoint and integral boundary conditions. We establish the requirements under which there exists a unique solution and show how to construct it in closed-form by decomposing and utilizing the extension method described in [1].

The decomposition, or factorization, method for problems embracing integro-differential operators and unperturbed conventional boundary conditions is studied in [4]. Therefore, the current work can be seen also as an advancement of [4] where perturbed boundary conditions are considered. Factorization techniques find applications in several areas in sciences and engineering, see, for example, in [5, 6].

The organization of the paper is as follows. We first describe the decomposition-extension method and then we apply the method to solve second and fourth-order differential and integro-differential problems which can be formulated as composite squares of first and second-order problems, respectively. Lastly, some conclusions are quoted.

Decomposition-extension method

Let X be a complex Banach space, usually $X = \mathbb{C}[a, b]$ (or $X = L_p(a, b)$, $p \geq 1$), and $A: X \rightarrow X$ an m^{th} order linear ordinary differential operator, namely:

$$Au = a_0u^{(m)} + a_1u^{(m-1)} + \dots + a_mu, \quad (1)$$

where $a_i \in \mathbb{R}$ and $u = u(x) \in X_A^m$, where $X_A^m = \mathbb{C}^m[a, b]$ (or $X_A^m = \mathbf{W}_p^m(a, b)$). Let the space $\ker A$ be finite dimensional and $\mathbf{z} = (z_1, \dots, z_m)$ be a basis of it. Let \hat{A} be a correct restriction of A , specifically $\hat{A}u = Au$ for all u in

$$\mathcal{D}(\hat{A}) = \{u \in \mathcal{D}(A) : \Phi(u) = 0\}, \quad (2)$$

where $\Phi = \text{col}(\Phi_1, \dots, \Phi_m)$ is a vector of m bounded linear functionals on X_A^m , which are biorthogonal to z_1, \dots, z_m and describe some boundary conditions.

Consider the integro-differential operator $B: X \rightarrow X$:

$$Bu = Au - \mathbf{gF}(Au), \quad \mathcal{D}(B) = \{u \in \mathcal{D}(A) : \Phi(u) = \mathbf{N}\Psi(u)\}, \quad (3)$$

and the more complex integro-differential operator $B_1: X \rightarrow X$:

$$B_1u = A^2u - \mathbf{qF}(Au) - \mathbf{gF}(A^2u), \quad \mathcal{D}(B_1) = \{u \in \mathcal{D}(A^2) : \Phi(u) = \mathbf{N}\Psi(u), \Phi(Au) = \mathbf{D}\mathbf{F}(Au) + \mathbf{N}\Psi(Au)\}, \quad (4)$$

where A^2 is meant to be the composite product $A^2 = A(A)$, $\Psi = \text{col}(\Psi_1, \dots, \Psi_n)$ is a vector of n bounded linear functionals on X_A^m , $\mathbf{F} = \text{col}(F_1, \dots, F_n)$ is a vector of n bounded linear functionals on X , $\mathbf{g} = (g_1, \dots, g_n)$, $\mathbf{q} = (q_1, \dots, q_n) \in X^n$, q_1, \dots, q_n are linearly independent vectors, and \mathbf{D}, \mathbf{N} are $m \times n$ constant matrices. The equations $\Phi(u) = \mathbf{N}\Psi(u)$ and $\Phi(Au) = \mathbf{D}\mathbf{F}(Au) + \mathbf{N}\Psi(Au)$ symbolize general boundary conditions including multipoint and integral boundary conditions.

The boundary value problems $Bu = f$ and $B_1u = f$, for any $f \in X$, were studied and solved exactly by utilizing the extension method in [1].

We contemplate here the special case of the boundary value problem $B_1 = f$, $\forall f \in X$, when $B_1 = B^2$; B^2 is understood to be the composite product: $B^2 = B(B)$. For this case we prove the following theorem which provides solvability conditions and describes the decomposition-extension procedure for obtaining the solution in closed form.

Theorem. (i) The operator B_1 is decomposed in $B_1 = B^2$ in the case if

$$\mathbf{g} \in \mathcal{D}(A)^n, \mathbf{q} = \mathbf{A}\mathbf{g} - \mathbf{gF}(\mathbf{A}\mathbf{g}), \mathbf{D} = \Phi(\mathbf{g}) - \mathbf{N}\Psi(\mathbf{g}). \quad (5)$$

The operator B^2 is defined by

$$B^2u = A^2u - [\mathbf{A}\mathbf{g} - \mathbf{gF}(\mathbf{A}\mathbf{g})]\mathbf{F}(Au) - \mathbf{gF}(A^2u), \quad \mathcal{D}(B^2) = \{u \in \mathcal{D}(A^2) : \Phi(u) = \mathbf{N}\Psi(u), \Phi(Au) = [\Phi(\mathbf{g}) - \mathbf{N}\Psi(\mathbf{g})]\mathbf{F}(Au) + \mathbf{N}\Psi(Au)\}. \quad (6)$$

(ii) If the vectors \mathbf{q}, \mathbf{g} and matrices \mathbf{D}, \mathbf{N} satisfy (5), then the operator B_1 is injective if and only if

$$\det \mathbf{V} = \det[\mathbf{I}_n - \Psi(\mathbf{z})\mathbf{N}] \neq 0; \quad \det \mathbf{W} = \det[\mathbf{I}_n - \mathbf{F}(\mathbf{g})] \neq 0. \quad (7)$$

(iii) If the vectors \mathbf{q}, \mathbf{g} and matrices \mathbf{D}, \mathbf{N} satisfy (5) and $\det \mathbf{V} \neq 0$, $\det \mathbf{W} \neq 0$, then the operator B_1 is correct and the unique solution of the problem $B_1u = f$ is

$$u = B_1^{-1}f = \hat{A}^{-2}f + \mathbf{YF}(\hat{A}^{-1}f) + \mathbf{zNV}^{-1}\Psi(\hat{A}^{-2}f) + \left[\hat{A}^{-1}\mathbf{Y} + \mathbf{YF}(\mathbf{Y}) + \mathbf{zNV}^{-1}\Psi(\hat{A}^{-1}\mathbf{Y}) \right] \mathbf{F}(f) + \left[\hat{A}^{-1}\mathbf{z} + \mathbf{YF}(\mathbf{z}) + \mathbf{zNV}^{-1}\Psi(\hat{A}^{-1}\mathbf{z}) \right] \mathbf{NV}^{-1}\Psi(\hat{A}^{-1}f), \quad (8)$$

or

$$u = B_1^{-1}f = \hat{A}^{-1}\tilde{f} + \mathbf{YF}(\tilde{f}) + \mathbf{zNV}^{-1}\Psi(\hat{A}^{-1}\tilde{f}), \quad (9)$$

where

$$\tilde{f} = \hat{A}^{-1}f + \mathbf{YF}(f) + \mathbf{zNV}^{-1}\Psi(\hat{A}^{-1}f), \quad \mathbf{Y} = \left[\hat{A}^{-1}\mathbf{g} + \mathbf{zNV}^{-1}\Psi(\hat{A}^{-1}\mathbf{g}) \right] \mathbf{W}^{-1}. \quad (10)$$

Proof: (i) First we prove the second formula in (6). Denote by

$$\tilde{\mathcal{D}} = \left\{ u \in \mathcal{D}(A^2) : \Phi(u) = \mathbf{N}\Psi(u), \Phi(Au) = [\Phi(\mathbf{g}) - \mathbf{N}\Psi(\mathbf{g})]\mathbf{F}(Au) + \mathbf{N}\Psi(Au) \right\}.$$

Let $u \in \mathcal{D}(B^2)$ and $\mathbf{g} \in \mathcal{D}(A)^n$. Then by definition, $u \in \mathcal{D}(B)$ and $Bu \in \mathcal{D}(B)$, which since (3) implies $u \in \mathcal{D}(A)$, $\Phi(u) = \mathbf{N}\Psi(u)$ and $Bu \in \mathcal{D}(A)$, $\Phi(Bu) = \mathbf{N}\Psi(Bu)$. From $Bu = Au - \mathbf{gF}(Au) \in \mathcal{D}(A)$ it follows that $u \in \mathcal{D}(A^2)$. Further from the equation $\Phi(Bu) = \mathbf{N}\Psi(Bu)$ is implied that $u \in \tilde{\mathcal{D}}$.

Conversely, let $u \in \tilde{\mathcal{D}}$, then $u \in \mathcal{D}(A^2)$, $\Phi(u) = \mathbf{N}\Psi(u)$ and $\Phi(Au) - [\Phi(\mathbf{g}) - \mathbf{N}\Psi(\mathbf{g})]\mathbf{F}(Au) = \mathbf{N}\Psi(Au)$. Then $u \in \mathcal{D}(B)$, $Bu \in \mathcal{D}(A)$ and $\Phi(Au) - \Phi(\mathbf{g})\mathbf{F}(Au) = \mathbf{N}\Psi(Au) + \mathbf{N}\Psi(\mathbf{g})\mathbf{F}(Au)$, which implies $\Phi(Bu) = \mathbf{N}\Psi(Bu)$ or $Bu \in \mathcal{D}(B)$. Hence $u \in \mathcal{D}(B^2)$, and so (6) holds. Now we prove the first formula in (6). Let $u \in \mathcal{D}(B^2)$, $y = Bu$, $\mathbf{g} \in \mathcal{D}(A)^n$. Then

$$B^2u = By = Ay - \mathbf{gF}(Ay) = ABu - \mathbf{gF}(ABu) = A[Au - \mathbf{gF}(Au)] - \mathbf{gF}(A[Au - \mathbf{gF}(Au)]) =$$

$$= A^2u - AgF(Au) - gF(A^2u) + gF(Ag)F(Au).$$

Hence, $B_1u = B^2u$.

(ii) Let (5) holds and $\det V, \det W \neq 0$. By statement (i), $B_1 = B^2$ and so $\mathcal{D}(B_1) = \mathcal{D}(B^2)$. Since $\Phi(z) = I_m$, the relations in (6) can be written as

$$\Phi(u - zN\Psi(u)) = 0,$$

$$\Phi(Au - [g - zN\Psi(g)]F(Au) - zN\Psi(Au)) = 0,$$

which taking into account (2) imply

$$u - zN\Psi(u) \in \mathcal{D}(\hat{A}),$$

$$Au - [g - zN\Psi(g)]F(Au) - zN\Psi(Au) \in \mathcal{D}(\hat{A}).$$

Then from (6), since $z \in [\ker A]^m$, we obtain

$$\hat{A}(Au - gF(Au) + zN[\Psi(g)F(Au) - \Psi(Au)]) + g[F(Ag)F(Au) - F(A^2u)] = f,$$

$$Au - gF(Au) + zN[\Psi(g)F(Au) - \Psi(Au)] + \hat{A}^{-1}g[F(Ag)F(Au) - F(A^2u)] = \hat{A}^{-1}f,$$

$$\hat{A}(u - zN\Psi(u) - gF(Au) + zN[\Psi(g)F(Au) - \Psi(Au)]) + \hat{A}^{-1}g[F(Ag)F(Au) - F(A^2u)] = \hat{A}^{-1}f,$$

and hence

$$u - zN\Psi(u) - \hat{A}^{-1}gF(Au) + \hat{A}^{-1}zN[\Psi(g)F(Au) - \Psi(Au)] + \hat{A}^{-2}g[F(Ag)F(Au) - F(A^2u)] = \hat{A}^{-2}f,$$

and then

$$\begin{aligned} A^2u &= [Ag - gF(Ag)]F(Au) + gF(A^2u) + f, \\ Au &= gF(Au) - zN[\Psi(g)F(Au) - \Psi(Au)] - \\ &\quad - \hat{A}^{-1}g[F(Ag)F(Au) - F(A^2u)] + \hat{A}^{-1}f, \\ u &= zN\Psi(u) + \hat{A}^{-1}gF(Au) - \\ &\quad - \hat{A}^{-1}zN[\Psi(g)F(Au) - \Psi(Au)] - \\ &\quad - \hat{A}^{-2}g[F(Ag)F(Au) - F(A^2u)] + \hat{A}^{-2}f. \end{aligned}$$

Acting by the vectors F and Ψ we get

$$\begin{aligned} F(Au) &= F(g)F(Au) - F(z) \times \\ &\quad \times N[\Psi(g)F(Au) - \Psi(Au)] - F(\hat{A}^{-1}g) \times \\ &\quad \times [F(Ag)F(Au) - F(A^2u)] + F(\hat{A}^{-1}f), \end{aligned}$$

$$\begin{aligned} \Psi(Au) &= \Psi(g)F(Au) - \Psi(z) \times \\ &\quad \times N[\Psi(g)F(Au) - \Psi(Au)] - \Psi(\hat{A}^{-1}g) \times \\ &\quad \times [F(Ag)F(Au) - F(A^2u)] + \Psi(\hat{A}^{-1}f), \end{aligned}$$

$$\begin{aligned} V\Psi(u) &= \\ &= [\Psi(\hat{A}^{-1}g) - \Psi(\hat{A}^{-1}z)N\Psi(g) - \Psi(\hat{A}^{-2}g)F(Ag)] \times \\ &\quad \times F(Au) + \Psi(\hat{A}^{-1}z)N\Psi(Au) + \\ &\quad + \Psi(\hat{A}^{-2}g)F(A^2u) + \Psi(\hat{A}^{-2}f), \\ F(A^2u) &= [F(Ag) - F(g)F(Ag)]F(Au) + \\ &\quad + F(g)F(A^2u) + F(f), \end{aligned}$$

or

$$\begin{aligned} &[I_n - F(g) + F(z)N\Psi(g) + F(\hat{A}^{-1}g)F(Ag)] \times \\ &\quad \times F(Au) - F(\hat{A}^{-1}g)F(A^2u) - \\ &\quad - F(z)N\Psi(Au) = F(\hat{A}^{-1}f), \\ V\Psi(Au) &- [V\Psi(g) - \Psi(\hat{A}^{-1}g)F(Ag)]F(Au) - \\ &\quad - \Psi(\hat{A}^{-1}g)F(A^2u) = \Psi(\hat{A}^{-1}f), \\ V\Psi(u) &- \Psi(\hat{A}^{-1}z)N\Psi(Au) - \\ &- [\Psi(\hat{A}^{-1}g) - \Psi(\hat{A}^{-1}z)N\Psi(g) - \Psi(\hat{A}^{-2}g)F(Ag)] \times \\ &\quad \times F(Au) - \Psi(\hat{A}^{-2}g)F(A^2u) = \Psi(\hat{A}^{-2}f), \\ &[F(Ag) - F(g)F(Ag)]F(Au) + \\ &\quad + [F(g) - I_n]F(A^2u) = -F(f). \end{aligned}$$

Denoting

$$\begin{aligned} D_1 &= -V\Psi(g) + \Psi(\hat{A}^{-1}g)F(Ag); \\ D_2 &= W + F(z)N\Psi(g) + F(\hat{A}^{-1}g)F(\hat{A}g); \\ D_3 &= -\Psi(\hat{A}^{-1}g) + \Psi(\hat{A}^{-1}z)N\Psi(g) + \\ &\quad + \Psi(\hat{A}^{-2}g)F(Ag), \end{aligned}$$

we get

$$\begin{pmatrix} 0_n & -F(z)N & D_2 & -F(\hat{A}^{-1}g) \\ 0_n & V & D_1 & -\Psi(\hat{A}^{-1}g) \\ V & -\Psi(\hat{A}^{-1}z)N & D_3 & -\Psi(\hat{A}^{-2}g) \\ 0_n & 0_n & WF(Ag) & -W \end{pmatrix} \times$$

$$\times \begin{pmatrix} \Psi(u) \\ \Psi(Au) \\ \mathbf{F}(Au) \\ \mathbf{F}(A^2u) \end{pmatrix} = \begin{pmatrix} \mathbf{F}(\hat{A}^{-1}f) \\ \Psi(\hat{A}^{-1}f) \\ \Psi(\hat{A}^{-2}f) \\ -\mathbf{F}(f) \end{pmatrix}. \quad (11)$$

Designating the matrix the left by \mathbf{L}_2 , we have

$$\begin{aligned} \det \mathbf{L}_2 &= \det \begin{pmatrix} \mathbf{0}_n & -\mathbf{F}(z)\mathbf{N} & \mathbf{D}_2 & -\mathbf{F}(\hat{A}^{-1}\mathbf{g}) \\ \mathbf{0}_n & \mathbf{V} & \mathbf{D}_1 & -\Psi(\hat{A}^{-1}\mathbf{g}) \\ \mathbf{V} & -\Psi(\hat{A}^{-1}z)\mathbf{N} & \mathbf{D}_3 & -\Psi(\hat{A}^{-2}\mathbf{g}) \\ \mathbf{0}_n & \mathbf{0}_n & \mathbf{WF}(Ag) & -\mathbf{W} \end{pmatrix} = \\ &= \det \left[\begin{pmatrix} \mathbf{V}^{-1} & \mathbf{0}_n & \mathbf{0}_n & \mathbf{0}_n \\ \mathbf{0}_n & \mathbf{V}^{-1} & \mathbf{0}_n & \mathbf{0}_n \\ \mathbf{0}_n & \mathbf{0}_n & \mathbf{V}^{-1} & \mathbf{0}_n \\ \mathbf{0}_n & \mathbf{0}_n & \mathbf{0}_n & \mathbf{W}^{-1} \end{pmatrix} \begin{pmatrix} \mathbf{0}_n & -\mathbf{F}(z)\mathbf{N} & \mathbf{D}_2 & -\mathbf{F}(\hat{A}^{-1}\mathbf{g}) \\ \mathbf{0}_n & \mathbf{V} & \mathbf{D}_1 & -\Psi(\hat{A}^{-1}\mathbf{g}) \\ \mathbf{V} & -\Psi(\hat{A}^{-1}z)\mathbf{N} & \mathbf{D}_3 & -\Psi(\hat{A}^{-2}\mathbf{g}) \\ \mathbf{0}_n & \mathbf{0}_n & \mathbf{WF}(Ag) & -\mathbf{W} \end{pmatrix} \right] |\mathbf{V}|^3 |\mathbf{W}| = \\ &= \det \begin{pmatrix} \mathbf{0}_n & -\mathbf{V}^{-1}\mathbf{F}(z)\mathbf{N} & \mathbf{V}^{-1}\mathbf{D}_2 & -\mathbf{V}^{-1}\mathbf{F}(\hat{A}^{-1}\mathbf{g}) \\ \mathbf{0}_n & \mathbf{I}_n & \mathbf{V}^{-1}\mathbf{D}_1 & -\mathbf{V}^{-1}\Psi(\hat{A}^{-1}\mathbf{g}) \\ \mathbf{I}_n & -\mathbf{V}^{-1}\Psi(\hat{A}^{-1}z)\mathbf{N} & \mathbf{V}^{-1}\mathbf{D}_3 & -\mathbf{V}^{-1}\Psi(\hat{A}^{-2}\mathbf{g}) \\ \mathbf{0}_n & \mathbf{0}_n & \mathbf{F}(Ag) & -\mathbf{I}_n \end{pmatrix} |\mathbf{V}|^3 |\mathbf{W}| = \pm \det \begin{pmatrix} -\mathbf{V}^{-1}\mathbf{F}(z)\mathbf{N} & \mathbf{V}^{-1}\mathbf{D}_2 & -\mathbf{V}^{-1}\mathbf{F}(\hat{A}^{-1}\mathbf{g}) \\ \mathbf{I}_n & \mathbf{V}^{-1}\mathbf{D}_1 & -\mathbf{V}^{-1}\Psi(\hat{A}^{-1}\mathbf{g}) \\ \mathbf{0}_n & \mathbf{F}(Ag) & -\mathbf{I}_n \end{pmatrix} |\mathbf{V}|^3 |\mathbf{W}|. \end{aligned}$$

Multiplying from the right the third column of the determinant by the matrix $\mathbf{F}(Ag)$ and adding to the second column, we obtain

$$\begin{aligned} \det \mathbf{L}_2 &= \pm \det \begin{pmatrix} -\mathbf{V}^{-1}\mathbf{F}(z)\mathbf{N} & \mathbf{V}^{-1}\mathbf{D}_2 - \mathbf{V}^{-1}\mathbf{F}(\hat{A}^{-1}\mathbf{g})\mathbf{F}(Ag) & -\mathbf{V}^{-1}\mathbf{F}(\hat{A}^{-1}\mathbf{g}) \\ \mathbf{I}_n & \mathbf{V}^{-1}\mathbf{D}_1 - \mathbf{V}^{-1}\Psi(\hat{A}^{-1}\mathbf{g})\mathbf{F}(Ag) & -\mathbf{V}^{-1}\Psi(\hat{A}^{-1}\mathbf{g}) \\ \mathbf{0}_n & \mathbf{0}_n & -\mathbf{I}_n \end{pmatrix} |\mathbf{V}|^3 |\mathbf{W}| = \\ &= \det \begin{pmatrix} -\mathbf{V}^{-1}\mathbf{F}(z)\mathbf{N} & \mathbf{V}^{-1}[\mathbf{W} + \mathbf{F}(z)\mathbf{N}\Psi(\mathbf{g})] & -\mathbf{V}^{-1}\mathbf{F}(\hat{A}^{-1}\mathbf{g}) \\ \mathbf{I}_n & -\Psi(\mathbf{g}) & -\mathbf{V}^{-1}\Psi(\hat{A}^{-1}\mathbf{g}) \\ \mathbf{0}_n & \mathbf{0}_n & -\mathbf{I}_n \end{pmatrix} |\mathbf{V}|^3 |\mathbf{W}| = \\ &= \det \begin{pmatrix} -\mathbf{V}^{-1}\mathbf{F}(z)\mathbf{N} & \mathbf{V}^{-1}[\mathbf{W} + \mathbf{F}(z)\mathbf{N}\Psi(\mathbf{g})] \\ \mathbf{I}_n & -\Psi(\mathbf{g}) \end{pmatrix} |\mathbf{V}|^3 |\mathbf{W}|. \end{aligned}$$

Finally, multiplying from the right the first column of the determinant by the matrix $\Psi(\mathbf{g})$ and adding to the second column we get

$$\begin{aligned} \det \mathbf{L}_2 &= \pm \det \begin{pmatrix} -\mathbf{V}^{-1}\mathbf{F}(z)\mathbf{N} & \mathbf{V}^{-1}\mathbf{W} \\ \mathbf{I}_n & \mathbf{0}_n \end{pmatrix} |\mathbf{V}|^3 |\mathbf{W}| = \\ &= \pm |\mathbf{V}^{-1}| |\mathbf{W}| |\mathbf{V}|^3 |\mathbf{W}| = \pm |\mathbf{V}|^2 |\mathbf{W}|^2. \end{aligned}$$

So $\det L_2 = \pm |V|^2 |W|^2 \neq 0$. Let $u \in \ker B_1$. Then in (11) $f = 0$ and $L_2 \text{col}(\Psi(u), \Psi(Au), F(Au), F(A^2u)) = 0$, which since $\det L_2 \neq 0$, yields $\Psi(u) = \Psi(Au) = F(Au) = F(A^2u) = 0$. Substitution of these values into (6) imply $B_1u = B^2u = A^2u = 0$, $\Phi(u) = \Phi(Au) = 0$. Taking into account (2) we acquire $u \in \mathcal{D}(\hat{A}^2)$ and $B_1u = \hat{A}^2u = 0$.

By hypothesis \hat{A} is correct and so $u = 0$. Thus $\ker B_1 = \{0\}$ and B_1 is injective.

Conversely, let $\det V = 0$, then there exists a vector $c = \text{col}(c_1, \dots, c_n) \neq 0$ such that $Vc = 0$. Note that $u_0 = zNc \neq 0$, otherwise, since the components of z are linearly independent, we have $Nc = 0$ and from $Vc = 0$ follows that $c = 0$ which contradicts the hypothesis $c \neq 0$. Substituting u_0 into the first boundary condition (6), we get $\Phi(u_0) - N\Psi(u_0) = Nc - N\Psi(z)Nc = N[I_n - \Psi(z)N]c = NVc = 0$. Substituting u_0 into the second boundary condition, we obtain $\Phi(Au_0) - [\Phi(g) - N\Psi(g)]F(Au_0) - N\Psi(Au_0) = 0$, since $z \in \ker A$. So $u_0 \in \mathcal{D}(B^2)$. It is evident that $u_0 \in \ker B^2$. Hence $u_0 \in \mathcal{D}(B^2)$ and $u_0 \in \ker B^2$. So $\ker B^2 = \ker B_1 \neq \{0\}$ and $B^2 = B_1$ is not injective.

Let now $\det V \neq 0$, but $\det W = 0$. Then there exists a vector $c = \text{col}(c_1, \dots, c_n) \neq 0$ such that $Wc = 0$. Note that $gc \neq 0$ because of g_1, \dots, g_n is a linearly independent set and that the element $u_0 = [\hat{A}^{-1}g + zNV^{-1}\Psi(\hat{A}^{-1}g)]c \neq 0$, otherwise $c = 0$.

For u_0 we obtain

$$\begin{aligned} \Phi(u_0) - N\Psi(u_0) &= NV^{-1}\Psi(\hat{A}^{-1}g)c - N\Psi(\hat{A}^{-1}g)c - \\ &\quad - N\Psi(z)NV^{-1}\Psi(\hat{A}^{-1}g)c = \\ &= N[I_n - \Psi(z)N]V^{-1}\Psi(\hat{A}^{-1}g)c - N\Psi(\hat{A}^{-1}g)c = \\ &= [N\Psi(\hat{A}^{-1}g) - N\Psi(\hat{A}^{-1}g)]c = 0. \end{aligned}$$

So u_0 satisfies the first boundary condition (6). For u_0 we also obtain

$$\begin{aligned} \Phi(Au_0) - [\Phi(g) - N\Psi(g)]F(Au_0) - N\Psi(Au_0) &= \\ = \Phi(g)c - [\Phi(g) - N\Psi(g)]F(g)c - N\Psi(g)c &= \\ = \Phi(g)[I_n - F(g)]c - N\Psi(g)[I_n - F(g)]c &= \\ = \Phi(g)Wc - N\Psi(g)Wc = 0. \end{aligned}$$

So $u_0 \in \mathcal{D}(B^2)$. Moreover

$$\begin{aligned} B^2u_0 = A^2u_0 - [Ag - gF(Ag)]F(Au_0) - gF(A^2u_0) &= \\ = Agc - [Ag - gF(Ag)]F(g)c - gF(Ag)c &= \\ = Ag[I_n - F(g)]c - gF(Ag)[I_n - F(g)]c &= \\ = AgWc - gF(Ag)Wc = 0. \end{aligned}$$

Hence $u_0 \in \mathcal{D}(B^2)$ and $u_0 \in \ker B^2$. So $\ker B^2 \neq \{0\}$ and B^2 is not injective. So we proved that B_1 is injective if and only if $\det V \neq 0, \det W \neq 0$. The statement (ii) holds.

(iii) Let the vectors q, g, v, w and matrices D, N satisfy (5) and $\det V \neq 0, \det W \neq 0$. Then, by the statement (ii), the operator $B_1 = B^2$ is injective and the problem $B_1u = f$ has a unique solution. We recall that by Theorem [1] the unique solution of the equation $Bu = f$ for all $f \in X$ is given by

$$\begin{aligned} u = B^{-1}f = \hat{A}^{-1}f + \\ + [\hat{A}^{-1}g + zNV^{-1}\Psi(\hat{A}^{-1}g)]W^{-1}F(f) + \\ + zNV^{-1}\Psi(\hat{A}^{-1}f). \end{aligned} \quad (12)$$

Let $B_1u = B^2u = f$, where $f \in X$. Denoting by $Y = [\hat{A}^{-1}g + zNV^{-1}\Psi(\hat{A}^{-1}g)]W^{-1}$ and $\tilde{f} = Bu$, we get $B\tilde{f} = f$. Then by means of (12) the solution of this equation is given by

$$\tilde{f} = B^{-1}f = \hat{A}^{-1}f + YF(f) + zNV^{-1}\Psi(\hat{A}^{-1}f). \quad (13)$$

Applying again (12) we find the solution of the problem $Bu = \tilde{f}$ or $B_1u = f$

$$u = B_1^{-1}f = B^{-1}\tilde{f} = \hat{A}^{-1}\tilde{f} + YF(\tilde{f}) + zNV^{-1}\Psi(\hat{A}^{-1}\tilde{f}), \quad (14)$$

which is equation (9). Substituting the value of \tilde{f} from (13) into (14), we get

$$\begin{aligned} u = B_1^{-1}f = B^{-2}f = B^{-1}\tilde{f} = \hat{A}^{-2}f + \hat{A}^{-1}YF(f) + \\ + \hat{A}^{-1}zNV^{-1}\Psi(\hat{A}^{-1}f) + Y[F(\hat{A}^{-1}f) + F(Y)F(f) + \\ + F(z)NV^{-1}\Psi(\hat{A}^{-1}f)] + zNV^{-1}[\Psi(\hat{A}^{-2}f) + \\ + \Psi(\hat{A}^{-1}Y)F(f) + \Psi(\hat{A}^{-1}z)NV^{-1}\Psi(\hat{A}^{-1}f)] = \\ = \hat{A}^{-2}f + YF(\hat{A}^{-1}f) + zNV^{-1}\Psi(\hat{A}^{-2}f) + \\ + [\hat{A}^{-1}Y + YF(Y) + zNV^{-1}\Psi(\hat{A}^{-1}Y)]F(f) + \\ + [\hat{A}^{-1}z + YF(z) + zNV^{-1}\Psi(\hat{A}^{-1}z)]NV^{-1}\Psi(\hat{A}^{-1}f), \end{aligned}$$

which is the solution (8). In the above solutions f is arbitrary, consequently, $R(B_1) = X$. Since the operators $\hat{A}^{-2}, \hat{A}^{-1}$ and the functionals F and Ψ are bounded, from (8) or (9) follows the boundedness of $B_1^{-1} = B^{-2}$, i. e. the operator B_1 is correct. The theorem is proved.

The next corollary follows from the above theorem in the case $q = g = 0$ and it is useful for solving a class of differential equations with multipoint and nonlocal boundary conditions.

Corollary. The differential operator $B_1: X \rightarrow X$ be defined by

$$\begin{aligned} B_1u = A^2u = f, \\ \mathcal{D}(B_1) = \{u \in \mathcal{D}(A^2): \Phi(u) = N\Psi(u), \\ \Phi(Au) = N\Psi(Au)\}. \end{aligned} \quad (15)$$

Then the operator B_1 is correct if and only if

$$\det \mathbf{V} = \det[\mathbf{I}_n - \Psi(\mathbf{z})\mathbf{N}] \neq 0 \quad (16)$$

and the unique solution of (15) is given by

$$u = B_1^{-1}f = \hat{A}^{-2}f + \mathbf{zNV}^{-1}\Psi(\hat{A}^{-2}f) + [\hat{A}^{-1}\mathbf{z} + \mathbf{zNV}^{-1}\Psi(\hat{A}^{-1}\mathbf{z})]\mathbf{NV}^{-1}\Psi(\hat{A}^{-1}f). \quad (17)$$

Proof: For $\mathbf{g} = \mathbf{0}$ from (7) immediately follows $\det \mathbf{W} = \det \mathbf{I}_n = 1 \neq 0$. Then by Theorem, the operator B_1 is correct if and only if $\det \mathbf{V} \neq 0$. From (8) for $\mathbf{g} = \mathbf{q} = \mathbf{0}$ follows the solution of this problem.

Examples

In this section we consider some examples boundary value problems to explain the application of the decomposition-extension method and to demonstrate its efficiency.

First, we recall some known results. The problem

$$\begin{aligned} \hat{A}u &= u^{(m)}(x) = f(x), \\ \mathcal{D}(\hat{A}) &= \{u \in \mathbb{C}^m[0, 1] : u(0) = \\ &= u'(0) = \dots = u^{(m-1)}(0) = 0\}, \end{aligned}$$

is correct and its exact solution is given by

$$u(x) = \hat{A}^{-1}f(x) = \frac{1}{(m-1)!} \int_0^x (x-t)^{m-1} f(t) dt. \quad (18)$$

If the function $u(x) \in \mathbb{C}^m[a, b]$ and $x_0 \in [a, b]$, then the functionals $T_k(u) = u^{(k-1)}(x_0)$, $k = 1, \dots, m$, and $T(u) = \sum_{k=1}^m a_k u^{(k-1)}(x_0)$ are linear and bounded on $\mathbb{C}^k[a, b]$ and $\mathbb{C}^m[a, b]$, respectively.

Example 1

Consider the differential boundary value problem $u''(x) = f(x)$, $x \in [0, 1]$:

$$u(0) = vu(1), u'(0) = vu'(1). \quad (19)$$

By taking $X = \mathbb{C}[0, 1]$ and

$$Au = u'(x), \mathcal{D}(A) = \{u \in \mathbb{C}^1[0, 1]\},$$

$$A^2u = u''(x), \mathcal{D}(A^2) = \{u \in \mathbb{C}^2[0, 1]\};$$

$$\hat{A}u = Au, \mathcal{D}(\hat{A}) = \{u \in \mathcal{D}(A) : u(0) = 0\},$$

$$\hat{A}^2u = A^2u, \mathcal{D}(\hat{A}^2) = \{u \in \mathcal{D}(A^2) : u(0) = u'(0) = 0\},$$

we can put the problem (19) as in Corollary, equation (15), namely:

$$B_1u = A^2u = f,$$

$$\mathcal{D}(B_1) = \{u \in \mathcal{D}(A^2) : u(0) = vu(1), u'(0) = vu'(1)\},$$

where $\Phi(u) = (u(0))$, $\Psi(u) = (u(1))$ and $\mathbf{N} = (v)$. Then $\Phi(Au) = (u'(0))$ and $\Psi(Au) = (u'(1))$. Let $\mathbf{z} = (1)$ and notice that $\Phi(\mathbf{z}) = \mathbf{z}(0) = 1$. If $\det \mathbf{V} = \det(1 - v) \neq 0$ then the unique solution $u = B_1^{-1}f$ follows from Corollary where

$$\hat{A}^{-1}f = \int_0^x f(t) dt, \hat{A}^{-2}f = \int_0^x (x-t)f(t) dt,$$

by means of (18).

Example 2

Let the fourth-order differential boundary value problem with multipoint boundary conditions $u^{(4)}(x) = f(x)$, $x \in [0, 1]$:

$$\begin{aligned} u(0) &= v_{11}u\left(\frac{1}{2}\right) + v_{12}u(1); \\ u'(0) &= v_{21}u\left(\frac{1}{2}\right) + v_{22}u(1); \\ u''(0) &= v_{11}u''\left(\frac{1}{2}\right) + v_{12}u''(1); \\ u'''(0) &= v_{21}u''\left(\frac{1}{2}\right) + v_{22}u''(1). \end{aligned} \quad (20)$$

We recast the problem (20) into the form (15) where $X = \mathbb{C}[0, 1]$ and

$$Au = u''(x), \mathcal{D}(A) = \{u \in \mathbb{C}^2[0, 1]\},$$

$$A^2u = u^{(4)}(x), \mathcal{D}(A^2) = \{u \in \mathbb{C}^4[0, 1]\},$$

$$\hat{A}u = Au, \mathcal{D}(\hat{A}) = \{u \in \mathcal{D}(A) : u(0) = u'(0) = 0\},$$

$$\hat{A}^2u = A^2u,$$

$$\mathcal{D}(\hat{A}^2) = \{u \in \mathcal{D}(A^2) : u(0) = u'(0) = u''(0) = u'''(0) = 0\};$$

$$\Phi(u) = \begin{pmatrix} u(0) \\ u'(0) \end{pmatrix}; \Psi(u) = \begin{pmatrix} u\left(\frac{1}{2}\right) \\ u(1) \end{pmatrix};$$

$$\Phi(Au) = \begin{pmatrix} u''(0) \\ u'''(0) \end{pmatrix}; \Psi(Au) = \begin{pmatrix} u''\left(\frac{1}{2}\right) \\ u''(1) \end{pmatrix};$$

$$\Phi(u) = \mathbf{N}\Psi(u) = \begin{pmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{pmatrix} \Psi(u); \Phi(Au) = \mathbf{N}\Psi(Au).$$

Let $\mathbf{z} = (1, x)$ and notice that $\Phi(\mathbf{z}) = \mathbf{z}(0) = \mathbf{I}_2$. If

$$\det \mathbf{V} = \det \left(\mathbf{I}_2 - \begin{pmatrix} 1 & 1/2 \\ 1 & 1 \end{pmatrix} \mathbf{N} \right) \neq 0,$$

then the unique solution $u = B_1^{-1}f$ follows from Corollary by making use of (18).

Example 3

Contemplate the fourth-order Fredholm integro-differential boundary value problem with general integral boundary conditions

$$\begin{aligned}
 &u^{(4)}(x) + (x^2 - 1) \int_0^1 x u''(x) dx - (x^2 + 1) \times \\
 &\quad \times \int_0^1 x u^{(4)}(x) dx = x - 2, \\
 &\quad x \in [0, 1], \\
 &u(0) = \frac{15}{22} \int_0^1 u(x) dx; \\
 &u'(0) = 0; \\
 &u''(0) = \frac{1}{11} \int_0^1 x u''(x) dx + \frac{15}{22} \int_0^1 u''(x) dx; \\
 &u'''(0) = 0.
 \end{aligned} \tag{21}$$

We formulate the problem (21) as in (4). We take $X = \mathbb{C}[0, 1]$ and

$$\begin{aligned}
 &Au = u''(x), \mathcal{D}(A) = \{u \in \mathbb{C}^2[0, 1]\}, \\
 &A^2u = u^{(4)}(x), \mathcal{D}(A^2) = \{u \in \mathbb{C}^4[0, 1]\}, \\
 &\hat{A}u = Au, \mathcal{D}(\hat{A}) = \{u \in \mathcal{D}(A) : u(0) = u'(0) = 0\}, \\
 &\quad \hat{A}^2u = A^2u, \\
 &\mathcal{D}(\hat{A}^2) = \{u \in \mathcal{D}(A^2) : u(0) = u'(0) = u''(0) = u'''(0) = 0\}; \\
 &\quad f(x) = x - 2; \\
 &\quad \mathbf{q} = (1 - x^2); \mathbf{g} = (x^2 + 1); \\
 &\mathbf{F}(Au) = \left(\int_0^1 x u''(x) dx \right); \mathbf{F}(A^2u) = \left(\int_0^1 x u^{(4)}(x) dx \right); \\
 &\quad \Phi(u) = \begin{pmatrix} u(0) \\ u'(0) \end{pmatrix}; \Psi(u) = \left(\int_0^1 u(x) dx \right); \\
 &\quad \Phi(Au) = \begin{pmatrix} u''(0) \\ u'''(0) \end{pmatrix}; \Psi(Au) = \left(\int_0^1 u''(x) dx \right);
 \end{aligned}$$

References

1. Vassiliev N. N., Parasidis I. N., Providas E. Exact solution method for Fredholm integro-differential equations with multipoint and integral boundary conditions. Part 1. Extension method. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 6, pp. 14–23. doi:10.31799/1684-8853-2018-6-14-23
2. Parasidis I. N., Providas E. *Extension operator method for the exact solution of integro-differential Equations*. In: Pardalos P., Rassias T. (eds). *Contributions*

$$\Phi(u) = \mathbf{N}\Psi(u) = \begin{pmatrix} 15/22 \\ 0 \end{pmatrix} \Psi(u);$$

$$\begin{aligned}
 \Phi(Au) &= \mathbf{D}\mathbf{F}(Au) + \mathbf{N}\Psi(Au) = \\
 &= \begin{pmatrix} 1/11 \\ 0 \end{pmatrix} \mathbf{F}(Au) + \mathbf{N}\Psi(Au).
 \end{aligned}$$

Observe that $\mathbf{g} = (x^2 + 1) \in \mathcal{D}(A)$ and $\mathbf{q} = A\mathbf{g} - \mathbf{g}\mathbf{F}(A\mathbf{g})$,

$$\mathcal{D} = \Phi(\mathbf{g}) - \mathbf{N}\Psi(\mathbf{g}). \tag{22}$$

Let $\mathbf{z} = (1, x)$. Then $\Phi(\mathbf{z}) = \mathbf{I}_2$, $\Psi(\mathbf{z}) = \left(1, \frac{1}{2}\right)$ and

$$\det \mathbf{V} = \det(\mathbf{I}_1 - \Psi(\mathbf{z})\mathbf{N}) = \frac{7}{22} \neq 0;$$

$$\det \mathbf{W} = \det(\mathbf{I}_1 - \mathbf{F}(\mathbf{g})) = \frac{1}{4} \neq 0. \tag{23}$$

Because of (22) and (23), Theorem applies. Hence the operator B_1 is correct which means that the problem (21) admits a unique solution. By substituting into (8) or (9) and making use of (18), we get

$$\begin{aligned}
 u(x) &= -\frac{1}{317520} (2352x^6 - 2646x^5 + \\
 &+ 212387x^4 + 1169427x^2 + 926103).
 \end{aligned}$$

Conclusion

By means of decomposition and the extension method, we provided a ready to use formula for constructing the solution in closed form of boundary value problems involving the composite square of an m^{th} order integro-differential operator of Fredholm type and nonlocal boundary conditions such as appropriate multipoint and integral conditions. The method is also applicable to boundary value problems for the composite squared m^{th} order linear ordinary differential operators.

- in *Mathematics and Engineering*. Springer, Cham, 2016, pp. 473–496. doi:10.1007/978-3-319-31317-7_23
3. Parasidis I. N., Providas E. *On the exact solution of nonlinear integro-differential equations*. In: Rassias T. (eds). *Applications of Nonlinear Analysis*. Series: *Springer Optimization and Its Applications*, Springer, Cham, 2018, vol. 134, pp. 591–609. doi:10.1007/978-3-319-89815-5_21
4. Parasidis I. N., Providas E., Tsekrekos P. C. Factorization of linear operators and some eigenvalue problems of special operators. *Vestnik Bashkirskogo universiteta*, 2012, vol. 17, no. 2, pp. 830–839.

5. Barkovskii L. M., Furs A. N. Factorization of integro-differential equations of optics of dispersive anisotropic media and tensor integral operators of wave packet velocities. *Optics and Spectroscopy*, 2001, vol. 90, iss. 4, pp. 561–567. doi:10.1134/1.1366751

6. Barkovsky L. M., Furs A. N. Factorization of integro-differential equations of the acoustics of dispersive viscoelastic anisotropic media and the tensor integral operators of wave packet velocities. *Acoustical Physics*, 2002, vol. 48, iss. 2, pp. 128–132. doi:10.1134/1.1460945

УДК 338.984

doi:10.31799/1684-8853-2019-2-2-9

Метод нахождения точных решений для интегро-дифференциальных уравнений Фредгольма с мультиточечными и интегральными краевыми условиями. Часть 2. Метод разложения-расширения квадратичных операторов

Н. Н. Васильев^{а,б}, канд. физ.-мат. наук, старший научный сотрудник,

orcid.org/0000-0002-0841-1168, vasiliev@pdmi.ras.ru

И. Н. Парасидис^в, PhD, доцент, paras@teilar.gr

Е. Провидас^г, PhD, доцент, providas@teilar.gr

^аСанкт-Петербургское отделение Математического института им. В. А. Стеклова РАН, наб. р. Фонтанки, 27, Санкт-Петербург, 191023, РФ

^бСанкт-Петербургский государственный электротехнический университет «ЛЭТИ», Санкт-Петербург, Профессора Попова ул., 5, Санкт-Петербург, 197376, РФ

^вКафедра электротехники, Технологический институт Фессалии, 41110, Лариса, Греция

^гКафедра машиностроения, Технологический институт Фессалии, 41110, Лариса, Греция

Введение: в первой части статьи представлен прямой метод исследования проблемы разрешимости и единственности и получения в замкнутой форме решения краевых задач, включающих линейный обыкновенный интегро-дифференциальный оператор Фредгольма или дифференциальный оператор m -го порядка, а также многоточечные и интегральные граничные условия. Здесь мы сосредоточимся на специальном классе краевых задач, включающих квадрат интегро-дифференциального оператора и соответствующих нелокальных граничных условий. **Цель:** исследование построения единственного решения краевых задач 2-го порядка в частном случае оператора, который может быть представлен в виде композиции квадратов операторов более низких порядков, а также разработка алгоритма построения точного решения в этом частном случае. **Результаты:** с помощью декомпозиции и метода расширения, описанного в первой части, нами разработан алгоритм для получения точного решения краевых задач для квадрато-интегро-дифференциальных операторов или дифференциальных операторов с многоточечными и интегральными граничными условиями. Этот метод прост в использовании и может быть легко имплементирован в большинство современных систем компьютерной алгебры.

Ключевые слова — дифференциальные и фредгольмовы интегро-дифференциальные уравнения, многоточечные и нелокальные интегральные граничные условия, разложение операторов, корректные операторы, точные решения.

Для цитирования: Vassiliev N. N., Parasidis I. N., Providas E. Exact solution method for Fredholm integro-differential equations with multipoint and integral boundary conditions. Part 2. Decomposition-extension method for squared operators. *Информационно-управляющие системы*, 2019, № 2, с. 2–9. doi:10.31799/1684-8853-2019-2-2-9

For citation: Vassiliev N. N., Parasidis I. N., Providas E. Exact solution method for Fredholm integro-differential equations with multipoint and integral boundary conditions. Part 2. Decomposition-extension method for squared operators. *Informatsionno- upravliaiushchie sistemy* [Information and Control Systems], 2019, no. 2, pp. 2–9. doi:10.31799/1684-8853-2019-2-2-9

Авторы статьи Vassiliev N. N., Parasidis I. N., Providas E. Exact solution method for Fredholm integro-differential equations with multipoint and integral boundary conditions. Part 1. Extension method, 2018, № 6 приносят извинения за опечатки.

Страница	Столбец	Строка	Напечатано	Следует читать
15	правый	8 снизу	of (5) is	of (5) for $\alpha_0 = 1, \alpha_1 = \dots = \alpha_m = 0, a = 0$ is
15	правый	6 снизу	$f(x) \in C[a, b]$.	$f(x) \in C[0, b]$.
16	левый	9 снизу	X^{m-1} and respectively	$[X^{m-1}]^*$ and X^* respectively

УДК 519.21:537.86

doi:10.31799/1684-8853-2019-2-10-18

Вероятностный анализ двумерных векторных случайных процессов на фазовой плоскости

В. И. Хименко^а, доктор техн. наук, профессор, orcid.org/0000-0003-3005-052X, vih.13@yandex.ru
^аОАО «НИО ЦИТ «Петрокомета», Б. Морская ул., 67, Санкт-Петербург, 190000, РФ

Введение: для моделей векторных случайных процессов характерно многообразие практических применений и, одновременно с этим, сложность исследований детальной вероятностной структуры. Такие особенности достаточно часто приводят к рассмотрению лишь отдельных компонент векторных функций и вынужденному переходу к анализу скалярных процессов. Все это существенно снижает общую информативность исследований и повышает актуальность поиска перспективных подходов к вероятностному анализу векторных процессов. **Цель:** представление векторных случайных процессов в фазовом пространстве состояний и использование методов общей теории выбросов случайных функций в исследовании вероятностной структуры фазовых траекторий. **Результаты:** выделены характерные особенности отображения случайных процессов на фазовой плоскости, введены числовые характеристики для описания и анализа вероятностного поведения выборочных функций. Выполнен вероятностный анализ характеристик типа «пересечений уровней» для векторных процессов при различном определении областей допустимых значений. На примере анализа типовых моделей двумерных гауссовых процессов и моделей класса «сигнал плюс шум» показаны аналитические зависимости вероятностной структуры фазовых траекторий от заданных пороговых уровней, основных параметров распределений и спектрально-корреляционных свойств исследуемых процессов. **Практическая значимость:** повышена информативность исследований и наглядность отображения информации при анализе вероятностного поведения векторных случайных процессов. Объединены общие методы фазового пространства векторных процессов, методы фазовой плоскости при анализе скалярных процессов и общие методы теории выбросов случайных функций. Классические подходы визуального описания фазовых траекторий дополнены методами количественного анализа детальной вероятностной структуры случайных функций.

Ключевые слова – случайные процессы, фазовое пространство, фазовые траектории, вероятностный анализ, выбросы случайных функций, характеристики пересечений уровней, обработка информации.

Для цитирования: Хименко В. И. Вероятностный анализ двумерных векторных случайных процессов на фазовой плоскости. *Информационно-управляющие системы*, 2019, № 2, с. 10–18. doi:10.31799/1684-8853-2019-2-10-18

For citation: Khimenko V. I. Probabilistic analysis of two-dimensional vector random processes on the phase plane. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2019, no. 2, pp. 10–18 (In Russian). doi:10.31799/1684-8853-2019-2-10-18

Введение

Многие задачи приема, обработки и анализа информации, задачи теории управления и теории надежности, вопросы исследования устойчивости сложных динамических систем и задачи теории рискованных ситуаций по своему содержанию приводят к необходимости рассмотрения вероятностной структуры векторных случайных процессов [1–4]. Возникают подобные задачи и при пространственно-временной обработке сигналов в радиофизических приложениях, при исследованиях вибраций и колебательных процессов в технике, при изучении электромагнитных полей в акустике, биофизике и геофизике [5–8].

Как правило, математическое описание и исследование векторных процессов заметно сложнее вероятностного анализа скалярных случайных функций. Эта особенность достаточно часто приводит к тому, что исследования векторных процессов выполняются лишь на уровне вероятностного анализа отдельных компонент, т. е. решаемые задачи сводятся к рассмотрению скалярных случайных процессов.

Такой подход упрощает вероятностный анализ, однако при этом заметно снижает информа-

тивность, ограничивает общность исследований и, соответственно, сокращает потенциальные возможности практического использования моделей векторных процессов.

Целью данной работы является рассмотрение особенностей представления двумерных случайных процессов на фазовой плоскости, введение числовых характеристик для вероятностного описания фазовых траекторий и общий анализ вероятностной структуры векторных процессов в фазовом пространстве состояний.

Фазовое пространство случайного векторного процесса

Обычно состояние сложных динамических систем характеризуется несколькими параметрами. Эти параметры могут изменяться случайным образом. Мгновенное состояние исследуемой системы можно при этом рассматривать как точку некоторого фазового пространства (пространства состояний). Изменения во времени состояния системы описываются при таком подходе функцией $\zeta(t)$ времени t , значения которой меняются в пространстве состояний. Функция $\zeta(t)$ представляет

собой случайный процесс, и если пространство состояний является многомерным, то процесс $\xi(t)$ относится к классу векторных случайных процессов.

Рассмотрим одну из наиболее простых и наиболее распространенных ситуаций, когда процесс $\xi(t)$ является двумерным векторным процессом: $\xi(t) = \{\xi_1(t), \xi_2(t)\}$. Компоненты этого процесса $\xi_1(t)$ и $\xi_2(t)$ будем считать непрерывными случайными функциями, значения которых изменяются на вещественной оси $\xi_i(t) \in (-\infty, \infty)$, $i = 1, 2$. Пространство состояний для такого процесса $\xi(t)$ является двумерным, и, следовательно, рассмотрение вероятностной структуры векторного процесса $\xi(t)$ может быть выполнено на основе его исследования на фазовой плоскости $(\xi_1, \xi_2) = (\xi_1(t), \xi_2(t))$.

Если при анализе случайного процесса $\xi(t)$ компоненты $\xi_1(t)$ и $\xi_2(t)$ интерпретировать как координаты точки на плоскости (ξ_1, ξ_2) , то векторный процесс $\xi(t)$ геометрически можно представить в виде семейства траекторий $\{\xi_1(t), \xi_2(t), t \in T\}$. Плоскость (ξ_1, ξ_2) является в данном случае фазовым пространством (фазовой плоскостью) двумерного векторного процесса $\xi(t)$. Изменения значений $\xi(t)$ во времени $t \in T$ описываются движением фазовой (или изображающей) точки $(\xi_1(t), \xi_2(t))$ по некоторой кривой — фазовой траектории $L(\xi_1, \xi_2; t)$ на плоскости (ξ_1, ξ_2) .

При анализе векторных процессов $\xi(t) = \{\xi_1(t), \xi_2(t)\}$ фазовая плоскость (ξ_1, ξ_2) отображает информацию о совместных изменениях мгновенных значений двух случайных функций $\xi_1(t)$ и $\xi_2(t)$. Эти функции, в зависимости от решаемых задач, могут иметь различные вероятностные характеристики, могут быть независимыми или функционально связанными. Как частный случай, пространство состояний $(\xi_1(t), \xi_2(t))$ векторных процессов $\xi(t) = \{\xi_1(t), \xi_2(t)\}$ при $\xi_1(t) = \xi(t)$ и $\xi_2(t) = \xi'(t) = d\xi(t)/dt$ может включать в себя и рассмотрение фазовых траекторий $(\xi_1(t), \xi_2(t)) = (\xi(t), \xi'(t))$ скалярных случайных процессов $\xi(t)$ на фазовой плоскости классического вида $(\xi_1, \xi_2) = (\xi, \xi')$.

Характеристики вероятностной структуры процессов

Исследования векторных процессов во многих прикладных задачах связаны с необходимостью детального анализа вероятностных характеристик отдельных траекторий $(\xi_1(t), \xi_2(t))$ на фазовой плоскости (ξ_1, ξ_2) . При этом обычно задается некоторая область допустимых значений Ω и рассматриваются вероятностные характеристики выходов траектории $\xi(t)$ за границы заданной области Ω . Примерами таких задач могут являться задачи исследования надежности, эффективности и безопасности работы сложных систем, задачи оценки

рисковых ситуаций, задачи технической и медицинской диагностики, задачи исследования экстремальных состояний и предельно допустимых режимов работы систем [2–4, 9–12]. Моменты выходов в подобных задачах обычно отождествляются с моментами отказов системы, потерей устойчивости, наступлением рискованных ситуаций или переходом системы в режимы критических состояний.

Рассмотрим особенности описания и основные характеристики траекторий двумерного векторного случайного процесса $\xi(t) = \{\xi_1(t), \xi_2(t)\}$ при его представлении на плоскости (ξ_1, ξ_2) (рис. 1). В диапазоне возможных изменений компонент $\xi_1(t)$ и $\xi_2(t)$ зададим некоторую область допустимых значений Ω , границы которой определяются из условий

$$\Omega: \xi_1(t) \in [-H_1, H_1], \xi_2(t) \in [-H_2, H_2]. \quad (1)$$

Выходы траектории $(\xi_1(t), \xi_2(t))$ из области Ω через границы $[-H_1, H_1]$ (см. рис. 1) могут происходить только за счет выходов компоненты $\xi_1(t)$ за уровни $\pm H_1$ при условии, что вторая компонента $\xi_2(t)$ находится в интервале $[-H_2, H_2]$. Аналогично выходы траектории $(\xi_1(t), \xi_2(t))$ из области Ω через границы $[-H_2, H_2]$ являются следствием выходов компоненты $\xi_2(t)$ за уровни $\pm H_2$ при условии $\xi_1(t) \in [-H_1, H_1]$. Учитывая эти особенности, для среднего числа выходов $N_\xi(\Omega, T)$ двумерного процесса $\xi(t)$ из заданной области (1) на интервале $t \in [t_0, t_0 + T]$ можно записать общую формулу [13, 14]

$$\begin{aligned} N_\xi(\Omega, T) = & \left[N_1^+(H_1, T) + N_1^-(-H_1, T) \right] \times \\ & \times P\{\xi_2(t) \in [-H_2, H_2]\} + \\ & + \left[N_2^+(H_2, T) + N_2^-(-H_2, T) \right] \times \\ & \times P\{\xi_1(t) \in [-H_1, H_1]\}, \quad (2) \end{aligned}$$

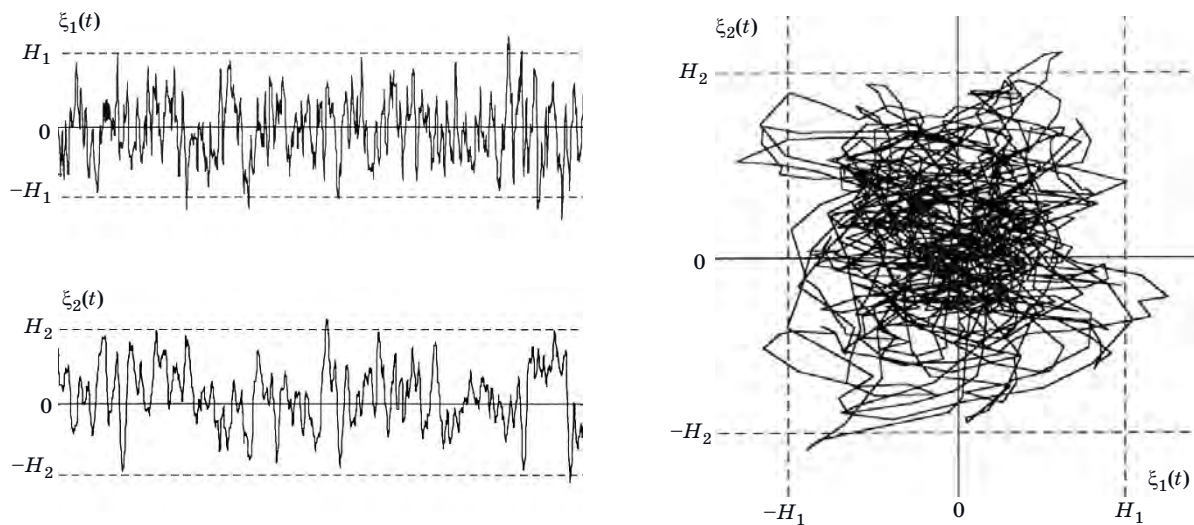
где $N_i^\pm(\pm H_i, T)$ — среднее число положительных или отрицательных выбросов одномерного процесса $\xi_i(t)$, $i = 1, 2$ на интервале $t \in [t_0, t_0 + T]$ за соответствующий пороговый уровень $\pm H_i$.

Пребывание траектории $(\xi_1(t), \xi_2(t))$ процесса $\xi(t)$ в области Ω эквивалентно совместному выполнению условий

$$\xi_1(t) \in [-H_1, H_1] \text{ и } \xi_2(t) \in [-H_2, H_2].$$

Следовательно, для среднего значения $T_\xi(\Omega)$ относительной длительности и вероятности пребывания процесса $\xi(t)$ в заданной области Ω будет справедливо равенство

$$\begin{aligned} T_\xi(\Omega) = & P\{\xi(t) \in \Omega\} = \\ = & P\{\xi_1(t) \in [-H_1, H_1], \xi_2(t) \in [-H_2, H_2]\}. \quad (3) \end{aligned}$$



■ **Рис. 1.** Представление отдельной реализации двумерного векторного процесса на фазовой плоскости
 ■ **Fig. 1.** Representation of a sample function of a two-dimensional vector process on the phase plane

В свою очередь, если известны значения $T_{\xi}(\Omega)$ и $N_{\xi}(\Omega, T)$, то может быть найдена и средняя длительность выброса двумерного процесса $\xi(t)$ из области Ω :

$$\bar{\tau}_{\xi}(\Omega) = \frac{1 - P\{\xi(t) \in \Omega\}}{(1/T)N_{\xi}(\Omega, T)} = \frac{1 - T_{\xi}(\Omega)}{(1/T)N_{\xi}(\Omega, T)}. \quad (4)$$

Поскольку компоненты $\xi_1(t)$ и $\xi_2(t)$ в общем случае могут быть взаимно зависимыми, то для вычисления характеристик (2)–(4) необходимо знать совместную плотность вероятностей $p(\xi_1, \xi_2, \xi'_1, \xi'_2)$ для значений компонент $\xi_1(t), \xi_2(t)$ и значений их первых производных $\xi'_1(t)$ и $\xi'_2(t)$ в один и тот же момент времени.

Если $\xi_1(t)$ и $\xi_2(t)$ обладают свойством стационарности и взаимной независимости, то для входящих в формулы (2)–(4) вероятностей можно записать

$$P\{\xi_i(t) \in [-H_i, H_i]\} = \int_{-H_i}^{H_i} p(\xi_i; t) d\xi_i, \quad i = 1, 2;$$

$$P\{\xi_1(t) \in [-H_1, H_1], \xi_2(t) \in [-H_2, H_2]\} =$$

$$= \int_{-H_1}^{H_1} p(\xi_1; t) d\xi_1 \int_{-H_2}^{H_2} p(\xi_2; t) d\xi_2. \quad (5)$$

Для определения среднего числа выбросов, согласно результатам [13–15], при стационарности компонент будут справедливы общие выражения

$$N_i^+(H_i, T) = T \int_0^{\infty} \xi'_i p(H_i, \xi'_i; t) d\xi'_i;$$

$$N_i^-(-H_i, T) = -T \int_{-\infty}^0 \xi'_i p(-H_i, \xi'_i; t) d\xi'_i, \quad i = 1, 2. \quad (6)$$

В приведенных формулах (5) и (6) значения $p(\xi_i; t)$ соответствуют одномерной плотности вероятностей процесса $\xi_i(t)$, $i = 1, 2$, а функции $p(\pm H_i, \xi'_i; t)$ — совместным плотностям вероятностей $p(\xi_i, \xi'_i; t)$ для значений рассматриваемой компоненты $\xi_i(t)$ и значений ее производной $\xi'_i(t)$ в один и тот же момент времени при условии $\xi_i(t) = \pm H_i$, $i = 1, 2$.

Результаты (1)–(6) показывают возможность общего исследования вероятностной структуры двумерных векторных случайных процессов $\{\xi_1(t), \xi_2(t)\}$ на фазовой плоскости (ξ_1, ξ_2) . Они позволяют определять вероятностные характеристики, связанные с выходами траекторий отдельных компонент $\xi_1(t), \xi_2(t)$ и в целом рассматриваемых процессов $\xi(t) = \{\xi_1(t), \xi_2(t)\}$ за границы заданной области Ω .

Анализ вероятностной модели «сигнал плюс шум»

Рассмотрим теперь возможности анализа векторного случайного процесса $\xi(t) = \{\xi_1(t), \xi_2(t)\}$ при несколько иной постановке исходных задач. Выделим при этом особенности исследования одной из наиболее распространенных на практике модели — вероятностной модели аддитивной смеси гармонического колебания и флюктуационного гауссова шума.

Будем считать, что компоненты $\xi_1(t)$ и $\xi_2(t)$ векторного процесса $\xi(t)$ представляют собой ортогональные гармонические функции вида $A_0 \cos \omega_0 t$

и $A_0 \sin \omega_0 t$, изменяющиеся с некоторой частотой ω_0 и имеющие постоянную амплитуду A_0 . Будем, кроме того, считать, что на эти гармонические функции накладывается аддитивный флюктуационный гауссов шум $n(t)$ с математическим ожиданием $m_n = M\{n(t)\} = 0$ и некоторой дисперсией $\sigma_n^2 = M\{[n(t) - m_n]^2\} = M\{n^2(t)\}$.

Из общей теории колебаний хорошо известно, что при переходе к фазовой плоскости (x, y) гармонические компоненты $\cos \omega_0 t$ и $\sin \omega_0 t$ будут формировать в фазовом пространстве состояний (ξ_1, ξ_2) окружность с единичным радиусом. Воздействия на гармонические составляющие помеховой компоненты $n(t)$ приводят к искажениям формы фазовых траекторий $\{\xi_1(t), \xi_2(t), t \in T\}$. Степень таких искажений зависит от величины дисперсии шума или, при иной трактовке, от отношения сигнал/шум.

Характерный вид отдельных составляющих $\xi_1(t)$ и $\xi_2(t)$ векторного процесса $\xi(t) = \{\xi_1(t), \xi_2(t)\}$ и отображение траекторий $(\xi_1(t), \xi_2(t)), t \in T$ такого процесса $\xi(t)$ на фазовой плоскости (ξ_1, ξ_2) показаны на рис. 2 и 3. Для наглядности сравнения представленные результаты получены при одинаковых отношениях сигнал/шум, и основными отличиями является здесь спектральный состав

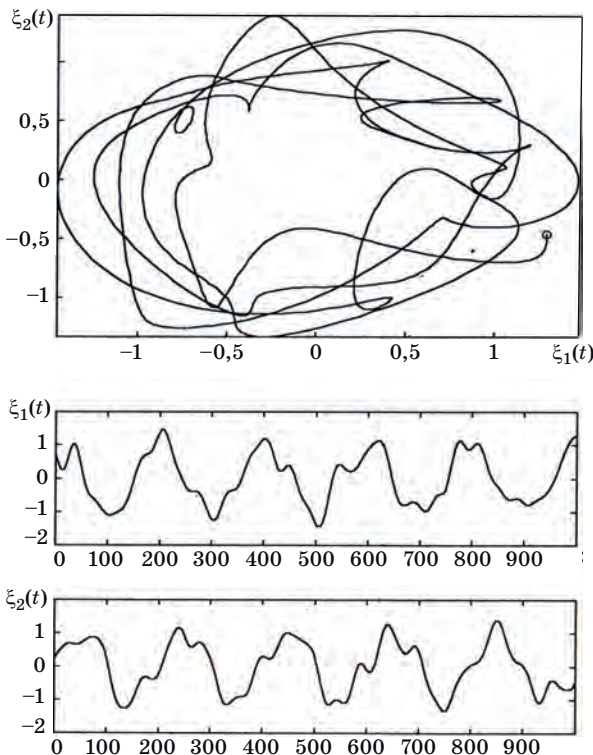
помеховых воздействий $n(t)$. Флюктуации $n(t)$ на рис. 2 относятся к классу «медленных» и узкополосных. На рис. 3 показан характер воздействий «быстрых» флюктуационных помех $n(t)$ с более широкой спектральной плотностью.

Необходимость исследования детальной вероятностной структуры подобного класса процессов $\xi(t) = \{\xi_1(t), \xi_2(t)\}$ возникает во многих областях физики и техники. Характерными примерами могут здесь являться задачи приема и обработки радиолокационных, гидроакустических и сейсмических сигналов [6, 8, 16]. Такие задачи характерны для исследований работы автогенераторов при учете собственных флюктуационных шумов [6, 17], для анализа каналов передачи информации [1, 18], вероятностного анализа систем наведения и целеуказания [19–21], анализа поляризационных эффектов и рассмотрения эллиптической поляризации электромагнитных излучений [14, 22].

При рассмотрении особенностей векторного процесса $\xi(t) = \{\xi_1(t), \xi_2(t)\}$ на фазовой плоскости (ξ_1, ξ_2) зададим область допустимых значений Ω уравнением

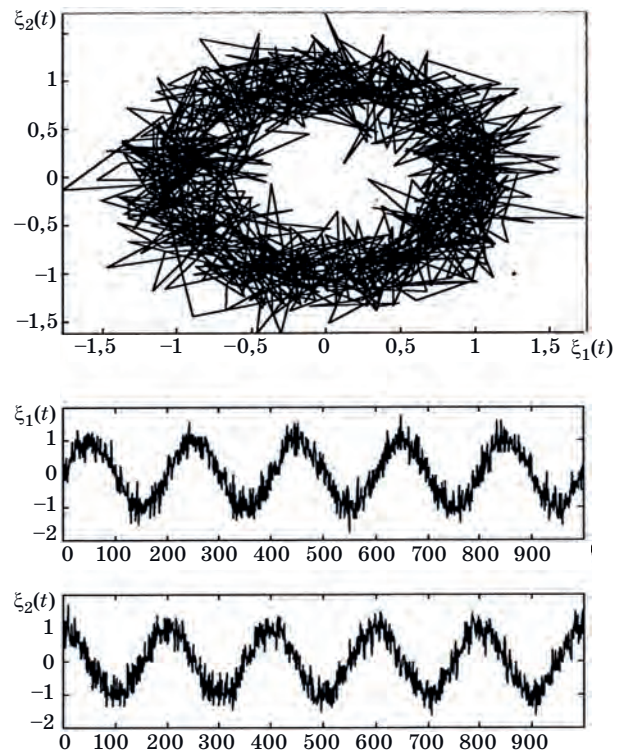
$$\Omega: (\xi_1^2/\sigma_1^2) + (\xi_2^2/\sigma_2^2) = H^2, H > 0. \quad (7)$$

В отличие от условий (1) такое определение показывает, что границы заданной области Ω



■ **Рис. 2.** Фазовые траектории и отдельные компоненты двумерного векторного процесса при воздействии «медленных» флюктуаций

■ **Fig. 2.** Phase trajectories and individual components of a two-dimensional vector process under the influence of “slow” fluctuations



■ **Рис. 3.** Влияние «быстрых» широкополосных флюктуаций на структуру векторного процесса

■ **Fig. 3.** Effect of “fast” broadband fluctuations on the structure vector process

в общем случае имеют форму эллипса или, при одинаковых слагаемых в уравнении (7), форму окружности.

В подобных задачах для вероятностного анализа фазовых траекторий $(\xi_1(t), \xi_2(t)) = L(\xi_1, \xi_2; t)$ удобно дополнительно ввести определение случайного процесса

$$\eta(t) = \left[\frac{\xi_1^2(t)}{\sigma_1^2} + \frac{\xi_2^2(t)}{\sigma_2^2} \right]^{\frac{1}{2}}, \quad \eta(t) \geq 0. \quad (8)$$

Из этого определения видно, что процесс $\eta(t)$ отражает поведение модуля двумерного вектора $\xi(t) = \{\xi_1(t), \xi_2(t)\}$ с нормированными компонентами $\xi_1(t)/\sigma_1$ и $\xi_2(t)/\sigma_2$. При прохождении фазовой траектории $L(\xi_1, \xi_2; t)$ через границу заданной области (7) для функции $\eta(t)$ выполняется равенство $\eta(t) = H$, и, следовательно, выходы векторного процесса $\xi(t)$ из области допустимых значений Ω эквивалентны выходам случайного процесса $\eta(t)$ за заданный уровень H .

С учетом такой взаимосвязи можно определить общее выражение для вероятности нахождения исследуемого векторного процесса $\xi(t)$ в области допустимых значений Ω :

$$\begin{aligned} P\{\xi(t) \in \Omega\} &= P\{\eta(t) \leq H\} = \\ &= \int_0^H p_\eta(\eta) d\eta = F_\eta(H). \end{aligned} \quad (9)$$

Вероятность выхода случайного процесса $\xi(t)$ из области Ω совпадает при этом с вероятностью события $\eta(t) > H$ и определяется как

$$\begin{aligned} P\{\xi(t) \notin \Omega\} &= P\{\eta(t) \geq H\} = \\ &= \int_H^\infty p_\eta(\eta) d\eta = 1 - F_\eta(H). \end{aligned} \quad (10)$$

Записанные результаты (9) и (10) позволяют достаточно просто вычислять соответствующие вероятности $P\{\xi(t) \in \Omega\}$ и $P\{\xi(t) \notin \Omega\}$ при условии известной плотности вероятностей $p_\eta(\eta)$ или известной функции распределения $F_\eta(\eta)$ модуля $\eta(t)$ исследуемого процесса $\xi(t) = \{\xi_1(t), \xi_2(t)\}$. При рассмотрении общего подхода к решению подобных задач нахождение функций $p_\eta(\eta)$ и $F_\eta(\eta)$ должно выполняться по совместной плотности вероятностей $p_\xi(\xi_1, \xi_2; t)$ исходного процесса $\xi(t) = \{\xi_1(t), \xi_2(t)\}$ и введенного определения (8) для процесса $\eta(t)$.

Применительно к рассматриваемой модели векторного процесса $\xi(t)$, порожденного аддитивной смесью $s(t) + n(t)$ гармонического колебания $s(t)$ и независимого гауссова шума $n(t)$, известно, что случайный процесс $\eta(t)$ будет характеризоваться обобщенным распределением Рэлея или распределением Райса [5, 6, 23]. Плотность веро-

яностей $p_\eta(\eta)$ процесса $\eta(t)$ будет при этом иметь вид

$$\begin{aligned} p_\eta(\eta) &= \frac{\eta}{\sigma_n^2} \exp\left(-\frac{\eta^2 + A_0^2}{2\sigma_n^2}\right) I_0\left(\frac{A_0\eta}{\sigma_n^2}\right), \\ \eta &= \eta(t) \geq 0, \end{aligned} \quad (11)$$

где $I_0(x)$ — модифицированная функция Бесселя нулевого порядка, а величина A_0/σ_n характеризует отношение сигнал/шум.

Полезно здесь еще раз подчеркнуть, что записанное распределение (11) по своей сути характеризует вероятностное поведение длины двумерного случайного вектора $\eta(t)$, составляющие которого $\xi_1(t)$ и $\xi_2(t)$ ортогональны и представляют собой квадратурные компоненты (см. рис. 2 и 3) смеси $s(t) + n(t)$ гармонического колебания $s(t)$ и флюктуационного гауссова шума $n(t)$.

Если теперь для исследуемого векторного процесса $\xi(t) = \{\xi_1(t), \xi_2(t)\}$ рассмотреть события, связанные с выходами фазовых траекторий $L(\xi_1, \xi_2; t)$ за границы (7) области допустимых значений Ω , то среднее число таких выходов $N_L(\Omega, T)$ на интервале $t \in [t_0, t_0 + T]$ длительностью T можно характеризовать средним числом положительных выбросов $N_\eta^+(H, T)$ процесса $\eta(t)$ за заданный уровень H .

По аналогии с выражениями (6) здесь можно воспользоваться общими результатами [13–15] и записать

$$N_L(\Omega, T) = N_\eta^+(H, T) = T \int_0^\infty \eta' p(H, \eta'; t) d\eta', \quad (12)$$

где $p(H, \eta'; t)$ — совместная плотность вероятностей $p(\eta, \eta'; t)$ для значений процесса $\eta(t)$ и значений его производной $\eta'(t)$ в совпадающие моменты времени t при условии $\eta(t) = H$.

Следовательно, нахождение среднего числа выходов $N_L(\Omega, T)$ связано здесь с предварительным определением совместной плотности вероятностей $p(\eta, \eta'; t)$. Для процессов $\eta(t)$ с распределением Райса (11) эта функция имеет вид [13]

$$\begin{aligned} p(\eta, \eta'; t) &= \frac{\eta}{\sigma_n^2} \exp\left(-\frac{\eta^2 + A_0^2}{2\sigma_n^2}\right) I_0\left(\frac{A_0\eta}{\sigma_n^2}\right) \times \\ &\times \frac{1}{\sqrt{2\pi D[\eta']}} \exp\left\{-\frac{(\eta')^2}{2D[\eta']}\right\} = p_\eta(\eta) p_{\eta'}(\eta'), \end{aligned} \quad (13)$$

где $p_\eta(\eta)$ — одномерная плотность вероятностей (11) случайного процесса $\eta(t)$, а $p_{\eta'}(\eta')$ — гауссова плотность вероятностей производной $\eta'(t)$. Записанное выражение (13) показывает, что рассматриваемый процесс $\eta(t)$ относится к классу случайных процессов со статистически независимой производной $\eta'(t)$ в совпадающие момен-

ты времени. Кроме того, важно подчеркнуть, что хотя сам процесс $\eta(t)$ характеризуется негауссовым распределением (11), его производная $\eta'(t)$ описывается гауссовой одномерной плотностью вероятностей с математическим ожиданием $m_{\eta'} = M\{\eta'(t)\} = 0$ и дисперсией $D[\eta'] = M\{[\eta'(t)]^2\}$.

Если теперь распределение (13) подставить в общую формулу (12) и выполнить простое интегрирование, то получим

$$N_L(\Omega, T) = N_{\eta}^+(H, T) = \frac{T}{\sqrt{2\pi}} D^{\frac{1}{2}}[\eta'] \left(\frac{H}{\sigma_n} \right) \exp\left(-\frac{H^2}{2\sigma_n^2} - \frac{a^2}{2} \right) I_0\left(a \frac{H}{\sigma_n} \right),$$

$$a = A_0/\sigma_n, H \geq 0. \quad (14)$$

Этот результат является достаточно общим для модели «сигнал плюс шум» и позволяет при различных значениях параметра $a = A_0/\sigma_n$ находить среднее число выходов $N_L(\Omega, T)$ фазовой траектории $L(\xi_1, \xi_2; t)$ за границы (7) заданной области Ω . В целом результаты (7)–(14) данного раздела показывают возможности анализа вероятностной структуры векторного случайного процесса $\xi(t) = \{\xi_1(t), \xi_2(t)\}$ при его представлении на фазовой плоскости (ξ_1, ξ_2) и задании области допустимых значений Ω в виде эллипса (7).

Структура двумерных гауссовых процессов

В целях конкретизации общих выводов и более детального анализа фазовых траекторий $L(\xi_1, \xi_2; t)$ рассмотрим теперь практически важный частный случай векторного процесса $\xi(t) = \{\xi_1(t), \xi_2(t)\}$ с ортогональными гауссовыми компонентами $\xi_1(t)$ и $\xi_2(t)$. Вероятностные модели подобных процессов $\xi(t)$ настолько важны для приложений, что заслуживают специального, отдельного рассмотрения и с точки зрения анализа фазовых траекторий.

Будем считать, что компоненты $\xi_1(t)$ и $\xi_2(t)$ векторного процесса $\xi(t) = \{\xi_1(t), \xi_2(t)\}$ имеют одинаковые вероятностные характеристики и описываются гауссовым распределением с математическим ожиданием $m_{\xi_i} = m_{\xi_2} = M\{\xi_i(t)\} = 0$, дисперсией $\sigma^2 = M\{\xi_i^2(t)\}$ и некоторой корреляционной функцией общего вида

$$R_i(\tau) = M\{\xi_i(t)\xi_i(t + \tau)\} = R(\tau) = \sigma^2 r(\tau), i = 1, 2. \quad (15)$$

При этих условиях модуль $\eta(t)$ двумерного векторного процесса $\xi(t) = \{\xi_1(t), \xi_2(t)\}$ будет характеризоваться распределением Рэлея [5, 23]

$$\eta(t) = [\xi_1^2(t) + \xi_2^2(t)]^{1/2};$$

$$p_{\eta}(\eta) = \eta \sigma^{-2} \exp(-\eta^2 / 2\sigma^2), \eta = \eta(t) \in [0, \infty). \quad (16)$$

Такой результат можно рассматривать как описание χ -процесса с двумя степенями свободы [13, 23] или как частный случай распределения (11), когда параметр $a = A_0/\sigma_n = 0$ и функция $I_0(x)$ принимает значение $I_0(x) = I_0(0) = 1$.

Если при анализе векторного процесса $\xi(t) = \{\xi_1(t), \xi_2(t)\}$ область допустимых значений Ω рассматривать в виде эллипса (7), то с учетом плотности вероятностей (16), на основе результатов (9) и (10) можно определить вероятность нахождения исследуемого процесса $\xi(t)$ в заданной области Ω

$$P\{\xi(t) \in \Omega\} = P\{\eta(t) \leq H\} = 1 - \exp(-H^2 / 2\sigma^2) \quad (17)$$

и вероятность выхода векторного процесса $\xi(t)$ за границы области допустимых значений Ω

$$P\{\xi(t) \notin \Omega\} = P\{\eta(t) > H\} = \exp(-H^2 / 2\sigma^2). \quad (18)$$

Такие результаты наглядно показывают, что вероятности (17) и (18) зависят в данном случае лишь от порогового уровня H и дисперсии $\sigma^2 = \sigma_{\xi_1}^2 = \sigma_{\xi_2}^2$ гауссовых компонент $\xi_1(t)$ и $\xi_2(t)$.

При анализе особенностей детальной вероятностной структуры случайных процессов важную роль играет поведение производной и характер совместного распределения для значений исследуемого процесса и значений его производной в совпадающие моменты времени [13, 24]. Рассматривая процессы с распределением Рэлея (16), на основе выражения (13) можно определить совместную плотность вероятностей для значений $\eta(t)$ и $\eta'(t) = d\eta(t)/dt$:

$$p(\eta, \eta'; t) = \frac{\eta}{\sigma^2 \sqrt{2\pi(-R_0'')}} \exp\left\{ -\frac{\eta^2}{2\sigma^2} - \frac{(\eta')^2}{2(-R_0'')} \right\} = p_{\eta}(\eta) \frac{1}{\sqrt{2\pi(-R_0'')}} \exp\left\{ -\frac{(\eta')^2}{2(-R_0'')} \right\}, \eta = \eta(t) \geq 0. \quad (19)$$

В этом выражении функция $p_{\eta}(\eta)$ соответствует плотности вероятностей Рэлея (16), а значение $(-R_0'')$ характеризует дисперсию производной $\eta'(t)$ и определяется по корреляционной функции (15) как

$$-R_0'' = -d^2R(\tau)/d\tau^2|_{\tau=0} = -\sigma^2 r_0''. \quad (20)$$

Случайная функция $\eta(t)$ относится к классу процессов, порожденных исходным векторным процессом $\xi(t) = \{\xi_1(t), \xi_2(t)\}$ с гауссовыми компонентами $\xi_1(t)$ и $\xi_2(t)$. На рис. 4 для процесса $\eta(t)$ показан характер фазовой траектории $L(\eta, \eta'; t)$ в пространстве состояний вида (η, η') . По своей

сути подобные траектории $L(\eta, \eta'; t)$ отражают поведение выборочной функции двумерного векторного процесса $\{\eta(t), \eta'(t)\}$ на фазовой плоскости классического вида (η, η') .

Поведение фазовых траекторий (см. рис. 4) хорошо согласуется с характером распределения (19). Здесь наглядно видно, что ориентация выборочных функций $L(\eta, \eta'; t)$ относительно координатных осей на плоскости (η, η') отражает свойство некоррелированности значений $\eta(t)$ и $\eta'(t)$ в совпадающие моменты времени t и свойство нулевого математического ожидания $M\{\eta'(t)\} = 0$. Эти свойства характерны для всего класса стационарных процессов [14, 24]. В дополнение к этому совместное распределение (19) показывает, что значения $\eta(t)$ и $\eta'(t)$ для вероятностной модели (16) не только некоррелированные, но и взаимно независимы.

Нужно отметить, что поведение фазовых траекторий $L(\eta, \eta'; t)$ на рис. 4 наглядно отражает особенности одномерных распределений отдельных компонент векторного процесса $\{\eta(t), \eta'(t)\}$. Так, в частности, общие изменения значений $\eta(t)$ по горизонтальной оси хорошо согласуются с особенностями функции $p_\eta(\eta)$ плотности вероятностей Рэлея (16). Изменения значений второй компоненты $\eta'(t)$ по вертикальной оси фазовой плоскости (η, η') отражают характер гауссовой составляющей $\eta'(t)$ векторного процесса $\{\eta(t), \eta'(t)\}$.

Рассмотрим теперь вероятностные характеристики, связанные с выходами исследуемого двумерного гауссова процесса $\xi(t) = \{\xi_1(t), \xi_2(t)\}$ за границы (7) области допустимых значений Ω . Для нахождения среднего числа выходов траектории $L(\xi_1, \xi_2; t)$ из области Ω на интервале времени $t \in [t_0, t_0 + T]$ здесь можно воспользоваться общими формулами (12) и (19) или из выражения

(14), как частный случай, при условии $a = 0$ получить простой результат

$$N_L(\Omega, T) = N_\eta^+(H, T) = T \left(\frac{-r_0''}{2\pi} \right)^{\frac{1}{2}} \frac{H}{\sigma} \exp\left(-\frac{H^2}{2\sigma^2} \right), \quad H \geq 0. \quad (21)$$

Если в дополнение к этому учесть значение вероятности выхода (18), то можно получить простое выражение для средней длительности $\tau^+(\Omega)$ выходов векторного процесса $\xi(t)$ за границы заданной области Ω

$$\tau^+(\Omega) = \tau_\eta^+(H) = \frac{P\{\xi(t) \notin \Omega\}}{N_L(\Omega, 1)} = \left(\frac{2\pi}{-r_0''} \right)^{\frac{1}{2}} \left(\frac{\sigma}{H} \right). \quad (22)$$

Параметр $(-r_0'')$, входящий в формулы (20)–(22), определяется по нормированной корреляционной функции $r(\tau)$ и непосредственно связан со спектрально-корреляционными свойствами гауссовых компонент $\xi_1(t)$ и $\xi_2(t)$. Во многих практических задачах этот параметр удобно представить в виде [13, 14]

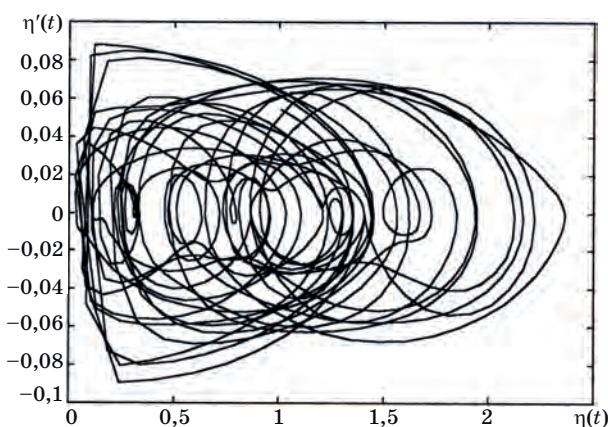
$$-r_0'' = -d^2 r(\tau) / d\tau^2 |_{\tau=0} = \kappa^2 \Delta\omega_\Omega^2, \quad (23)$$

где κ — коэффициент формы спектральной плотности гауссовых компонент исследуемого векторного процесса $\xi(t)$; $\Delta\omega_\Omega$ — эффективная ширина спектра. Такое соотношение позволяет оценивать влияние спектральных и корреляционных характеристик на поведение выборочных функций $L(\xi_1, \xi_2; t)$ в фазовом пространстве (ξ_1, ξ_2) .

В зависимости от содержания решаемых задач и требуемой детальности исследований на основе представленных результатов (15)–(23) могут рассматриваться вероятности или относительные длительности нахождения случайных процессов в заданной области, вероятности выходов фазовых траекторий за границы области допустимых значений, среднее число таких выходов и средние длительности выходов.

Заключение

Основной отличительной особенностью данной работы является рассмотрение и анализ детальной вероятностной структуры двумерных векторных случайных процессов на основе фазовых отображений — представлений случайных процессов в фазовом пространстве состояний. По существу, здесь объединяются общие методы фазового пространства векторных процессов, методы фазовой плоскости при анализе скалярных процессов и общие методы теории выбросов случайных функций. Такое объединение позволяет



■ **Рис. 4.** Характер фазовых траекторий случайного процесса с распределением Рэлея

■ **Fig. 4.** Character of phase trajectories of random process with distribution Rayleigh's

получать наглядное представление информации о вероятностной структуре двумерных процессов, накапливать данные на фазовой плоскости при увеличении длительности выборочных функций и наглядно отображать вероятностные зависимости при исследованиях векторных процессов «в целом» на рассматриваемом интервале наблюдений.

Представленные в работе результаты не только показывают особенности визуального, качественного рассмотрения, но и дают основу для количественного описания и анализа вероят-

ностных свойств векторных случайных процессов при их представлении на фазовой плоскости. Рассмотренные здесь вероятностные модели и результаты их анализа являются достаточно общими и могут быть полезны при решении многих, различных по своему содержанию, прикладных задач. Типовыми примерами являются при этом задачи теории надежности и теории рисков, задачи вероятностного анализа систем пространственно-временного сканирования, систем целеуказания, наведения, автоматического сопровождения и управления.

Литература

1. **Boccuzzi J.** *Signal processing for wireless communications*. N. Y., McGraw-Hill, 2008. 672 p.
2. **Simiu E.** *Chaotic transitions in deterministic and stochastic dynamical systems*. Princeton and Oxford, Princeton University Press, 2002. 208 p.
3. **Гнеденко Б. В., Беляев Ю. К., Соловьев А. Д.** *Математические методы в теории надежности*. М., URSS, 2017. 584 с.
4. *Proceedings of the International Scientific School Modelling and Analysis of Safety and Risk in Complex Systems (MASR-2016)*. Saint-Petersburg, 2016. 212 p.
5. **Ахманов С. А., Дьяков Ю. Е., Чиркин А. С.** *Статистическая радиофизика и оптика*. М., Физматлит, 2010. 428 с.
6. **Шахтарин Б. И.** *Случайные процессы в радиотехнике. Нелинейные преобразования*. М., Горячая линия-Телеком, 2018. 376 с.
7. **Moore F.** *Chaotic vibrations*. N. Y., Wiley, 1992. 312 p.
8. **Гордиенко В. А.** *Векторно-фазовые методы в акустике*. М., Физматлит, 2007. 480 с.
9. **Махутов Н. А., Гаденин М. М.** Научные исследования и подготовка специалистов по обеспечению защищенности критически важных объектов. *Машиностроение и инженерное образование*, 2004, № 1, с. 19–32.
10. **Абрамов О. В.** Анализ и прогнозирование техногенных рисков. *Надежность и техническая диагностика*, 2012, № 3, с. 97–105.
11. *Методы и приборы космической кардиологии на борту Международной космической станции* / под ред. Р. М. Баевского, О. И. Орлова. М., Техносфера, 2016. 368 с.
12. **Светлицкий В. А.** *Статистическая механика и теория надежности*. М., Изд-во МГТУ им. Н. Э. Баумана, 2004. 504 с.
13. **Тихонов В. И., Хименко В. И.** *Выбросы траекторий случайных процессов*. М., Наука, 1987. 304 с.
14. **Хименко В. И.** *Случайные данные: структура и анализ*. М., Техносфера, 2017. 424 с.
15. **Tikhonov V. I., Khimenko V. I.** Level-crossing problems for stochastic processes in physics and radio engineering: An overview. *Journal of Communications Technology and Electronics*, 1998, vol. 43, no. 5, pp. 501–523.
16. **Skolnik M. I.** *Radar handbook*. N. Y., McGraw-Hill, 2008. 680 p.
17. **Малахов А. Н.** *Флуктуации в автоколебательных системах*. М., Наука, 1967. 660 с.
18. **Яковлев О. И.** *Распространение радиоволн*. М., URSS, 2017. 496 с.
19. **Лысенко Л. Н.** *Наведение и навигация баллистических ракет*. М., Изд-во МГТУ им. Н. Э. Баумана, 2007. 672 с.
20. **Kolobov M. I.** *Quantum imaging*. N. Y., Springer, 2007. 328 p.
21. **Афонин А. А., Никитин В. Н.** *Автоматические системы наведения лазерных пучков*. М., Физматлит, 2016. 279 с.
22. **Saleh B., Teich M.** *Fundamentals of photonics*. N. Y., Wiley, 2007. 760 p.
23. **Тихонов В. И.** *Статистическая радиотехника*. М., Радио и связь, 1982. 624 с.
24. **Хименко В. И.** Вероятностная структура фазовых траекторий случайных процессов. *Информационно-управляющие системы*, 2016, № 3, с. 85–94. doi:10.15217/issn 1684-8853.2016.3.85

UDC 519.21:537.86

doi:10.31799/1684-8853-2019-2-10-18

Probabilistic analysis of two-dimensional vector random processes on the phase plane

V. I. Khimenko^a, Dr. Sc., Tech., Professor, orcid.org/0000-0003-3005-052X, vih.13@yandex.ru

^aJSC «The Scientific Research and Experimental Centre of Intelligent Technologies «Petrokometa», 67, B. Morskaja St., 190000, Saint-Petersburg, Russian Federation

Introduction: The models of vector random processes are characterized by a variety of practical applications and, at the same time, by the complexity of studying the detailed probabilistic structure. Such features often lead to considering only separate components of vector functions, forcing the shift to scalar process analysis. All this significantly reduces the overall information content of the research, increasing the relevance of the search for promising approaches to probabilistic analysis of vector processes. **Purpose:** To present vector random processes in a phase space of states, and to apply the general theory of level crossings of random functions to the study of probabilistic structure of phase trajectories. **Results:** We have identified the characteristic features of displaying random processes on a phase plane, and introduced numerical characteristics for the description and analysis of probabilistic behavior of sample functions. A probabilistic analysis has been performed for the characteristics of the «level crossing» type for vector processes with various definitions of the areas of acceptable values. We used the typical models of two-dimensional Gaussian processes and «signal plus noise» models to demonstrate how the probabilistic structure of the phase trajectories depends on the given threshold levels, major distribution parameters and spectral-correlation properties of the studied processes. **Practical relevance:** The article contributes to the research information content and its visualization in the analysis of probabilistic behavior of vector random processes. It combines general methods of phase space of vector processes, methods of phase plane in the analysis of scalar processes and the theory of level crossings of random functions. The classical approaches to the visual description of phase trajectories are supplemented by the methods of quantitative analysis of detailed probabilistic structure of random functions.

Keywords — stochastic processes, phase space, phase trajectory, probabilistic analysis, random functions, characteristics of intersections of levels of processing information.

For citation: Khimenko V. I. Probabilistic analysis of two-dimensional vector random processes on the phase plane. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2019, no. 2, pp. 10–18 (In Russian). doi:10.31799/1684-8853-2019-2-10-18

References

1. Bocuzzi J. *Signal processing for wireless communications*. New York, McGraw-Hill, 2008. 672 p.
2. Simiu E. *Chaotic transitions in deterministic and stochastic dynamical systems*. Princeton and Oxford, Princeton University Press, 2002. 208 p.
3. Gnedenko B. V., Belyaev Y. K., Soloviev A. D. *Matematicheskie metody v teorii nadezhnosti* [Mathematical methods in the theory of reliability]. Moscow, URSS Publ., 2017. 584 p. (In Russian).
4. *Proceedings of the International Scientific School Modeling and Analysis of Safety and Risk in Complex Systems (MA SR-2016)*. Saint-Petersburg, 2016. 212 p.
5. Akhmanov S. A., Dyakov Yu. E., Charkin A. S. *Statisticheskaja radiofizika i optika* [Statistical radiophysics and optics]. Moscow, Fizmatlit Publ., 2010. 428 p. (In Russian).
6. Shakhhtar B. I. *Slychaynye processy v radiotekhnike. Nelineynye preobrazovaniy* [Random processes in radio engineering. Nonlinear transformation]. Moscow, Goriachaia liniia-Telekom Publ., 2018. 376 p. (In Russian).
7. Moon F. *Chaotic vibrations*. New York, Wiley, 1992. 312 p.
8. Gordienko V. A. *Vektorno-fazovye metody v akustike* [Vector-phase methods in acoustics]. Moscow, Fizmatlit Publ., 2007. 480 p. (In Russian).
9. Makhutov N. A., Gadenin M. M. Research and training of specialists to ensure the security of critical facilities. *Mashinostroenie i inzhenernoe obrazovanie*, 2004, no. 1, pp. 19–32 (In Russian).
10. Abramov O. V. Analysis and forecasting of technogenic risks. *Nadezhnost' i tekhnicheskaya diagnostika* [Reliability and Technical Diagnostics], 2012, no. 3, pp. 97–105 (In Russian).
11. *Metody i pribory kosmicheskoy kardiologii na bortu Mezhunarodnoj kosmicheskoy stantsii* [Methods and devices in space cardiology on board the International space station]/ editors R. M. Baevsky, O. I. Orlov. Moscow, Technosfera Publ., 2016. 368 p. (In Russian).
12. Svetlitskiy V. A. *Statisticheskaja mekhanika i teoriia nadezhnosti* [Statistical mechanics and the theory of reliability]. Moscow, Moskovskij gosudarstvennyj tekhnicheskij universitet im. N. E.H. Baumana Publ., 2004. 504 p. (In Russian).
13. Tikhonov V. I., Khimenko V. I. *Vybrosy traektorii sluchainykh protsessov* [Trajectories emissions of stochastic processes]. Moscow, Nauka Publ., 1987. 304 p. (In Russian).
14. Khimenko V. I. *Sluchainye dannye: struktura i analisis* [Random data: structure and analysis]. Moscow, Technosfera Publ., 2017. 424 p. (In Russian).
15. Tikhonov V. I., Khimenko V. I. Level-crossing problems for stochastic processes in physics and radio engineering: An overview. *Journal of Communications Technology and Electronics*, 1998, vol. 43, no. 5, pp. 501–523.
16. Skolnik M. I. *Radar handbook*. New York, McGraw-Hill, 2008. 680 p.
17. Malakhov A. N. *Flyktyatsii v avtokolebatelnykh sistemakh* [Fluctuations in self-oscillatory systems]. Moscow, Nauka Publ., 1967. 660 p. (In Russian).
18. Yakovlev O. I. *Rasprostranenie radiovoln* [Propagation of radio waves]. Moscow, URSS Publ., 2017. 496 p. (In Russian).
19. Lysenko L. N. *Navedenie i navigatsiia ballisticheskikh raket* [Guidance and navigation of ballistic missiles]. Moscow, Moskovskij gosudarstvennyj tekhnicheskij universitet im. N. E.H. Baumana Publ., 2007. 672 p. (In Russian).
20. Kolobov M. I. *Quantum imaging*. New York, Springer, 2007. 328 p.
21. Afonin A. A., Nikitin V. N. *Avtomaticheskije sistemy navedeniia lasernykh puchkov* [Automatic laser beam guidance systems]. Moscow, Fizmatlit Publ., 2016. 279 p. (In Russian).
22. Saleh B., Teich M. *Fundamentals of photonics*. New York, Wiley, 2007. 760 p.
23. Tikhonov V. I. *Statisticheskaja radiotekhnika* [Statistical radiotechnics]. Moscow, Radio i sviaz' Publ., 1982. 624 p. (In Russian).
24. Khimenko V. I. Probability structure of phase trajectories of random processes. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2016, no. 3, pp. 85–94 (In Russian). doi:10.15217/issn 1684-8853.2016.3.85

Команда роботов-футболистов для соревнований RoboCup в лиге SSL: система и алгоритмы

П. А. Коновалов^а, студент, orcid.org/0000-0002-4910-9184

А. В. Корнилова^б, инженер-программист, orcid.org/0000-0002-2267-9689

Д. М. Королев^в, инженер, orcid.org/0000-0002-6673-2998

Г. В. Ренева^а, студент, orcid.org/0000-0003-2322-4608

А. Л. Фрадков^{а, г}, доктор техн. наук, профессор, orcid.org/0000-0002-5633-0944, fradkov@mail.ru

И. Ю. Широколов^{б, д}, инженер, orcid.org/0000-0003-4346-5433

Д. С. Ярош^а, студент, orcid.org/0000-0002-2017-8979

^аСанкт-Петербургский государственный университет, Университетская наб., 7–9, Санкт-Петербург, 199034, РФ

^бООО «Кибертех», Железнодорожная ул., 68, кв. 26, Пушкин, Санкт-Петербург, 196606, РФ

^вЦентральный научно-исследовательский и опытно-конструкторский институт робототехники и технической кибернетики, Тихорецкий пр., 21, Санкт-Петербург, 194064, РФ

^гИнститут проблем машиноведения РАН, Большой пр. В. О., 61, Санкт-Петербург, 199178, РФ

^дПрезидентский физико-математический лицей № 239, Кирочная ул., 8, Санкт-Петербург, 191028, РФ

Введение: международные соревнования по футболу роботов RoboCup проводятся с середины 1990-х годов и привлекают внимание университетов всего мира. Их целью является развитие научных и инженерных исследований в областях робототехники, искусственного интеллекта, технического зрения, навигации, группового взаимодействия роботов и мехатронных устройств. **Цель:** создание исследовательского стенда нового поколения, позволяющего разрабатывать, тренировать и оценивать роботов-футболистов для подготовки команд, а также разработка базовых алгоритмов для управления как отдельным роботом-футболистом, так и группой таких роботов. **Результаты:** разработан стенд для создания и исследования алгоритмов управления группой роботов и навигации в высокодинамичной среде при централизованном управлении, организуемом следующим образом: вычислительный центр получает данные от системы видеонаблюдения, использует полученные сведения о положении роботов и мяча и информацию о состоянии их датчиков для вычисления управляющих команд (вектора скорости, угловой скорости, сигнала об ударе), которые затем передаются роботам для исполнения. Базовые алгоритмы управления для роботов-футболистов, такие как движение в точку, поворот на точку, движение по окружности, движение с объездом препятствий, прием паса, поведение вратаря, удар в заданную точку, а также модель игры двух полевых игроков против вратаря были реализованы и использованы для тестирования работоспособности стенда, а также для записи квалификационного видео и подготовки документации при подаче заявки на участие команды Санкт-Петербургского государственного университета в финале RoboCup Small Size League 2019 года в Сиднее, Австралия.

Ключевые слова – групповое взаимодействие, футбол роботов, исследовательский стенд, управление роботами, обход препятствий.

Для цитирования: Коновалов П. А., Корнилова А. В., Королев Д. М., Ренева Г. В., Фрадков А. Л., Широколов И. Ю., Ярош Д. С. Команда роботов-футболистов для соревнований RoboCup в лиге SSL: система и алгоритмы. *Информационно-управляющие системы*, 2019, № 2, с. 19–25. doi:10.31799/1684-8853-2019-2-19-25

For citation: Konovalov P. A., Kornilova A. V., Korolev D. M., Reneva G. V., Fradkov A. L., Shirokolobov I. Yu., Yarosh D. S. A team of soccer robots for RoboCup competitions in SSL league: system and algorithms. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2019, no. 2, pp. 19–25 (In Russian). doi:10.31799/1684-8853-2019-2-19-25

Введение

Международные соревнования по футболу роботов RoboCup проводятся с середины 1990-х годов и привлекают внимание университетов всего мира [1–10]. Их целью является развитие научных и инженерных исследований в областях робототехники, искусственного интеллекта, технического зрения, навигации, группового взаимодействия роботов и мехатронных устройств. Соревнования RoboCup проходят в нескольких лигах, среди которых одной из самых доступных и в то же время зрелищных является Small

Size League (SSL). В начале 2000-х годов исследования в области моделирования виртуального футбола роботов велись и в России [2, 11, 12], но полноценного участия в соревнованиях в физических лигах российские роботы до сих пор не принимали. Следует отметить, что разработанный в 2011–2015 гг. в СПбГУ стенд RoboCup [13–16] успешно использовался для научных исследований и популяризации соревнований управляемого футбола роботов среди школьных команд [17]. В настоящей статье описываются исследовательский стенд нового поколения, позволяющий разрабатывать, тренировать и оценивать робо-

тов-футболистов, а также базовые алгоритмы для управления как отдельным роботом-футболистом, так и группой таких роботов. Разработка алгоритмов велась в соответствии с правилами международных соревнований [18].

Исследовательский стенд для футбола роботов

Стенд состоит из двух основных частей: централизованной системы видеонаблюдения для определения положения роботов и централизованной системы управления для расчета управляющих сигналов для роботов (вектора скорости, угловой скорости, сигнала об ударе).

Централизованная система видеонаблюдения SSL-Vision [19] оценивает текущие позиции роботов с помощью камеры над полем и отправляет эти данные в централизованную систему управления через сеть. Положение роботов определяется по специальной маркировке на крышке робота, которая позволяет установить номер робота, его команду, ориентацию и координаты на поле.

Централизованная система управления получает данные от централизованной системы видеонаблюдения и роботов (например, от их встроенных датчиков) и передает эти данные в систему исполнения алгоритмов MATLAB, где эта информация может использоваться для вычисления сигналов управления роботами. После расчета управляющие сигналы экспортируются из системы MATLAB и отправляются роботам.

Программное обеспечение было разработано с использованием C++ в среде Qt. Этот выбор был мотивирован кроссплатформенностью данного фреймворка. Программная реализация алгоритмов была выполнена на языке MATLAB отчасти из-за его удобства на этапе прототипирования.

Взаимодействие между всеми модулями иллюстрирует рис. 1. Соединение с сервером SSL-Vision устанавливается через модуль SSL-клиента, который распределяет полученные данные среди всех остальных модулей. Передача команд на роботов происходит посредством модуля связи с роботами. Эти команды вычисляются средствами MATLAB в модуле выполнения MATLAB-скриптов. Модуль пользовательского интерфейса (UI-модуль) отвечает за графический интерфейс пользователя, который отображает ситуацию на поле и позволяет оператору запускать некоторые алгоритмы в тестовом режиме или управлять роботами вручную.

С помощью библиотеки алгоритмов MATLAB можно анализировать ситуацию в игровых условиях и назначать текущие роли роботам на основе этого анализа. Эта библиотека также исполь-



■ Рис. 1. Взаимодействие модулей стенда
 ■ Fig. 1. Interaction of framework modules

зуется движком MATLAB для расчета управляющих сигналов для каждого робота с учетом его назначенной роли.

Роботы-футболисты

Основой робота-футболиста является четырехколесная омниплатформа, позволяющая перемещаться в любом направлении (рис. 2). Робот оснащен двумя ударными устройствами (кикерами): для прямого удара и навеса мяча. Для связи с роботом используется Wi-Fi. С торцевой сторо-



■ Рис. 2. Внешний вид робота
 ■ Fig. 2. Robot appearance

ны установлены лазерные датчики для определения мяча перед роботом. Верхняя часть роботов маркирована цветными кругами для идентификации на поле (координаты и направление) с помощью камеры и SSL-Vision. ЖК-экран установлен на материнской плате робота и отображает название текущей сети Wi-Fi и IP-адрес робота в этой сети. Робот оснащен сервисным USB-портом, интерфейсом для изменения внутренних параметров и ограничений. Каждое устройство контролируется отдельной платой, что упрощает процесс диагностики и ремонта.

Основные технические характеристики: диаметр корпуса робота 180 мм, высота 146 мм, вес 3500 г, максимальная линейная скорость 1,5 м/с, максимальная угловая скорость 3 об/с, максимальная скорость удара 3 м/с, высота навеса 250 мм.

Робот оснащен съемным аккумулятором емкостью 3000 мА·ч и средним напряжением 26 В. Время работы без подзарядки 30 мин со средним потреблением тока 6 А. Для того чтобы зарядить аккумулятор, его следует предварительно удалить из робота. Аккумулятор оснащен балансировочным устройством, а также устройством защиты от перегрева, короткого замыкания, переразряда и перезаряда. Он также имеет индикатор уровня заряда на передней панели.

Робот приводится в движение четырьмя бесколлекторными двигателями для омниколес. Реализовано векторное управление двигателем с контролем тока в обмотках и ограничением на максимальный крутящий момент на роторе. Каждый двигатель обслуживается отдельной платой с микроконтроллером. Также на лицевой части робота расположены «дрибблер» — устройство для удержания мяча перед роботом и два кикера — механизмы, выполняющие прямой удар и удар навесом.

Алгоритмы

Приведенные ниже алгоритмы апробированы во время игры по правилам RoboCup SSL.

Алгоритм движения в точку управляет модулем скорости робота при перемещении в направлении цели. Он реализован как П-регулятор, где в качестве регулируемой величины выступает расстояние между роботом и целью, которое поддерживается около нуля.

Алгоритм поворота на точку контролирует модуль угловой скорости робота при повороте на цель. Он реализован как ПД-регулятор, где в качестве регулируемой величины выступает наименьший угол между направлением робота и вектором, соединяющим позицию робота и цель, этот угол поддерживается вблизи нуля.

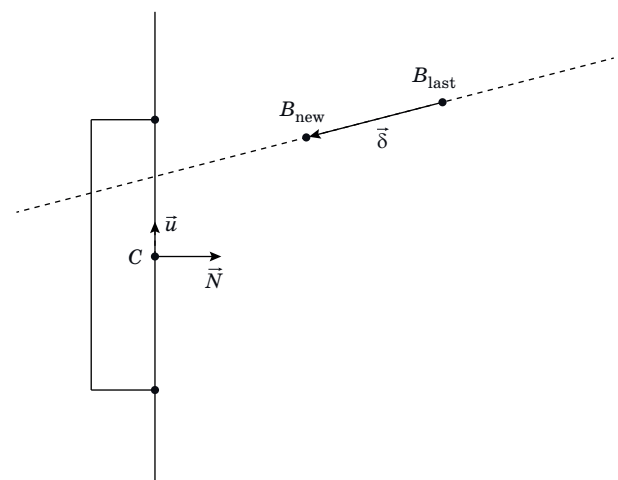
В алгоритме движения по окружности вектор скорости робота складывается из касательного к окружности вектора фиксированной длины и ортогонального ему вектора, модуль которого контролируется алгоритмом движения вокруг точки. Данный алгоритм реализован как П-регулятор, где в качестве регулируемой величины выступает расстояние между позицией робота и центром окружности, которое поддерживается около некоторого заданного значения. Угловая скорость устанавливается при помощи алгоритма поворота на точку таким образом, чтобы робот смотрел на центр окружности. Далее считая, что все перемещения выполняются при помощи этих базовых алгоритмов.

Алгоритм приема паса работает в случае, если мяч движется в сторону робота и траектория движения мяча проходит в некоторой окрестности робота. В такой ситуации робот движется в точку на траектории движения мяча и гасит скорость движения мяча отъездом назад, чтобы мяч не отскочил далеко. С использованием алгоритма построения маршрута, описанного в статье [19], был реализован алгоритм движения с объездом препятствий.

Модели поведения

Вратарь. Вратарь двигается вдоль ворот по прямой. Если прогнозируемая траектория движения мяча пересекает ворота, то робот движется к точке их пересечения, иначе стоит напротив центра ворот. Траектория мяча оценивается по двум точкам B_{new} и B_{last} (рис. 3).

Для перемещения к точке пересечения используется алгоритм движения в точку с ПД-регулятором.



■ Рис. 3. Алгоритм вратаря
■ Fig. 3. Goalkeeper algorithm

Нападающий. В этом алгоритме каждая комбинация положения робота, мяча и целевой позиции принадлежит одному состоянию из четырех возможных. В зависимости от того, какое состояние описывает текущее расположение этих объектов, робот принимает решение о дальнейших действиях.

Состояние 1: робот далеко от мяча. Это означает, что расстояние от него до мяча больше определенной константы. В этом случае робот движется к мячу, используя алгоритм движения в точку, пока не подъедет достаточно близко.

Состояние 2: робот не находится в состоянии 1 и мяч не находится на линии между роботом и целью. В этом случае используется алгоритм движения вокруг точки.

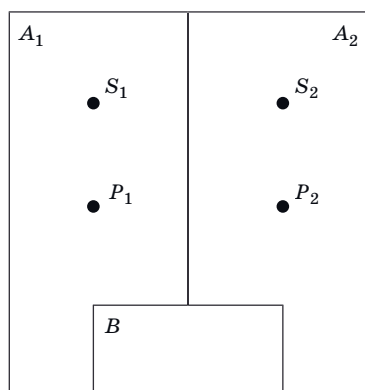
Состояние 3: робот не находится в состоянии 2, мяч находится на линии между роботом и целью, робот находится рядом с мячом. В этом случае робот подъезжает к мячу вплотную. В состоянии 1 робот не приближается плотно к мячу, чтобы не касаться его при прицеливании.

Состояние 4: робот касается мяча и прицеливается. Чтобы проверить, что мяч касается робота, используются данные от датчика наличия мяча. В этом состоянии робот пинает мяч.

Модель игры

Для подачи на квалификацию RoboCup SSL была разработана модель игры двух атакующих против одного вратаря. В данной модели поле делится на несколько зон: A_1 и A_2 — зоны атакующих, B — зона вратаря, S_1 и S_2 — начальные позиции атакующих, P_1 и P_2 — позиции, в которых атакующие принимают пас (рис. 4). Каждому роботу сопоставляется некоторое число — его приоритет (вратарь имеет наибольший приоритет).

Вначале определяется владелец мяча. Им является робот, который находится на расстоянии



■ Рис. 4. Два атакующих против вратаря
 ■ Fig. 4. Two attackers against goalkeeper



■ Рис. 5. Фрагмент игры «Два нападающих против вратаря»
 ■ Fig. 5. Frame from the game “Two attackers against goalkeeper”

от мяча не дальше заданного значения. В случае если на поле несколько таких роботов, то владельцем назначается робот с большим приоритетом. Если все роботы расположены далеко от мяча, то владельцем назначается робот, в зоне которого оказался мяч. Если мяч вылетел за пределы поля, то владелец не определен и роботы ожидают возвращения мяча на поле.

Если владелец вратарь, то он отбивает мяч в ближайшего атакующего робота (для того чтобы игра была цикличной), в это время атакующие роботы перемещаются в стартовые позиции своих игровых зон (пока вратарь подъезжает и нацеливается для удара) и ловят мяч (когда он приходит в движение). Если владельцем является атакующий робот, то он либо пасует мяч другому роботу, либо пробивает по воротам — решение принимается случайно равновероятно. В это время другой атакующий робот занимает позицию для приема паса в своей игровой зоне, а вратарь защищает ворота.

В настоящее время проходят тренировки команды и совершенствование алгоритмов игры на тактическом уровне. На рис. 5 изображен фрагмент игры «Два нападающих против вратаря».

Заключение

В ходе исследования был разработан стенд, предназначенный для создания и изучения алгоритмов управления несколькими роботами и навигации в высокодинамичной среде с централизованной системой управления. Стенд был протестирован на алгоритмах для задач RoboCup SSL и продолжает активно тестироваться в целях совершенствования алгоритмов планирования пути и обхода препятствий. Описанные в работе базовые алгоритмы управления роботом-футбо-

листов были использованы для записи квалификационного видео [20] и подготовки квалификационной статьи TDP (Team Description Paper) команды URoboRus при подаче заявки на участие в финале RoboCup SSL, который пройдет в июле 2019 года в Сиднее, Австралия.

Работа выполнена при частичной поддержке ООО «Кибертех» и Центра технологий распре-

деленных реестров СПбГУ. Авторы благодарят Президентский физико-математический лицей № 239 за предоставление набора роботов для проведения совместных исследований. Авторы пользуются случаем выразить признательность Р. М. Лучину и С. А. Филиппову за энтузиазм и поддержку футбола роботов в Санкт-Петербурге, а также А. С. Матвееву за полезные дискуссии.

Литература

1. *Robocup*. <https://robocup.org/> (дата обращения: 29.03.2019).
2. Ронжин А. Л., Станкевич Л. А., Шандаров Е. С. Международные соревнования роботов по футболу RoboCup и перспективы участия в них российских команд. *Робототехника и техническая кибернетика*, 2015, № 2, с. 24–29.
3. Bruce J., Zickler S., Licitra M., Veloso M. CMDragons: Dynamic passing and strategy on a champion robot soccer team. *2008 IEEE Intern. Conf. on Robotics and Automation*, Pasadena, CA, 2008, pp. 4074–4079. doi:10.1109/ROBOT.2008.4543837
4. Fu W., Lin K., Shih C. Key-frame selection for multi-robot simultaneous localization and tracking in robot soccer field. *2018 IEEE/RSJ Intern. Conf. on Intelligent Robots and Systems (IROS)*, Madrid, 2018, pp. 109–116. doi:10.1109/IROS.2018.8593785
5. Hwang K.-S., Tan S.-W., Chen C.-C. Cooperative strategy based on adaptive Q-learning for robot soccer systems. *IEEE Transactions on Fuzzy Systems*, Aug. 2004, vol. 12, no. 4, pp. 569–576. doi:10.1109/TFUZZ.2004.832523
6. Kim J.-H., Shim H.-S., Kim H.-S., Jung M.-J., Choi I.-H., Kim J.-O. A cooperative multi-agent system and its real time application to robot soccer. *Proc. of Intern. Conf. on Robotics and Automation*, Albuquerque, NM, USA, 1997, vol. 1, pp. 638–643. doi:10.1109/ROBOT.1997.620108
7. Lee B.-J., Park G.-T. A robot in intelligent environment: soccer robot. *1999 IEEE/ASME Intern. Conf. on Advanced Intelligent Mechatronics (Cat. No.99TH8399)*, Atlanta, GA, USA, 1999, pp. 73–78. doi:10.1109/AIM.1999.803145
8. Li X., Wang M., Zell A. Dribbling control of omnidirectional soccer robots. *Proc. 2007 IEEE Intern. Conf. on Robotics and Automation*, Roma, 2007, pp. 2623–2628. doi:10.1109/ROBOT.2007.363861
9. Shieh M.-Y. Visual servo strategy for robot soccer systems. *Sensors and Materials*, 2018, vol. 30, no. 4, pp. 893–906.
10. Weigel T., Gutmann J.-S., Dietl M., Kleiner A., Nebel B. CS Freiburg: coordinating robots for successful soccer playing. *IEEE Transactions on Robotics and Automation*, Oct. 2002, vol. 18, no. 5, pp. 685–699. doi:10.1109/TRA.2002.804041
11. Охочимский Д. Е., Павловский В. Е., Плахов А. Г., Туганов А. Н. Моделирование игры роботов-футболистов в пакете «виртуальный футбол». *Мехатроника*, 2002, № 1, с. 2–5.
12. Котова О. А., Павловский В. Е. *Теоретико-игровые кинематические модели виртуального футбола*. М., ИПМ им. М. В. Келдыша, 2010. 20 с. (Препринт/ ИПМ им. М. В. Келдыша; № 77). <http://library.keldysh.ru/preprint.asp?id=2010-77> (дата обращения: 29.03.2019)
13. Matveev A. S., Colin Hoy M., Ovchinnikov K. S., Anisimov A., Savkin A. V. Robot navigation for monitoring unsteady environmental boundaries without field gradient estimation. *Automatica*, 2015, vol. 62, pp. 227–235. doi:10.1016/j.automatica.2015.09.003
14. Липкович М. М., Лучин Р. М., Сохарев А. Ю. Учебно-исследовательский стенд для мультиагентных систем управления (на примере футбола роботов). *Theoretical and Applied Aspects of Cybernetics: Proc. of the Intern. Scientific Conf. of Students and Young Scientists*, Киев, Bukrek, 2011, pp. 161–164.
15. Лучин Р. М., Широколов И. Ю., Овчинников К. С. Исследование алгоритмов мультиагентного взаимодействия с помощью робототехнического комплекса. *5-я Российская мультиконференция по проблемам управления*, 2012, с. 213–215.
16. Shirokolobov I., Filippov S., Luchin R., Ovchinnikov K., Fradkov A. *Control engineering at high schools and universities: project-based learning*. In: *Handbook of Research on Estimation and Control Techniques in E-Learning Systems*. IGI Global, 2016, pp. 141–170. doi:10.4018/978-1-4666-9489-7.ch011
17. Фестиваль «Робофинист». *Управляемый футбол 4x4*. <https://robofinist.ru/event/info/competitions/id/213#kind1399> (дата обращения: 29.03.2019).
18. *Robocup SSL Rules*. <https://ssl.robocup.org/rules/> (дата обращения: 29.03.2019).
19. Rodríguez S., Rojas E., Pérez K., López J., Quintero C., Calderón J. *Fast path planning algorithm for the RoboCup small size league*. In: *RoboCup 2014: Robot World Cup XVIII. Lecture Notes in Computer Science* / R. Bianchi, H. Akin, S. Ramamoorthy, K. Sugiura (eds.), 2015, vol. 8992, pp. 407–418. doi:10.1007/978-3-319-18615-3_33
20. *URoboRus Qualification Video 2019*. <https://www.youtube.com/watch?v=W6JNiaXGRRk> (дата обращения: 29.03.2019)

UDC 621.865

doi:10.31799/1684-8853-2019-2-19-25

A team of soccer robots for RoboCup competitions in SSL league: system and algorithmsP. A. Konovalov^a, Student, orcid.org/0000-0002-4910-9184A. V. Kornilova^b, Programmer Engineer, orcid.org/0000-0002-2267-9689D. M. Korolev^c, Engineer, orcid.org/0000-0002-6673-2998G. V. Reneva^a, Student, orcid.org/0000-0003-2322-4608A. L. Fradkov^{a,d}, Dr. Sc., Tech., Professor, orcid.org/0000-0002-5633-0944, fradkov@mail.ruI. Yu. Shirokolobov^{b,e}, Engineer, orcid.org/0000-0003-4346-5433D. S. Yarosh^a, Student, orcid.org/0000-0002-2017-8979^aSaint-Petersburg State University, 7–9, Universitetskaya Emb., 199034, Saint-Petersburg, Russian Federation^bООО «Cybertech», 68, ap. 26, Zheleznodorozhnaya St., 196606, Pushkin, Saint-Petersburg, Russian Federation^cRussian State Scientific Center for Robotics and Technical Cybernetics, 21, Tikhoretsky Pr., 194064, Saint-Petersburg, Russian Federation^dInstitute of Problems of Mechanical Engineering of RAS, 61, Bol'shoi Pr. V. O., 199178, Saint-Petersburg, Russian Federation^ePresidential Physics and Mathematics Lyceum No. 239, 8, Kirochnaya St., 191028, Saint-Petersburg, Russian Federation

Introduction: The international RoboCup soccer tournaments for robots have been held since the mid-1990s and have attracted a lot of attention from universities all over the world. Their goal is to develop scientific and engineering research in the areas of robotics, artificial intelligence, computer vision, navigation, group and multi-agent interaction of robots and mechatronic devices. **Purpose:** To create a new generation research bench which assists in developing, training and evaluating soccer robots, as well as developing basic algorithms for managing an individual soccer robot or a group of them. **Results:** A stand was designed to develop and analyze algorithms for control over a group of robots and navigation in a highly dynamic environment with centralized control organized as follows: the computing center receives data from a video surveillance system, uses the obtained information about the positions of the robots and the ball and about the state of their sensors to calculate control commands (velocity vector, angular velocity, impact signal) which are then transmitted to the robots for execution. Basic control algorithms for robot soccer players, such as: moving to a point, turning on a point, movement around a circle, obstacle avoidance, taking a pass, goalkeeper actions, hitting a set point, as well as a game model for two field players acting against a goalkeeper were implemented and used to test the stand performance, to record the qualification video and to prepare documents for the participation of St. Petersburg State University team in the RoboCup Small Size League final, 2019, in Sydney, Australia.

Keywords — group interaction, soccer robots, research stand, robot control, obstacle avoidance.

For citation: Konovalov P. A., Kornilova A. V., Korolev D. M., Reneva G. V., Fradkov A. L., Shirokolobov I. Yu., Yarosh D. S. A team of soccer robots for RoboCup competitions in SSL league: system and algorithms. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2019, no. 2, pp. 19–25 (In Russian). doi:10.31799/1684-8853-2019-2-19-25

References

1. *Robocup*. Available at: <https://robocup.org/> (accessed 29 March 2019).
2. Ronzhin A. L., Stankevich L. A., Shandarov E. S. International competitions of robots football RoboCup and the prospects for participation in them of the Russian teams. *Robototekhnika i tekhnicheskaya kibernetika*, 2015, no. 2, pp. 24–29 (In Russian).
3. Bruce J., Zickler S., Licitra M., Veloso M. CMDragons: Dynamic passing and strategy on a champion robot soccer team. *2008 IEEE Intern. Conf. on Robotics and Automation*, Pasadena, CA, 2008, pp. 4074–4079. doi:10.1109/ROBOT.2008.4543837
4. Fu W., Lin K., Shih C. Key-frame selection for multi-robot simultaneous localization and tracking in robot soccer field. *2018 IEEE/RSJ Intern. Conf. on Intelligent Robots and Systems (IROS)*, Madrid, 2018, pp. 109–116. doi:10.1109/IROS.2018.8593785
5. Hwang K.-S., Tan S.-W., Chen C.-C. Cooperative strategy based on adaptive Q-learning for robot soccer systems. *IEEE Transactions on Fuzzy Systems*, Aug. 2004, vol. 12, no. 4, pp. 569–576. doi:10.1109/TFUZZ.2004.832523
6. Kim J.-H., Shim H.-S., Kim H.-S., Jung M.-J., Choi I.-H., Kim J.-O. A cooperative multi-agent system and its real time application to robot soccer. *Proc. of Intern. Conf. on Robotics and Automation*, Albuquerque, NM, USA, 1997, vol. 1, pp. 638–643. doi:10.1109/ROBOT.1997.620108
7. Lee B.-J., Park G.-T. A robot in intelligent environment: soccer robot. *1999 IEEE/ASME Intern. Conf. on Advanced Intelligent Mechatronics (Cat. No.99TH8399)*, Atlanta, GA, USA, 1999, pp. 73–78. doi:10.1109/AIM.1999.803145
8. Li X., Wang M., Zell A. Dribbling control of omnidirectional soccer robots. *Proc. 2007 IEEE Intern. Conf. on Robotics and Automation*, Roma, 2007, pp. 2623–2628. doi:10.1109/ROBOT.2007.363861
9. Shieh M.-Y. Visual servo strategy for robot soccer systems. *Sensors and Materials*, 2018, vol. 30, no. 4, pp. 893–906.
10. Weigel T., Gutmann J.-S., Dietl M., Kleiner A., Nebel B. CS Freiburg: coordinating robots for successful soccer playing. *IEEE Transactions on Robotics and Automation*, Oct. 2002, vol. 18, no. 5, pp. 685–699. doi:10.1109/TRA.2002.804041
11. Ohocimskij D. E., Pavlovskij V. E., Plahov A. G., Tuganov A. N. Simulation of the game of robot soccer players in the “virtual football” package. *Mekhatronika*, 2002, no. 1, pp. 2–5 (In Russian).
12. Kotova O. A., Pavlovskij V. E. *Teoretiko-igrovie kinematicheskie modeli virtual'nogo futbola* [Game-theoretic kinematic models of virtual football]. Moscow, Institut prikladnoj matematiki imeni M. V. Keldysha RAN Publ., 2010. 20 p. (Preprint. Institut prikladnoj matematiki imeni M. V. Keldysha RAN, no. 77). Available at: <http://library.keldysh.ru/preprint.asp?id=2010-77> (accessed 29 March 2019) (In Russian).
13. Matveev A. S., Colin Hoy M., Ovchinnikov K. S., Anisimov A., Savkin A. V. Robot navigation for monitoring unsteady environmental boundaries without field gradient estimation. *Automatica*, 2015, vol. 62, pp. 227–235. doi:10.1016/j.automatica.2015.09.003
14. Lipkovich M. M., Luchin R. M., Soharev A. Yu. Educational and research stand for multi-agent control systems (using the example of robot football). *Proc. of the Intern. Scientific Conf. of Students and Young Scientists “Theoretical and Applied Aspects of Cybernetics”*, Kiev, Bukrek, 2011, pp. 161–164 (In Russian).

15. Luchin R. M., Shirokolobov I. Yu., Ovchinnikov K. S. Investigation of multi-agent interaction algorithms using a robotic complex. *5-ya Rossijskaya mul'tikonferenciya po problemam upravleniya* [5th Russian Multiconference on Control Problems], 2012, pp. 213–215 (In Russian).
16. Shirokolobov I., Filippov S., Luchin R., Ovchinnikov K., Fradkov A. *Control engineering at high schools and universities: project-based learning*. In: *Handbook of Research on Estimation and Control Techniques in E-Learning Systems*. IGI Global, 2016, pp. 141–170. doi:10.4018/978-1-4666-9489-7.ch011
17. Festival' "Robofinist". *Upravlyaemyj futbol 4x4* [Festival "robinist". Controlled football 4x4]. Available at: <https://robofinist.ru/event/info/competitions/id/213#kind1399> (accessed 29 March 2019).
18. *Robocup SSL Rules*. Available at: <https://ssl.robocup.org/rules/> (accessed 29 March 2019).
19. Rodríguez S., Rojas E., Pérez K., López J., Quintero C., Calderón J. *Fast path planning algorithm for the RoboCup small size league*. In: *RoboCup 2014: Robot World Cup XVIII. Lecture Notes in Computer Science*, Bianchi R., Akin H., Ramamoorthy S., Sugiura K. (eds.), 2015, vol. 8992, pp. 407–418. doi:10.1007/978-3-319-18615-3_33
20. *URoboRus Qualification Video 2019*. Available at: <https://www.youtube.com/watch?v=W6JN1aXGRRk> (accessed 29 March 2019).

Научный журнал
«ИНФОРМАЦИОННО-УПРАВЛЯЮЩИЕ СИСТЕМЫ»
выходит каждые два месяца.

Стоимость годовой подписки (6 номеров) для подписчиков России — 6000 рублей, для подписчиков стран СНГ — 6600 рублей, включая НДС 20%, таможенные и почтовые расходы.

Подписку на печатную версию журнала можно оформить в любом отделении связи по каталогу:

«Пресса России»: № 15385 — полугодовой индекс,

а также через посредство подписных агентств:

«Северо-Западное агентство „Прессинформ“»

Санкт-Петербург, тел.: (812) 335-97-51, 337-23-05,

эл. почта: press@crp.spb.ru, zajavka@crp.spb.ru,

сайт: <http://www.pinform.spb.ru>

«МК-Периодика» (РФ + 90 стран)

Москва, тел.: (495) 681-91-37, 681-87-47,

эл. почта: export@periodicals.ru, сайт: <http://www.periodicals.ru>

«Деловая пресса»

Москва, тел.: (495) 962-11-11, эл. почта: podpiska@delpress.ru,

сайт: <http://delpress.ru/contacts.html>

«Коммерсант-Курьер»

Казань, тел.: (843) 291-09-99, 291-09-47, эл. почта: kazan@komcur.ru,

сайт: <http://www.komcur.ru/contacts/kazan/>

«Урал-Пресс» (филиалы в 40 городах РФ)

Сайт: <http://www.ural-press.ru>

«Идея» (Украина)

Сайт: <http://idea.com.ua>

«ВТЛ» (Узбекистан)

Сайт: <http://btl.sk.uz/ru/cat17.html> и др.

На электронную версию нашего журнала (все выпуски, годовая подписка, один выпуск, одна статья)

вы можете подписаться на сайтах НЭБ: <http://elibrary.ru>; РУКОИТ: <http://www.rucont.ru>;

ИВИС: <http://www.ivis.ru>; Некс-Медиа: <http://biblioclub.ru/index.php?page=news&id=11196>

Полнотекстовые версии журнала за 2002–2017 гг.

в свободном доступе на сайте журнала (<http://www.i-us.ru>),

НЭБ (<http://www.elibrary.ru>)

и Киберленинки (<http://cyberleninka.ru/journal/n/informatsionno-upravlyayuschiesistemy>).

Исследование влияния высокоскоростных видеоданных на точность распознавания аудиовизуальной речи

Д. В. Иванько^{а, б, в}, аспирант, orcid.org/0000-0003-0412-7765, denis.ivanko11@gmail.com

Д. А. Рюмин^{а, б}, аспирант, orcid.org/0000-0002-7935-0569

А. А. Карпов^{а, б}, доктор техн. наук, доцент, orcid.org/0000-0003-3424-652X

М. Железны^г, доктор техн. наук, доцент, orcid.org/0000-0003-1695-4370

^аСанкт-Петербургский институт информатики и автоматизации РАН, 14-я линия В. О., 39, Санкт-Петербург, 199178, РФ

^бСанкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, Кронверкский пр., 49, Санкт-Петербург, 197101, РФ

^вУльмский университет, Гельмгольцштрассе, 16, 89081, Ульм, Германия

^гЗападночешский университет, Университетская ул., 2732/8, 30100, Пльзень, Чехия

Введение: эффективность работы современных систем автоматического распознавания речи в тихих акустических условиях достаточно высока и в среднем достигает 90–95 %. Однако в неконтролируемой среде зачастую происходит искажение звукового сигнала, что сильно снижает результирующую точность распознавания. В подобных условиях представляется целесообразным использовать визуальную информацию о речи, так как она не подвержена влиянию акустического шума. На настоящий момент не существует исследований, объективно показывающих зависимость точности распознавания визуальной речи от частоты кадров видео. Также отсутствуют соответствующие аудиовизуальные базы данных для обучения моделей. **Цель:** сбор представительной базы данных, разработка и исследование автоматической системы аудиовизуального распознавания слитной русской речи. **Методы:** для распознавания речевых сигналов применяются методы на основе двоянных скрытых марковских моделей. Для параметрического представления акустических и визуальных сигналов применяются методы на основе мел-частотных кепстральных коэффициентов и пиксельные признаки, использующие анализ главных компонент. **Результаты:** исследовались видеоданные с пятью различными скоростями следования кадров: 25, 50, 100, 150 и 200 кадров в секунду. Эксперименты показали положительный эффект от использования высокоскоростной видеокамеры: удалось добиться абсолютного прироста точности на 1,48 % для бимодальной и 3,10 % для одномодальной системы по сравнению со стандартной скоростью записи 25 кадров в секунду. В результате экспериментов с зашумленными данными удалось установить, что бимодальное распознавание речи превосходит по точности распознавания одномодальное, особенно для низких значений ОСШ < 15 дБ. При очень низких значениях ОСШ < 5 дБ акустическая информация становится неинформативной, и наилучшие результаты показывает одномодальная система видеораспознавания речи. **Практическая значимость:** использование высокоскоростной камеры позволяет улучшить точность и робастность системы распознавания слитной русской речи.

Ключевые слова – высокоскоростная видеокамера, аудиовизуальное распознавание речи, шумовые условия, визуальное, многомодальное взаимодействие, чтение речи по губам диктора.

Для цитирования: Иванько Д. В., Рюмин Д. А., Карпов А. А., Железны М. Исследование влияния высокоскоростных видеоданных на точность распознавания аудиовизуальной речи. *Информационно-управляющие системы*, 2019, № 2, с. 26–34. doi:10.31799/1684-8853-2019-2-26-34

For citation: Ivanko D. V., Ryumin D. A., Karpov A. A., Zelezny M. Measuring the effect of high-speed video data on the audio-visual speech recognition accuracy. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2019, no. 2, pp. 26–34 (In Russian). doi:10.31799/1684-8853-2019-2-26-34

Введение

Визуальная речь, т. е. видеозапись речевых звуков, играет важную роль в улучшении свойств устойчивости автоматического распознавания речи (Automatic Speech Recognition — ASR).

Эффективность работы современных систем ASR в тихих акустических условиях достаточно высока и в среднем достигает 90–95 %. Однако в неконтролируемой среде зачастую происходит искажение звукового сигнала, что сильно влияет на результирующую точность ASR. В настоящее время использование мультимедийных данных в коммуникационных технологиях получило широкое распространение, и развился новый

подход к повышению производительности ASR в акустически неблагоприятных условиях. Он исходит из того, что речь является бимодальной (аудиовизуальной).

Многие эксперименты продемонстрировали, что способность человека понимать речь находится в зависимости от окружающего шума и снижается при его наличии. Для того чтобы повысить разборчивость речи в шумной обстановке, люди начинают читать движения губ собеседника и объединяют эту информацию с речевыми сигналами. После изучения этого феномена идея применения визуальной информации в ASR была изучена в ряде исследований [1, 2]. Видеозапись речевых звуков может быть крайне полезна для

систем ASR в неконтролируемых шумных условиях, так как визуальная модальность не подвержена влиянию акустического шума. Однако, несмотря на серьезные успехи, достигнутые глубоким обучением в области ASR, такие системы по-прежнему не могут приблизиться к уровню компетентности и шумоустойчивости распознавания речи человеком. Существует ряд ограничений, тормозящих прогресс в данной области, в основном связанных с отсутствием аудиовизуальных корпусов и необходимостью сочетания двух областей знаний: автоматического распознавания речи и компьютерного зрения.

Большинство из представленных в литературе экспериментов по аудиовизуальному распознаванию речи было проведено с использованием небольших корпусов, таких как GRID и CUAVE, или неопубликованных корпусов, таких как IBM ViaVoice™ [3].

В лаборатории речевых и мультимодальных интерфейсов СПИИРАН также проводятся исследования по аудиовизуальному распознаванию русской речи. Проведенные эксперименты на собранном мультимедийном корпусе аудиовизуальных данных (RusAVSpeechCorpus, гос. регистрация № 2011620085) показали, что использование стандартной скорости записи видеоданных 25 кадров в секунду зачастую недостаточно для захвата быстрой динамики движений области губ во время слитной речи. То есть некоторые звуки просто смазываются или сливаются на видеоданных, что в конечном итоге приводит к тому, что видеомодальность не только не улучшает точность работы всей системы, но даже ухудшает ее, внося искажения в гипотезы распознавания. Это происходит при достаточно быстром темпе речи говорящего, однако именно этот аспект (распознавание слитной речи) является наиболее важным с практической точки зрения и заслуживает пристального изучения.

Зарубежные исследования по вопросу чтения речи по губам для английского и голландского языков [4] подтверждают эту точку зрения и показывают, что существует некая зависимость между точностью распознавания визуальной речи и частотой следования видеокадров записи. Однако ответов на то, как именно выражается эта зависимость и какую скорость следования видеокадров считать оптимальной для задачи чтения речи по губам, данные исследования не дают.

Повышенная скорость записи видеокадров (высокоскоростные видеокамеры) давно и успешно используются в ряде задач человеко-машинного взаимодействия. К таким задачам можно отнести распознавание микровыражений лиц, распознавание эмоций, определение заболеваний глаз путем отслеживания характера моргания. Тем не менее для задач распознавания речи высокоско-

ростные камеры по-прежнему считаются слишком ресурсозатратными. Однако уже сейчас многие современные смартфоны оснащены высокоскоростными видеокамерами. К примеру, Apple iPhone X имеет скоростную видеокамеру, позволяющую записывать короткое видео со скоростью 240 кадров в секунду при разрешении 720×480 пикселей. А новый смартфон Sony Xperia имеет скорость записи до 960 кадров в секунду. Продолжающийся технологический прогресс позволяет с уверенностью предположить, что уже в ближайшем будущем внедрение систем аудиовизуального распознавания речи с высокоскоростными видеокамерами будет являться новой тенденцией в многомодальном распознавании речи. Вышесказанное определяет актуальность и значимость научного исследования, направленного на создание робастной системы аудиовизуального распознавания русской речи с использованием высокоскоростных видеоданных.

Аудиовизуальная база данных русской речи

Большинство современных систем ASR основаны на вероятностных моделях и методах обработки [5]. Для обучения подобных систем (обучения акустических и визуальных моделей в рамках статистического подхода к распознаванию речи) необходимо иметь представительную речевую базу данных. Сегодня существует целый ряд аудиовизуальных баз данных русской речи, находящихся как в коммерческом, так и в открытом доступе [6, 7]. Тем не менее все они были записаны с использованием стандартной скорости видеозаписи 25 кадров в секунду и поэтому не подходят для нашего исследования. Высокоскоростных баз данных аудиовизуальной русской речи на момент исследования не существовало.

По этим причинам было разработано программное обеспечение и собран корпус аудиовизуальной русской речи с высокоскоростными видеозаписями HAVRUS [8] (High-Speed Recordings of Audio-Visual Russian Speech) (табл. 1), который включает в себя записи 20 дикторов (10 мужчин и 10 женщин). Средний возраст дикторов составляет 22 года, все они являются нормативными носителями русского языка. Каждый из дикторов произнес 200 подобранных фраз на русском языке: 130 фраз для обучения были взяты из фонетически представительных текстов и являлись общими для всех говорящих, 70 фраз для тестирования отличались для каждого диктора и являлись номерами телефонов.

Продолжительность речевых данных каждого диктора составляет приблизительно 15–20 мин

■ **Таблица 1.** Общие характеристики речевой базы данных HAVRUS

■ **Table 1.** General characteristics of the HAVRUS database

Параметр	Значение
Количество дикторов	20
Количество фраз у каждого диктора	200
Аудиоданных на каждого диктора, мин	15–20
Общая длительность аудиоданных, ч	≈6
Процент чистой речи в записях, %	>80
Частота дискретизации, кГц	44,1
Квантование сигнала, бит	16
Отношение сигнал/шум (ОСШ), дБ	>35
Тип микрофона	Oktava MK-012
Тип камеры	JAI Pulnix RMC-6740
Тип объектива	KOWA LM6NCM
Разрешение изображения, пикселей	640 × 480
Общий объем данных, ТБ	≈5,5

слитной речи. Видеоданные имеют частоту кадров 200 кадров в секунду, несжатые данные в RAW-формате.

Обработка аудиовизуальной русской речи

Архитектура системы аудиовизуального распознавания русской речи

Архитектура разработанной системы аудиовизуального распознавания русской речи на базе микрофона и высокоскоростной видеокамеры показана на рис. 1. Выходными данными является результат распознавания в текстовой форме, который выводится в окне ПО AVSpeechRecognition. В режиме офлайн-распознавания также в окне программы будет выведено время работы программы, длительность видео, скорость обработки видеофайлов и точность распознавания.

Методы аудиовизуальной обработки речи

На этапе моделирования и объединения аудиовизуальных данных используется несколько распространенных методов, таких как метод опорных векторов, байесовские сети доверия, скрытые марковские модели (СММ), нейронные сети, оценочные алгоритмы и пр. [9–11].

В настоящей работе применяется хорошо себя зарекомендовавшая технология сдвоенных скры-

тых марковских моделей (ССММ) по причине того, что они способны учитывать естественную асинхронность движений губ и речевой информации, обеспечивая синхронизацию аудиовизуальной речи на границах слов.

Объединение аудиовизуальных модальностей речи.

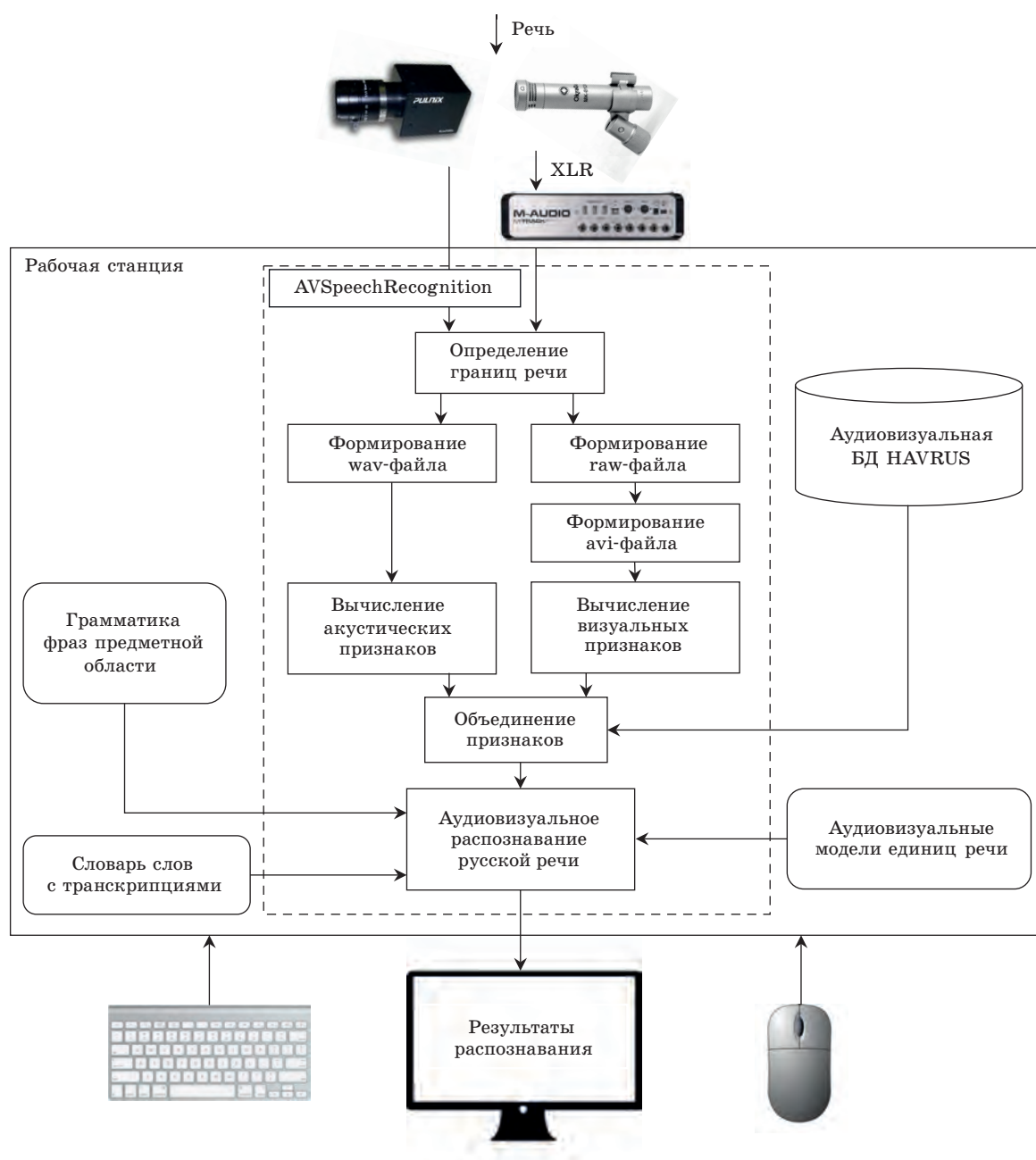
Скрытую марковскую модель можно рассматривать как простую форму байесовской сети доверия, которая представляет распределения вероятностей по последовательностям наблюдений. Как и байесовские сети доверия, СММ широко эксплуатируются в задачах обработки речи и видео. Одной из форм СММ являются многопоточные модели, которые имеют в своей структуре два отдельных потока для аудио- и видеонаблюдений, объединяя эти наблюдения на каждом кадре. Сложность декодирования такого алгоритма является линейной и зависит от количества потоков. В ССММ параллельные потоки моделируются с использованием параллельных СММ, где каждое состояние СММ может проходить в асинхронной области, но должно оставаться привязанным к границам модели [12]. Основная проблема ССММ заключается в том, что их алгоритмы обучения становятся неразрешимыми при наличии более двух потоков (модальностей) [13].

В работе [14] был предложен метод адаптации сдвоенных СММ, применение которого позволяет увеличить точность распознавания визем по сравнению со стандартным методом совместного обучения аудиовизуальных СММ. Его суть заключается в раздельном обучении акустических и визуальных моделей, причем акустическая модель может быть обучена в том числе и с привлечением сторонних баз данных. Объем аудиоданных в современных аудиовизуальных базах не является достаточно большим, и привлечение дополнительных аудиоданных позволяет создать лучшие модели фонем. В итоге это приводит к лучшему выравниванию видеокадров. Этот подход обеспечивает не только лучшие акустические модели для конечных аудиовизуальных моделей, но и лучшую согласованность, благодаря которой лучшие визуальные модели будут добавлены к акустической модели. Данный подход хорошо себя показал при использовании в зашумленных условиях.

Параметрическое представление аудиовизуальной речи

Для параметрического описания аудиосигнала рассчитываются мел-частотные кепстральные коэффициенты (Mel-Frequency Cepstral Coefficients) с их первой и второй производными.

В отличие от аудио, не существует стандартизированного набора визуальных признаков.



■ **Рис. 1.** Архитектура автоматической системы многомодального распознавания русской речи
 ■ **Fig. 1.** Architecture of the automatic system of multimodal Russian speech recognition

Большинство исследователей применяют различные алгоритмы для их извлечения [15–17]. В качестве визуальных признаков, описывающих форму губ человека (визем), были использованы пиксельные визуальные признаки на основе метода анализа главных компонент (Principal Component Analysis) визуальной области губ человека. Визуальные признаки вычисляются в несколько этапов. Во-первых, на видеокдрах происходит поиск области лица человека каскадным класси-

фикатором по методу Adaptive Boosting (адаптивное усиление классификаторов), который основан на алгоритме Виолы — Джонса. Результатом подобной обработки является обнаружение области интереса (области рта диктора). После чего выполняется цифровая обработка изображения, включающая в себя следующие этапы: нормализацию области губ до изображения размером 32×32 пикселей и отображение в 32-мерный вектор признаков по методу анализа главных компонент [18].

Классы визем русской речи

Согласно нашим предыдущим исследованиям [19], наибольшая точность распознавания русской речи достигалась при использовании 20 визуально различных единиц речи (визем) (табл. 2).

Этот параметр зависит от языка [20, 21], и для русского использовалось 10–14 визем в различных работах. Однако при внедрении высокоскоростной видеокамеры удается намного лучше отследить быструю динамику движения губ в слитной речи.

Результаты экспериментов

В целях проведения экспериментов было реализовано несколько автоматических систем распознавания речи, как многомодальных (аудиовизуальных), так и одномодальных. Все системы являлись дикторозависимыми с малым словарем распознавания. Затем была проведена их экспериментальная проверка с применением собранного аудиовизуального корпуса русской речи с высокоскоростными видеозаписями HAVRUS. При этом ОСШ изменялось в пределах от 0 до 40 дБ. Модели сравнивались по количественному показателю точности распознавания слов слитной речи WRR.

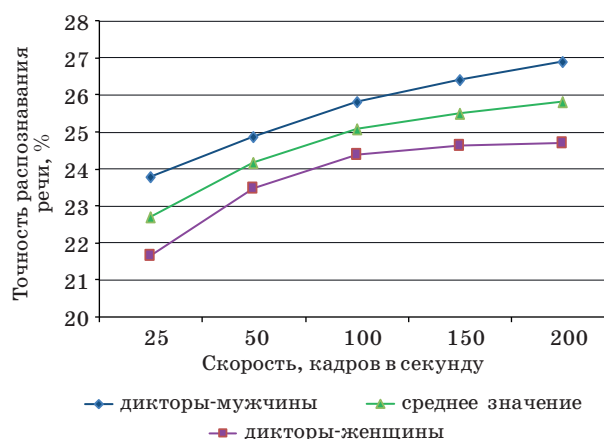
Распознавание визуальной речи

В настоящих экспериментах исследовались видеоданные с пятью различными скоростями следования кадров: 25, 50, 100, 150 и 200 кадров в секунду.

■ **Таблица 2.** Классы визем и их соответствие фонемам русской речи

■ **Table 2.** Viseme classes and their correspondence to the phonemes of the Russian speech

Класс виземы	Соответствующие фонемы русской речи	Класс виземы	Соответствующие фонемы русской речи
V1	Тишина (пауза)	V11	э!
V2	а, а!	V12	ы, ы!
V3	и, и!	V13	у, у!
V4	о!	V14	э
V5	б, б', п, п'	V15	с, с', з, з', ц
V6	ф, ф', в, в'	V16	й
V7	ш, щ	V17	х, х'
V8	л, л', р, р'	V18	ч
V9	д, д', т, т', н, н'	V19	м, м'
V10	г, г', к, к'	V20	ж



■ **Рис. 2.** Изменение пословной точности распознавания речи при увеличении скорости следования видеокадров в системе чтения русской речи по губам

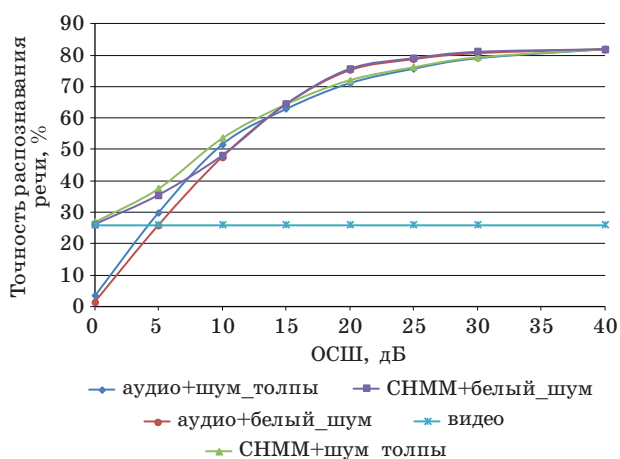
■ **Fig. 2.** Word recognition rate on increasing speed of video frames rate in the Russian lip reading system

Пословная точность распознавания речи (Word Recognition Rate — WRR) для всех 20 дикторов показана в зависимости от частоты кадров (рис. 2). Как видно из рисунка, WRR значительно увеличивается (около 1,5 %) при увеличении частоты следования видеокадров от 25 до 50. При дальнейшем увеличении частоты до 100 кадров в секунду WRR продолжает расти, однако темп роста замедляется (всего 0,91 % за дополнительные 50 кадров в секунду). Эта тенденция продолжает сохраняться и при увеличении до 150 кадров в секунду — WRR увеличивается еще на 0,42 %, но при увеличении до 200 кадров в секунду прирост точности составил всего 0,3 %.

Аудиовизуальное распознавание речи

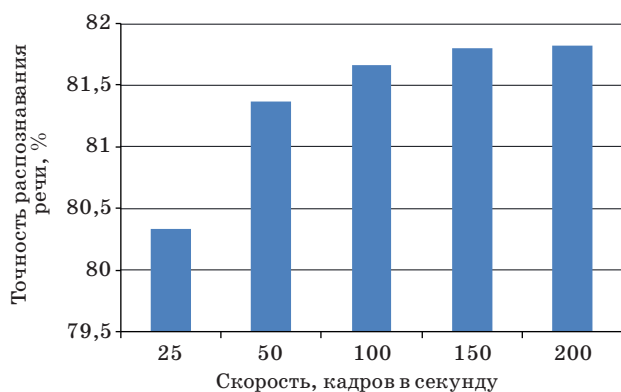
В ходе экспериментов тестовые данные для всех дикторов зашумлялись двумя видами шумов: широкополосным белым шумом и «шумом толпы» — одновременным говорением нескольких людей [22]. Интенсивность аддитивного шума изменялась в пределах от 0 до 40 дБ с шагом 5 дБ.

Результаты экспериментов по проверке трех систем распознавания (одномодальных систем с распознаванием только аудиосигнала или видеосигнала, бимодальной системы ССММ) представлены на рис. 3. Анализ результатов показывает, что бимодальное распознавание речи превосходит по точности распознавания слов одномодальное распознавание, что особенно очевидно для низких значений ОСШ < 15 дБ. При очень низких значениях ОСШ < 5 дБ акустическая информация становится неинформативной, и наилучшие результаты показывает одномодальная система видеораспознавания речи.



■ **Рис. 3.** Зависимость пословной точности распознавания речи от ОСШ аудиосигнала для различных конфигураций системы аудиовизуального распознавания речи

■ **Fig. 3.** The dependence of word recognition rate on SNR for various configurations of the audio-visual speech recognition system



■ **Рис. 4.** Пословная точность аудиовизуального распознавания русской речи при различных значениях скорости следования видеоданных

■ **Fig. 4.** WRR values for audio-visual Russian speech recognition with various fps of video data

Наилучшие результаты распознавания достигаются объединенной бимодальной системой. Так, в сильно зашумленных акустических условиях вес аудиомодальности минимизируется, и система полагается только на результат видеомодальности (при ОСШ < 10 дБ). Однако при ОСШ > 10 дБ видеомодальность уже не может обеспечить значительного прироста точности распознавания, поэтому ее вес уменьшается, а аудиомодальности — увеличивается.

При правильном использовании этого метода аудиовизуальная система становится более робастной к акустическим шумам, и появляется возможность добиться наилучших результатов распознавания речи при любых акустических

условиях путем изменения весов модальностей. Для повышения качества распознавания речи в реальных условиях применения этот метод следует сочетать с различными методами шумоочистки.

Другая серия проведенных экспериментов была ориентирована на исследование влияния скорости следования видеок кадров на точность распознавания бимодальной системы. На рис. 4 представлены результаты работы такой системы при различных уровнях скорости следования кадров в секунду.

Пословная точность распознавания речи достигает своего максимального значения 81,82 % при 200 кадрах в секунду. Это на 1,48 % выше, чем при использовании стандартной скорости записи видеоданных 25 кадров в секунду.

Заключение

В работе представлены результаты экспериментов по распознаванию русской речи и чтению речи по губам, полученные при помощи аудиовизуальной системы распознавания речи на основе сдвоенных скрытых марковских моделей. Эксперименты показали положительный эффект от использования высокоскоростной видеокамеры: при использовании высокоскоростной камеры удалось добиться абсолютного прироста точности на 1,48 % для бимодальной и 3,10 % для одномодальной системы по сравнению со стандартной скоростью записи 25 кадров в секунду. Эксперименты с системой видеорасознавания речи показали улучшение WRR до 7,28 % для некоторых дикторов.

Основываясь на проведенных исследованиях, можно сделать вывод, что внедрение высокоскоростной камеры позволяет улучшить точность и робастность системы распознавания слитной русской речи. Согласно нашим наблюдениям, белый шум уменьшает точность распознавания системы сильнее, чем «шум толпы». Однако белый шум достаточно легко поддается шумоочистке, тогда как задача очистки «шума толпы» по-прежнему остается большой и до конца нерешенной проблемой. В силу того, что «шум толпы» довольно часто встречается в реальных условиях применения, очень важно иметь надежную систему распознавания речи для таких условий. По нашему мнению, это может быть достигнуто с использованием аудиовизуальной системы распознавания речи на основе высокоскоростной видеокамеры.

Данное исследование проводится при поддержке фонда РФФИ (проекты № 18-37-00306, 16-37-60100, 18-07-01407), Правительства РФ (грант № 08-08), а также бюджетной темы № 0073-2019-0005.

Литература

1. Katsaggelos K., Bahaadini S., Molina R. Audiovisual fusion: challenges and new approaches. *Proc. of the IEEE*, 2015, vol. 103, no. 9, pp. 1635–1653.
2. Zhou Z., Zhao G., Hong X., Pietikainen M. A review of recent advances in visual speech decoding. *Proc. of the Image and Vision Computing*, 2014, vol. 32, pp. 590–605.
3. Ivanko D., Karpov A., Ryumin D., Kipyatkova I., Saveliev A., Budkov V., Zelezny M. Using a high-speed video camera for robust audio-visual speech recognition in acoustically noisy conditions. *Intern. Conf. on Speech and Computer (SPECOM)*, 2017, pp. 757–766.
4. Chitu A. G., Driel K., Rothkrantz L. J. M. Automatic lip reading in the Dutch language using active appearance models on high speed recordings. *Text, Speech and Dialogue, Springer LNCS (LNAI)*, 2010, vol. 6231, pp. 259–266.
5. Rajavel R., Sathidevi P. S. Adaptive reliability measure and optimum integration weight for decision fusion audio-visual speech recognition. *Journal of Signal Processing Systems*, 2012, vol. 68, no. 1, pp. 83–93.
6. Stewart D., Seymour R., Pass A., Ming J. Robust audio-visual speech recognition under noisy audio-visual conditions. *IEEE Transactions on Cybernetics*, Feb. 2014, vol. 44, no. 2, pp. 175–184.
7. Abhishek N., Prasanta K. G. PRAV: a phonetically rich audio visual corpus. *Proc. of the Interspeech*, 2017, pp. 3747–3751.
8. Verkhodanova V., Ronzhin A., Kipyatkova I., Ivanko D., Karpov A., Železný M. HAVRUS corpus: high-speed recordings of audio-visual russian. *Intern. Conf. on Speech and Computer (SPECOM)*, 2016, vol. 9811, pp. 338–345.
9. Shivappa S. T., Trivedi M. M., Rao B. D. Audiovisual information fusion in human-computer interfaces and intelligent environments: A survey. *Proc. of IEEE*, 2010, vol. 98, no. 10, pp. 1692–1715.
10. Abdelaziz A. H., Kolossa D. Dynamic stream weight estimation in coupled HMM-based audio-visual speech recognition using multilayer perceptrons. *Proc. of the Interspeech*, 2014, pp. 1144–1148.
11. Huang J., Kingsbury B. Audio-visual deep learning for noise robust speech recognition. *Proc. of the IEEE Intern. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)*, 2013, pp. 7596–7599.
12. Graves A., Mohamed A., Hinton G. Speech recognition with deep recurrent neural networks. *Proc. of the IEEE Intern. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, 2013, pp. 6645–6649.
13. Shiell D. J., Terry L. H., Aleksic P. S., Katsaggelos A. K. Audio-visual and visual-only speech and speaker recognition: Issues about theory, system design and implementation. In: *Visual Speech Recognition: Lip Segmentation and Mapping*. IGI Global, 2009. Pp. 1–38.
14. Abdelaziz A. H., Zeiler S., Kolossa D. A new EM estimation of dynamic stream weights for coupled-HMM-based audio-visual ASR. *Proc. of IEEE Intern. Conf. on Acoustic Speech and Signal Processing (ICASSP)*, 2014, pp. 54–62.
15. Kumar S., Bhuyan M. K., Chakraborty B. K. Extraction of texture and geometrical features from informative facial regions for sign language recognition. *Journal of Multimodal User Interfaces (JMUI)*, 2017, vol. 11, no. 2, pp. 227–239.
16. Lan Y., Theobald B., Harvey E., Ong E., Bowden R. Improving visual features for lip-reading. *Proc. of Auditory-Visual Speech Processing (AVSP)*, 2010, pp. 142–147.
17. Хафизов Р. Г., Яранцева Т. В. Оценка геометрических искажений контуров изображений губ в системах визуального ввода информации. *Информационно-управляющие системы*, 2017, № 4, с. 2–6. doi:10.15217/issn1684-8853.2017.4.2
18. Кухарев Г. А., Каменская Е. И., Матвеев Ю. Н., Щеголева Н. Л. Методы обработки и распознавания изображений лиц в задачах биометрии. СПб., Политехника, 2013. 388 с.
19. Ivanko D., Karpov A., Fedotov D., Kipyatkova I., Ryumin D., Ivanko Dm., Minker W., Zelezny M. Multimodal speech recognition: increasing accuracy using high speed video data. *Journal of Multimodal User Interfaces*, 2018, vol. 12, no. 4, pp. 319–328.
20. Иванько Д., Кипяткова И. С., Ронжин А. Л., Карпов А. А. Анализ методов многомодального объединения информации для аудиовизуального распознавания речи. *Научно-технический вестник информационных технологий, механики и оптики*, 2016, т. 16, № 3(103), с. 387–401.
21. Estellers V., Gurban M., Thiran J. On dynamic stream weighting for audio-visual speech recognition. *IEEE Transactions on Audio, Speech and Language Processing*, 2012, vol. 20, no. 4, pp. 1145–1157.
22. Stewart D., Seymour R., Pass A., Ming J. Robust audio-visual speech recognition under noisy audio-visual conditions. *IEEE Transactions on Cybernetics*, 2013, vol. 44, no. 2, pp. 175–184.

UDC 004.522

doi:10.31799/1684-8853-2019-2-26-34

Measuring the effect of high-speed video data on the audio-visual speech recognition accuracyD. V. Ivanko^{a,b,c}, Post-Graduate Student, orcid.org/0000-0003-0412-7765, denis.ivanko11@gmail.comD. A. Ryumin^{a,b}, Post-Graduate Student, orcid.org/0000-0002-7935-0569A. A. Karpov^{a,b}, Dr. Sc., Tech., Associate Professor, orcid.org/0000-0003-3424-652XM. Zelezny^d, PhD, Associate Professor, orcid.org/0000-0003-1695-4370^aSaint-Petersburg Institute for Informatics and Automation of the RAS, 39, 14 Line, V. O., 199178, Saint-Petersburg, Russian Federation^bSaint-Petersburg National Research University of Information Technologies, Mechanics and Optics, 49, Kronverkskii Pr., 197101, Saint-Petersburg, Russian Federation^cUniversität Ulm, 16, Helmholtzstraße, 89081, Ulm, Germany^dUniversity of West Bohemia (UWB), 2732/8, Univerzitní ul., 301 00, Plzeň, Czech Republic

Introduction: The effectiveness of modern automatic speech recognition systems in quiet acoustic conditions is quite high and reaches 90–95%. However, in noisy uncontrolled environment, acoustic signals are often distorted, which greatly reduces the resulting recognition accuracy. In adverse conditions, it seems appropriate to use the visual information about the speech, as it is not affected by the acoustic noise. At the moment, there are no studies which objectively reflect the dependence of visual speech recognition accuracy on the video frame rate, and there are no relevant audio-visual databases for model training. **Purpose:** Improving the reliability and accuracy of the automatic audio-visual Russian speech recognition system; collecting representative audio-visual database and developing an experimental setup. **Methods:** For audio-visual speech recognition, we used coupled hidden Markov model architectures. For parametric representation of audio and visual features, we used mel-frequency cepstral coefficients and principal component analysis-based pixel features. **Results:** In the experiments, we studied 5 different rates of video data: 25, 50, 100, 150, and 200 fps. Experiments have shown a positive effect from the use of a high-speed video camera: we achieved an absolute increase in accuracy of 1.48% for a bimodal system and 3.10% for a unimodal one, as compared to the standard recording speed of 25 fps. During the experiments, test data for all speakers were added with two types of noise: wide-band white noise and “babble noise”. Analysis shows that bimodal speech recognition exceeds unimodal in accuracy, especially for low SNR values <15 dB. At very low SNR values <5 dB, the acoustic information becomes non-informative, and the best results are achieved by a unimodal visual speech recognition system. **Practical relevance:** The use of a high-speed camera can improve the accuracy and robustness of a continuous audio-visual Russian speech recognition system.

Keywords — high-speed video camera, audio-visual speech recognition, noisy conditions, visemes, multimodal processing, lip-reading.

For citation: Ivanko D. V., Ryumin D. A., Karpov A. A., Zelezny M. Measuring the effect of high-speed video data on the audio-visual speech recognition accuracy. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2019, no. 2, pp. 26–34 (In Russian). doi:10.31799/1684-8853-2019-2-26-34

References

- Katsaggelos K., Bahaadini S., Molina R. Audiovisual Fusion: Challenges and New Approaches. *Proc. of the IEEE*, 2015, vol. 103, no. 9, pp. 1635–1653.
- Zhou Z., Zhao G., Hong X., Pietikainen M. A review of recent advances in visual speech decoding. *Proc. of the Image and Vision Computing*, 2014, vol. 32, pp. 590–605.
- Ivanko D., Karpov A., Ryumin D., Kipyatkova I., Saveliy A., Budkov V., Zelezny M. Using a high-speed video camera for robust audio-visual speech recognition in acoustically noisy conditions. *Intern. Conf. on Speech and Computer (SPECOM)*, 2017, pp. 757–766.
- Chitu A. G., Driel K., Rothkrantz L. J. M. Automatic lip reading in the Dutch language using active appearance models on high speed recordings. *Text, Speech and Dialogue, Springer LNCS (LNAI)*, 2010, vol. 6231, pp. 259–266.
- Rajavel R., Sathidevi P. S. Adaptive reliability measure and optimum integration weight for decision fusion audio-visual speech recognition. *Journal of Signal Processing Systems*, 2012, vol. 68, no. 1, pp. 83–93.
- Stewart D., Seymour R., Pass A., Ming J. Robust audio-visual speech recognition under noisy audio-video conditions. *IEEE Transactions on Cybernetics*, Feb. 2014, vol. 44, no. 2, pp. 175–184.
- Abhishek N., Prasanta K. G. PRAV: a phonetically rich audio visual corpus. *Proc. of the Interspeech*, 2017, pp. 3747–3751.
- Verkhodanova V., Ronzhin A., Kipyatkova I., Ivanko D., Karpov A., Zelezny M. HAVRUS Corpus: High-Speed Recordings of Audio-Visual Russian Speech. *Intern. Conf. on Speech and Computer (SPECOM)*, 2016, vol. 9811, pp. 338–345.
- Shivappa S. T., Trivedi M. M., Rao B. D. Audiovisual information fusion in human-computer interfaces and intelligent environments: A survey. *Proc. of IEEE*, 2010, vol. 98, no. 10, pp. 1692–1715.
- Abdelaziz A. H., Kolossa D. Dynamic stream weight estimation in coupled HMM-based audio-visual speech recognition using multilayer perceptrons. *Proc. of the Interspeech*, 2014, pp. 1144–1148.
- Huang J., Kingsbury B. Audio-visual deep learning for noise robust speech recognition. *Proc. of the IEEE Intern. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)*, 2013, pp. 7596–7599.
- Graves A., Mohamed A., Hinton G. Speech recognition with deep recurrent neural networks. *Proc. of the IEEE Intern. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)*, 2013, pp. 6645–6649.
- Shiell D. J., Terry L. H., Aleksic P. S., Katsaggelos A. K. *Audio-visual and visual-only speech and speaker recognition: Issues about theory, system design and implementation*. In: *Visual Speech Recognition: Lip Segmentation and Mapping*. IGI Global, 2009. Pp. 1–38.
- Abdelaziz A. H., Zeiler S., Kolossa D. A new EM estimation of dynamic stream weights for coupled-HMM-based audio-visual ASR. *Proc. of IEEE Intern. Conf. on Acoustic Speech and Signal Processing (ICASSP)*, 2014, pp. 54–62.
- Kumar S., Bhuyan M. K., Chakraborty B. K. Extraction of texture and geometrical features from informative facial regions for sign language recognition. *Journal of Multimodal User Interfaces (JMUI)*, 2017, vol. 11, no. 2, pp. 227–239.
- Lan Y., Theobald B., Harvey E., Ong E., Bowden R. Improving visual features for lip-reading. *Proc. of Auditory-Visual Speech Processing (AVSP)*, 2010, pp. 142–147.
- Khafizov R. G., Yaranceva T. V. Estimation of geometrical distortions of lip contours in visual input systems. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2017, no. 4, pp. 2–6 (In Russian). doi:10.15217/issn1684-8853.2017.4.2

18. Kukharev G. A., Kamenskaya E. I., Matveev Yu. N., Schegoleva N. L. *Metody obrabotki i raspoznavaniya izobrazhenii lits v zadachakh biometrii* [Methods of Facial Images Processing and Recognition in Biometrics]. Saint-Petersburg, Politehnika Publ., 2013. 388 p. (In Russian).
19. Ivanko D., Karpov A., Fedotov D., Kipyatkova I., Ryumin D., Ivanko Dm., Minker W., Zelezny M. Multimodal speech recognition: increasing accuracy using high speed video data. *Journal of Multimodal User Interfaces*, 2018, vol. 12, no. 4, pp. 319–328.
20. Ivanko D. V., Kipyatkova I. S., Ronzhin A. L., Karpov A. A. Analysis of multimodal fusion techniques for audio-visual speech recognition. *Nauchno-tehnicheskij vestnik informacionnyh tehnologij, mehaniki i optiki* [Scientific and Technical Journal of Information Technologies, Mechanics and Optics], 2016, vol. 16, no. 3, pp. 387–401 (In Russian).
21. Estellers V., Gurban M., Thiran J. On dynamic stream weighting for audio-visual speech recognition. *IEEE Transactions on Audio, Speech and Language Processing*, 2012, vol. 20, no. 4, pp. 1145–1157.
22. Stewart D., Seymour R., Pass A., Ming J. Robust audio-visual speech recognition under noisy audio-video conditions. *IEEE Transactions on Cybernetics*, 2013, vol. 44, no. 2, pp. 175–184.

Уважаемые авторы!

При подготовке рукописей статей необходимо руководствоваться следующими рекомендациями.

Статьи должны содержать изложение новых научных результатов. Название статьи должно быть кратким, но информативным. В названии недопустимо использование сокращений, кроме самых общепринятых (РАН, РФ, САПР и т. п.).

Объем статьи (текст, таблицы, иллюстрации и библиография) не должен превышать эквивалента в 20 страниц, напечатанных на бумаге формата А4 на одной стороне через 1,5 интервала Word шрифтом Times New Roman размером 13, поля не менее двух сантиметров.

Обязательными элементами оформления статьи являются: индекс УДК, заглавие, инициалы и фамилия автора (авторов), ученая степень, звание (при отсутствии — должность), полное название организации, аннотация и ключевые слова на русском и английском языках, ORCID и электронный адрес одного из авторов. При написании аннотации не используйте аббревиатур и не делайте ссылок на источники в списке литературы. Предоставляйте подрисовочные подписи и названия таблиц на русском и английском языках.

Статьи авторов, не имеющих ученой степени, рекомендуется публиковать в соавторстве с научным руководителем, наличие подписи научного руководителя на рукописи обязательно; в случае самостоятельной публикации обязательно предоставляйте заверенную по месту работы рекомендацию научного руководителя с указанием его фамилии, имени, отчества, места работы, должности, ученого звания, ученой степени.

Формулы набирайте в Word, не используя формульный редактор (Mathtype или Equation), при необходимости можно использовать формульный редактор; для набора одной формулы не используйте два редактора; при наборе формул в формульном редакторе знаки препинания, ограничивающие формулу, набирайте вместе с формулой; для установки размера шрифта никогда не пользуйтесь вкладкой Other..., используйте заводские установки редактора, не подгоняйте размер символов в формулах под размер шрифта в тексте статьи, не растягивайте и не сжимайте мышью формулы, вставленные в текст; в формулах не отделяйте пробелами знаки: + = -.

Для набора формул в Word никогда не используйте Конструктор (на верхней панели: «Работа с формулами» — «Конструктор»), так как этот ресурс предназначен только для внутреннего использования в Word и не поддерживается программами, предназначенными для изготовления оригинал-макета журнала.

При наборе символов в тексте помните, что символы, обозначаемые латинскими буквами, набираются светлым курсивом, русскими и греческими — светлым прямым, векторы и матрицы — прямым полужирным шрифтом.

Иллюстрации предоставляются отдельными исходными файлами, поддающимися редактированию:

— рисунки, графики, диаграммы, блок-схемы предоставляются в виде отдельных исходных файлов, поддающихся редактированию, используя векторные программы: Visio (*.vsd, *.vsdx); Coreldraw (*.cdr); Excel (*.xls); Word (*.docx); Adobe Illustrator (*.ai); AutoCad (*.dxf); Matlab (*.ps, *.pdf или экспорт в формат *.ai);

— если редактор, в котором Вы изготавливаете рисунок, не позволяет сохранить в векторном формате, используйте функцию экспорта (только по отношению к исходному рисунку), например, в формат *.ai, *.esp, *.wmf, *.emf, *.svg;

— фото и растровые — в формате *.tif, *.png с максимальным разрешением (не менее 300 pixels/inch).

Наличие подрисовочных подписей и названий таблиц на русском и английском языках обязательно (желательно не повторяющих дословно комментарии к рисункам в тексте статьи).

В редакцию предоставляются:

— сведения об авторе (фамилия, имя, отчество, место работы, должность, ученое звание, учебное заведение и год его окончания, ученая степень и год защиты диссертации, область научных интересов, количество научных публикаций, домашний и служебный адреса и телефоны, e-mail), фото авторов: анфас, в темной одежде на белом фоне, должны быть видны плечи и грудь, высокая степень четкости изображения без теней и отблесков на лице, фото можно представить в электронном виде в формате *.tif, *.png с максимальным разрешением — не менее 300 pixels/inch при минимальном размере фото 40×55 мм;

— экспертное заключение.

Список литературы составляется по порядку ссылок в тексте и оформляется следующим образом:

— для книг и сборников — фамилия и инициалы авторов, полное название книги (сборника), город, издательство, год, общее количество страниц;

— для журнальных статей — фамилия и инициалы авторов, полное название статьи, название журнала, год издания, номер журнала, номера страниц;

— ссылки на иностранную литературу следует давать на языке оригинала без сокращений;

— при использовании web-материалов указывайте адрес сайта и дату обращения.

Список литературы оформляйте двумя отдельными блоками по образцам lit.dot на сайте журнала (<http://i-us.ru/paperrules>): Литература и References.

Более подробно правила подготовки текста с образцами изложены на нашем сайте в разделе «Правила для авторов».

Контакты

Куда: 190000, Санкт-Петербург,
Б. Морская ул., д. 67, ГУАП, РИЦ

Кому: Редакция журнала «Информационно-управляющие системы»

Тел.: (812) 494-70-02

Эл. почта: i-us.spb@gmail.com

Сайт: www.i-us.ru

УДК 004.75

doi:10.31799/1684-8853-2019-2-35-43

Кластеризация данных в распределенных системах мониторинга

А. Н. Рукавицын^а, аспирант, orcid.org/0000-0001-5382-0465, rkvtstn@gmail.com^аСанкт-Петербургский государственный электротехнический университет «ЛЭТИ»,
Профессора Попова ул., 5, Санкт-Петербург, 197376, РФ

Введение: традиционные способы анализа распределенных источников данных обычно используют централизованные хранилища данных и имеют ряд недостатков, связанных с конфиденциальностью, высокой стоимостью централизованного хранения данных, ограниченной пропускной способностью и высокой нагрузкой на телекоммуникационные сети. Методики, по которым выполняется децентрализованный анализ, не учитывают вид распределения данных и особенности выбранного алгоритма. Это снижает производительность и точность анализа или может быть причиной невыполнимости его в заданных условиях. **Цель:** обзор и анализ особенностей работы распределенных систем мониторинга и алгоритмов интеллектуального анализа данных. **Результаты:** для проведения кластеризации на основе распределенных источников данных установлены требования к алгоритму в системах распределенного мониторинга: односторонность, поддержка разных типов входных данных, работа онлайн-режима, адаптация к данным при изменении среды, масштабирование больших объемов данных, выполнение анализа без предположений о распределении входных данных, анализ данных на источниках информации без их передачи третьей стороне. Определены два основных способа распределения данных на источниках в гетерогенных системах: вертикальный и горизонтальный. Выполнена классификация методов в соответствии с их основным принципом разграничения кластеров. Классификация включает основные алгоритмы кластеризации, их принцип работы, достоинства и недостатки. Обзор и анализ существующих методов кластеризации выявил, что в распределенных системах мониторинга наиболее эффективными являются алгоритмы на основе нейронных сетей Кохонена. Декомпозирован алгоритм самоорганизующихся карт Кохонена и определены блоки работы с данными: вычисление нейрона-победителя и настройка весов нейронов. Предложены две стратегии кластеризации распределенных данных. **Практическая значимость:** предложенные стратегии позволяют выполнять кластеризацию в системах с распределенными источниками с учетом характеристик среды без передачи всех данных.

Ключевые слова — интеллектуальный анализ данных, кластеризация, датчики, распределенные источники данных, самоорганизующиеся карты Кохонена, распределенные системы обработки данных, виды распределения, системы мониторинга.

Для цитирования: Рукавицын А. Н. Кластеризация данных в распределенных системах мониторинга. *Информационно-управляющие системы*, 2019, № 2, с. 35–43. doi:10.31799/1684-8853-2019-2-35-43

For citation: Rukavitsyn A. N. Data clustering in distributed monitoring systems. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2019, no. 2, pp. 35–43 (In Russian). doi:10.31799/1684-8853-2019-2-35-43

Введение

Развитие информационных технологий привело к появлению новых видов устройств и методов их взаимодействия, например, облачной виртуализированной среды или интернета вещей. Данные хранятся на разных, независимо работающих устройствах, которые могут быть связаны друг с другом через локальные или глобальные сети, например, в охранных системах, мобильных или корпоративных сетях, где датчики расположены географически в разных местах. Распределенные системы мониторинга — это системы, состоящие из разного вида датчиков с возможностями обнаружения, связи и ограниченным количеством вычислительной мощности. Системы мониторинга применяются в различных сферах жизнедеятельности [1–3]: в медицине, для защиты окружающей среды, обеспечения безопасности и др.

В ряде исследований в области интеллектуального анализа распределенных данных чаще

всего рассматривают алгоритмы на основе вероятностного подхода без учета различных видов распределения.

Методики, как правило, сводятся к двум этапам [4–9]:

1) кластеризации на источниках данных, где установлены модели для кластеризации и передачи результатов на следующий этап;

2) объединению результатов и группировке кластеризации.

Объединение может быть выполнено иерархически [4], и тогда каждый вычислительный узел является листом иерархии, который объединяет свои результаты с соседними листьями на узлах до корня полученного дерева. С таким подходом сложность выполнения растет с ростом количества узлов и установленных слоев. При этом выделяется дополнительное время на передачу данных между соседними узлами.

В работе Ф. Л. Горгоньо (F. L. Gorgonio) [8] описан подход, использующий ансамбль моделей (нейронных сетей), построенных SOM-алго-

ритмом (Self-Organized Maps). В нем на обоих этапах используется один и тот же алгоритм кластеризации, но для разных задач. На первом этапе SOM определяет нейронную сеть с кластерами и передает глобальной модели. На втором этапе используется процесс кластеризации карт для определения кластеров из переданных карт. Такой метод может привести к увеличению степени неопределенности.

Одним из наиболее распространенных подходов анализа распределенных данных является централизация в хранилище данных, к которым применяются традиционные методы интеллектуального анализа [10–12]. Хранилище данных — популярная технология, которая объединяет данные из нескольких источников в один, чтобы эффективно выполнять сложные аналитические запросы [13]. Однако, несмотря на широкое распространение, этот подход может быть непрактичным или невозможным по следующим причинам [14–16]:

- огромное количество данных генерируется на разных источниках, а из-за стоимости централизации данные не могут масштабироваться для связи, хранения и вычислений;

- владельцы данных не могут или не хотят передавать информацию, например, защищая конфиденциальность или потому, что раскрытие такой информации может привести к преимуществу конкурентов или значительной коммерческой добавленной стоимости;

- ограниченная пропускная способность и высокая нагрузка при передаче данных.

Соответственно, подход для интеллектуального анализа распределенных данных должен учитывать:

- адаптацию алгоритма построения модели в распределенной сети;
- объединение данных с минимальной нагрузкой сети и достаточной точностью анализа;
- вид распределения.

Распределенная модель системы мониторинга

Рассмотрим систему мониторинга с использованием сенсорных сетей, где узлы датчиков могут быть оснащены звуковыми, вибрационными, температурными и отражающими зондами. Допустим, датчики контролируют географический регион и должны следить за данными, обрабатывать их и обмениваться данными друг с другом, чтобы отслеживать и идентифицировать объекты, интересующие пользователя. Наблюдения обычно представляют собой данные в виде временных рядов. Система мониторинга может контролировать множество объектов X :

$$X = \{x_1, x_2, \dots, x_i, \dots, x_n\}.$$

Каждый объект характеризуется набором атрибутов A :

$$A = \{a_1, a_2, \dots, a_j, \dots, a_m\}.$$

При этом система мониторинга следит за значением атрибута x каждого объекта в каждый момент времени t :

$$x_i(t) = \{a_{1i}(t), a_{2i}(t), \dots, a_{ji}(t), \dots, a_{mi}(t)\}.$$

Анализируемые данные для системы мониторинга можно представить в виде таблицы (табл. 1).

Стоит обратить внимание, что системы могут быть однородными (т. е. каждый узел наблюдает общее подмножество данных) или гетерогенными (т. е. каждый узел наблюдает соответствующее подмножество данных). Гетерогенная система рассматривает два вида распределения данных: горизонтальное и вертикальное. При горизонтальном распределении данные делятся на источники по строкам или для временных рядов по времени либо по объектам из X . При вертикальном распределении данные делятся на источники по столбцам или по атрибутам из A .

Одной из главных задач систем мониторинга является сегментация сцен для дальнейших

■ **Таблица 1.** Формальное представление данных для системы мониторинга

■ **Table 1.** Formal presentation of data for the monitoring system

t	X	A_1	A_2	...	A_j	...	A_m	M
t_1	$x_1(t_1)$	$a_{11}(t_1)$	$a_{21}(t_1)$...	$a_{j1}(t_1)$...	$a_{m1}(t_1)$	M_1
	$x_2(t_1)$	$a_{12}(t_1)$	$a_{22}(t_1)$...	$a_{j2}(t_1)$...	$a_{m2}(t_1)$	
	
	$x_i(t_1)$	$a_{1i}(t_1)$	$a_{2i}(t_1)$...	$a_{ji}(t_1)$...	$a_{mi}(t_1)$	
	
	$x_n(t_1)$	$a_{1n}(t_1)$	$a_{2n}(t_1)$...	$a_{jn}(t_1)$...	$a_{mn}(t_1)$	
...								
t_2	$x_1(t_2)$	$a_{11}(t_2)$	$a_{21}(t_2)$...	$a_{j1}(t_2)$...	$a_{m1}(t_2)$	M_2
	$x_2(t_2)$	$a_{12}(t_2)$	$a_{22}(t_2)$...	$a_{j2}(t_2)$...	$a_{m2}(t_2)$	
	
	$x_i(t_2)$	$a_{1i}(t_2)$	$a_{2i}(t_2)$...	$a_{ji}(t_2)$...	$a_{mi}(t_2)$	
	
	$x_n(t_2)$	$a_{1n}(t_2)$	$a_{2n}(t_2)$...	$a_{jn}(t_2)$...	$a_{mn}(t_2)$	

идентификации интересующих объектов (например, транспортные средства) и последующей их классификации (например, седан). Стандартным подходом сегментации является кластеризация [17]. Кластеризация данных — наиболее используемый метод интеллектуального анализа данных в распределенных системах мониторинга [18]. Цель этого метода заключается в разложении или разбиении набора данных на группы путем минимизации межгруппового несходства и максимизации внутргруппового сходства. При этом требуется организовать передачу данных от датчиков в единое хранилище для последующей кластеризации. В сенсорных системах мониторинга это сложно выполнить по таким причинам, как ограниченная пропускная способность связи или вычислительная мощность. Традиционная структура централизованных алгоритмов кластеризации плохо масштабируется в распределенных приложениях, так как предполагается передача данных на базовый узел. При этом интенсивная передача данных по ограниченному каналу полосы пропускания может привести к снижению времени отклика всей системы.

Еще одной типичной задачей для интеллектуального анализа данных в системах мониторинга является выявление выбросов в данных, например обнаружение химического разлива или вторжений. Кластерный анализ можно отнести к основным подходам обнаружения выбросов. Для таких задач использование централизованных методов кластеризации имеет ранее перечисленные недостатки. Поэтому важное значение приобретает разработка эффективных алгоритмов распределенной кластеризации, которые требуют небольшой полосы пропускания при построении одноранговой среды сетей датчиков.

Основные трудности при решении такой проблемы заключаются в следующем:

- количество обрабатываемых измерений чрезвычайно велико — 10^6 и более;
- пространство объектов, используемое для группировки входных данных в регионы, имеет множество измерений;
- нет предварительной информации о количестве и местонахождении искомым регионов;
- данные распределены на нескольких источниках и не могут быть объединены в центральном хранилище.

Из описанных особенностей работы распределенных систем мониторинга вытекают нижеперечисленные требования к методам кластеризации:

- однопроходность по данным — алгоритм должен выполнять кластеризацию за один проход по всем данным;
- поддержка разных типов входных данных — данные могут быть дискретными, непрерывными, категориальными и др.;

- визуализация кластеров, связей и возможность интерпретации полученной визуализации — одно из главных требований, позволяющее на основе полученных результатов производить анализ и диагностику специалистами для выявления причин выбросов в данных;

- поддержка онлайн-режима и адаптация к данным при изменении среды;

- проецирование многомерного пространства в пространство с более низкой размерностью, что увеличивает скорость обработки многомерных данных, получаемых от множества датчиков системы;

- выделение наиболее значимых атрибутов;

- масштабирование больших объемов данных;

- отсутствие ограничений входных параметров априорными знаниями, например количеством кластеров;

- выполнение анализа без предположений о распределении входных данных;

- обнаружение аномалий;

- анализ данных на источниках информации без их передачи третьей стороне;

- поддержка анализа распределенных данных.

В соответствии с сформулированными требованиями необходимо определить наиболее подходящий метод кластеризации для анализа в системах мониторинга с распределенными источниками данных.

Методы кластерного анализа

С появлением интеллектуального анализа данных было предложено множество методов кластеризации. Кластеризацию можно классифицировать по нескольким категориям [19–22] в соответствии с принципом формирования кластеров.

Иерархическая кластеризация

Иерархическая кластеризация направлена на создание иерархии вложенных разбиений исходного множества объектов. Иерархическая кластеризация представляет иерархию, т. е. структуру, которая является более информативной, чем неструктурированный набор плоских кластеров. Алгоритмы имеют в основном высокие требования по памяти и по времени, что делает их слишком медленными даже для средних наборов данных.

При *агломеративном* подходе («снизу-вверх») новые кластеры создаются путем объединения более мелких кластеров — дерево иерархии строится от листьев к стволу.

В случае *дивизимного* подхода («сверху-вниз») новые кластеры формируются путем деления бо-

лее крупных кластеров на мелкие — дерево иерархии строится от ствола к листьям.

Одним из популярных алгоритмов иерархической кластеризации является BIRCH (Balanced Iterative Reducing and Clustering Using Hierarchies). В отличие от своих аналогов, он был специально разработан для минимизации количества операций ввода-вывода. Алгоритм особенно подходит для больших объемов данных [23]. Он динамически сегментирует входящие многомерные метрические точки данных, чтобы создать кластер с наилучшим качеством и доступными ресурсами (доступной памятью и временными ограничениями) [24]. После первой кластеризации может улучшить качество с помощью нескольких дополнительных сканирований. Он также является первым алгоритмом кластеризации, предложенным в области базы данных для эффективного управления «шумом».

Преимущества:

- 1) гибкость в отношении уровня детализации;
- 2) удобство представления для иерархических данных;

- 3) быстрота обработки.

Недостатки:

- 1) невозможность внести исправления после разделения/слияния;

- 2) отсутствие интерпретируемости дескрипторов кластера;

- 3) нечеткость критерия завершения;

- 4) невозможность работать с асферическими кластерами разного размера;

- 5) неэффективность в высокоразмерных пространствах из-за проклятия размерности;

- 6) чувствительность порядка данных.

Центроидная кластеризация

Задачу центроидной кластеризации [25] можно формально определить следующим образом: необходимо найти центры k кластеров, назначить объекты ближайшему центру кластера и минимизировать квадраты расстояний.

Основным представителем центроидного метода является алгоритм K-Means. Он может организовывать более точные кластеры, чем при иерархической кластеризации. Модель может изменять кластеры (переходить на другой кластер), когда центроиды пересчитываются. При работе с этим методом трудно предсказать количество k кластеров. Начальная инициализация и порядок данных оказывают сильное влияние на результаты. Алгоритм имеет чувствительность к масштабируемости — нормализация и стандартизация влияют на результаты.

Преимущества:

- 1) относительно масштабируемый и простой;

- 2) хорошо подходит для наборов данных с компактными сферическими кластерами, которые качественно разделены.

Недостатки:

- 1) сильное снижение эффективности в высокоразмерных пространствах, поскольку почти все пары точек находятся примерно так же далеко, как и средние;

- 2) плохие дескрипторы кластера;

- 3) опора на экспертные знания пользователя;

- 4) высокая чувствительность к начальным значениям, шуму и выбросам;

- 5) частые остановки на локальных оптимумах;

- 6) низкая точность анализа данных с невыпуклыми кластерами разного размера и плотности.

Плотностная кластеризация

С помощью методов, основанных на плотности, определяются кластеры относительно плотностных регионов [26]. Известны алгоритмы с одним сканированием, например DBSCAN (Density-Based Spatial Clustering of Applications with Noise) и OPTICS (Ordering Points to Identify the Clustering Structure). Ключевым недостатком таких алгоритмов является то, что они ожидают снижения плотности для обнаружения границ кластера.

Преимущества:

- 1) возможность определения кластеров произвольной формы с разным размером;

- 2) сопротивляемость шуму и выбросам.

Недостатки:

- 1) высокая чувствительность к настройке входных параметров;

- 2) плохие дескрипторы кластера;

- 3) низкая эффективность при анализе высокоразмерных наборов данных из-за явления проклятия размерности.

Кластеризация на основе нейронных сетей Кохонена

Эти методы используют нейронные сети, основным элементом которых является слой Кохонена. Слой Кохонена представляет собой адаптивные линейные сумматоры. Выходные сигналы слоя обычно обрабатываются по правилу [27]: наибольший сигнал превращается в единичные, остальные — в ноль. В результате нейроны, обученные на входных данных, образуют области, соответствующие кластерам в этих данных. Примерами такого типа алгоритмов являются SOM или GNG (Growing Neural Gas). GNG относится к алгоритмам [28] Topology Representing Networks [29]. Он использует процесс роста данных для постепенного создания сетевой модели прототипа входного пространства, где результирующая сетевая структура формирует индуцированную триангуляцию Делоне пространства. Это отличает алгоритм GNG от аналогичных, таких как SOM [30], который использует фиксированную топологию сети. Это позволяет GNG

более подробно отражать структуру входного пространства. GNG может создавать непересекающиеся серии, которые указывают возможные классы во входных данных. GNG [31] и другие растущие нейронные сети (такие как Growing Cell Structures [32], Growing Grid [33] и т. д.) используют механизм постепенной адаптации и самостоятельной корректировки своих собственных размеров. GNG-алгоритм имеет сложность $O(n^2)$, где n — количество нейронов (т. е. промежуточных кластеров), и не зависит от количества входных данных. GNG может работать, когда входные данные подчиняются непрерывному распределению. Этот факт является значительным преимуществом, например, по сравнению с алгоритмом K-Means, и позволяет использовать алгоритм для больших объемов входных данных.

В случае применения алгоритма SOM в качестве средства предобработки данных для дальнейшей кластеризации результаты показывают высокую эффективность для сжатия многомерных наборов данных, что делает интерпретацию и диагностику намного проще и очевиднее.

Преимущества:

- 1) определение кластеров произвольной формы с разным размером;
- 2) кластеризация без начальных знаний о данных;
- 3) обработка больших объемов данных;
- 4) сжатие размерности данных.

Недостатки:

- 1) интерпретация результатов кластеризации требует преобразования результатов, например U-matrix [34];
- 2) точность определения кластеров ниже, чем у предопределенных (например, K-Means, при условии верно подобранного количества кластеров k).

По результатам сравнительной характеристики (табл. 2) наиболее эффективными для выполнения кластеризации в распределенных системах мониторинга являются алгоритмы, построенные на основе нейронных сетей Кохонена. Для анализа данных системы мониторинга этот класс методов обладает такими преимуществами, как поддержка разных типов входных данных и работа с высокоразмерными данными. Кроме того, метод позволяет производить визуализацию кластеров и их связей. Кластеризация на основе нейронных сетей Кохонена не требует ввода начальных значений и априорных знаний об исследуемом объекте. Метод адаптируется к данным и может автоматически определять количество кластеров, что делает процесс кластеризации универсальным. Еще одним преимуществом является комбинирование с другими методами кластеризации и уменьшение размерности данных.

■ **Таблица 2.** Сравнительная характеристика методов кластеризации для систем мониторинга

■ **Table 2.** Comparative characteristics of clustering methods for monitoring systems

Требование	Методы			
	Иерархические	Центроидные	Плотностные	Нейронные сети Кохонена
Однопроходность	–	+	+	+
Работа с высокоразмерными данными	–	–	–	+
Поддержка разных типов входных данных	–	+	+	+
Поддержка онлайн-режима (адаптация к данным при их изменении)	–	+	+	+
Работа с большим количеством данных	–	+	+	+
Визуализация связей	–	+	–	++
Визуализация высокоразмерных данных	–	–	–	++
Графическая интерпретация данных	–	–	–	++
Отсутствует необходимость предварительных данных	+	–	+	+
Работа с распределенными данными	++	++	++	++

* — требует предобработку специальными методами.

Стратегии распределенного анализа данных

Для кластеризации распределенных данных необходимо декомпозировать алгоритм на основные блоки [35, 36] — выделить блоки работы с данными [37] и определить требования относительно горизонтального и вертикального распределения данных.

В алгоритмах кластеризации на основе нейронных сетей Кохонена можно выделить общую последовательность блоков:

- 1) инициализация нейронов;
- 2) для каждого вектора:
 - определение нейрона-победителя;
 - вычисление окрестности;
 - корректировка весов и конфигурации сети (также вычисление ошибок и соединений нейро-

нов, корректировка топологии путем добавления и удаления нейронов).

Следующие блоки предполагают работу с данными:

- определение нейрона-победителя (вычисление расстояния от каждого вектора до нейрона);
- корректировку весов (в вычислениях участвуют значения атрибутов векторов).

Соответственно, описанные блоки должны выполняться на узлах источников. Возможны следующие стратегии распределенного выполнения кластеризации:

1) с промежуточной синхронизацией нейрона-победителя, когда результаты отправляются на общий узел, объединяются и обобщенная модель рассылает на узлы источники для выполнения шага;

2) со слиянием нейронных сетей после обработки всех или части входных данных, когда карты строятся отдельно на каждом узле.

Перечисленные стратегии обладают разными характеристиками по времени выполнения и нагрузке на сеть.

Первая стратегия является медленной и создает высокую нагрузку на сеть, так как для каждого вектора выполняется два взаимодействия узлов источников с общим узлом, что замедляет работу алгоритма и увеличивает сетевой трафик.

Вторая стратегия оптимальна и требует одного или нескольких взаимодействий. При этом могут получиться разные результаты в зависимости от частоты слияния n . Так, при $n = 1$ слияние нейронных сетей выполняется для каждого вектора. Это повышает точность модели, но увеличивает число передач данных по сети и, как следствие, время анализа и сетевой трафик. При такой стратегии необходимо учитывать время завершения обработки самого медленного источника. При значении параметра n , равного количеству векторов входных

данных z , снижается количество сетевого трафика, так как слияние результатов выполняется после анализа всех входных векторов данных. Возможны разные промежуточные варианты $1 < n < z$, определяющие слияние нейронных сетей:

- количество обработанных входных векторов данных;
- интервал времени между каждым слиянием (каждую минуту, каждые 10 минут и т. д.);
- промежуток времени после получения входного вектора и т. п.

Выбор стратегии и параметров слияния зависит от условий среды выполнения кластеризации и способа распределения данных.

Заключение

В работе были сформулированы основные требования методов кластеризации для распределенных систем мониторинга. Проведена классификация и обзор методов кластеризации. Выделены основные преимущества и недостатки алгоритмов и методов кластеризации.

Обзор методов кластеризации данных показал, что наиболее эффективным для анализа распределенных систем мониторинга является метод на основе нейронных сетей Кохонена. Однако для его применения к решению задачи требуется его адаптация к распределенному выполнению.

Декомпозирован алгоритм самоорганизующихся карт Кохонена и определены блоки работы с данными: вычисление нейрона-победителя и настройка весов нейронов. Предложены две стратегии кластеризации распределенных данных. Дальнейшие работы будут связаны с исследованием и разработкой кластеризации распределенных данных на основе нейронных сетей с учетом различных видов распределения.

Литература

1. Catarinucci L., et al. An IoT-aware architecture for smart healthcare systems. *IEEE Internet of Things Journal*, 2015, vol. 2, no. 6, pp. 515–526.
2. Andreu-Perez J., et al. From wearable sensors to smart implants — toward pervasive and personalized healthcare. *IEEE Transactions on Biomedical Engineering*, 2015, vol. 62, no. 12, pp. 2750–2762.
3. Del Vecchio P., et al. Creating value from social big data: Implications for smart tourism destinations. *Information Processing & Management*, 2018, vol. 54, no. 5, pp. 847–860.
4. Bendeche M., Kechadi M. T., Le-Khac N. A. Efficient large scale clustering based on data partitioning. *Data Science and Advanced Analytics (DSAA), 2016 IEEE Intern. Conf.*, 2016, pp. 612–621.
5. Aouad L. M., Le-Khac N. A., Kechadi T. M. Lightweight clustering technique for distributed data mining applications. *Industrial Conf. on Data Mining*, Springer, Berlin, Heidelberg, 2007, pp. 120–134.
6. Le-Khac N. A., Aouad L. M., Kechadi M. T. A new approach for distributed density based clustering on grid platform. *British National Conf. on Databases*, Springer, Berlin, Heidelberg, 2007, pp. 247–258.
7. Januzaj E., Kriegel H. P., Pfeifle M. DBDC: Density based distributed clustering. *Intern. Conf. on Extending Database Technology*, Springer, Berlin, Heidelberg, 2004, pp. 88–105.
8. Gorgonio F. L., Costa J. A. F. Parallel self-organizing maps with application in clustering distributed data. *IEEE Intern. Joint Conf. Neural Networks (IJCNN 2008)*, 2008, pp. 3276–3283.

9. Kim M., Jung S., Park M. A distributed self-organizing map for DoS attack detection. *Seventh Intern. Conf. Ubiquitous and Future Networks (ICUFN)*, 2015, pp. 19–22.
10. Ali A. A., et al. Distributed data mining systems: Techniques, approaches and algorithms. *22nd Intern. Conf. on Circuits, Systems, Communications and Computers (CSCC 2018)*, Majorca, Spain, July 14–17, 2018, vol. 210, p. 04038.
11. Song C., et al. Updating strategy of master-slave data consistency based on message. *Computer Engineering*, 2004, vol. 1, pp. 035.
12. Robert H. A majority consensus approach to concurrency control for multiple copy databases. *ACM Trans. on Database System*, 2003, vol. 4, no. 2, pp. 543–549.
13. Moro G., Sartori C. Incremental maintenance of multi-source views. *Proc. of the 12th Australasian Database Conf.*, IEEE Computer Society, 2001, pp. 13–20.
14. Oussous A., et al. Big data technologies: A survey. *Journal of King Saud University-Computer and Information Sciences*, 2018, vol. 30, no. 4, pp. 431–448.
15. Li T., et al. Differentially private naive bayes learning over multiple data sources. *Information Sciences*, 2018, vol. 444, pp. 89–104.
16. Assunção M. D., et al. Big data computing and clouds: Trends and future directions. *Journal of Parallel and Distributed Computing*, 2015, vol. 79, pp. 3–15.
17. Duda R. O., Hart P. E., Stork D. G. *Pattern classification*. John Wiley & Sons, 2012. 688 p.
18. Bendeche M., Kechadi M. T. Distributed clustering algorithm for spatial data mining. *2nd IEEE Intern. Conf. on Spatial Data Mining and Geographical Knowledge Services (ICSMD)*, IEEE, 2015, pp. 60–65.
19. Saraswathi S., Sheela M. I. A comparative study of various clustering algorithms in data mining. *International Journal of Computer Science and Mobile Computing*, 2014, vol. 11, no. 11, pp. 422–428.
20. Sisodia D., et al. Clustering techniques: a brief survey of different clustering algorithms. *International Journal of Latest Trends in Engineering and Technology (IJLTET)*, 2012, vol. 1, no. 3, pp. 82–87.
21. Qian W., Zhou A. Y. Analyzing popular clustering algorithms from different viewpoints. *Journal of Software*, 2002, vol. 13, no. 8, pp. 1382–1394.
22. Барсегян А. А., Куприянов М. С., Степаненко В. В., Холод И. И. *Технологии анализа данных: Data Mining, Visual Mining, Text Mining, OLAPM*. СПб., БХВ-Петербург, 2007. 384 с.
23. Bindra K., et al. Effective Data Clustering Algorithms. *Soft Computing: Theories and Applications: Proc. of SoCTA 2017*, Singapore, Springer, 2019, pp. 419–432.
24. Zhang T., Ramakrishnan R., Livny M. BIRCH: an efficient data clustering method for very large databases. *ACM Sigmod Record*, ACM, 1996, vol. 25, no. 2, pp. 103–114.
25. Uppada S. K. Centroid based clustering algorithms-A clarion study. *International Journal of Computer Science and Information Technologies*, 2014, vol. 5, no. 6, pp. 7309–7313.
26. Shah H., Napanda K., D'mello L. Density based clustering algorithms. *International Journal of Computer Sciences and Engineering*, 2015, vol. 3, no. 11, pp. 54–57.
27. Kerdels J., Peters G. A sparse representation of high-dimensional input spaces based on an augmented growing neural gas. *GCAI*, 2016, pp. 303–313.
28. Fritzke B. A growing neural gas network learns topologies. *Proc. of the Conf. "Advances in neural information processing systems"*, 1995, pp. 625–632.
29. Martinetz T., Schulten K. Topology representing networks. *Neural Networks*, 1994, vol. 7, no. 3, pp. 507–522.
30. Kohonen T. Self-organized formation of topologically correct feature maps. *Biological Cybernetics*, 1982, vol. 43, no. 1, pp. 59–69.
31. Fišer D., Faigl J., Kulich M. Growing neural gas efficiently. *Neurocomputing*, 2013, vol. 104, pp. 72–82.
32. Fritzke B. Growing cell structures — a self-organizing network in k dimensions. *Proc. of the 1992 Intern. Conf. on Artificial Neural Networks (ICANN-92)*, Brighton, UK, 4–7 September, 1992, 1992, pp. 1051–1056.
33. Fritzke B. Growing grid — a self-organizing network with constant neighborhood range and adaptation strength. *Neural Processing Letters*, 1995, vol. 2, no. 5, pp. 9–13.
34. Ultsch A. Self organized feature maps for monitoring and knowledge aquisition of a chemical process. *ICANN'93*, Springer, London, 1993, pp. 864–867.
35. Kholod I., Kuprianov M., Petukhov I. Distributed data mining based on actors for internet of things. *5th Mediterranean Conf. on Embedded Computing (MECO)*, June 12–16, 2016, Bar, Montenegro, 2016, pp. 480–484. doi:10.1109/MECO.2016.7525698
36. Kholod I., Petuhov I., Malyskin V. Creation of data mining algorithms as functional expression for parallel and distributed execution. *Parallel Computing Technologies, LNCS*, Springer, 2015, vol. 9251, pp. 62–68.
37. Kholod I., Kupriyanov M., Shorov A. Decomposition of data mining algorithms into unified functional blocks. *Mathematical Problems in Engineering*, 2016, vol. 2016, p. 11.

UDC 004.75

doi:10.31799/1684-8853-2019-2-35-43

Data clustering in distributed monitoring systemsA. N. Rukavitsyn^a, Post-Graduate Student, orcid.org/0000-0001-5382-0465, rkvtsn@gmail.com^aSaint-Petersburg Electrotechnical University «LETI», 5, Prof. Popov St., 197376, Saint-Petersburg, Russian Federation

Introduction: The traditional methods of analyzing distributed data sources typically use centralized data warehouses and have several disadvantages associated with confidentiality, a high cost of centralized data storage, a limited bandwidth and a high load on telecommunications networks. Techniques for a decentralized analysis do not take into account the data distribution type and the features of the selected algorithm. This reduces the performance and accuracy, or can be impracticable under certain conditions. **Purpose:** Studying and analyzing the features of distributed monitoring systems and data mining algorithms. **Results:** For clustering based on distributed data sources, requirements were set for an algorithm in distributed monitoring systems: single pass, support of different types of input data, online operation and adaptation to data when the environment changes, large data scaling, analysis without assumptions about the input data distribution, analysis of data on information sources without involving a third party. Two main ways of data distribution on sources in heterogeneous systems are defined: vertical and horizontal ones. The methods are classified according to their basic principle of cluster delimitation. The classification includes the main clustering algorithms, their operation principles, advantages and disadvantages. A review of the existing clustering methods has shown that in distributed monitoring systems, Kohonen's neural networks are most efficient. The algorithm of Kohonen self-organizing maps was decomposed, and blocks for data processing were determined: the calculation of a winner neuron and the adjustment of neuron weights. Two strategies have been proposed for clustering distributed data. **Practical relevance:** The proposed strategies allow you to perform clustering in systems with distributed sources taking into account the characteristics of the environment, without transferring all the data.

Keywords — data mining, clustering, sensors, distributed data sources, Kohonen self-organizing maps, distributed data processing systems, types of distribution, monitoring systems.

For citation: Rukavitsyn A. N. Data clustering in distributed monitoring systems. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2019, no. 2, pp. 35–43 (In Russian). doi:10.31799/1684-8853-2019-2-35-43

References

- Catarinucci L., et al. An IoT-aware architecture for smart healthcare systems. *IEEE Internet of Things Journal*, 2015, vol. 2, no. 6, pp. 515–526.
- Andreu-Perez J., et al. From wearable sensors to smart implants — toward pervasive and personalized healthcare. *IEEE Transactions on Biomedical Engineering*, 2015, vol. 62, no. 12, pp. 2750–2762.
- Del Vecchio P., et al. Creating value from social big data: Implications for smart tourism destinations. *Information Processing & Management*, 2018, vol. 54, no. 5, pp. 847–860.
- Bendechache M., Kechadi M. T., Le-Khac N. A. Efficient large scale clustering based on data partitioning. *Data Science and Advanced Analytics (DSAA), 2016 IEEE Intern. Conf.*, 2016, pp. 612–621.
- Aouad L. M., Le-Khac N. A., Kechadi M. T. Lightweight clustering technique for distributed data mining applications. *Industrial Conf. on Data Mining*, Springer, Berlin, Heidelberg, 2007, pp. 120–134.
- Le-Khac N. A., Aouad L. M., Kechadi M. T. A new approach for distributed density based clustering on grid platform. *British National Conf. on Databases*, Springer, Berlin, Heidelberg, 2007, pp. 247–258.
- Januzaj E., Kriegel H. P., Pfeifle M. DBDC: Density based distributed clustering. *Intern. Conf. on Extending Database Technology*, Springer, Berlin, Heidelberg, 2004, pp. 88–105.
- Gorgonio F. L., Costa J. A. F. Parallel self-organizing maps with application in clustering distributed data. *IEEE Intern. Joint Conf. Neural Networks (IJCNN 2008)*, 2008, pp. 3276–3283.
- Kim M., Jung S., Park M. A distributed self-organizing map for DoS attack detection. *Seventh Intern. Conf. Ubiquitous and Future Networks (ICUFN)*, 2015, pp. 19–22.
- Ali A. A., et al. Distributed data mining systems: Techniques, approaches and algorithms. *22nd Intern. Conf. on Circuits, Systems, Communications and Computers (CSCC 2018)*, Majorca, Spain, July 14–17, 2018, vol. 210, p. 04038.
- Song C., et al. Updating strategy of master-slave data consistency based on message. *Computer Engineering*, 2004, vol. 1, pp. 035.
- Robert H. A majority consensus approach to concurrency control for multiple copy databases. *ACM Trans. on Database System*, 2003, vol. 4, no. 2, pp. 543–549.
- Moro G., Sartori C. Incremental maintenance of multi-source views. *Proc. of the 12th Australasian Database Conf.*, IEEE Computer Society, 2001, pp. 13–20.
- Oussous A., et al. Big data technologies: A survey. *Journal of King Saud University-Computer and Information Sciences*, 2018, vol. 30, no. 4, pp. 431–448.
- Li T., et al. Differentially private naive bayes learning over multiple data sources. *Information Sciences*, 2018, vol. 444, pp. 89–104.
- Assunção M. D., et al. Big data computing and clouds: Trends and future directions. *Journal of Parallel and Distributed Computing*, 2015, vol. 79, pp. 3–15.
- Duda R. O., Hart P. E., Stork D. G. *Pattern classification*. John Wiley & Sons, 2012. 688 p.
- Bendechache M., Kechadi M. T. Distributed clustering algorithm for spatial data mining. *2nd IEEE Intern. Conf. on Spatial Data Mining and Geographical Knowledge Services (ICSDM)*, IEEE, 2015, pp. 60–65.
- Saraswathi S., Sheela M. I. A comparative study of various clustering algorithms in data mining. *International Journal of Computer Science and Mobile Computing*, 2014, vol. 11, no. 11, pp. 422–428.
- Sisodia D., et al. Clustering techniques: a brief survey of different clustering algorithms. *International Journal of Latest Trends in Engineering and Technology (IJLTET)*, 2012, vol. 1, no. 3, pp. 82–87.
- Qian W., Zhou A. Y. Analyzing popular clustering algorithms from different viewpoints. *Journal of Software*, 2002, vol. 13, no. 8, pp. 1382–1394.
- Barsegyan A. A., Kupriyanov M. S., Stepanenko V. V., Kholod I. I. *Tekhnologii analiza dannykh. Data Mining, Visual Mining, Text Mining, OLAP* [Technologies of data analysis. Data Mining, Visual Mining, Text Mining, OLAP]. Saint-Petersburg, BKhV-Peterburg Publ., 2007. 384 p. (In Russian).
- Bindra K., et al. Effective Data Clustering Algorithms. *Soft Computing: Theories and Applications (SoCTA 2017)*, Springer, Singapore, 2019, pp. 419–432.
- Zhang T., Ramakrishnan R., Livny M. BIRCH: an efficient data clustering method for very large databases. *ACM Sigmod Record*, ACM, 1996, vol. 25, no. 2, pp. 103–114.
- Uppada S. K. Centroid based clustering algorithms-A clarification study. *International Journal of Computer Science and Information Technologies*, 2014, vol. 5, no. 6, pp. 7309–7313.
- Shah H., Napanda K., D'mello L. Density based clustering algorithms. *International Journal of Computer Sciences and Engineering*, 2015, vol. 3, no. 11, pp. 54–57.

27. Kerdels J., Peters G. A sparse representation of high-dimensional input spaces based on an augmented growing neural gas. *GCAI*, 2016, pp. 303–313.
28. Fritzke B. A growing neural gas network learns topologies. *Proc. of the Conf. "Advances in neural information processing systems"*, 1995, pp. 625–632.
29. Martinetz T., Schulten K. Topology representing networks. *Neural Networks*, 1994, vol. 7, no. 3, pp. 507–522.
30. Kohonen T. Self-organized formation of topologically correct feature maps. *Biological Cybernetics*, 1982, vol. 43, no. 1, pp. 59–69.
31. Fišer D., Faigl J., Kulich M. Growing neural gas efficiently. *Neurocomputing*, 2013, vol. 104, pp. 72–82.
32. Fritzke B. Growing cell structures — a self-organizing network in k dimensions. *Proc. of the 1992 Intern. Conf. on Artificial Neural Networks (ICANN-92)*, Brighton, UK, 4–7 September, 1992, 1992, pp. 1051–1056.
33. Fritzke B. Growing grid — a self-organizing network with constant neighborhood range and adaptation strength. *Neural Processing Letters*, 1995, vol. 2, no. 5, pp. 9–13.
34. Ultsch A. Self organized feature maps for monitoring and knowledge acquisition of a chemical process. *ICANN'93*, Springer, London, 1993, pp. 864–867.
35. Kholod I., Kuprianov M., Petukhov I. Distributed data mining based on actors for internet of things. *5th Mediterranean Conf. on Embedded Computing (MECO)*, June 12–16, 2016, Bar, Montenegro, 2016, pp. 480–484. doi:10.1109/MECO.2016.7525698
36. Kholod I., Petuhov I., Malyshkin V. Creation of data mining algorithms as functional expression for parallel and distributed execution. *Parallel Computing Technologies, LNCS*, Springer, 2015, vol. 9251, pp. 62–68.
37. Kholod I., Kupriyanov M., Shorov A. Decomposition of data mining algorithms into unified functional blocks. *Mathematical Problems in Engineering*, 2016, vol. 2016, p. 11.

УВАЖАЕМЫЕ АВТОРЫ!

Научные базы данных, включая SCOPUS и Web of Science, обрабатывают данные автоматически. С одной стороны, это ускоряет процесс обработки данных, с другой — различия в транслитерации ФИО, неточные данные о месте работы, области научного знания и т. д. приводят к тому, что в базах оказывается несколько авторских страниц для одного и того же человека. В результате для всех по отдельности считаются индексы цитирования, снижая рейтинг ученого.

Для идентификации авторов в сетях Thomson Reuters проводит регистрацию с присвоением уникального индекса (ID) для каждого из авторов научных публикаций.

Процедура получения ID бесплатна и очень проста, есть возможность провести регистрацию на 12-ти языках, включая русский (чтобы выбрать язык, кликните на зеленое поле вверх справа на стартовой странице): <https://orcid.org>

УДК 628.3

doi:10.31799/1684-8853-2019-2-44-51

Математическая модель анаэробного биореактора с закрепленной биомассой как объекта управления

А. А. Ключарёв^а, канд. техн. наук, доцент, orcid.org/0000-0001-5345-1209А. А. Фоменкова^а, ассистент, orcid.org/0000-0002-5570-6743, a.a.fomenkova@mail.ru^аСанкт-Петербургский государственный университет аэрокосмического приборостроения, Б. Морская ул., 67, Санкт-Петербург, 190000, РФ

Постановка проблемы: от эффективности работы биоочистных сооружений зависят производительность и экологическая безопасность предприятий. Решение задачи управления процессами анаэробной биологической очистки сточных вод существенно ограничивается сложностью описания физических и биохимических процессов, происходящих в биореакторах. **Цель:** разработка обобщенной математической модели анаэробного биореактора с закрепленной биомассой как объекта системы автоматического управления, позволяющей учитывать конструктивные и технологические особенности системы очистки сточных вод. **Методы:** имитационное моделирование процессов массопереноса и биохимических преобразований в анаэробном биореакторе и разработка на его основе математической модели, применимой для синтеза системы управления. **Результаты:** на основе уравнений массопереноса получена обобщенная математическая модель анаэробного биореактора с закрепленной биомассой. Эта модель дополнена компонентами, учитывающими биохимические преобразования. На базе конструктивных и технологических параметров биореактора сделаны необходимые допущения, которые дают возможность сформулировать граничные и начальные условия с достаточной для инженерной практики точностью. Обоснован выбор управляемых параметров технологического процесса в анаэробном биореакторе, обеспечивающих достижение необходимой степени очистки при наименьших затратах. Полученная обобщенная математическая модель анаэробного биореактора с закрепленной биомассой позволяет выполнить имитационное моделирование биореактора в заданных условиях эксплуатации. Для этих условий построена аппроксимирующая модель, благодаря которой возможно выполнение синтеза системы управления. **Практическая значимость:** разработанные алгоритмы позволяют применить полученные результаты к широкому классу действующих анаэробных биореакторов с закрепленной биомассой и построить системы управления, существенно повышающие эффективность их работы, в том числе на этапах модернизации систем очистки.

Ключевые слова — система автоматической стабилизации, анаэробный биореактор, математическая модель, имитационное моделирование.

Для цитирования: Ключарёв А. А., Фоменкова А. А. Математическая модель анаэробного биореактора с закрепленной биомассой как объекта управления. *Информационно-управляющие системы*, 2019, № 2, с. 44–51. doi:10.31799/1684-8853-2019-2-44-51

For citation: Klucharev A. A., Fomenkova A. A. Mathematical model of a fixed-biomass anaerobic bioreactor as a control object. *Information and Control Systems*, 2019, no. 2, pp. 44–51 (In Russian). doi:10.31799/1684-8853-2019-2-44-51

Введение

Биологическая очистка сточных вод, являясь основной технологией обеспечения экологической безопасности промышленных предприятий, широко используется как в промышленности, так и в локальных бытовых очистных сооружениях. Известна высокая эффективность применения анаэробных биореакторов [1, 2], работа которых сопровождается выделением биогаза, используемого, например, для обеспечения оптимального температурного режима.

При проектировании биореактора [3–5] центральное внимание уделяется его конструктивным и технологическим параметрам на основе заданных характеристик обрабатываемых сточных вод и допустимой концентрации загрязнений в очищенной воде $S_{\text{вых}}$.

Зачастую промышленные предприятия характеризуются неравномерным во времени выходом продукции, а следовательно, неравномерным во времени объемом и составом сточных вод. Это обу-

словлено цикличностью технологического процесса, сезонностью спроса на некоторые виды продукции, необходимостью санитарного и технического обслуживания технологических линий и другими причинами. Очистные сооружения должны обеспечить заданную санитарными нормами степень очистки при изменяющихся концентрации загрязнений на входе S_0 и объеме сточных вод Q , поступающих с предприятия на очистку в единицу времени. При этом следует учитывать, что эффективность работы биореактора существенно зависит от климатических условий, в частности, температуры окружающей среды.

В современных условиях с учетом накопленного опыта проектирования и эксплуатации анаэробных биореакторов актуальна задача разработки и внедрения автоматизированных систем контроля и управления работой очистных сооружений [6–9]. Задачей управления биореактором является обеспечение допустимого уровня концентрации загрязнений на выходе биореактора $S_{\text{вых}}$ при изменениях начальной концентрации загрязнений

сточной воды S_0 , расхода жидкости в биореакторе Q и температуры окружающей среды $t_{окр}$. Решение таких задач позволяет уменьшить эксплуатационные затраты, обеспечить эффективность системы очистки при пиковых нагрузках и залповых выбросах загрязнений, повысить безопасность системы очистки. Рассматриваемая система относится к классу непрерывных систем автоматической стабилизации с входными параметрами S_0, Q . Заданный режим работы биореактора обеспечивается автоматической регулировкой рабочей температуры очищаемой жидкости t .

Объем жидкости Q , поступающей в биореактор в единицу времени, определяет время пребывания сточной воды на очистке и, следовательно, производительность биореактора по сточной воде. Увеличение расхода жидкости приводит к уменьшению времени контакта сточной воды с микроорганизмами биореактора и, как следствие, к снижению показателей качества очистки. Этот эффект компенсируется увеличением скорости биохимических реакций при повышении температуры сбраживания. Однако дополнительный обогрев биореактора существенно увеличивает затраты на очистку.

При решении задач управления биореактором возникает необходимость разработки математической модели объекта управления, учитывающей влияние технологических и возмущающих факторов на эффективность его функционирования. Для ряда конструкций биореакторов такие модели известны, и строятся они на основе описания процессов массопереноса и микробиологической кинетики преобразования органического загрязнения [3, 4, 6]. В конструкции, описанной в [10], повышение эффективности работы биореактора достигается использованием нескольких секций с микроорганизмами, закрепленными на носителях биомассы. Подобные биореакторы могут быть использованы для очистки сточных вод, содержащих липофильные вещества [11]. Такое техническое решение требует уточнения математической модели для решения задачи синтеза алгоритмов управления.

Математическое описание процессов массопереноса в анаэробном биореакторе с закрепленной биомассой

В основе математического описания процессов распределения и преобразования веществ в биореакторе лежат уравнения массопереноса, в частности, уравнения конвективной диффузии вида [12]

$$\frac{\partial S}{\partial \tau} = D \left(\frac{\partial^2 S}{\partial x^2} + \frac{\partial^2 S}{\partial y^2} + \frac{\partial^2 S}{\partial z^2} \right) -$$

$$- \left(W_x \frac{\partial S}{\partial x} + W_y \frac{\partial S}{\partial y} + W_z \frac{\partial S}{\partial z} \right) + r_1 + r_2, \quad (1)$$

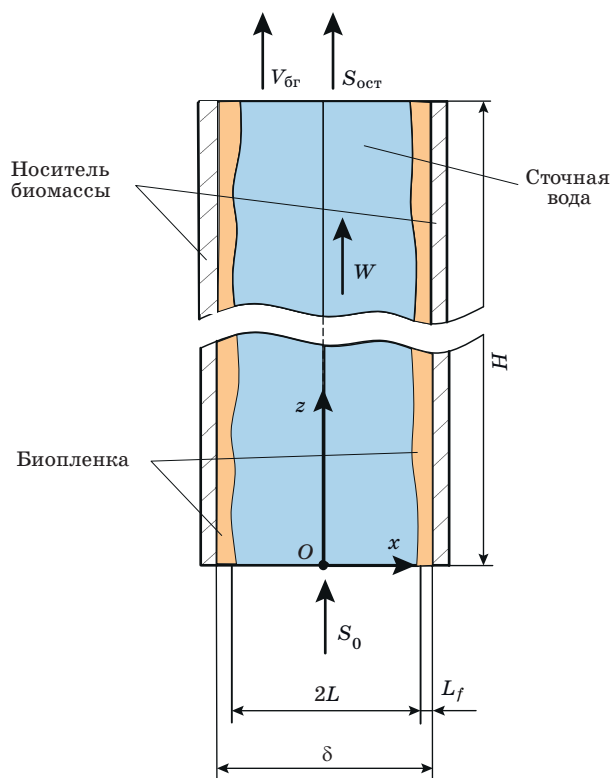
где $S = S(x, y, z)$ — концентрация рассматриваемого вещества; D — коэффициент молекулярной диффузии вещества в жидкости; $W = W(x, y, z)$ — скорость потока жидкости; r_1, r_2 — скорости соответственно образования и разложения рассматриваемого вещества анаэробной микрофлорой.

Процессы анаэробного разложения органического загрязнения сточной воды с достаточной степенью точности можно представить в виде двух стадий [13–15] — кислотогенеза и метаногенеза. На стадии кислотогенеза исходное загрязнение, выраженное обобщенной концентрацией глюкозы $S_{гр}$, преобразуется микроорганизмами в ацетат (уксусную кислоту) с обобщенной концентрацией S_y . В результате метаногенеза из уксусной кислоты образуется биогаз с концентрацией метана $S_{бр}$. Концентрации перечисленных веществ изменяются в соответствии с уравнением (1), а обобщенная модель процессов массопереноса имеет вид

$$\left\{ \begin{aligned} \frac{\partial S_{гр}}{\partial \tau} &= D_{гр} \left(\frac{\partial^2 S_{гр}}{\partial x^2} + \frac{\partial^2 S_{гр}}{\partial y^2} + \frac{\partial^2 S_{гр}}{\partial z^2} \right) - \\ &- \left(W_x \frac{\partial S_{гр}}{\partial x} + W_y \frac{\partial S_{гр}}{\partial y} + W_z \frac{\partial S_{гр}}{\partial z} \right) + r_{2гр}; \\ \frac{\partial S_y}{\partial \tau} &= D_y \left(\frac{\partial^2 S_y}{\partial x^2} + \frac{\partial^2 S_y}{\partial y^2} + \frac{\partial^2 S_y}{\partial z^2} \right) - \\ &- \left(W_x \frac{\partial S_y}{\partial x} + W_y \frac{\partial S_y}{\partial y} + W_z \frac{\partial S_y}{\partial z} \right) + r_1 + r_{2y}; \\ \frac{\partial S_{бр}}{\partial \tau} &= D_{бр} \left(\frac{\partial^2 S_{бр}}{\partial x^2} + \frac{\partial^2 S_{бр}}{\partial y^2} + \frac{\partial^2 S_{бр}}{\partial z^2} \right) - \\ &- \left(W_x \frac{\partial S_{бр}}{\partial x} + W_y \frac{\partial S_{бр}}{\partial y} + W_z \frac{\partial S_{бр}}{\partial z} \right) + r_{1бр}. \end{aligned} \right. \quad (2)$$

Различия в конструкции и принципе действия биореактора определяют специфику граничных и начальных условий для решения системы уравнений (2). В частности, показанная на рис. 1 расчетная схема для биореактора с закрепленной биомассой позволяет сформулировать начальные и граничные условия для системы уравнений (2). Биомасса в предложенном биореакторе закреплена на носителях, представляющих собой листы полимерного материала. Листы расположены параллельно друг другу и образуют щелевидные каналы, вдоль которых восходящим потоком протекает очищаемая жидкость.

Рассмотрим щелевидный канал, образованный двумя соседними листами-носителями биомассы. В нижнее сечение канала поступает жид-



■ **Рис. 1.** Расчетная схема канала между двумя листами-носителями биомассы

■ **Fig. 1.** The design scheme of the channel between two sheets of biomass carriers

кость с расходом Q и средней концентрацией загрязнений S_0 . По мере продвижения по каналу со скоростью W жидкость взаимодействует с микроорганизмами в биопленке, за счет чего концентрация загрязнений уменьшается и на выходе из канала равна $S_{огр}$. Микроорганизмы биопленки преобразуют загрязнения сточной воды в биогаз объемом $V_{огр}$, который отводится из канала. Рабочая ширина канала $2L$ определяется расстоянием δ между листами-носителями биомассы и толщиной биопленки L_f как $2L = \delta - 2L_f$. За счет идентичности процессов в каждом канале такое представление может быть положено в основу общей модели анаэробного биореактора с закрепленной биомассой.

Для определения начальных и граничных условий при решении системы уравнений (2) применительно к конкретной конструкции биореактора будем полагать:

— по окончании периода запуска биореактора процесс можно считать стационарным и $\frac{dS}{dt} = 0$;

— перенос вещества осуществляется по осям Ox и Oz , составляющие концентрации $\frac{\partial S}{\partial y}$ и скорости W_y равны нулю;

— при низких скоростях движения жидкости поток имеет ламинарный характер и $W_z \gg W_x$. Следовательно, по оси Oz массоперенос осуществляется посредством молекулярной диффузии, а

по оси Ox — посредством конвекции, и $D \frac{d^2 S}{dz^2} = 0$,

$$W_x \frac{dS}{dx} = 0.$$

Для ламинарного потока жидкости в щелевидном канале вертикальная компонента скорости W_z определяется из уравнения Навье — Стокса [12] как

$$W_z = -\frac{3W}{2L^2}x^2 + \frac{3}{2}W, \quad (3)$$

где $W = \frac{Q}{2Ll}$ — средняя скорость потока жидкости; Q — объемный расход жидкости в канале, m^3/c ; l — длина канала, м.

Тогда, с учетом сделанных допущений, уравнение (1) примет вид

$$\frac{\partial^2 S(x, z)}{\partial x^2} - \xi(x) \frac{\partial S(x, z)}{\partial z} + \frac{1}{D}(r_1 + r_2) = 0, \quad (4)$$

где $\xi(x) = k(x^2 - b)$, $k = -\frac{3W}{2DL^2}$, $b = -\frac{1}{L^2}$.

В анаэробном биореакторе с биомассой, закрепленной на носителях, образование и разложение вещества происходит преимущественно в биопленке. На стадии кислотогенеза изменения концентрации глюкозы в ходе микробиологической реакции описываются уравнениями [13, 16, 17]

$$\begin{aligned} r_{1г} &= 0; \\ r_{2г} &= r_{2г}(X_г, S_{г_{огр}}(z), K_г) = \\ &= -X_г \left(\frac{\mu_г}{Y_{X_г}} + K_{SX_г} + K_{mX_г} \frac{S_{г_{огр}}(z)}{K_{S_г} + S_{г_{огр}}(z)} \right). \end{aligned} \quad (5)$$

Концентрация уксусной кислоты на стадии кислотогенеза увеличивается в соответствии с уравнением

$$\begin{aligned} r_{1у} &= r_{1у}(X_г, S_{г_{огр}}(z), K_г) = \\ &= Y_{S_г} X_г \left(K_{SX_г} + K_{mX_г} \frac{S_{г_{огр}}(z)}{K_{S_г} + S_{г_{огр}}(z)} \right). \end{aligned} \quad (6)$$

Для стадии метаногенеза уменьшение концентрации уксусной кислоты и соответствующее увеличение концентрации биогаза описываются уравнениями

$$r_{2y} = r_{2y}(X_y, S_{y_{6н}}(z), K_y) = -X_y \left(\frac{\mu_y}{Y_{X_y}} + K_{SX_y} + K_{mX_y} \frac{S_{y_{6н}}(z)}{K_{S_y} + S_{y_{6н}}(z)} \right); \quad (7)$$

$$r_{16г} = r_{16г}(X_y, S_{y_{6н}}(z), K_y) = V_{m \max} X_y \frac{S_{y_{6н}}(z)}{K_m + S_{y_{6н}}(z)} \frac{K_{im}}{K_{im} + S_{y_{6н}}(z)}. \quad (8)$$

В формулах (6)–(8) приняты следующие обозначения: $X_{\Gamma} = X_{\Gamma}(S_{\Gamma_{6н}}(z), I_{\Gamma}(z), \text{pH}, t)$ и $X_y = X_y(S_{y_{6н}}(z), I_y(z), \text{pH}, t)$ — соответственно концентрация кислотогенной и метаногенной биомассы в биопленке; $\mu_y, Y_{X_{\Gamma}}, K_{SX_{\Gamma}}, K_{mX_{\Gamma}}, K_{S_{\Gamma}}, Y_{S_{\Gamma}}, \mu_y, Y_{X_y}, K_{SX_y}, K_{mX_y}, K_{S_y}, Y_{S_y}, V_{m \max}, K_m, K_{im}$ — кинетические параметры процесса, которые характеризуют особенности жизнедеятельности кислотогенной и метаногенной популяций микроорганизмов. Подробное описание кинетических параметров и их значения, определенные опытным путем, приведены в работах [18–21].

Таким образом, расчетная модель процессов массопереноса в анаэробном биореакторе с закрепленной биомассой примет вид

$$\begin{cases} \frac{\partial^2 S_{\Gamma}(x, z)}{\partial x^2} - \xi_{\Gamma}(x) \frac{\partial S_{\Gamma}(x, z)}{\partial z} + \\ + \frac{1}{D} r_{\Gamma 2}(X_{\Gamma}, S_{\Gamma_{6н}}(z), K_{\Gamma}) = 0; \end{cases} \quad (9.1)$$

$$D_{\Gamma_{6н}} \frac{\partial^2 S_{\Gamma_{6н}}(z)}{\partial x^2} + r_{2\Gamma}(X_{\Gamma}, S_{\Gamma_{6н}}(z), K_{\Gamma}) = 0; \quad (9.2)$$

$$\begin{cases} D_{y_{6н}} \frac{\partial^2 S_{y_{6н}}(z)}{\partial x^2} + r_{1y}(X_y, S_{y_{6н}}(z), K_y) + \\ + r_{2y}(X_y, S_{y_{6н}}(z), K_y) = 0; \end{cases} \quad (9.3)$$

$$\begin{cases} \frac{\partial^2 S_y(x, z)}{\partial x^2} - \xi_y(x) \frac{\partial S_y(x, z)}{\partial z} + \\ + \frac{1}{D} \left(r_{1y}(X_y, S_{y_{6н}}(z), K_y) + \right. \\ \left. + r_{2y}(X_y, S_{y_{6н}}(z), K_y) \right) = 0; \end{cases} \quad (9.4)$$

$$\left(\frac{\partial S_{\Gamma_{6н}}}{\partial t} \right)_R = r_{16г}(X_y, S_{y_{6н}}(z), K_y). \quad (9.5)$$

Для уравнения (9.1), описывающего изменение концентрации глюкозы в жидкости, граничные условия определяются концентрацией глюкозы на границе раздела фаз биопленка — жидкость и в середине канала, где наблюдается экстремум и

$$\frac{\partial S_{\Gamma}}{\partial z} \Big|_{x=0, z} = 0, \text{ и для уравнения (9.2), описывающе}$$

го преобразование глюкозы в биопленке, граничными условиями являются концентрации глюкозы на границах раздела фаз биопленка — жидкость и биопленка — подложка, где $S_{\Gamma_{6н}} \Big|_{x=\delta, z} = 0$. Так как закон изменения концентрации глюкозы на границе раздела фаз биопленка — жидкость заранее неизвестен, при численном решении с достаточно малым шагом сетки предлагается определять концентрацию вещества в биопленке на i -м шаге решения из концентрации вещества в жидкости на $(i-1)$ -м шаге. При условии, что входящий поток жидкости идеально гомогенизирован, начальные условия определяются концентрацией загрязнений в стоке, подающемся на очистку: $S_{\Gamma} \Big|_{x, z=0} = S_0$.

Аналогично определяются начальные и граничные условия для вычисления концентрации уксусной кислоты [уравнения (9.3) и (9.4)] и биогаза [уравнение (9.5)].

В целом система уравнений (9.1)–(9.5) может рассматриваться как обобщенная модель биореактора, однако перечисленные особенности ограничивают возможности аналитического решения системы уравнений (9.1)–(9.5) и приводят к необходимости разработки алгоритмов имитационного моделирования процессов в биореакторе.

Концентрации биомассы X_y и X_{Γ} в общем случае являются сложными функциями, зависящими от концентрации субстрата S , концентрации ингибирующих веществ I , кислотности pH, температуры t [22, 23]. Уравнения системы (9.1)–(9.5) содержат ряд кинетических параметров процесса, которые характеризуют особенности жизнедеятельности кислотогенной и метаногенной популяций микроорганизмов. Эти параметры, так же как и концентрация биомассы в биопленке, определяются опытным путем, зависят от типа микроорганизмов и от температуры, при которой развивается популяция. Так, при повышении рабочей температуры среды интенсифицируются процессы роста микробной популяции, ускоряются биохимические превращения, следовательно, повышается эффективность очистки стока.

Для улучшения эксплуатационных характеристик анаэробных биореакторов используются секционные конструкции [10]. Каждая секция содержит блоки загрузки носителей биомассы, образующие одинаковые вертикальные каналы, процессы в которых идентичны и описываются системой уравнений (9.1)–(9.5). Секции связаны между собой перетоками, где происходит выравнивание концентраций веществ и подогрев очищаемой жидкости до рабочей температуры. В каждой секции обеспечивается отвод образованного биогаза, который может быть использован для обогрева биореактора. Рассмотрим особенности имитационного моделирования секционных биореакторов подобной конструкции [10].

Имитационное моделирование и численные расчеты секционного биореактора

В процессе технической эксплуатации анаэробного биореактора с помощью специальных датчиков должны контролироваться значения таких показателей процесса анаэробной очистки, как концентрация загрязнений на входе S_0 и выходе $S_{\text{ВЫХ}}$ биореактора, температура t_0 поступающей на очистку жидкости и рабочая температура t в биореакторе, расход поступающей в биореактор жидкости Q . Параметр Q позволяет рассчитать среднюю скорость жидкости в канале $W_0 = Q/(F - nL_{\text{л}}y)$, где F — площадь сечения секции; n — количество листов-носителей биомассы; $L_{\text{л}}$ — толщина листа; y — ширина листа. Концентрация загрязнений на выходе биореактора $S_{\text{ВЫХ}}$ является регулируемым параметром, значение которого необходимо стабилизировать на заданном уровне. Рабочую температуру процесса t можно рассматривать как управляющее воздействие, а значение концентрации загрязнений на входе в биореактор S_0 и скорость жидкости в канале W_0 — как входные параметры. Для реализации процессов управления в целях стабилизации параметра $S_{\text{ВЫХ}}$ на заданном уровне должна быть известна зависимость $S_{\text{ВЫХ}} = S(t, S_0, W_0)$. Система уравнений (9.1)–(9.5) с принятыми начальными и граничными условиями описывает данную зависимость, однако не позволяет получить аналитические выражения для $S_{\text{ВЫХ}} = S(t, S_0, W_0)$ в форме, приемлемой для синтеза системы управления. Поэтому для рабочих диапазонов заданных пара-

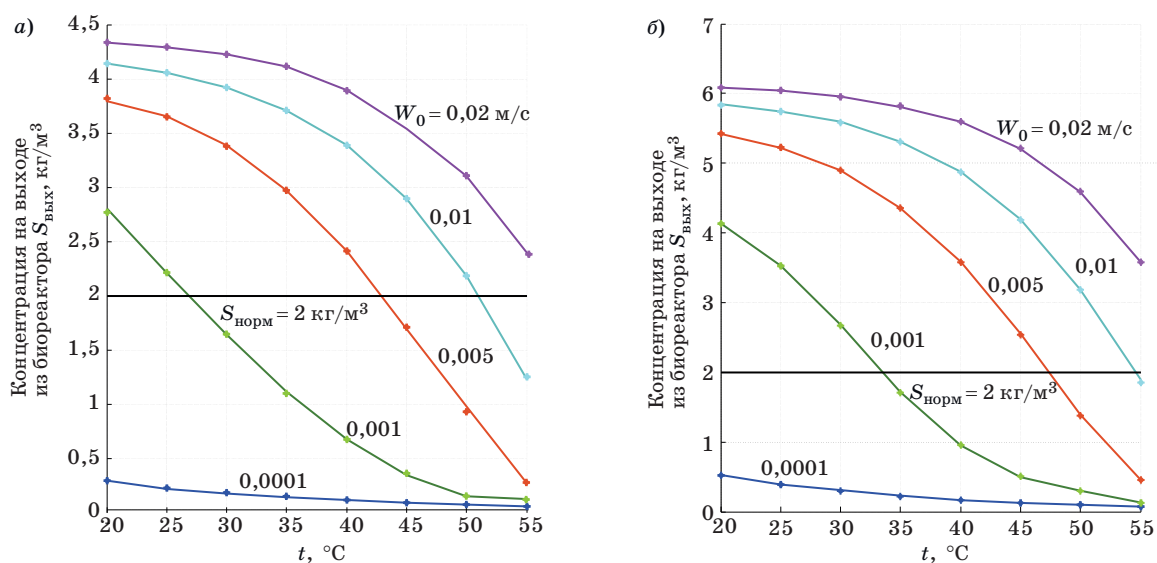
метров на основе численного решения системы уравнений (9.1)–(9.5) были найдены зависимости, аппроксимирующие поведение функции $S_{\text{ВЫХ}} = S(t, S_0, W_0)$ для различных рабочих условий.

Аппроксимирующие зависимости $S_{\text{ВЫХ}} = S(t, S_0, W_0)$ для синтеза системы управления строятся на основе имитационной модели, обеспечивающей численное решение системы уравнений (9.1)–(9.5) для заданных входных параметров и текущего состояния реактора. Модель реализована в системе MATLAB [3].

Исследования показали, что зависимость конечной концентрации $S_{\text{ВЫХ}}$ от начальной концентрации S_0 носит линейный характер, в то время как для описания зависимости $S_{\text{ВЫХ}}$ от рабочей температуры процесса t достаточно ограничиться полиномом 3-й степени. Также имеют место парные и квадратичные эффекты. В результате имитационного моделирования для изменений температур в диапазоне $t = 20 \div 55$ °С, начальной концентрации загрязнений стока $S_0 = 4 \div 12$ кг_{ХПК}/м³ и скорости жидкости в канале $W_0 = 0,1 \div 20$ мм/с получена аппроксимирующая функция, устанавливающая взаимосвязь этих параметров с показателем качества очистки $S_{\text{ВЫХ}}$ со средней квадратической погрешностью аппроксимации 5 %. Аппроксимирующая функция может быть при этом записана в виде

$$S = a_1(W_0) + a_2(W_0)t + a_3(W_0)S_0 + a_4(W_0)tS_0 + a_5(W_0)W_0t^2 + a_6(W_0)t^3 + a_7(W_0)S_0t^3 + a_8(W_0)t^2W_0S_0, \quad (10)$$

где:



■ **Рис. 2.** Сравнение результатов имитационного эксперимента и аппроксимации для начальной концентрации $S_0 = 6$ кг/м³ (а) и $S_0 = 10$ кг/м³ (б)

■ **Fig. 2.** Simulation experiment results and approximation comparison for an initial concentration of $S_0 = 6$ kg/m³ (a) and $S_0 = 10$ kg/m³ (b)

— для диапазона температур $t = 20 \div 30$ °C

$$a_i = b_{i1}W_0^3 + b_{i2}W_0^2 + b_{i3}W_0 + b_{i4}, \quad i = 1, \dots, 8;$$

— для диапазона температур $t = 30 \div 45$ °C

$$a_i = c_{i1}W_0^4 + c_{i2}W_0^3 + c_{i3}W_0^2 + c_{i4}W_0 + c_{i5}, \quad i = 1, \dots, 8;$$

— для диапазона температур $t = 45 \div 55$ °C

$$a_i = d_{i1}W_0^3 + d_{i2}W_0^2 + d_{i3}W_0 + d_{i4}, \quad i = 1, \dots, 8.$$

Результаты расчетов представлены на рис. 2 в виде зависимости концентрации загрязнений на выходе из биореактора от рабочей температуры жидкости в биореакторе $S(t)$ для начальной концентрации загрязнений $6 \text{ кг}_{\text{ХПК}}/\text{м}^3$ (рис. 2, а) и $10 \text{ кг}_{\text{ХПК}}/\text{м}^3$ (рис. 2, б). Зависимости показаны для различных значений скорости жидкости в канале W_0 . Точками отмечены результаты расчета, полученные при имитационном моделировании, сплошными линиями — по аппроксимирующей модели. Заданное качество очистки стока $S_{\text{норм}} =$

$= 2 \text{ кг}_{\text{ХПК}}/\text{м}^3$ представляет собой значение, которое концентрация загрязнений сточной воды на выходе из биореактора не должна превышать.

Заключение

Выполненное в работе математическое описание основных процессов функционирования анаэробного биореактора позволяет отразить конструктивные и технологические особенности системы очистки сточных вод. Такое описание может рассматриваться как обобщенная модель анаэробного секционного биореактора с закрепленной биомассой.

На основе имитационной модели и результатов численных расчетов в работе получены аппроксимирующие зависимости, описывающие взаимосвязь эффективности работы биореактора от входных и возмущающих воздействий. Представленные результаты могут использоваться в задачах синтеза алгоритмов управления технологическими процессами в анаэробном биореакторе.

Литература

1. Баженов В. И., Эпов А. Н., Носкова И. А. Математическое моделирование объекта очистки сточных вод. *Экологический вестник России*, 2011, № 4, с. 30–35.
2. Баженов В. И., Эпов А. Н., Носкова И. А. Математическое моделирование объекта очистки сточных вод. *Экологический вестник России*, 2011, № 5, с. 38–42.
3. Ключарев А. А., Фоменкова А. А. Проектирование секционного анаэробного биореактора. *Известия Санкт-Петербургского государственного технологического института (технического университета)*, 2018, № 34 (60), с. 95–100.
4. Saeed M., Fawzy S., El-Saadawi M. Modeling and simulation of biogas-fueled power system. *International Journal of Green Energy*, 2019, no. 16(2), pp. 125–151. doi:10.1080/15435075.2018.1549997
5. Khalekuzzaman M., Hasan M., Haque R., Alamgir M. Hydrodynamic performance of a hybrid anaerobic baffled reactor (HABR): Effects of number of chambers, hydraulic retention time, and influent temperature. *Water Science and Technology*, 2018, no. 78(4), pp. 968–981. doi:10.2166/wst.2018.379
6. Баженов В. И., Эпов А. Н., Носкова И. А. Использование комплексов имитационного моделирования для технологий очистки сточных вод. *Водоснабжение и санитарная техника*, 2014, № 2, с. 62–72.
7. Sbarciog M., Giovannini G., Chamy R., Wouwer A. V. Control and estimation of anaerobic digestion processes using hydrogen and volatile fatty acids measurements. *Water Science and Technology*, 2018, vol. 78, no. 10, pp. 2027–2035.
8. Draa K. C., Zemouche A., Alma M., Voos H., Darouach M. Nonlinear observer-based control with application to an anaerobic digestion process. *European Journal of Control*, 2019, no. 45, pp. 74–84. doi:10.1016/j.ejcon.2018.09.012
9. Li L., Peng X., Wang X., Wu D. Anaerobic digestion of food waste: A review focusing on process stability. *Bioresource Technology*, 2018, no. 248, pp. 20–28. doi:10.1016/j.biortech.2017.07.012
10. Пат. UA 93476 У Украины, МПК (2014.01) C02F 11/00. Анаэробный биореактор для очистки сточной воды, Л. И. Ружинская, А. А. Фоменкова: заявитель и патентообладатель: Национальный технический университет Украины «Киевский политехнический институт». № u 2013 14720; заявл. 16.12.13; опубл. 10.10.14, Бюл. № 19. — 3 с.
11. *Технологии для воды Enviro-chemie*. <http://www.enviro-chemie.ru/biomar/index.htm> (дата обращения: 15.10.2018).
12. Касаткин А. Г. *Основные процессы и аппараты химической технологии*. М., Химия, 1971. 784 с.
13. Moletta R., Verrier D., Albagnac G. Dynamic modeling of anaerobic digestion. *Water Research*, 1986, vol. 20, no. 4, pp. 427–434.
14. Moletta R. Methanisation: General aspects. [La méthanisation : Aspects généraux]. *Eau, l'Industrie, Les Nuisances*, 2013, no. 365, pp. 9–12.
15. Duan Z., Cruz Bournazou M. N., Kravaris C. Dynamic model reduction for two-stage anaerobic digestion processes. *Chemical Engineering Journal*,

- 2017, no. 327, pp. 1102–1116. doi:10.1016/j.cej.2017.06.110
16. Batstone D. J., Keller J., Angelidaki R. I., Kalyuzhnyi S. V., Pavlostathis S. G., Rozzi A., Sanders W. T. M., Siegrist H., Vavilin V. A. *Anaerobic Digestion Model No1 (ADM1)*. London, UK, IWA Publishing, 2002. 77 p.
 17. Eberl H., Morgenroth E., Noguera D., Picioreanu C., Rittmann B., Loosdrecht M., Wanner O. *Mathematical Modeling of Biofilms*. London, UK, IWA Publishing, 2006. 208 p.
 18. Ружинская Л. И., Фоменкова А. А. Математическое моделирование процессов анаэробного сбраживания органического субстрата. Обзор. *Scientific Journal «ScienceRise»*, 2014, № 4/2(4), с. 52–59. doi:10.15587/2313-8416.2014.28767
 19. Zhao X., Li L., Wu D., Xiao T., Ma Y., Peng X. Modified anaerobic digestion model no. 1 for modeling methane production from food waste in batch and semi-continuous anaerobic digestions. *Bioresource Technology*, 2019, no. 271, pp. 109–117. doi:10.1016/j.biortech.2018.09.091
 20. Li L., He Q., Zhao X., Wu D., Wang X., Peng X. Anaerobic digestion of food waste: Correlation of kinetic parameters with operational conditions and process performance. *Biochemical Engineering Journal*, 2018, no. 130, pp. 1–9. doi:10.1016/j.bej.2017.11.003
 21. Wang P., Wang H., Qiu Y., Ren L., Jiang B. Microbial characteristics in anaerobic digestion process of food waste for methane production—A review. *Bioresource Technology*, 2018, no. 248, pp. 29–36. doi:10.1016/j.biortech.2017.06.152
 22. Вавилин В. А. Математическое моделирование динамики сообществ анаэробных микроорганизмов. *Известия Самарского научного центра Российской академии наук*, 2009, т. 11, № 1(7), с. 1615–1619.
 23. Carotenuto C., Guarino G., Morrone B., et al. Temperature and pH effect on methane production from buffalo manure anaerobic digestion. *International Journal of Heat and Technology*, 2016, no. 34(2), pp. 425–429.

UDC 628.3

doi:10.31799/1684-8853-2019-2-44-51

Mathematical model of a fixed-biomass anaerobic bioreactor as a control object

A. A. Klucharev^a, PhD, Tech., Associate Professor, orcid.org/0000-0001-5345-1209

A. A. Fomenkova^a, Assistant, orcid.org/0000-0002-5570-6743, a.a.fomenkova@mail.ru

^aSaint-Petersburg State University of Aerospace Instrumentation, 67, B. Morskaja St., 190000, Saint-Petersburg, Russian Federation

Introduction: The efficiency and environmental safety of industrial objects depend on the efficiency of their bioremediation facilities. Managing the purification in bioreactors is significantly limited by the complexity of describing the physical and biochemical processes. **Purpose:** Developing a generalized mathematical model of an anaerobic bioreactor with fixed biomass as an object of an automatic control system, in order to take into account the structural and technological features of a wastewater treatment system. **Methods:** Simulation of mass transfer and biochemical transformations in an anaerobic bioreactor, and development of a mathematical model applicable for the synthesis of a control system. **Results:** A generalized mathematical model of an anaerobic bioreactor with fixed biomass was obtained on the basis of mass transfer equations. This model is supplemented with components which take into account biochemical transformations. On the basis of the bioreactor design and technological parameters, necessary assumptions were made which allowed us to formulate boundary and initial conditions with an accuracy sufficient for engineering practice. The choice of controlled parameters of the technological process in an anaerobic bioreactor is substantiated, ensuring the best purification at the lowest cost. The obtained generalized mathematical model of an anaerobic bioreactor with fixed biomass allows you to perform a simulation of a bioreactor under specified operating conditions. For these conditions, an approximating model was constructed which can be used for a control system synthesis. **Practical relevance:** The developed algorithms allow you to apply the obtained results to a wide class of the existing anaerobic bioreactors with fixed biomass, and to build control systems of a much higher efficiency, including the cleaning system upgrade stages.

Keywords — automatic stabilization system, anaerobic bioreactor, mathematical model, simulation.

For citation: Klucharev A. A., Fomenkova A. A. Mathematical model of a fixed-biomass anaerobic bioreactor as a control object. *Informacionno-upravliaiushchie sistemy* [Information and Control Systems], 2019, no. 2, pp. 44–51 (In Russian). doi:10.31799/1684-8853-2019-2-44-51

References

1. Bazhenov V. I., Jepov A. N., Noskova I. A. Mathematical modeling of the wastewater treatment object. *Jekologicheskij vestnik Rossii* [Environmental Bulletin of Russia], 2011, no. 4, pp. 30–35 (In Russian).
2. Bazhenov V. I., Jepov A. N., Noskova I. A. Mathematical modeling of the wastewater treatment object. *Jekologicheskij vestnik Rossii* [Environmental Bulletin of Russia], 2011, no. 5, pp. 38–42 (In Russian).
3. Klucharev A. A., Fomenkova A. A. Sectional anaerobic bioreactor design. *Izvestija Sankt-Peterburgskogo gosudarstvennogo tehnologicheskogo instituta (tehničeskogo universiteta)* [Bulletin of the Saint Petersburg State Institute of Technology (Technical University)], 2018, no. 34 (60), pp. 95–100 (In Russian).
4. Saeed M., Fawzy S., El-Saadawi M. Modeling and simulation of biogas-fueled power system. *International Journal*

- of Green Energy, 2019, no. 16(2), pp. 125–151. doi:10.1080/15435075.2018.1549997
5. Khalekuzzaman M., Hasan M., Haque R., Alamgir M. Hydrodynamic performance of a hybrid anaerobic baffled reactor (HABR): Effects of number of chambers, hydraulic retention time, and influent temperature. *Water Science and Technology*, 2018, no. 78(4), pp. 968–981. doi:10.2166/wst.2018.379
 6. Bazhenov V. I., Epov A. N., Noskova I. A. The use of simulation modeling packages for wastewater treatment technologies. *Vodosnabzhenie i sanitarnaja tehnika* [Water Supply and Sanitary Technique], 2014, no. 2, pp. 62–72 (In Russian).
 7. Sbarciog M., Giovannini G., Chamy R., Wouwer A. V. Control and estimation of anaerobic digestion processes using hydrogen and volatile fatty acids measurements. *Water Science and Technology*, 2018, vol. 78, no. 10, pp. 2027–2035.
 8. Draa K. C., Zemouche A., Alma M., Voos H., Darouach M. Nonlinear observer-based control with application to an anaerobic digestion process. *European Journal of Control*, 2019, no. 45, pp. 74–84. doi:10.1016/j.ejcon.2018.09.012
 9. Li L., Peng X., Wang X., Wu D. Anaerobic digestion of food waste: A review focusing on process stability. *Bioresource Technology*, 2018, no. 248, pp. 20–28. doi:10.1016/j.biortech.2017.07.012
 10. Ruzhinskaja L. I., Fomenkova A. A. *Anajerobnyj bioreaktor dlja ochistki stochnoj vody* [Anaerobic bioreactor for wastewater treatment]. Patent Ukraine, UA 93476 U, 2014.
 11. *Tehnologii dlja vody Enviro-chemie* [Enviro-chemie water technology]. Available at: <http://www.enviro-chemie.ru/bi-omar/index.htm> (accessed 15 October 2018).
 12. Kasatkin A. G. *Osnovnye processy i apparaty himicheskoj tehnologii* [Basic Processes and Apparatuses of Chemical Technology]. Moscow, Himija Publ., 1971. 784 p. (In Russian).
 13. Moletta R., Verrier D., Albagnac G. Dynamic modeling of anaerobic digestion. *Water Research*, 1986, vol. 20, no. 4, pp. 427–434.
 14. Moletta R. Methanisation: General aspects. [La méthanisation : Aspects généraux]. *Eau, l'Industrie, Les Nuisances*, 2013, no. 365, pp. 9–12 (In French).
 15. Duan Z., Cruz Bournazou M. N., Kravaris C. Dynamic model reduction for two-stage anaerobic digestion processes. *Chemical Engineering Journal*, 2017, no. 327, pp. 1102–1116. doi:10.1016/j.cej.2017.06.110
 16. Batstone D. J., Keller J., Angelidaki R. I., Kalyuzhnyi S. V., Pavlostathis S. G., Rozzi A., Sanders W. T. M., Siegrist H., Vavilin V. A. *Anaerobic Digestion Model No1 (ADM1)*. IWA Publishing, London, UK, 2002. 77 p.
 17. Eberl H., Morgenroth E., Noguera D., Picioreanu C., Rittmann B., Loosdrecht M., Wanner O. *Mathematical Modeling of Biofilms*. IWA Publishing, London, UK, 2006. 208 p.
 18. Ruzhinskaya L. I., Fomenkova A. A. Mathematical modeling of organic substrate anaerobic digestion. Review. *Scientific Journal "ScienceRise"*, 2014, no. 4/2(4), pp. 52–59 (In Russian). doi:10.15587/2313-8416.2014.28767
 19. Zhao X., Li L., Wu D., Xiao T., Ma Y., Peng X. Modified anaerobic digestion model no. 1 for modeling methane production from food waste in batch and semi-continuous anaerobic digestions. *Bioresource Technology*, 2019, no. 271, pp. 109–117. doi:10.1016/j.biortech.2018.09.091
 20. Li L., He Q., Zhao X., Wu D., Wang X., Peng X. Anaerobic digestion of food waste: Correlation of kinetic parameters with operational conditions and process performance. *Biochemical Engineering Journal*, 2018, no. 130, pp. 1–9. doi:10.1016/j.bej.2017.11.003
 21. Wang P., Wang H., Qiu Y., Ren L., Jiang B. Microbial characteristics in anaerobic digestion process of food waste for methane production—A review. *Bioresource Technology*, 2018, no. 248, pp. 29–36. doi:10.1016/j.biortech.2017.06.152
 22. Vavilin V. A. Dynamics of Anaerobic Microorganisms Community: Experience in Mathematical Modeling. *Izvestija Samarского nauchnogo centra Rossijskoj akademii nauk*, 2009, vol. 11, no. 1(7), pp. 1615–1619 (In Russian).
 23. Carotenuto C., Guarino G., Morrone B., et al. Temperature and pH effect on methane production from buffalo manure anaerobic digestion. *International Journal of Heat and Technology*, 2016, no. 34(2), pp. 425–429.

ПАМЯТКА ДЛЯ АВТОРОВ

Поступающие в редакцию статьи проходят обязательное рецензирование.

При наличии положительной рецензии статья рассматривается редакционной коллегией. Принятая в печать статья направляется автору для согласования редакторских правок. После согласования автор представляет в редакцию окончательный вариант текста статьи.

Процедуры согласования текста статьи могут осуществляться как непосредственно в редакции, так и по e-mail (ius.spb@gmail.com).

При отклонении статьи редакция представляет автору мотивированное заключение и рецензию, при необходимости доработать статью — рецензию.

Редакция журнала напоминает, что ответственность за достоверность и точность рекламных материалов несут рекламодатели.

УДК 519.688

doi:10.31799/1684-8853-2019-2-52-56

Структура каталога как фильтр поиска записей службы каталогов OpenLDAP

А. В. Гордеев^а, доктор техн. наук, профессор, orcid.org/0000-0001-5390-3498, ff2avg@mail.ru

А. В. Андреев^а, аспирант, orcid.org/0000-0001-5743-0229

^аСанкт-Петербургский государственный университет аэрокосмического приборостроения,
Б. Морская ул., 67, Санкт-Петербург, 190000, РФ

Введение: количество публикаций по исследованию структур каталогов ограничено видимой простотой закрытых, проприетарных, реализаций. Эта закрытость накладывает ограничения на возможности исследования служб каталогов, особенно вопросов, касающихся изменения структуры каталога. Открытые решения, такие как OpenLDAP, получили популярность относительно недавно, и большинство исследований сводятся к вопросам интеграции открытых решений в существующую инфраструктуру. **Цель:** определение влияния структуры хранения данных каталога при различных параметрах индексирования на время отклика на запрос к каталогу. **Результаты:** основными параметрами, влияющими на эффективность использования индексирования при работе со службами каталогов, определены количество записей каталога и типы индексов, используемые в фильтре поиска. Предложен новый подход формирования структуры службы каталогов с учетом параметров поиска информации. Выявлено, что при наличии индексируемых атрибутов у большого количества записей каталога наименее эффективными являются индексы типов присутствия и подстроки. Дополнительные ветки каталога, созданные по определенным критериям записей, были заданы как исходные ветки поиска записей и использовались как дополнительный фильтр поиска. Данная модификация структуры каталога позволила дополнительно уменьшить общее число записей при изначальной выборке. Такое решение существенно улучшило время отклика на запрос к каталогу при использовании индексов присутствия и подстроки в фильтре поиска, а также сократило фильтр поиска. **Практическая значимость:** формирование структуры каталога с учетом параметров записей и способов поиска параметров записей позволяет уменьшить время отклика каталога на запрос.

Ключевые слова – служба каталогов, структура каталога, OpenLDAP, индексирование, тип индекса.

Для цитирования: Гордеев А. В., Андреев А. В. Структура каталога как фильтр поиска записей службы каталогов OpenLDAP. *Информационно-управляющие системы*, 2019, № 2, с. 52–56. doi:10.31799/1684-8853-2019-2-52-56

For citation: Gordeyev A. V., Andreyev A. V. OpenLDAP directory service structure as a search filter. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2019, no. 2, pp. 52–56 (In Russian). doi:10.31799/1684-8853-2019-2-52-56

Введение

Служба каталогов (Directory Service) — средство иерархического представления различных ресурсов сети и хранения информации об этих ресурсах. В качестве ресурсов могут выступать персонал, различные информационно-вычислительные ресурсы, сетевые ресурсы [1] и т. д. Информация об определенном объекте (ресурсе) хранится в виде набора значений атрибутов этого объекта. Самые распространенные службы каталогов — Active Directory, OpenLDAP, 389-DS и др. [2]. Различные службы каталогов широко используются малыми, средними и крупными предприятиями, где служба каталогов не только применяется для хранения учетных записей и аутентификации пользователей, но и выступает в роли механизма авторизованного доступа к ресурсам сети [3]. Многие продукты, например почтовый сервер Zimbra, в основе своей архитектуры используют службу каталогов [4]. Также существуют облачные реализации служб каталогов, например, JumpCloud предлагает облачный продукт Directory-as-a-Service, использующий OpenLDAP [5].

Количество исследований и, соответственно, публикаций по данной тематике не так велико и

лимитировано популярностью, а также видимой простотой закрытых, проприетарных, реализаций. Закрытость таких реализаций накладывает ограничения на возможности исследования служб каталогов, особенно вопросов, касающихся изменения структуры каталога. Компании, реализующие закрытые решения, не раскрывают архитектуру построения служб каталогов и являются основным источником информации. Например, компания Microsoft является разработчиком популярной закрытой реализации службы каталогов Active Directory. Актуальную информацию об Active Directory компания распространяет через свой веб-сайт. Открытые решения, такие как OpenLDAP, получили популярность относительно недавно, и большинство исследований сводится к вопросам интеграции открытых решений в существующую инфраструктуру.

Службы каталогов используют структуру хранения данных В-дерево и работают по протоколу LDAP [6].

Одной из основных характеристик работы службы каталогов является время отклика на запрос поиска записи или записей.

Производительность каталога зависит от множества факторов, например:

- аппаратных характеристик серверов и сетевой инфраструктуры;
- видов записей;
- количества записей в базе данных службы каталога.

Ранее проводимые исследования рассматривают различные аспекты внедрения службы каталогов в инфраструктуру вычислительной сети, где структура каталога формируется по организационной принадлежности объекта [7]. С увеличением общего количества сервисов, поддерживающих работу по протоколу LDAP, увеличивается и количество применяемых атрибутов объектов, что накладывает ограничения на подходы в целях оптимизации служб каталогов.

Один из таких подходов — использование индексирования [8].

Параметры индексирования записей в службе каталога

Индексирование атрибутов записей служб каталогов — распространенная практика оптимизации поиска записей. Индексируемый атрибут используется в качестве фильтра поиска. Самый распространенный фильтр поиска — `(uid=%u)`. Различные службы сети могут создать разнообразные комбинации фильтров, оперируя логическими операциями И, ИЛИ, НЕ [9]. Например, служба файлового сервера, реализуемая средствами `samba`, пользуется по умолчанию фильтром поиска `(&(uid=%u)(objectClass=sambaAccount)`, но рекомендуется индексировать следующий набор атрибутов для оптимизации запросов: `uid`, `sn`, `sambaSID`, `gidNumber`, `uidNumber`, `displayName` [10]. Помимо файлового сервера, один каталог могут использовать разные службы: почтовый сервер, VPN-сервер, прокси-сервер, сервер репликации данных, web-серверы, серверы обмена сообщениями и различные внутренние службы сети. Все службы имеют свои наборы фильтров поиска и атрибутов хранения данных.

Индексирование имеет ряд ограничений, которые не позволяют индексировать все необходимые атрибуты для всех сервисов сети сразу:

1. Индексы занимают дисковое пространство. Для хранения данных каталога OpenLDAP по умолчанию используется бэкенд `hdb`, где данные хранятся в двух файлах: `dn2id.bdb` и `id2entry.bdb`. Для каждого индекса создается отдельный файл. Например, при индексировании атрибута `objectClass` будет создан файл `objectclass.bdb`. Размер файла зависит от выбранных типов индексов для атрибута.

2. Каждое добавление новой записи с индексированными атрибутами наравне с добавлением записи добавляет новые индексы, что влияет

на скорость добавления или обновления записей [11].

3. Рекомендуется индексировать атрибуты, которые часто используются, но в сетях со множеством различных служб, одни службы используются чаще других, а другие — реже. В таком случае индексирование атрибутов редко используемых служб является напрасной тратой аппаратных ресурсов [12].

4. Использование неиндексированных атрибутов вместе с индексированными в одном фильтре поиска отрицательно сказывается на производительности такого поиска. Также количество и уникальность атрибутов фильтра влияют на скорость поиска соответствующей записи [13].

Помимо перечисленных ограничений, также существуют ограничения, накладываемые типом индекса. Следует внимательно выбирать атрибуты с типами индексирования присутствия (`presence`) и подстроки (`substring`) [14]. Если некий атрибут существует у большинства записей, то все эти записи будут в конечной выборке, потому что они все подходят под фильтр поиска. Скорость поиска будет в данном случае зависеть от количества записей. Например, если большинство пользователей организации используют почтовый сервер, применение индекса присутствия для атрибутов почтового сервиса, содержащихся в фильтре поиска, мало влияет на скорость поиска записей, так как все записи будут в конечной выборке [15].

Структура каталога

В случае неэффективности индексирования из-за ограничений, накладываемых типами фильтров, предлагается использовать структуру каталога как фильтр поиска, где каждая ветка каталога содержит только записи, относящиеся к конкретной службе или выбранному критерию.

Рассмотрим каталог, где каждая запись-пользователь имеет набор общих атрибутов `objectClass`, а также один уникальный атрибут `objectClass`, который позволяет сгруппировать эти записи по заданному критерию. Все записи хранятся в одной ветке, что является распространенной практикой построения каталога [16]. Каталог содержит 100 000 записей: 50 131 запись вида `posixAccount`, 25 053 записи вида `shadowAccount`, 24 816 записей вида `account`.

Записи `posixAccount`:

```
dn: uid=User9999,ou=People,dc=example,dc=com
ou: PosixAccount
objectClass: top
objectClass: person
objectClass: organizationalPerson
objectClass: inetOrgPerson
```

```
objectClass: posixAccount
uid: User9999
cn: User9999_
sn: User9999--
userPassword:: c2VjcmV0
mail: User9999@example.com
uidNumber: 19999
gidNumber: 19999
homeDirectory: /home/User9999
loginShell: /bin/bash
```

Записи “shadowAccount”:

```
dn: uid=User10000,ou=People,dc=example,dc=com
objectClass: top
objectClass: person
objectClass: organizationalPerson
objectClass: shadowAccount
ou: People
uid: User10000
cn: User10000_
sn: User10000--
userPassword:: c2VjcmV0
description: Line for ShadowAccount
```

Записи “account”:

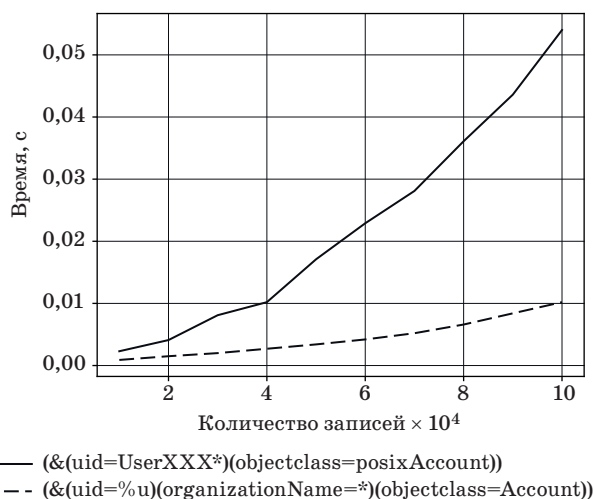
```
dn: uid=User9997,ou=People,dc=example,dc=com
ou: Account
objectClass: top
objectClass: account
uid: User9997
organizationName: User9997
```

Рассмотрим два фильтра поиска (&(uid=UserXXX*)(objectclass=posixAccount)) и (&(uid=%u)(organizationName=*)(objectclass=Account)), где XXX — сгенерированное случайное трехзначное значение от 1 до 999, %u — заданные uid, например User9997.

Первый фильтр использует индекс подстроки и осуществляет поиск заданного набора записей вида “posixAccount”. Второй фильтр использует индекс присутствия и осуществляет поиск заданной записи. Операция поиска повторяется 100 раз после добавления 10 000 записей, а по результатам высчитывается среднее значение. Результаты поиска записей для двух фильтров с использованием индексирования представлены на рис. 1.

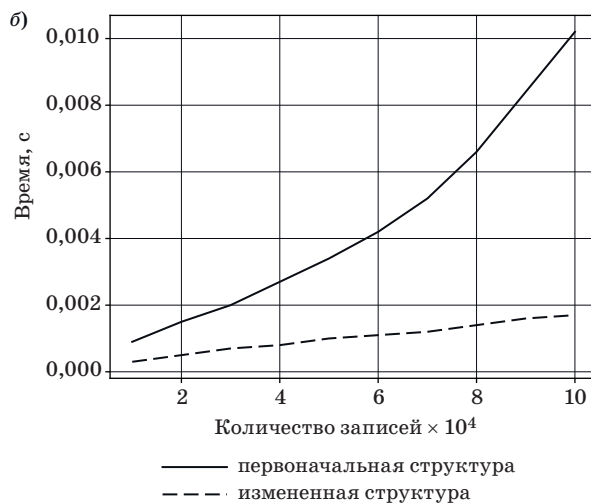
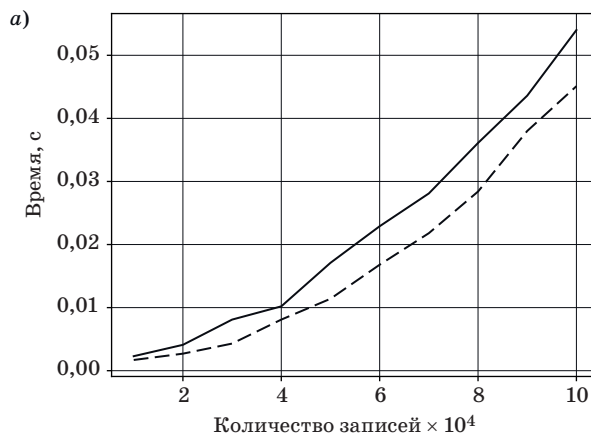
В данном случае структура каталога может быть изменена по условному критерию — типу записи. Записи “posixAccount” будут храниться в ou=PosixAccount, ou=People, dc=example, dc=com, записи “shadowAccount” — в ou=People, dc=example, dc=com, а записи “account” — в ou=Account, ou=People, dc=example, dc=com.

Такое перераспределение записей позволяет уменьшить общее количество записей поиска, а также использовать ветку поиска как фильтр поиска, сократив тем общее количество атрибутов в фильтре поиска до (uid=UserXXX*) и (&(uid=%u)(organizationName=*)). При этом сохраняется возможность поиска всех записей при



■ **Рис. 1.** Время поиска записей индексированного каталога

■ **Fig. 1.** The search time of records for the indexed directory



■ **Рис. 2.** Время поиска записей вида “posixAccount” (а) и “account” (б)

■ **Fig. 2.** The search time of records “posixAccount” (а) and “account” (б)

использовании поиска типа subtree и ветки поиска ou=People,dc=example,dc=com.

Результаты поиска записей индексированного каталога с первоначальной структурой и измененной для первого фильтра поиска изображены на рис. 2, а, для второго — на рис. 2, б.

Как видно по полученным данным, количество записей в ветке поиска, количество атрибутов в фильтре поиска и тип индекса влияют на скорость поиска записей.

Заключение

Производительность каталога зависит от множества факторов. Индексирование — хорошее решение для оптимизации службы каталогов, но данный подход имеет ряд ограничений, зависящих от того, как используется каталог и как используются данные, хранящиеся в данном каталоге.

Одними из таких ограничений являются тип фильтра и параметры его использования. Тип фильтра влияет на эффективность индексирования, например, индексы присутствия и подстроки. Одним из решений оптимизации такого каталога может быть новый подход формирования структуры службы каталогов, где ветка поиска заменяет основной фильтр поиска и уменьшает общее количество записей для выборки. Основной особенностью данного подхода является перенос заранее известного значения элемента логического выражения фильтра поиска в область, задающую начальный объект поиска записей службы каталогов. Данное изменение можно реализовать с помощью записей-ссылок, не изменяя основную структуру каталога, используя сетевую структуру хранения данных [17].

Проведенные эксперименты показали эффективность данного подхода, но конечный результат зависит от конкретного исследуемого каталога.

Литература

1. Kretschmer F., von Arnim C., Lechler A., Verl A. Persistent data backend for OPC UA namespaces in IT infrastructures. *Procedia CIRP*, 2018, vol. 72, pp. 174–178. doi:10.1016/j.procir.2018.03.233
2. *List of LDAP Software*. https://en.wikipedia.org/wiki/List_of_LDAP_software (дата обращения: 25.12.2018).
3. *Microsoft Active Directory*. <https://discovery.hgdata.com/product/microsoft-active-directory> (дата обращения: 25.12.2018).
4. *Zimbra Directory Service (LDAP)*. [https://wiki.zimbra.com/wiki/Zimbra_Directory_Service_\(LDAP\)](https://wiki.zimbra.com/wiki/Zimbra_Directory_Service_(LDAP)) (дата обращения: 25.12.2018).
5. *Top 5 Challenges with OpenLDAP*. <https://jumpcloud.com/blog/top-5-challenges-with-openldap/> (дата обращения: 25.12.2018).
6. *Introduction to OpenLDAP Directory Services*. <https://www.openldap.org/doc/admin24/intro.html> (дата обращения: 25.12.2018).
7. Andjarwirawan J., Palit H. N., Salim J. C. Linux PAM to LDAP authentication migration. *2017 Intern. Conf. on Soft Computing, Intelligent System and Information Technology (ICSIIIT)*, Denpasar, 2017, pp. 155–159. doi:10.1109/ICSIIIT.2017.66
8. *Tuning*. <https://www.openldap.org/doc/admin24/tuning.html> (дата обращения: 03.01.2019).
9. *Lightweight Directory Access Protocol (LDAP): String Representation of Search Filters*. <https://tools.ietf.org/search/rfc4515#page-2> (дата обращения: 03.01.2019).
10. Eckstein R., Carter G., Ts J. *Using Samba*. O'Reilly Media, 2007. 600 p.
11. *Indexing Attributes in the Directory*. https://docs.oracle.com/cd/E15217_01/doc.1014/e12490/perform.htm#CFHICJIE (дата обращения: 09.01.2019).
12. *Managing Indexes*. https://access.redhat.com/documentation/en-US/Red_Hat_Directory_Server/8.2/html/Administration_Guide/Managing_Indexes.html (дата обращения: 08.01.2019).
13. Lowe-Norris A., Desmond B., Richards J., Allen R. *Active Directory*. O'Reilly Media, 2008. 866 p.
14. *LDAP Setup and Configuration Guide*. <https://docs.oracle.com/cd/E19455-01/806-5580/6jej518pd/index.html> (дата обращения: 20.11.2018).
15. *LDAP Indexes*. <https://ldapwiki.com/wiki/LDAP%20Indexes> (дата обращения: 08.01.2019).
16. Butcher M. *Mastering OpenLDAP*. Packt Publishing, 2007. 484 p.
17. Андреев А. В. Сетевая модель данных службы каталогов. *Тр. МФТИ*, 2014, т. 6, № 3, с. 27–36.

UDC 519.688

doi:10.31799/1684-8853-2019-2-52-56

OpenLDAP directory service structure as a search filterA. V. Gordeyev^a, Dr. Sc., Tech., Professor, orcid.org/0000-0001-5390-3498, ff2avg@mail.ruA. V. Andreyev^a, Post-Graduate Student, orcid.org/0000-0001-5743-0229^aSaint-Petersburg State University of Aerospace Instrumentation, 67, B. Morskaya St., 190000, Saint-Petersburg, Russian Federation

Introduction: The number of directory structure studies is limited by the seeming simplicity of closed proprietary implementations. This closedness complicates studying directory services, especially the issues related to changing the directory structure. Open source solutions, such as OpenLDAP, gained popularity relatively recently, and most research is about integrating them into the existing infrastructure. **Purpose:** Determining the influence of directory data-storing structure on the directory query response time, under various indexing parameters. **Results:** The main parameters affecting the indexing efficiency are the number of records in the directory and the types of indexes used in the search filters. A new approach to building up a directory service structure is proposed, taking into account the information search parameters. It has been found out that when a large number of records have indexable attributes, the most efficient indices are those of presence and substring types. Additional subtrees created according to certain criteria of the records were specified as the starting points for the search and used as an additional search filter. This approach can additionally reduce the total amount of records for the initial search and the number of attributes in the main search filter. Besides, it can significantly improve the response time. **Practical relevance:** A directory structure which takes into account the record parameters and the ways of search for them can provide a shorter response time.

Keywords — directory service, directory structure, OpenLDAP, indexing, index type.

For citation: Gordeyev A. V., Andreyev A. V. OpenLDAP directory service structure as a search filter. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2019, no. 2, pp. 52–56 (In Russian). doi:10.31799/1684-8853-2019-2-52-56

References

1. Kretschmer F., von Arnim C., Lechler A., Verl A. Persistent data backend for OPC UA namespaces in IT infrastructures. *Procedia CIRP*, 2018, vol. 72, pp. 174–178. doi:10.1016/j.procir.2018.03.233
2. *List of LDAP Software*. Available at: https://en.wikipedia.org/wiki/List_of_LDAP_software (accessed 25 December 2018).
3. *Microsoft Active Directory*. Available at: <https://discovery.hgdata.com/product/microsoft-active-directory> (accessed 25 December 2018).
4. *Zimbra Directory Service (LDAP)*. Available at: [https://wiki.zimbra.com/wiki/Zimbra_Directory_Service_\(LDAP\)](https://wiki.zimbra.com/wiki/Zimbra_Directory_Service_(LDAP)) (accessed 25 December 2018).
5. *Top 5 Challenges with OpenLDAP*. Available at: <https://jumpcloud.com/blog/top-5-challenges-with-openldap/> (accessed 25 December 2018).
6. *Introduction to OpenLDAP Directory Services*. Available at: <https://www.openldap.org/doc/admin24/intro.html> (accessed 25 December 2018).
7. Andjarwirawan J., Palit H. N., Salim J. C. Linux PAM to LDAP authentication migration. *2017 Intern. Conf. on Soft Computing, Intelligent System and Information Technology (ICSIIIT)*, Denpasar, 2017, pp. 155–159. doi:10.1109/ICSIIIT.2017.66
8. *Tuning*. Available at: <https://www.openldap.org/doc/admin24/tuning.html> (accessed 3 January 2019).
9. *Lightweight Directory Access Protocol (LDAP): String Representation of Search Filters*. Available at: <https://tools.ietf.org/search/rfc4515#page-2> (accessed 3 January 2019).
10. Eckstein R., Carter G., Ts J. *Using Samba*. O'Reilly Media, 2007. 600 p.
11. *Indexing Attributes in the Directory*. https://docs.oracle.com/cd/E15217_01/doc.1014/e12490/perform.htm#CF-HICJIE (accessed 9 January 2019).
12. *Managing Indexes*. Available at: https://access.redhat.com/documentation/en-US/Red_Hat_Directory_Server/8.2/html/Administration_Guide/Managing_Indexes.html (accessed 8 January 2019).
13. Lowe-Norris A., Desmond B., Richards J., Allen R. *Active Directory*. O'Reilly Media, 2008. 866 p.
14. *LDAP Setup and Configuration Guide*. Available at: <https://docs.oracle.com/cd/E19455-01/806-5580/6jej518pd/index.html> (accessed 20 November 2018).
15. *LDAP Indexes*. Available at: <https://ldapwiki.com/wiki/LDAP%20Indexes> (accessed 8 January 2019).
16. Butcher M. *Mastering OpenLDAP*. Packt Publishing, 2007. 484 p.
17. Andreyev A. V. Network data model of directory services. *Trudy Moskovskogo fiziko-tekhnicheskogo instituta* [Proc. of MIPT], 2014, vol. 6, no. 2, pp. 27–36 (In Russian).

УДК 004.056

doi:10.31799/1684-8853-2019-2-57-67

Открытые задачи визуального анализа в системах управления информационной безопасностью

Е. С. Новикова^{а, б}, канд. техн. наук, доцент, orcid.org/0000-0003-2923-4954, novikova@comsec.spb.ru
И. В. Котенко^а, доктор техн. наук, профессор, orcid.org/0000-0001-6859-7120

^аСанкт-Петербургский институт информатики и автоматизации РАН, 14-я линия В. О., 39, Санкт-Петербург, 199178, РФ

^бСанкт-Петербургский государственный электротехнический университет «ЛЭТИ», Профессора Попова ул., 5, Санкт-Петербург, 197376, РФ

Введение: методики визуального анализа позволяют эффективно исследовать постоянно увеличивающиеся объемы данных, генерируемые сенсорами безопасности, и способствуют своевременному и обоснованному реагированию на угрозы. В современных системах управления информационной безопасностью реализованы различные решения по обработке больших объемов данных и интеграции различных источников, которые могут быть использованы для построения полноценной системы визуального анализа инцидентов безопасности. **Цель:** исследование методик визуального анализа, реализованных в системах управления информационной безопасностью и предназначенных для исследования событий безопасности в контексте основных задач визуального анализа, включая верификацию корректности работы моделей автоматического анализа данных. **Результаты:** выявлено существующее противоречие между возможностями систем управления информационной безопасностью по анализу данных и реализацией этих возможностей. Практически отсутствуют методики визуального анализа данных, поддерживающие визуальную корреляцию данных от сенсоров безопасности и визуальную валидацию автоматических моделей анализа, позволяющую оценить эффективность и корректность их функционирования. Возможным решением этого противоречия является применение методик, реализующих гибкий механизм настройки анализируемых атрибутов событий сетевых устройств и сенсоров. В работе рассмотрены основные подходы к построению таких методик, обсуждаются их достоинства и недостатки. Предложена панель управления, предназначенная для мониторинга поведения автоматической модели анализа сетевого трафика в системе облачных вычислений, которая позволяет контролировать функционирование модели анализа, выполнять визуальную корреляцию исследуемых параметров и отслеживать характер изменений в сетевых потоках. **Практическая значимость:** результаты работы могут быть использованы при проектировании инструментов визуального анализа для исследования событий безопасности, мониторинга потоков данных и поведения автоматических моделей анализа.

Ключевые слова — визуальная аналитика, управление информационной безопасностью, визуальная корреляция данных, визуальная валидация моделей анализа, матричное представление данных.

Для цитирования: Новикова Е. С., Котенко И. В. Открытые задачи визуального анализа в системах управления информационной безопасностью. *Информационно-управляющие системы*, 2019, № 2, с. 57–67. doi:10.31799/1684-8853-2019-2-57-67

For citation: Novikova E. S., Kotenko I. V. Open challenges in visual analytics for security information and event management. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2019, no. 2, pp. 57–67 (In Russian). doi:10.31799/1684-8853-2019-2-57-67

Введение

В настоящее время методики *визуальной аналитики* (visual analytics), в которых сочетаются достоинства графического представления многомерных данных и их автоматизированного анализа, считаются едва ли не единственным способом анализа и исследования данных большого объема [1]. Визуальный анализ данных, описывающих безопасность информационной системы, является относительно новым, но активно развивающимся и востребованным направлением визуальной аналитики [2]. Методы визуального анализа помогают эффективно исследовать постоянно увеличивающиеся объемы данных, генерируемые сенсорами безопасности и, как следствие, способствуют своевременному и обоснованному реагированию на угрозы.

Ниже приведены задачи информационной безопасности, для решения которых могут быть использованы методики визуальной аналитики: оперативный контроль периметра компьютерной сети; оценка уровня защищенности компьютерной сети; обнаружение внутренних нарушителей; исследование инцидентов безопасности и формирование сценариев атаки; верификация политик безопасности сетевых устройств; исследование вредоносного кода; верификация автоматических моделей анализа, используемых для управления информационной безопасностью.

Выделим шесть основных проблем проектирования методик визуального анализа: большой объем данных; многообразие источников данных; отсутствие связи (синхронизации) между источниками данных; низкое качество данных; необходимость формирования паттернов нор-

мального поведения информационной системы; сложность отслеживания развития информационных угроз и реагирования [3, 4].

Возможные решения по обработке больших объемов данных и интеграции различных источников для построения компонента визуального анализа систем управления информационной безопасностью могут быть найдены в технологии управления информацией и событиями безопасности (Security Information and Event Management — SIEM), которая дает возможность собирать и обрабатывать данные от различных сенсоров безопасности, например, систем обнаружения/предотвращения вторжений, сканеров уязвимости, межсетевых экранов, роутеров и т. д., решая указанные проблемы.

В настоящей работе рассмотрены основные методики визуального анализа, применяемые в системах управления информацией и событиями безопасности, выявлены основные открытые задачи и определены возможные пути их решения. Отличительной особенностью данной работы от других исследований [2, 5, 6] является оценка методик визуализации существующих SIEM-систем в контексте основных задач визуальной аналитики: мониторинг текущей ситуации (оперативный контроль); исследование данных (исторический анализ данных); формирование отчетов; верификация корректности работы моделей автоматического анализа [7]. В работе также представлен макет панели управления, предназначенной для валидации модели анализа, используемой для обнаружения атак в системе облачных вычислений OpenStack, который построен с помощью выявленных методик визуального анализа многомерных данных.

Обзор методик визуального анализа данных в SIEM-системах

Помимо широких возможностей по обработке больших потоков разнородных данных, SIEM-системы позволяют выполнять исторический анализ событий безопасности и их последствий, представлять трассы атак [8], моделировать и формировать новые правила обнаружения информационных угроз [9] и оценки уровня их критичности [10]. Они предоставляют механизмы по управлению вычислительными ресурсами информационной системы, выполняют ее проверку на соответствие нормативным документам в области безопасности и создают отчетную документацию [10].

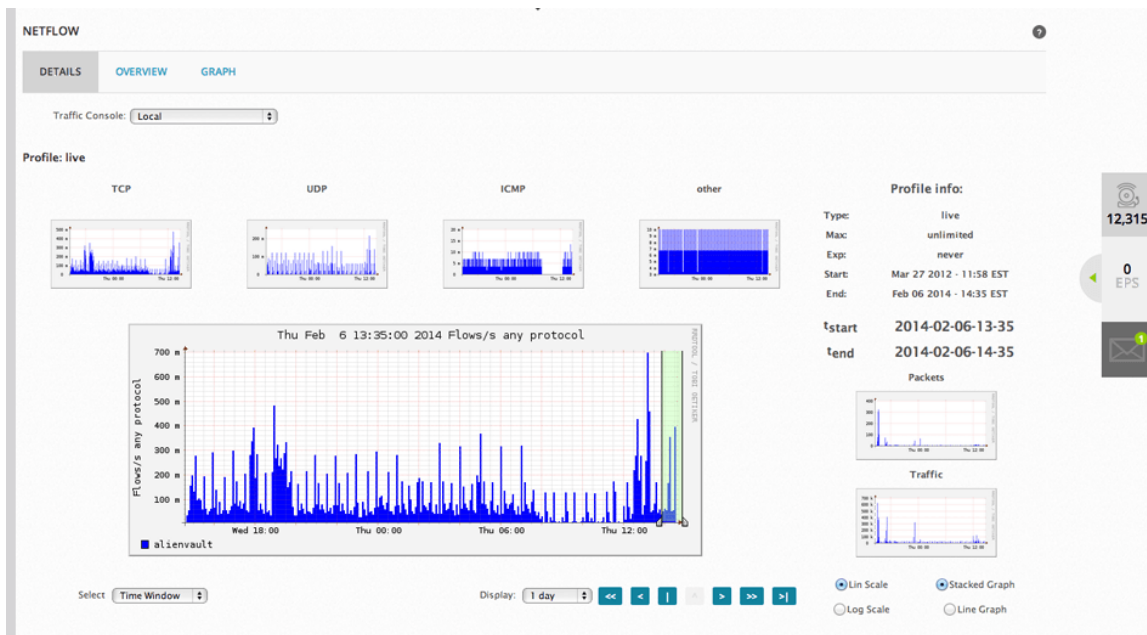
Для решения этих задач в SIEM-системах используются так называемые *панели управления* (dashboards), которые задают способ организации информации, необходимой для решения од-

ной или нескольких задач в графическом виде, обеспечивающем ее понимание с первого взгляда [11]. В зависимости от характера выполняемой задачи можно выделить *аналитические* и *оперативные* панели управления. Аналитические панели, как правило, содержат обобщающие результаты обработки и анализа данных о событиях, в то время как оперативные панели предоставляют низкоуровневые, часто необработанные данные, например записи межсетевого экрана.

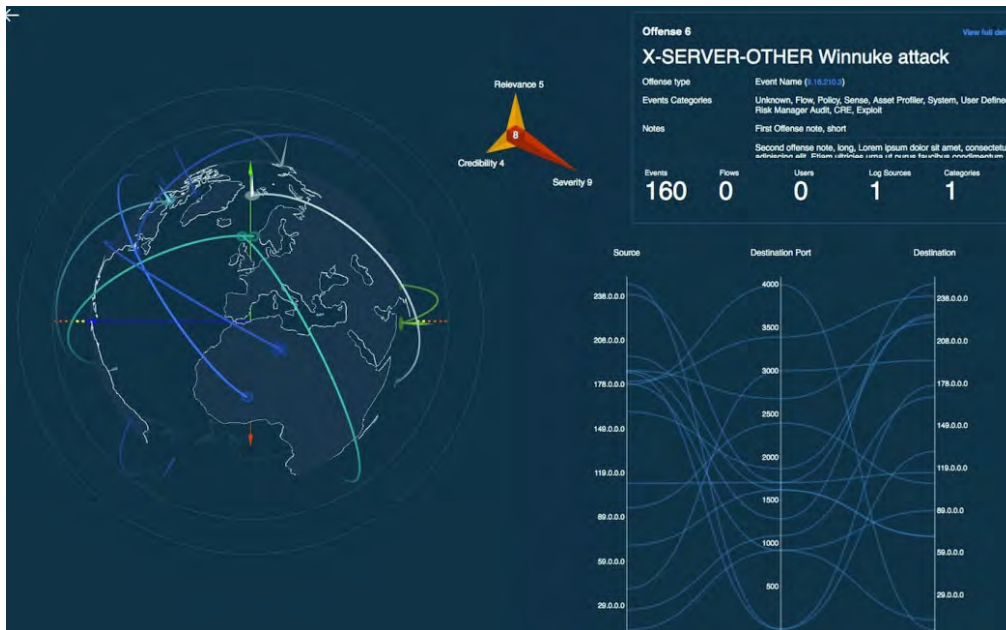
Оперативные панели управления, используемые для мониторинга событий безопасности или просмотра данных от сенсоров безопасности в режиме реального времени, чаще всего включают в себя табличное представление данных, в некоторых случаях дополненное моделями визуализации — столбчатыми диаграммами и линейными графиками. Например, на рис. 1 [12] представлена информация от сенсора учета сетевого трафика NetFlow. Очевидно, что наиболее интересными для аналитика характеристиками сетевого трафика являются входящий/исходящий объем сетевого трафика, его распределение по используемым протоколам, поэтому ключевым элементом данной панели является множество линейных графиков, показывающих изменение сетевого трафика во времени с учетом сетевого протокола.

Можно выделить два типа аналитических панелей управления: 1) панели, представляющие статистические данные, характеризующие функционирование SIEM-системы и ее компонентов на достаточно высоком уровне, и 2) панели управления, предназначенные для исследования инцидентов безопасности. Показательным примером аналитических панелей управления первого типа являются панели управления, отображающие статистику по сенсорам безопасности. Они обычно характеризуют количественную составляющую их функционирования, т. е. общее число срабатываний правил безопасности (сигналов тревоги), распределение срабатываний по типам (правилам), выявление объектов сети, наиболее подверженных атакам, основных источников атак и т. д. Для представления этих данных используются двухмерные стандартные модели визуализации (столбчатые или линейные диаграммы, круговые диаграммы, линейные графики с временной шкалой) и текстовые представления (таблицы).

Исследование инцидентов безопасности обычно выполняется с помощью интерактивных таблиц, элементы которых часто реализуются в виде гиперссылок, по которым осуществляется переход к другим данным, благодаря чему значительно упрощается поиск требуемой информации. Некоторые SIEM-системы используют более сложные модели визуализации для анализа событий безопасности и определения возможного



■ **Рис. 1.** Панель просмотра данных от NetFlow в SIEM-системе OSSIM
 ■ **Fig. 1.** The dashboard for monitoring Netflow data in the OSSIM SIEM



■ **Рис. 2.** Панель исследования событий безопасности в подключаемом компоненте QRadar Pulse SIEM-системы IBM QRadar
 ■ **Fig. 2.** The dashboard for security incident investigation in the QRadar Pulse, offense visualization application for the QRadar SIEM

сценария атак. Так, например, в SIEM-системе IBM QRadar для исследования распределения атак различного типа в пространстве предложена трехмерная модель земного шара, на который наносятся события разного типа (рис. 2) [13].

Для представления событий используется сложный глиф, характеризующий следующие аспекты: уровень критичности событий (высота графического элемента на глобусе); частоту событий (размер круга-основания графического

элемента на круге); тенденцию частоты новых событий (концентрические круги вокруг графического элемента). Связь между событиями показана дугами, соединяющими исходный IP-адрес и IP-адрес назначения. Детали событий безопасности отображаются с помощью графа на параллельных координатах, оси которого достаточно традиционны для исследования сетевых потоков: IP-адрес источника, номер порта получателя и IP-адрес получателя.

В ArcSight ESM имеется возможность представить последовательность событий в виде связанного графа, вершинами которого являются внешние и внутренние хосты компьютерной сети и сетевые порты, ребра графа связывают атакующий и атакуемый хосты. Цветом указывается статус хоста — источник атаки (красный цвет), атакующий хост, захваченный в процессе атаки (голубой цвет), и атакуемый на текущий момент хост (белый цвет).

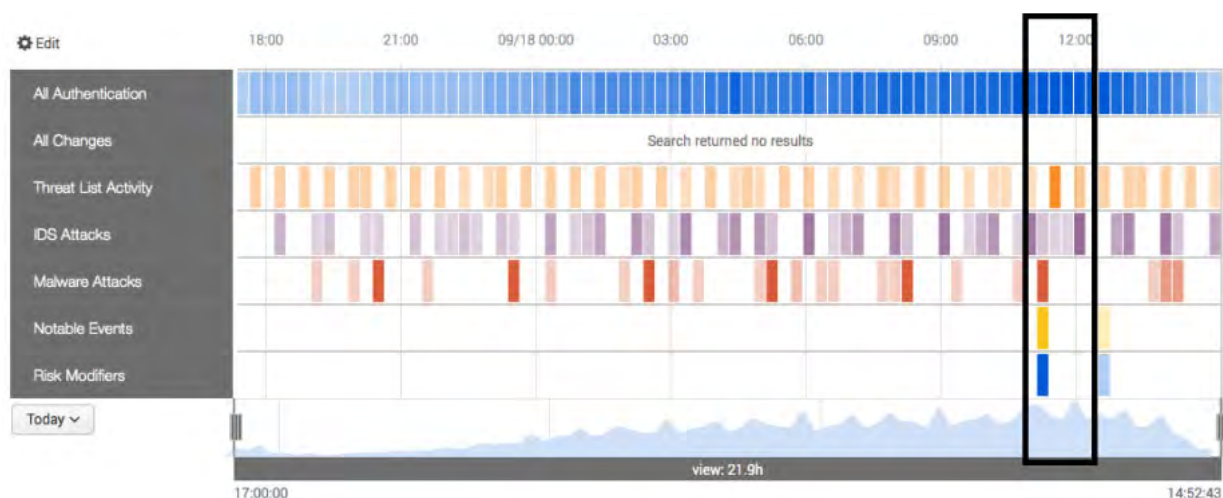
Интересная модель визуализации для анализа событий безопасности одного хоста на основе матричного представления предложена в системе Splunk (рис. 3) [14].

События безопасности сгруппированы по сенсорам безопасности, которые их генерируют: события системы аутентификации пользователей, события системы обнаружения вторжений, значимые события, определяемые по правилам пользователей, события антивирусных решений и т. д. События отображаются по оси Y. Тип события на графике кодируется цветом, причем интенсивность цвета указывает число событий в заданный период времени. По оси X откладывается время. Внизу матричного представления расположен обычный линейный график, который отображает общее количество событий во

времени. Отличительными характеристиками данной модели визуализации являются возможность установить временные связи между событиями разного типа и ее относительная простота. Например, на рис. 3 выделен интервал времени, которому соответствует максимальное число событий всех типов. При этом легко установить, что сначала от системы обнаружения вторжений было зафиксировано большое число событий, за которыми последовали события, классифицированные как проявление вредоносной активности. Они обозначены как значимые события, связанные с изменением уровня риска хоста. Затем были объявлены события, входящие в список информационных угроз.

Исследование SIEM-систем, являющихся лидерами рынка систем управления информационной безопасностью [15], показало, что основными методиками визуализации в SIEM-системах являются: интерактивные таблицы; стандартные двумерные и трехмерные методики визуализации, представленные столбчатыми и линейчатыми диаграммами, линейными графиками с временной шкалой и круговыми диаграммами; географические карты, в том числе трехмерные, используемые в сочетании с пиктограммами для отображения обобщенных показателей безопасности; связанные графы, также часто используемые наряду с пиктограммами для описания статуса объекта, обозначенного вершиной графа.

Стандартные двумерные модели визуализации в основном отображают различную статистическую информацию. Связные графы применяются для отображения топологии компьютерных сетей, информационных ресурсов сети, информационных потоков между узлами [12–14, 16]. Они также используются для графического



■ **Рис. 3.** Визуальный анализ событий безопасности одного хоста

■ **Fig. 3.** Visualization of the security events per computer network asset

представления связей между событиями безопасности, действиями пользователей в системе. Практически во всех SIEM-системах реализована поддержка географических карт для отображения географического расположения хостов контролируемой сети и событий безопасности. С их помощью пользователь имеет возможность отслеживать развитие атаки в пространстве. Кроме того, они помогают системному администратору определять границы ответственности подразделений организации.

Перечисленные выше методики визуализации предназначены для представления одномерных и двумерных данных, легко воспринимаются человеком, что делает их полезными для выполнения задач оперативного контроля и представления результатов исследования. Исследование событий безопасности в большинстве случаев заключается в изучении данных, представленных в разных таблицах, лишь в некоторых случаях имеется возможность визуально проанализировать события безопасности, построив граф событий. Вместе с тем именно SIEM-системы дают возможность построить полноценную систему визуального анализа состояния информационной безопасности благодаря уже реализованным механизмам сбора и подготовки данных от разнородных источников данных. Таким образом, очевидна необходимость использования в SIEM-системах моделей графического представления многомерных данных, которые позволяют выявлять скрытые зависимости как между атрибутами объектов, так и между данными от разных источников, что в конечном итоге позволит обнаруживать новые сценарии атак.

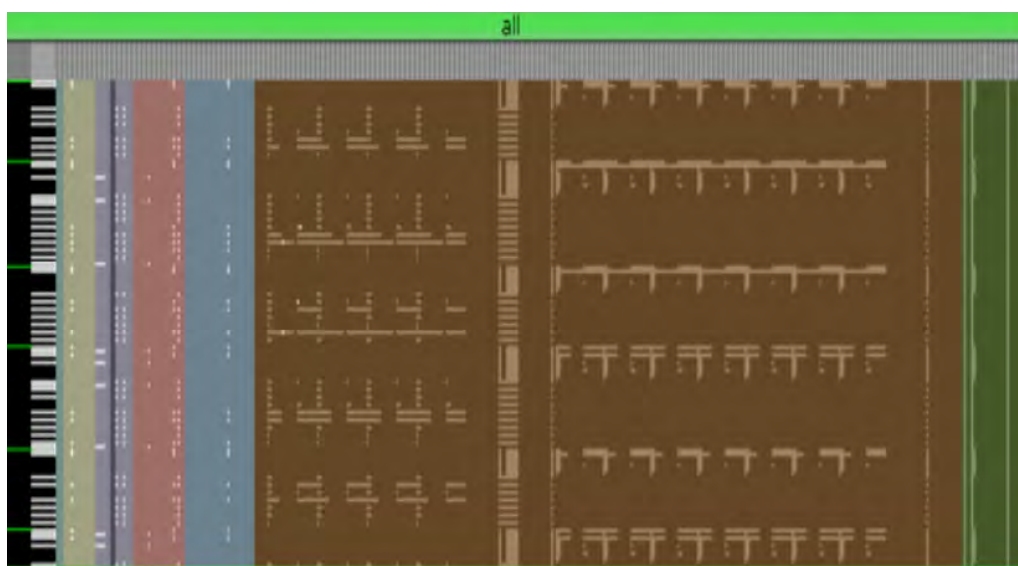
Определенные попытки для устранения этой проблемы предприняты ведущими поставщиками SIEM-систем, которые выпустили подключаемые модули, реализующие механизмы визуального анализа для исследования событий безопасности [13–15]. Однако предложенные механизмы визуального анализа опираются на определенные атрибуты сетевого трафика или событий безопасности, которые показали свою состоятельность при обнаружении традиционных сетевых атак и, соответственно, широко используются в системах обнаружения вторжений (СОВ). В связи с этим необходимость в использовании предложенных методик визуального анализа резко снижается, поскольку они, по сути, дублируют результат работы системы обнаружения вторжений. Одним из исключений является SIEM-система Splunk, в которой именно пользователь определяет исходные данные для анализа, задает модель их графического представления и методики взаимодействия с ней. Однако подобная свобода действий предполагает наличие у пользователей опыта по проектированию

моделей визуального анализа, панелей управления, а также навыков по разработке программ. Очевидно, что подобными навыками обладают далеко не все специалисты по информационной безопасности.

Вызовы и возможные решения

Анализ существующих методик обозначил противоречия между возможностями SIEM-систем по анализу данных и реализацией этих возможностей. С одной стороны, SIEM-системы являются удобной платформой для построения системы визуального анализа благодаря масштабируемым и эластичным технологиям сбора, предварительной обработке данных от разнородных источников, а особенности архитектурного решения SIEM-систем позволяют использовать подключаемые компоненты, реализующие методики визуального анализа. С другой стороны, представленные методики визуального анализа в основном предназначены для выполнения задач оперативного контроля и представления результатов. Практически отсутствуют методики визуального анализа, поддерживающие исследование событий безопасности, визуальную корреляцию данных, визуальную валидацию автоматических моделей анализа, позволяющую оценить эффективность и корректность их функционирования. Возможным решением этой проблемы является применение методик, осуществляющих анализ журналов сетевых устройств и сенсоров безопасности независимо от их природы и без привязки к значениям определенных атрибутов [4, 17–21]. Это позволяет обнаруживать качественно новые атаки и взаимосвязи между событиями. Данные методики можно условно разделить на две группы: 1) построение универсального графического анализатора журнала сетевых сенсоров безопасности и сетевых устройств, который выполняет анализ логов по некоторому множеству правил-фильтров, заданному пользователем [4, 19]; 2) выполнение визуальной корреляции атрибутов логов, генерируемых несколькими сенсорами безопасности, причем атрибуты определяются пользователем, исходя из содержимого логов устройства [18, 20].

Примером первого подхода является методика, представленная в работе [21]. Сырые сетевые пакеты, поступающие от различных сенсоров сетевого трафика и информационной безопасности, преобразуются в сообщения, имеющие иерархическую структуру и заданные с помощью XML-подобного языка Packet Detail Markup Language (PDML), который описывает содержимое пакета, учитывая различные атрибуты сетевых протоколов. Пакеты отображаются вдоль оси Y, а атрибуты сетевых па-



■ *Рис. 4.* Пиксель-ориентированное представление логов сетевых устройств

■ *Fig. 4.* Pixel-oriented visualization of the security devices' logs

кетов — вдоль оси X . Таким образом, каждое сообщение представляет собой последовательность пикселей. Яркость (i, j) пикселя обозначает частоту атрибута j в сообщении i . На рис. 4 представлена предложенная авторами модель визуализации логов сетевых устройств [21].

Важным моментом в использовании данного подхода является корректный выбор цветовой схемы кодирования значения параметра, отображаемого с помощью пикселя. Она должна позволять акцентировать внимание аналитика на аномалии в данных, например, на редко встречаемых и часто встречаемых элементах, при этом следует помнить, что редко встречаемые значения не всегда являются признаком аномальной активности. Следует также учитывать, что из-за мелких размеров пиксели разного оттенка, но одинаковой степени насыщенности, плохо различаются человеком [22], и в результате для анализа создаваемого графического представления аналитику приходится прилагать значительные усилия. Другим очевидным недостатком этого подхода является возможность анализировать данные только от одного источника, что не позволяет выявлять связи между показаниями разных источников.

Вторая группа инструментов дает возможность визуально оценивать совокупность логов от разных устройств, а анализ выполняется по множеству атрибутов записей журналов, выбираемых самим пользователем. Для их представления используются достаточно простые модели визуализации — линейные графики, графики рассеивания, столбиковые диаграммы. В работе [18] «привязка» событий от разных источников

данных осуществляется путем использования линейных графиков с общей временной шкалой. В [20] формируются временные срезы, в которых состояние информационной системы неизменно относительно параметров, заданных пользователями (число событий безопасности, число получателей/отправителей сетевых потоков, число запущенных программ и т. д.). Для каждого такого среза строится множество моделей визуализации, отражающих основные характеристики потоков данных за данный период времени. Таким образом, несмотря на то, что явной привязки событий к шкале времени нет, имеется возможность оценить общий характер потоков и увидеть определенные связи между их атрибутами. Использование временных срезов как средства агрегации характеристик потоковых данных может быть рассмотрено как механизм снижения когнитивной нагрузки на аналитика при анализе высокоскоростных потоков данных большого объема: графическое представление должно быть достаточно предсказуемым и постоянным, не зависящим от скорости, объема и разнообразия исходных данных. В качестве возможных недостатков предложенного подхода можно выделить использование достаточно сложных моделей визуализации (плоскостных карт деревьев, связанных графов), которые могут отличаться от одного временного среза к другому, что требует их тщательного изучения со стороны аналитика. Кроме того, модели визуализации отображают данные, представленные только в текущем срезе потока, что не позволяет оценить изменения параметров потока данных в контексте состояния всей информационной системы.

Панель управления автоматической моделью анализа сетевого трафика в инфраструктуре облачных вычислений

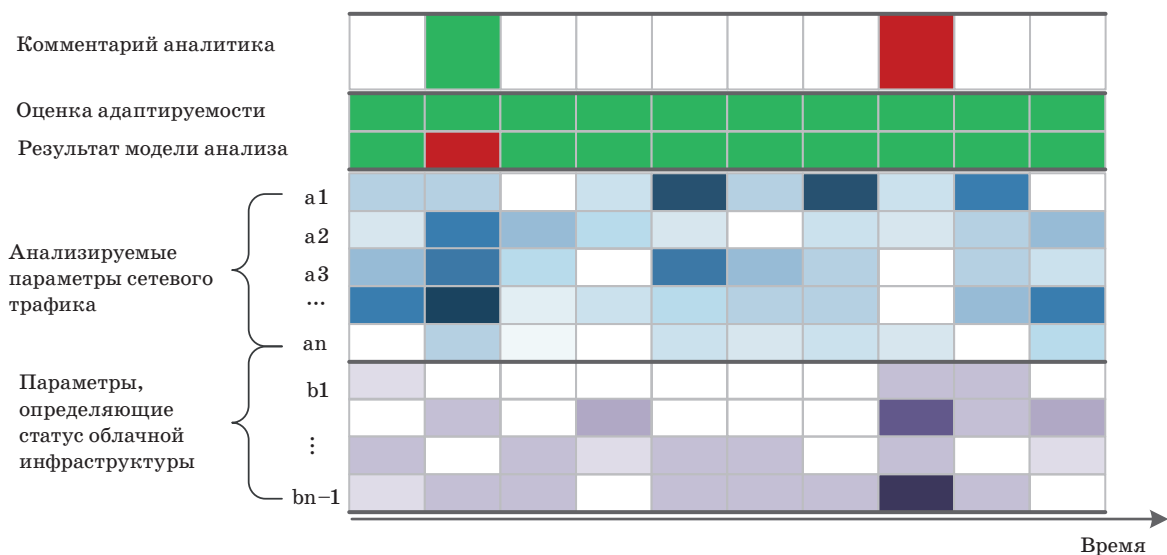
Задача визуальной валидации автоматических моделей анализа сетевого трафика и выявления аномалий связана с предоставлением механизмов мониторинга безопасности, используемых для выявления различных информационных угроз. Они должны позволять пользователю отслеживать результаты работы моделей анализа, оценивать уровень ее адаптируемости к изменениям в потоках данных и возможности переобучения и, в случае необходимости, отмечать ситуации ложных срабатываний, необходимые для последующего анализа и настройки или доработки.

Исходя из этих требований была разработана панель управления, предназначенная для мониторинга функционирования модели анализа сетевого трафика, применяемой для обнаружения атак в инфраструктуре облачных вычислений [23]. В ее основе лежат модели классификации. Механизм переобучения моделей, адаптирующий систему защиты к изменениям структуры и объема трафика в системе облачных вычислений, основан на оценке параметра уверенности (confidence) обученной модели в предсказании типа трафика. Авторы предлагают исследовать результат ее работы с учетом параметров сетевого трафика, на основе которых принимается решение о его легитимности, внутренней оценки адаптируемости модели анализа к изменениям в сетевом трафике, а также параметров, описывающих общий статус облачной инфраструктуры. Первые две группы параметров фиксированы,

а последняя задается пользователем системы обнаружения атак в системе облачных вычислений и предназначена для выявления косвенных признаков наличия необнаруженной вредоносной активности в системе.

Ключевым элементом разработанной панели управления является матричное представление данных (рис. 5). Выбор данной модели обусловлен тем фактом, что она позволяет выполнять визуальную корреляцию данных, а также выявлять определенные паттерны в них, например, изменения в сетевом трафике, периодически повторяющиеся во времени. Контролируемые параметры откладываются по оси Y, а время — по оси X. Во время функционирования модели каждые 5 секунд формируются векторы, характеризующие сетевой трафик. Однако анализировать такой объем информации затруднительно даже визуально, поэтому было решено отображать интервалы времени, в течение которых в данных не происходит значительных изменений, и модель анализа классифицирует их одинаково. Если характер данных меняется сильно, то формируется новый срез данных.

Для определения значимости изменений в данных и необходимости формировать новый срез данных используются принципы кластеризации на основе оценки плотности распределения объектов в пространстве. Анализируемый моделью вектор данных рассматривается как точка в многомерном пространстве, при появлении новой точки вычисляются расстояния до множества точек, уже входящих в текущий срез данных, и если максимальное расстояние не превышает некоторого порогового значения ϵ , задаваемого пользователем, то точка включается



■ **Рис. 5.** Представление данных для валидации модели анализа трафика
 ■ **Fig. 5.** Data visualization for traffic analysis model validation

в текущий срез данных, в противном случае формируется новый срез. В качестве метрики расстояния применяется евклидова метрика. Новый срез данных формируется также в том случае, если меняется результат анализа сетевых потоков.

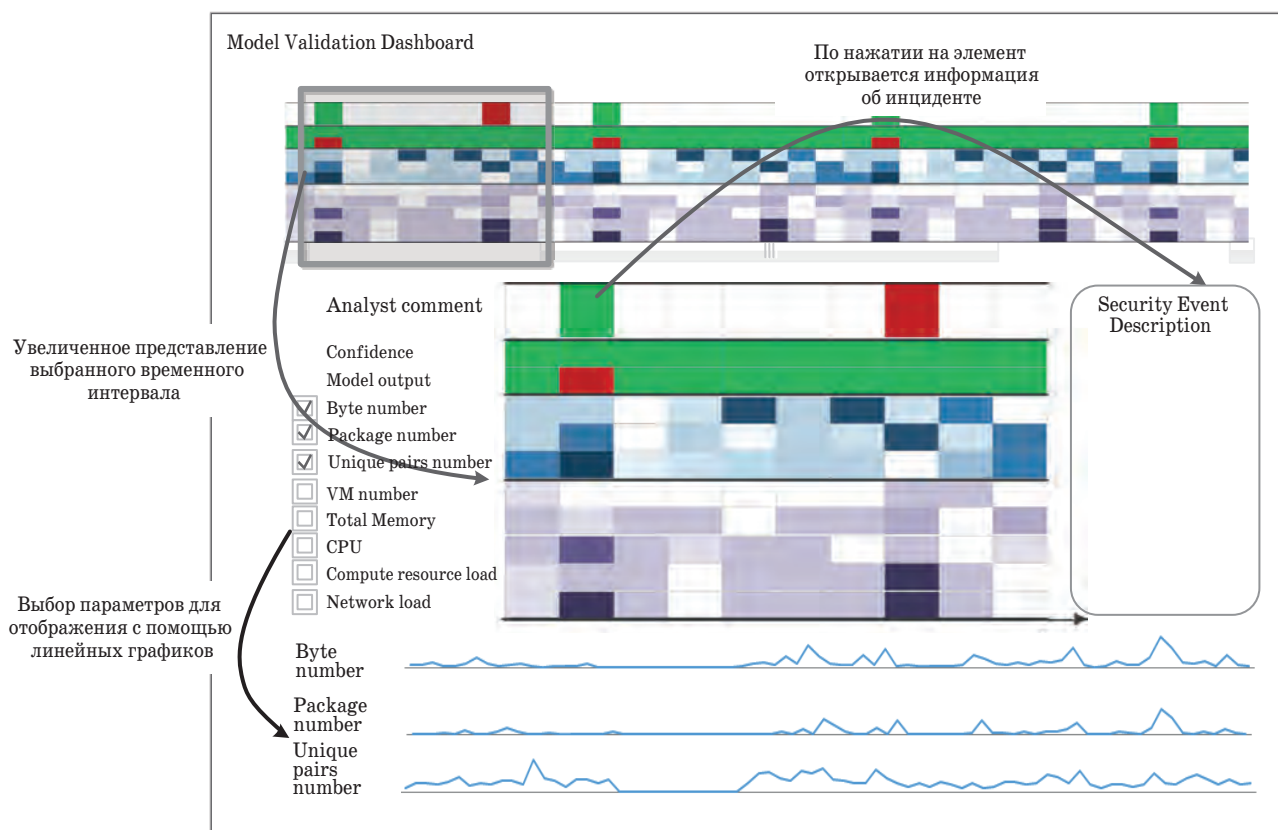
Откладываемые интервалы времени по шкале X могут иметь разную длительность, несмотря на то, что визуально они одинаковы. Элементом матрицы является среднее значение параметра, вычисленное для множества значений заданного среза данных. Полученное значение кодируется насыщенностью цвета. Чем насыщенней цвет элемента, тем выше значение параметра. Значение насыщенности цвета формируется путем нормализации значений в пределах интервала [0, 1]. Максимальное и минимальное значения параметра определяются или за весь исследуемый период времени (при использовании модели визуализации для исторического анализа данных), или за последний заданный интервал времени, например неделю, при мониторинге данных в режиме реального времени. Оттенок цвета используется для обозначения параметров, принадлежащих к одной группе.

Таким образом, пользователь, визуально сопоставляя значения параметров, генерируемых различными источниками, имеет возможность

выявить скрытые взаимосвязи между ними, построить графические паттерны функционирования облачных сервисов. Аналитик может также обнаружить тенденции в изменении сетевого трафика и выявить подозрительные ситуации, которые не были распознаны как подозрительные.

Отдельная цветовая схема применяется для представления результатов модели анализа и результатов исследования инцидентов, выполненных аналитиком вручную. Результат модели анализа кодируется следующим образом. Если трафик классифицирован как легитимный, то соответствующий элемент матрицы окрашен зеленым цветом. В противном случае элемент матрицы имеет красный цвет. Для обозначения уровня уверенности модели также используется двухцветная палитра: красным цветом кодируются значения ниже порогового значения, поскольку в этом случае инициируется переобучение модели, зеленым цветом — значения выше порогового. Результаты ручного анализа кодируются тоже зеленым и красным цветами: зеленым, если результат анализа аналитика совпадает с решением модели анализа, красным, если не совпадает.

Построенная модель визуализации позволяет анализировать характер изменений в сетевом трафике, выявлять периодические паттерны функ-



■ **Рис. 6.** Макет панели визуальной валидации модели анализа сетевого трафика

■ **Fig. 6.** Dashboard prototype for visual validation of the network analysis model

ционирования инфраструктуры облачных вычислений, выполнять визуальную корреляцию данных, отслеживать эффективность работы модели анализа, отображать результаты исследования инцидентов безопасности и сравнивать их с решениями, полученными на базе модели анализа.

Недостатком предложенной модели визуализации является нелинейность шкалы времени из-за особенности формирования срезов данных. Для устранения этого недостатка было предложено использовать дополнительные модели визуализации. При их выборе авторы руководствовались тем, что специалисты по информационной безопасности предпочитают работать со стандартными двумерными графическими представлениями (линейными графиками, гистограммами) [24], поэтому детальное представление данных осуществляется с помощью системы линейных графиков с общей шкалой времени. Они отображают изменения выбранных параметров во времени.

Матричное представление и система линейных графиков связаны между собой набором механизмов взаимодействия. На рис. 6 показан макет панели управления, поддерживающей процесс визуальной валидации модели анализа.

В верхней части панели управления расположено матричное представление параметров, оно формирует общую картину того, каким образом функционирует автоматическая модель анализа сетевых потоков. Кроме того, оно может быть использовано для мониторинга в режиме реального времени. В этом случае новые значения параметров добавляются справа. Поскольку данная модель является достаточно мелкой, для работы с ней предложен механизм увеличения, который позволяет выбрать часть матрицы и увеличить ее. Увеличенная часть матрицы находится в центре панели управления, и именно с ней в основном взаимодействует аналитик. Первая строка матрицы является редактируемой, аналитик может вносить свои результаты исследования инцидентов, отмечая достоверность результатов автоматической модели анализа. Кроме того, он может записывать свои комментарии в текстовую форму «*Security Event Description*». Параметры,

отображаемые с помощью линейных графиков, также задаются пользователем, они выбираются из множества параметров, для которых строится матричное представление.

Описанная панель управления реализована в системе управления облачной инфраструктурой OpenStack для мониторинга модели обнаружения атак в виде подключаемого компонента интерфейса пользователя Horizon.

Заключение

В статье исследованы известные методики визуального анализа, представленные в системах управления безопасностью, в частности в SIEM-системах, для исследования событий безопасности. Выделены противоречия между возможностями систем, связанных с механизмами подготовки исходных для анализа данных, и практической реализацией существующих решений.

Рассмотрены возможные пути решения представленных проблем, причем для каждой группы решений показаны потенциальные достоинства и недостатки, что позволит разработчикам в сфере информационной безопасности выбрать методику визуального анализа данных, подходящую для решения конкретной задачи. В работе также представлена панель управления, предназначенная для валидации автоматической модели выявления атак в инфраструктуре облачных вычислений, которая была разработана с учетом рассмотренных методик построения систем визуального анализа.

Дальнейшее направление исследований связано с выполнением оценки эффективности предложенной модели визуализации, ее адаптацией для контроля других моделей анализа, применяемых для решения различных задач информационной безопасности, а также с уточнением механизмов формирования срезов данных.

Работа выполнена при частичной финансовой поддержке РФФИ (проекты 16-29-09482, 18-37-20047, 18-07-01488 и 18-29-22034) и бюджетной темы 0073-2019-0002.

Литература

1. **Marty R.** *Visual Analytics — Delivering Actionable Security Intelligence*. <https://www.blackhat.com/us-16/training/visual-analytics-delivering-actionable-security-intelligence.html> (дата обращения: 16.01.2019).
2. **Guimarães V. T., Freitas C. M. D. S., Sadre R., Tarouco L. M. R., Granville L. Z.** A survey on information visualization for network and service management. *IEEE Communications Surveys & Tutorials*, 2016, vol. 18, no. 1, pp. 285–323. doi:10.1109/COMST.2015.2450538
3. **Best D. M., Endert A., Kidwell D.** 7 key challenges for visualization in cyber network defense. *Proc. of the Eleventh Workshop on Visualization for Cyber Security (VizSec 2014)*, ACM, 2014, pp. 33–40. doi:10.1145/2671491.2671497
4. **Stange J.-E., Dörk M., Landstorfer J., and Wettsch R.** Visual filter: graphical exploration of network security log files. *Proc. of the Eleventh Workshop on Visualization for Cyber Security (VizSec 2014)*, ACM, 2014, pp. 41–48. doi:10.1145/2671491.2671503

5. Котенко И. В., Новикова Е. С. Визуальный анализ защищенности компьютерных сетей. *Информационно-управляющие системы*, 2013, № 3, с. 56–61.
6. Zhang T., Wang X., Li Z., et al. A survey of network anomaly visualization. *Sci. China Inf. Sci.*, 2017, vol. 60, pp. 121101. doi:10.1007/s11432-016-0428-2
7. Krause J., Perer A., Bertini E. Using visual analytics to interpret predictive machine learning models. *Proc. of the 2016 ICML Workshop on Human Interpretability in Machine Learning (WHI 2016)*, New York, NY, USA, 2016. <https://arxiv.org/abs/1606.05685> (дата обращения 16.01.2019).
8. Котенко Д. И., Котенко И. В., Саенко И. Б. Методы и средства моделирования атак в больших компьютерных сетях: состояние проблемы. *Тр. СПИИРАН*, 2012, № 3 (22), с. 5–30.
9. Котенко И. В., Степашкин М. В. Анализ защищенности компьютерных сетей на основе моделирования действий злоумышленников и построения графа атак. *Тр. Института системного анализа Российской академии наук*, 2007, т. 31, с. 126–207.
10. Котенко И. В., Саенко И. Б. Архитектура системы интеллектуальных сервисов защиты информации в критически важных инфраструктурах. *Тр. СПИИРАН*, 2013, № 1 (24), с. 21–40.
11. Few S. *Information dashboard design the effective visual communication of data*. N. Y., O'Reilly, 2006. 224 p.
12. SIEM-система OSSIM. <https://www.alienvault.com/products/ossim> (дата обращения: 16.01.2019).
13. Компонент визуальной аналитики Qradar Pulse SIEM-системы IBM Radar. https://www.ibm.com/support/knowledgecenter/en/SS42VS_7.2.8/com.ibm.apps.doc/c_Qapps_Pulse_intro.html (дата обращения: 16.01.2019).
14. SIEM-система Splunk Enterprise. https://www.splunk.com/ru_ru/products/splunk-enterprise.html (дата обращения: 16.01.2019).
15. Gartner Magic Quadrant for Security Information and Event Management. <https://www.gartner.com/doc/3894573> (дата обращения: 16.01.2019).
16. Компонент визуальной аналитики Interactive Discovery SIEM-системы HP ArcSight. <https://community.softwaregrp.com/t5/ArcSight-Interactive-Discovery/tkb-p/arc-sight-interactive-discovery> (дата обращения: 16.01.2019).
17. Walton S., Maguire E., and Chen M. Multiple queries with conditional attributes (QCATs) for anomaly detection and visualization. *Proc. of the Eleventh Workshop on Visualization for Cyber Security (VizSec 2014)*, ACM, 2014. doi:10.1145/2671491.2671502
18. Humphries C., Prigent N., Bidan C., and Majorczyk F. CORGI: combination, organization and reconstruction through graphical interactions. *Proc. of the Eleventh Workshop on Visualization for Cyber Security (VizSec 2014)*, ACM, 2014, pp. 57–64. doi:10.1145/2671491.2671494
19. Landstorfer J., Herrmann I., Stange J., Dork M., and Wettach R. Weaving a carpet from log entries: A network security visualization built with cocreation. *Proc. of the Visual Analytics Science and Technology (VAST)*, Paris, 2014, pp. 73–82. doi:10.1109/VAST.2014.7042483
20. Fischer F., Keim D. A. NstreamAware: real-time visual analytics for data streams to enhance situational awareness. *Proc. of the Eleventh Workshop on Visualization for Cyber Security (VizSec 2014)*, ACM, 2014, pp. 65–72. doi:10.1145/2671491.2671495
21. Cappers B. C. M., van Wijk J. SNAPS: Semantic network traffic analysis through projection and selection. *Proc. of the IEEE Symposium on Visualization for Cyber Security (VizSec 2015)*, 2015, pp. 1–8. doi:10.1109/VIZSEC.2015.7312768
22. Stone M. In color perception, size matters. *IEEE Computer Graphics and Applications*, 2012, vol. 32(2), pp. 8–13. doi:10.1109/MCG.2012.37
23. Borisenko K., Smirnov A., Novikova E., and Shorov A. DDoS attacks detection in cloud computing using data mining techniques. *Proc. of the Industrial Conf. on Data Mining (ICDM'2016)*, LNAI, vol. 9728, pp. 1–15. doi:10.1007/978-3-319-41561-1_15
24. McKenna S., Staheli D., Fulcher C., Meyer M. BubbleNet: A cyber security dashboard for visualizing patterns. *Computer Graphics Forum*, 2016, vol. 35, pp. 281–290. doi:10.1111/cgf.12904

UDC 004.056

doi:10.31799/1684-8853-2019-2-57-67

Open challenges in visual analytics for security information and event management

E. S. Novikova^{a,b}, PhD, Tech, Associate Professor, orcid.org/0000-0003-2923-4954, novikova.evgenia123@gmail.com

I. V. Kotenko^a, Dr. Sc., Tech, Professor, orcid.org/0000-0001-6859-7120

^aSaint-Petersburg Institute for Informatics and Automation of the RAS, 39, 14 Line, V. O., 199178, Saint-Petersburg, Russian Federation

^bSaint-Petersburg Electrotechnical University «LETI», 5, Prof. Popov St., 197376, Saint-Petersburg, Russian Federation

Introduction: Visual analytics techniques support efficient analysis of the ever-growing amounts of data generated by security sensors and facilitate a timely and reasonable response to the threats. Modern security information and event management systems propose various solutions for processing large data streams and integrating heterogeneous sources which can be used as a framework

to construct a visual analytics system for security tasks. **Purpose:** The analysis of visual analytics techniques implemented in security information and event management systems and designed to support the studies on security incidents in the context of the main visual analytics problems, including the validation of automatic analysis models. **Results:** A contradiction has been detected between the capabilities of security information and event management systems in the visual analysis of security data and the implementation of these capabilities. Techniques for visual correlation of the data from different security sensors and for visual validation of automatic analysis models which would allow you to evaluate their accuracy and adaptability to the changes in data streams are almost missing. A possible way to resolve this contradiction is using techniques which support a flexible mechanism for adjusting the analyzed attributes of the network device events. The article presents the main approaches to the development of such techniques, discussing their advantages and disadvantages. We propose a dashboard for monitoring the behavior of an automated network traffic analysis model used in a cloud computing infrastructure. It allows you to monitor the analysis model behavior, perform a visual correlation of the analyzed parameters, and track changes in the network flows. **Practical relevance:** The results of the research can be used when designing security visual analytics tools for monitoring data flows and the behavior of automated analysis models.

Keywords — visual analytics, security information and event management, visual data correlation, visual validation of analysis models, matrix-base data visualization.

For citation: Novikova E. S., Kotenko I. V. Open challenges in visual analytics for security information and event management. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2019, no. 2, pp. 57–67 (In Russian). doi:10.31799/1684-8853-2019-2-57-67

References

- Marty R. *Visual Analytics — Delivering Actionable Security Intelligence*. Available at: <https://www.blackhat.com/us-16/training/visual-analytics-delivering-actionable-security-intelligence.html> (accessed 16 January 2019).
- Guimarães V. T., Freitas C. M. D. S., Sadre R., Tarouco L. M. R. and Granville L. Z. A survey on information visualization for network and service management. *IEEE Communications Surveys & Tutorials*, 2016, vol. 18, no. 1, pp. 285–323. doi:10.1109/COMST.2015.2450538
- Best D. M., Endert A., Kidwell D. 7 key challenges for visualization in cyber network defense. *Proc. of the Eleventh Workshop on Visualization for Cyber Security (VizSec 2014)*, ACM, 2014, pp. 33–40. doi:10.1145/2671491.2671497
- Stange J.-E., Dörk M., Landstorfer J., and Wettach R. Visual filter: graphical exploration of network security log files. *Proc. of the Eleventh Workshop on Visualization for Cyber Security (VizSec 2014)*, ACM, 2014, pp. 41–48. doi:10.1145/2671491.2671503
- Kotenko I. V., Novikova E. S. The visual analysis for computer network security assessment. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2013, no. 3, pp. 56–61 (In Russian).
- Zhang T., Wang X., Li Z., et al. A survey of network anomaly visualization. *Sci. China Inf. Sci.*, 2017, vol. 60, pp. 121101. doi:10.1007/s11432-016-0428-2
- Krause J., Perer A., Bertini E. Using visual analytics to interpret predictive machine learning models. *Proc. of the 2016 ICML Workshop on Human Interpretability in Machine Learning (WHI 2016)*, New York, NY, USA, 2016. Available at: <https://arxiv.org/abs/1606.05685> (accessed 16 January 2019).
- Kotenko D. I., Kotenko I. V., Saenko I. B. Methods and tools for attack modeling in large computer networks: state of the problem. *Trudy SPIIRAN* [SPIIRAS Proceedings], 2012, no. 3 (22), pp. 5–30 (In Russian).
- Kotenko I. V., Stepashkin M. V. Security analysis of computer networks based on modeling by malefactors actions and constructing attack graph. *Trudy Instituta sistemnogo analiza Rossijskoj akademii nauk* [Proceeding of the Institute for Systems Analysis of the Russian Academy of Science], 2007, vol. 31, pp. 126–207 (In Russian).
- Kotenko I. V., Saenko I. B. Architecture of the system of intelligent information security services in critical infrastructures. *Trudy SPIIRAN* [SPIIRAS Proceedings], 2013, no. 1 (24), pp. 21–40 (In Russian).
- Few S. *Information dashboard design the effective visual communication of data*. New York, O'Reilly, 2006. 224 p.
- OSSIM SIEM system. Available at: <https://www.alienvault.com/products/ossim> (accessed 16 January 2019).
- Qadar Pulse Visual analytics application of the IBM Radar SIEM-system. Available at: https://www.ibm.com/support/knowledgecenter/en/SS42VS_7.2.8/com.ibm.apps.doc/c_Qapps_Pulse_intro.html (accessed 16 January 2019).
- Splunk Enterprise SIEM system. Available at: https://www.splunk.com/ru_ru/products/splunk-enterprise.html (accessed 16 January 2019).
- Gartner Magic Quadrant for Security Information and Event Management. Available at: <https://www.gartner.com/doc/3894573> (accessed 16 January 2019).
- Interactive Discovery component for the ArcSight SIEM-system. Available at: <https://community.softwaregrp.com/t5/ArcSight-Interactive-Discovery/tkb-p/arc-sight-interactive-discovery> (accessed 16 January 2019).
- Walton S., Maguire E., and Chen M. Multiple queries with conditional attributes (QCATs) for anomaly detection and visualization. *Proc. of the Eleventh Workshop on Visualization for Cyber Security (VizSec 2014)*, ACM, 2014. doi:10.1145/2671491.2671502
- Humphries C., Prigent N., Bidan C., and Majorczyk F. CORGI: combination, organization and reconstruction through graphical interactions. *Proc. of the Eleventh Workshop on Visualization for Cyber Security (VizSec 2014)*, ACM, 2014, pp. 57–64. doi:10.1145/2671491.2671494
- Landstorfer J., Herrmann I., Stange J., Dörk M., and Wettach R. Weaving a carpet from log entries: A network security visualization built with cocreation. *Proc. of the Visual Analytics Science and Technology (VAST)*, Paris, 2014, pp. 73–82. doi:10.1109/VAST.2014.7042483
- Fischer F., Keim D. A. NstreamAware: real-time visual analytics for data streams to enhance situational awareness. *Proc. of the Eleventh Workshop on Visualization for Cyber Security (VizSec 2014)*, ACM, 2014, pp. 65–72. doi:10.1145/2671491.2671495
- Cappers B. C. M., van Wijk J. SNAPS: Semantic network traffic analysis through projection and selection. *Proc. of the IEEE Symposium on Visualization for Cyber Security (VizSec 2015)*, 2015, pp. 1–8. doi:10.1109/VIZSEC.2015.7312768
- Stone M. In Color Perception, Size Matters. *IEEE Computer Graphics and Applications*, 2012, vol. 32(2), pp. 8–13. doi:10.1109/MCG.2012.37
- Borisenko K., Smirnov A., Novikova E., and Shorov A. DDoS attacks detection in cloud computing using data mining techniques. *Proc. of the Industrial Conference on Data Mining (ICDM'2016)*, LNAI, vol. 9728, pp. 1–15. doi:10.1007/978-3-319-41561-1_15
- McKenna S., Staheli D., Fulcher C., Meyer M. BubbleNet: A cyber security dashboard for visualizing patterns. *Computer Graphics Forum*, 2016, vol. 35, pp. 281–290. doi:10.1111/cgf.12904

Организация многоуровневого хранения данных

Б. Я. Советов^а, доктор техн. наук, профессор, orcid.org/0000-0003-3116-8810

Т. М. Татарникова^б, доктор техн. наук, доцент, orcid.org/0000-0002-6419-0072, tm-tatarn@yandex.ru

Е. Д. Пойманова^б, старший преподаватель, orcid.org/0000-0002-7903-2480

^аСанкт-Петербургский государственный электротехнический университет «ЛЭТИ», Профессора Попова ул., 5, Санкт-Петербург, 197376, РФ

^бСанкт-Петербургский государственный университет аэрокосмического приборостроения, Б. Морская ул., 67, Санкт-Петербург, 190000, РФ

Введение: в условиях увеличивающегося объема данных практически все организации нуждаются в формировании развитой инфраструктуры их хранения. Инфраструктура реализуется с применением технологий многоуровневого хранения — проектированием систем хранения данных, включающих устройства хранения, разнородные по физической природе и архитектуре доступа. К файлам также предъявляются требования по соблюдению сроков гарантированного хранения, в том числе на законодательном уровне. Отсутствие алгоритмов эффективной организации хранения данных с разными требованиями к хранению актуализирует необходимость новых моделей управления ресурсами системы хранения данных. **Цель:** разработка новых моделей и алгоритмов управления ресурсами систем хранения данных с учетом многоуровневого принципа хранения данных. **Результаты:** предложен алгоритм многоуровневого хранения данных, позволяющий выполнять распределение файлов в системе хранения данных в соответствии с последовательным применением механизмов вертикального, горизонтального и динамического размещения файлов. Алгоритмы размещения и миграции данных по уровням иерархии системы хранения данных позволяют применять анализ метаданных, что дает возможность реализовать проактивное управление ресурсами систем хранения данных с учетом сроков гарантированного хранения. Представлена модель хранения данных в виде матрицы, размер которой соответствует количеству вертикальных и горизонтальных уровней системы хранения данных. Рекомендовано проводить распределение файлов по ячейкам матрицы с использованием аппарата нейронных сетей Кохонена. Эффект применения сети Кохонена заключается в возможности анализа метаданных и частоты обращения к файлам одним этапом. Это позволяет отказаться от последовательного выполнения механизмов размещения. Результаты эксперимента показали действенность аппарата сетей Кохонена при решении задачи размещения файлов в многоуровневой системе хранения данных. **Практическая значимость:** предложенные модель и алгоритмы могут найти применение при проектировании систем хранения данных с учетом многоуровневого принципа хранения данных.

Ключевые слова — система хранения данных, многоуровневое хранение данных, время хранения файла, метаданные, миграция данных, нейронная сеть Кохонена.

Для цитирования: Советов Б. Я., Татарникова Т. М., Пойманова Е. Д. Организация многоуровневого хранения данных. *Информационно-управляющие системы*, 2019, № 2, с. 68–75. doi:10.31799/1684-8853-2019-2-68-75

For citation: Sovetov B. Ya., Tatarnikova T. M., Poymanova E. D. Organization of multi-level data storage. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2019, no. 2, pp. 68–75 (In Russian). doi:10.31799/1684-8853-2019-2-68-75

Введение

Необходимость в хранении данных можно встретить на уровне отдельного предприятия, корпорации, государственных структур [1].

Для предприятий и корпоративного сектора потребность хранить большой объем данных определяется действующими бизнес-процессами, в государственном секторе — переходом к межведомственному электронному документообороту и созданием ведомственных аналитических ресурсов. Не менее мощный поток данных создают и обычные пользователи, которые выкладывают в Интернет свои фотографии, видеоролики и активно обмениваются мультимедийным контентом в социальных сетях [2].

Инженерным решением реализации инфраструктуры хранения являются системы хранения данных (СХД), которые в основном выделяют в отдельную подсистему вычислительного

комплекса, например, центра обработки данных [3].

Система хранения данных — это архитектурное решение для подключения внешних устройств хранения данных разной физической природы [4].

Основной задачей при проектировании инфраструктуры хранения данных является эффективное управление ресурсами СХД, в основном емкостными. Ее решение осложняется наличием следующих обстоятельств:

— разнородностью СХД: устройства хранения в СХД могут быть разными по физической природе (например, магнитные, оптические, твердотельные) и по архитектуре (СХД прямого или сетевого доступа) [5, 6];

— разными требованиями к хранению данных: критически важным транзакционным системам типа биллинговые, процессинговые, ERP и т. п. требуются высоконадежные и производительные

СХД; аналитическим системам — высокая производительность и низкая стоимость в расчете на единицу хранения; для работы с файлами — функциональность и низкая стоимость хранения [7];

— отсутствием алгоритмов эффективной многоуровневой организации хранения данных с разными требованиями к хранению. Современные СХД по-прежнему остаются одномерными, поскольку перемещение данных между уровнями СХД определяется одной метрикой — временем, прошедшим с момента последнего обращения к информации [8, 9].

В статье приводится предложенный авторами алгоритм распределения данных в СХД с многоуровневой организацией, учитывающий типы файлов данных, нормативные сроки их хранения и частоту обращения к ним.

Алгоритм организации многоуровневого хранения данных

Исходя из многоуровневой архитектуры СХД будем считать, что каждый уровень структуры предполагает свои технологии хранения: RAID-массивы, автоматизированные библиотеки и носители длительного хранения [10–12].

Предлагается распределять файлы данных по уровням СХД в соответствии с последовательным применением следующих механизмов:

- выбора уровня СХД в зависимости от времени хранения (вертикальное размещение);
- выбора логического тома уровня хранения СХД в зависимости от размера файла и длины логического блока данных (горизонтальное размещение);
- миграции данных по уровням СХД в зависимости от частоты обращения к ним (динамическое размещение).

Таким образом, СХД можно представить в виде матричной структуры, где каждая ее ячейка — это том соответствующего уровня СХД, предназначенного для хранения данных с определенными метрическими характеристиками (рис. 1).

RAID	ФС ₁	ФС ₂	...	ФС _n
Автоматизированные библиотеки	Носитель 1	Носитель 2	...	Носитель n
Носители длительного хранения	Носитель 1	Носитель 2	...	Носитель n

■ **Рис. 1.** Матричная структура СХД

■ **Fig. 1.** The structure of the storage system in the form of a matrix

Идея механизма выбора уровня СХД (вертикальное размещение файлов) основана на анализе организационных метаданных, содержащих сведения о типе данных. Например, можно указывать тип в имени файла через точку перед расширением *F.bck.txt*. Таким образом, при сохранении необходимо указывать атрибут файла *F*:

ind (initial data) — исходные данные, которые размещаются на уровень RAID;

bck (backups) — резервные копии, архивные данные, которые будут храниться на уровне автоматизированной библиотеки;

ngd (next generation data) — данные бессрочного хранения.

Во вторую очередь происходит горизонтальное распределение файлов. Идея горизонтального размещения основана на выборе файловой системы (ФС) для уровня RAID и по типам носителей на нижних уровнях СХД.

На уровне RAID предлагается проводить анализ размера сохраняемых файлов и в зависимости от его значения размещать файлы в разных томах RAID. Каждый том при этом может иметь свою ФС с определенным размером логического блока данных. Основное правило горизонтального размещения:

$$\text{Если } f \in (f_i; f_{i+1}], \text{ то } \rightarrow a_{i+1} \leftrightarrow F \rightarrow \text{Том}_{i+1},$$

где *f* — размер сохраняемого файла *F*; *f_i*, *f_{i+1}* — левая и правая границы соответственно размера файла *F*, с которым может работать ФС; *a_{i+1}* — размер логического блока данных, которым оперирует ФС; *Том_{i+1}* — номер тома RAID, управляемый соответствующей ФС.

На нижних уровнях СХД предлагается разделение емкости по типам носителей, например, стример, DVD, BD для уровня автоматизированных библиотек и М-диск, стеклянный диск, ДНК для носителей длительного хранения:

$$\text{Если } f \in (f_i; f_{i+1}], \text{ то } F \rightarrow al_{i+1} (lt_{i+1}),$$

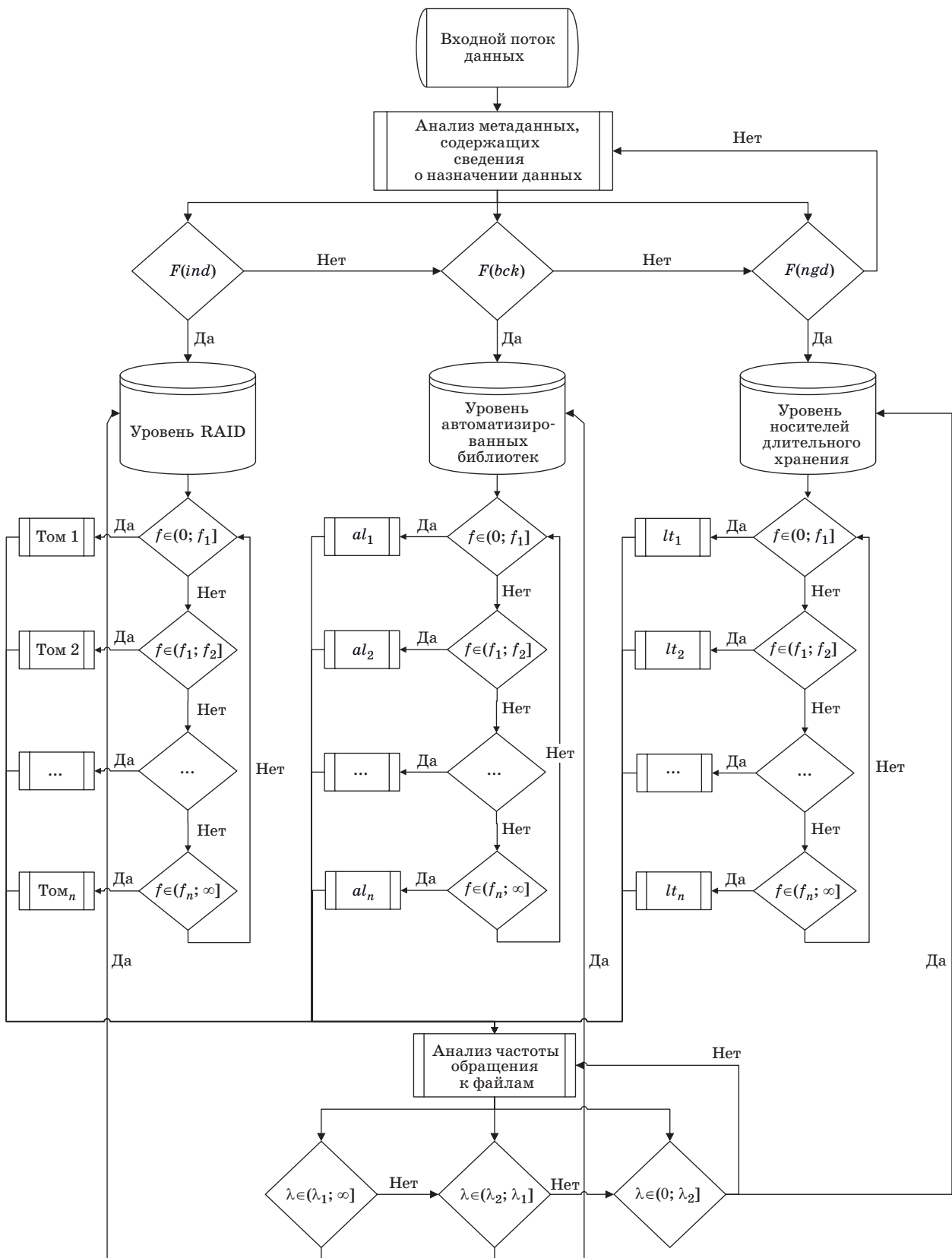
где *al_{i+1}* или *lt_{i+1}* — тип носителя на уровне автоматизированной библиотеки (*al*) или на уровне длительного хранения (*lt*).

Идея динамического размещения основана на миграции данных по уровням СХД в зависимости от частоты обращения к файлам данных [13]:

$$\text{Если } \lambda_F \in (\lambda_i; \lambda_{i+1}], \text{ то } F \rightarrow l,$$

где *λ_F* — частота запроса файла *F*; *λ_i*, *λ_{i+1}* — левая и правая границы соответственно частоты запроса файла; *l* — номер уровня СХД, на который мигрирует файл *F*.

Механизм миграции начинает функционировать только после первоначального распределе-



■ **Рис. 2.** Алгоритм управления емкостью СХД на основе анализа метаданных
 ■ **Fig. 2.** Storage capacity management algorithm based on metadata analysis

ния файлов при их записи в СХД. Введением механизма миграции можно преодолеть недостатки субъективного выбора типа сохраняемых файлов при реализации первого этапа (вертикального распределения).

Совокупность указанных механизмов позволяет управлять емкостью хранилища и осуществлять рациональное использование носителей. Отметим, что все указанные выше механизмы основаны на анализе метаданных сохраняемых файлов.

Блок-схема предлагаемого алгоритма управления емкостью СХД представлена на рис. 2.

Модель управления емкостью СХД

Работу механизмов вертикального и горизонтального размещения файлов в СХД в общем виде можно представить как матрицу размера $m \times n$, где m — это количество уровней хранения; n — количество физических или логических носителей. Элементами матрицы являются множества файлов, имеющие определенные значения характеристик: тип файла — $type$, размер файла — f (рис. 3). Частота обращения к файлам λ при их первоначальном размещении в СХД не учитывается, поскольку еще отсутствует статистика обращения к данным [14, 15].

В матрице всегда будет содержаться три строки, количество столбцов выбирается исходя из физической реализации каждого уровня СХД.

Последовательное выполнение алгоритмов размещения файлов требует больших энергетических затрат [16]. В связи с этим предлагается проводить распределение файлов по ячейкам матрицы рис. 3 на основе анализа метаданных файлов с использованием аппарата нейронных сетей Кохонена.

Выбор данного метода анализа обусловлен особенностями алгоритма, который реализует сеть Кохонена [17]:

	1	...	n
1	$type_1, f_1$...	$type_n, f_n$
...
m	...	$type_m, f_m$	$type_{m \times n}, f_{m \times n}$

■ Рис. 3. Матрица управления емкостью СХД

■ Fig. 3. Storage capacity matrix

1) используется неконтролируемое обучение, при котором правило обучения нейрона основано на информации о его расположении;

2) отсутствуют эталонные значения обучающего множества: каждому объекту из исходного множества соответствует строка таблицы, и исходное множество разбивается на классы в зависимости от векторов значений признаков его объектов;

3) результатом алгоритма является топологическая карта, в которой входные данные классифицируются на группы (кластеры).

В случае предлагаемого распределения файлов по ячейкам хранилища данных очевидно, что выбранные характеристики файлов задают векторы значений признаков объектов [18]. Таким образом, каждая ячейка получающейся карты должна соответствовать элементу матрицы управления емкостью СХД.

Описание и результаты эксперимента по распределению входящего потока файлов по ячейкам матрицы хранения

Задачей эксперимента являлась демонстрация метода нейронной сети Кохонена для распределения множества файлов по ячейкам матрицы хранения после проведенного обучения сети и нормализации исходных метаданных.

Работу предложенного метода проследим последовательно в соответствии с применяемыми механизмами:

1) вертикальное распределение файлов по уровням матрицы хранения в зависимости от предполагаемого времени хранения файлов;

2) горизонтальное распределение файлов по ячейкам уровня матрицы хранения в зависимости от типа и размера файлов;

3) динамическое распределение по уровням матрицы хранения в зависимости от частоты обращения к файлам.

В эксперименте участвовало 5000 файлов с различными характеристиками: тип файла, предполагаемое время его хранения, размер файла, частота обращения к данным. Поток сгенерирован на основе анализа 43 000 файлов, взятых с экспериментального файл-сервера [19]. Размер файлов не превышал 1 ГБ, частота обращения к данным является случайной величиной в интервале $[0; 300]$. Множество экспериментальных файлов имело следующий состав:

файлов типа ind — 4150 (83 % от общего количества файлов эксперимента);

файлов типа bck — 750 (15 % от общего количества файлов эксперимента);

файлов типа ngd — 100 (2 % от общего количества файлов эксперимента).

Кроме того, экспериментальное множество файлов было разбито на три части: 70 % всех файлов множества использовалось для обучения сети; 15 % — тестовое множество; 15 % — множество валидации.

Размещение файлов данных выполнялось в соответствии со структурой матрицы хранения размера 3×3 .

Нормализация типов файлов принята исходя из эксперимента по подбору значений, которые давали бы непересекающиеся классы. В результате получено следующее соответствие: тип файлов: *ind* соответствует значению 1000, *bck* — 2000, *ngd* — 3000.

Нормализация размера файлов выполнена в десятичных порядках и также принята исходя из эксперимента: при размере файла от 0 до 999 Б присваиваем значение атрибута 1000; от 1000 до 999 999 Б — 5000; от 1 000 000 до 999 999 999 Б — 10000.

Частота обращения рассматривалась в абсолютных единицах.

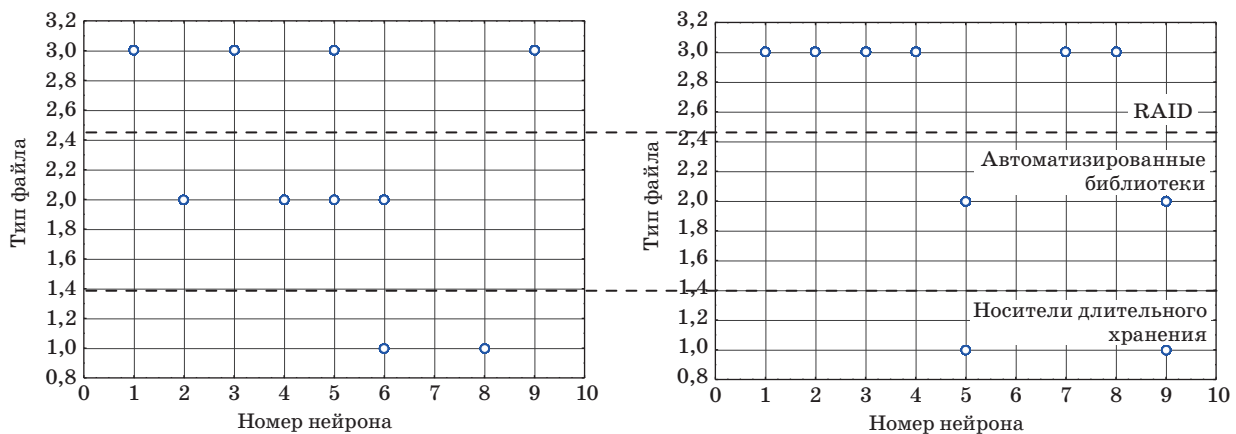
Количество тренировочных циклов — 1000.

Евклидово расстояние — 3.

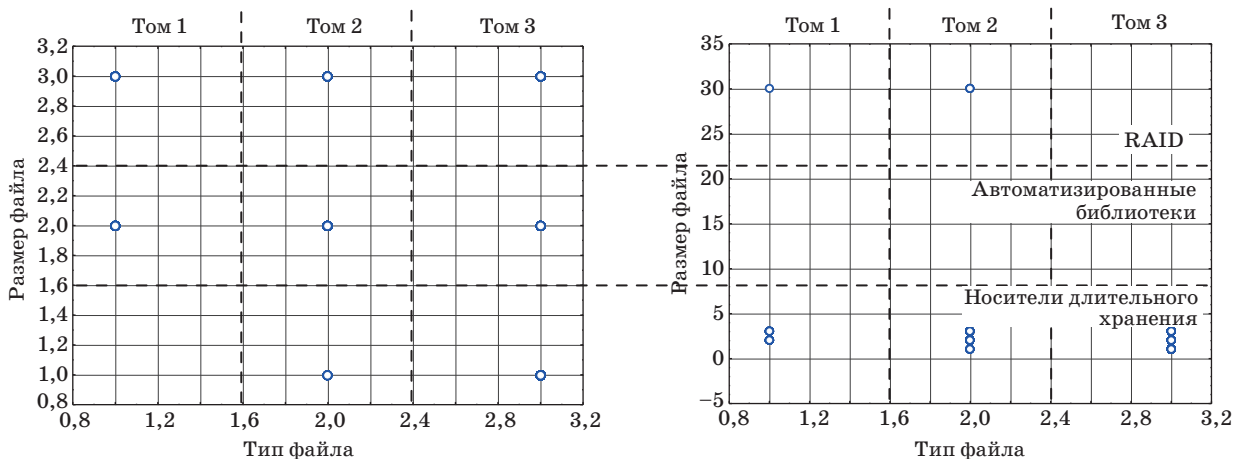
Две карты Кохонена при разных выборках экспериментальных файлов демонстрируют распределение файлов по уровням хранилища данных в зависимости от предполагаемого времени хранения файлов (рис. 4).

Две карты Кохонена при разных выборках экспериментальных файлов демонстрируют распределение файлов по ячейкам матрицы хранения в зависимости от типа и размера файлов (рис. 5).

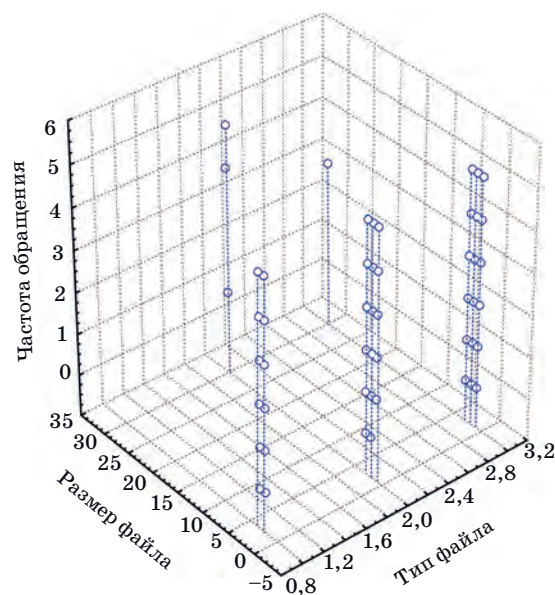
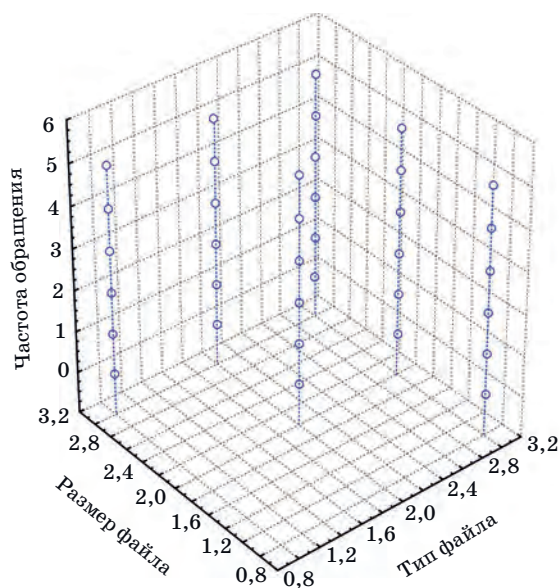
Две карты Кохонена при разных выборках экспериментальных файлов демонстрируют распределение файлов по ячейкам матрицы хранения в зависимости от типа, размера файлов и частоты обращения к файлам (рис. 6).



■ **Рис. 4.** Распределение файлов по уровням хранилища данных
 ■ **Fig. 4.** Distributing files by data storage level



■ **Рис. 5.** Распределение файлов в зависимости от типа и размера
 ■ **Fig. 5.** File distribution according to type and size



■ **Рис. 6.** Распределение файлов в зависимости от типа, размера и частоты обращения

■ **Fig. 6.** File distribution, depending on the type, size and frequency of treatment

Как показывают результаты эксперимента, аппарат нейронной сети Кохонена может быть инструментом решения задачи размещения объектов в соответствии с требуемыми параметрами. Основной сложностью при этом является выбор параметров классификации и нормализация параметров.

Заключение

В статье предложен алгоритм многоуровневого хранения данных, позволяющий распределять файлы в СХД в соответствии с последовательным применением механизмов:

— вертикального размещения по уровням системы хранения в зависимости от времени хранения файла;

— горизонтального размещения по логическим томам уровня хранения в зависимости от размера файла и длины логического блока данных;

— динамического размещения — миграции файлов по уровням СХД в зависимости от частоты обращения к ним.

Вертикальное и горизонтальное размещение основано на анализе организационных метаданных, приписанных файлам.

С учетом механизмов вертикального и горизонтального размещения файлов в системе хранения предложена модель структуры хранения в виде матрицы, размер которой соответствует количеству вертикальных и горизонтальных уровней СХД.

Предложено проводить распределение файлов по ячейкам матрицы с использованием аппарата нейронных сетей Кохонена, что позволяет отказаться от последовательного выполнения механизмов размещения.

Результаты эксперимента показали способность аппарата нейронной сети Кохонена как инструмента решения задачи размещения файлов в соответствии с требуемыми параметрами.

Литература

1. Проскуряков Н. Е., Ануфриева А. Ю. Анализ и перспективы современных систем хранения цифровых данных. *Изв. Тульского государственного университета. Технические науки*, 2013, вып. 3, с. 368–377.
2. Бурмистров В. Д., Заковряшин Е. М. Создание хранилища данных для распределенной системы. *Молодой ученый*, 2016, № 12, с. 143–147.
3. Farley M. *Building Storage Networks*. Osborne, McGraw-Hall, 2001. 576 p.

4. *Information Storage and Management*. 2nd Edition. New Jersey, John Wiley & Sons, 2016. 544 p.
5. Татарникова Т. М., Пойманова Е. Д. Технологии долговременного хранения данных. *Материалы Междунар. науч.-практ. конф. «Наука и образование в XXI веке»*, Тамбов, Бизнес-Наука-Общество, 2013, ч. 31, с. 136–137.
6. Landauer R. Information is physical. *Physics Today*, 1991, vol. 44, no. 5, pp. 23–29.
7. Kish L. B., Granqvist C. G. Does information have mass? *Proc. of the IEEE*, 2013, vol. 101, no. 9, pp. 1895–1899.

8. Buyya R., Broberg J., Goscinski A. *Cloud Computing. Principles and Paradigms*. New Jersey, John Wiley & Sons, 2011. 637 p.
9. Mesnier M., Ganger G., Riedel E. Object-based storage. *IEEE Communications Magazine*, 2003, vol. 41, no. 8, pp. 84–90.
10. Recommendation Y.3501: *Cloud computing framework and high-level requirements*. Geneva, ITU-T, 2013. <https://www.itu.int/ITU-T/recommendations/rec.aspx?rec=11917> (дата обращения: 29.01.2019).
11. Recommendation Y.3510: *Cloud computing infrastructure requirements*. Geneva, ITU-T, 2013. <https://www.itu.int/rec/T-REC-Y.3510-201305-S> (дата обращения: 29.01.2019).
12. Recommendation Y.3520: *Cloud computing framework for end-to-end resource management*. Geneva, ITU-T, 2015. <https://www.itu.int/rec/T-REC-Y.3520-201509-I> (дата обращения: 29.01.2019).
13. Hastic T., Tibshirani R., Friedman J. *The Elements of Statistical Learning*. Springer, 2003. 552 p.
14. Todman C. *Designing a Data Warehouse: Supporting Customer Relationship Management*. Prentice Hall, 2001. 414 p.
15. Stacey M., Salvatore J., Jorgensen A. *Visual Intelligence: Microsoft Tools and Techniques for Visualizing Data*. New Jersey, John Wiley & Sons, 2013. 432 p.
16. Kish L. B. Moore's law and the energy requirement of computing versus performance. *IEEE Proceedings: Circuits, Devices and Systems*, 2004, vol. 151, no. 2, pp. 190–194.
17. Kohonen T. *Self-Organizing Maps*. New York, 2001. 501 p.
18. Kutuzov O. I., Tatarnikova T. M. Model of a self-similar traffic generator and evaluation of buffer storage for classical and fractal queuing system. *Moscow Workshop on Electronic and Networking Technologies, MWENT*, 2018, pp. 1–3.
19. Morville P., Callender J. *Search Patterns: Design for Discovery*. O'Reilly, 2010. 192 p.

UDC 004.07

doi:10.31799/1684-8853-2019-2-68-75

Organization of multi-level data storageSovetov B. Ya.^a, Dr. Sc., Tech., Professor, orcid.org/0000-0003-3116-8810Tatarnikova T. M.^b, Dr. Sc., Tech., Associate Professor, orcid.org/0000-0002-6419-0072, tm-tatarn@yandex.ruPoymanova E. D.^b, Senior Lecturer, orcid.org/0000-0002-7903-2480^aSaint-Petersburg Electrotechnical University «LETI», 5, Prof. Popov St., 197376, Saint-Petersburg, Russian Federation^bSaint-Petersburg State University of Aerospace Instrumentation, 67, B. Morskaya St., 190000, Saint-Petersburg, Russian Federation

Introduction: In the face of a constantly increasing data volume, almost all organizations need to form a well-developed storage infrastructure. The infrastructure is implemented using multi-level storage technologies: designing systems of storage devices heterogeneous in physical nature and access architecture. On the other hand, the files are also subject to the requirements about the guaranteed storage time, even at the legislative level. The lack of efficient storage organization algorithms with different storage requirements actualizes new storage resource management models. **Purpose:** Developing new models and algorithms for managing data storage resources, taking into account the layered principle of data storage. **Results:** A multi-level data storage algorithm is proposed which allows you to distribute files in a data storage system in accordance with the sequential use of vertical, horizontal and dynamic file allocation. The algorithms for data placement and migration across the storage hierarchies differ in the use of metadata analysis, providing a way for proactive management of the storage resources based on guaranteed retention periods. A data storage model is proposed in the form of a matrix whose size corresponds to the number of vertical and horizontal levels in the system. It is suggested to distribute files to the cells of the matrix using the Kohonen neural network tools. The effect of using the Kohonen network is the ability to analyze metadata and the frequency of accessing files in one step, avoiding sequential implementation of the allocation procedures. The results of the experiment have demonstrated the effectiveness of the Kohonen network tools in solving the problem of placing files in a multi-level data storage system. **Practical relevance:** The proposed model and algorithms can be used in the design of data storage systems, taking into account the multi-level principle of data storage.

Keywords — data storage system, multi-level data storage, file storage time, metadata, data migration, Kohonen neural network.

For citation: Sovetov B. Ya., Tatarnikova T. M., Poymanova E. D. Organization of multi-level data storage. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2019, no. 2, pp. 68–75 (In Russian). doi:10.31799/1684-8853-2019-2-68-75

References

1. Proskuryakov N. E., Anufrieva A. Yu. Analysis and prospects of modern digital data storage systems. *Izvestiya Tul'skogo gosudarstvennogo universiteta. Tekhnicheskie nauki* [Proc. of the Tula State University. Technical Sciences], 2013, no. 3, pp. 368–377 (In Russian).
2. Burmistrov V. D., Zakovryashin E. M. Creating a data warehouse for a distributed system. *Molodoy uchenyy* [Young Scientist], 2016, no. 12, pp. 143–147 (In Russian).
3. Farley M. *Building Storage Networks*. Osborne, McGraw-Hall, 2001. 576 p.
4. *Information Storage and Management*. 2nd Edition. New Jersey, John Wiley & Sons, 2016. 544 p.
5. Tatarnikova T. M., Poymanova E. D. Long term storage technologies. *Materialy Mezhdunarodnoi nauchno-prakticheskoi konferentsii "Nauka i obrazovanie v XXI veke"* [Proc. of the Intern. Scientific and Practical Conf. "Science

- and Education in XXI Century”], Tambov, Biznes-Nauka-Obshchestvo Publ., 2013, vol. 31, pp. 136–137 (In Russian).
6. Landauer R. Information is physical. *Physics Today*, 1991, vol. 44, no. 5, pp. 23–29.
 7. Kish L. B., Granqvist C. G. Does information have mass? *Proc. of the IEEE*, 2013, vol. 101, no. 9, pp. 1895–1899.
 8. Buyya R., Broberg J., Goscinski A. *Cloud Computing. Principles and Paradigms*. New Jersey, John Wiley & Sons, 2011. 637 p.
 9. Mesnier M., Ganger G., Riedel E. Object-based storage. *IEEE Communications Magazine*, 2003, vol. 41, no. 8, pp. 84–90.
 10. *Recommendation Y.3501: Cloud computing framework and high-level requirements*. Geneva, ITU-T, 2013. Available at: <https://www.itu.int/ITU-T/recommendations/rec.aspx?rec=11917> (accessed 29 January 2019).
 11. *Recommendation Y.3510: Cloud computing infrastructure requirements*. Geneva, ITU-T, 2013. Available at: <https://www.itu.int/rec/T-REC-Y.3510-201305-S> (accessed 29 January 2019).
 12. *Recommendation Y.3520: Cloud computing framework for end-to-end resource management*. Geneva, ITU-T, 2015. Available at: <https://www.itu.int/rec/T-REC-Y.3520-201509-I> (accessed 29 January 2019).
 13. Hastie T., Tibshirani R., Friedman J. *The Elements of Statistical Learning*. Springer, 2003. 552 p.
 14. Todman C. *Designing a Data Warehouse: Supporting Customer Relationship Management*. Prentice Hall, 2001. 414 p.
 15. Stacey M., Salvatore J., Jorgensen A. *Visual Intelligence: Microsoft Tools and Techniques for Visualizing Data*. New Jersey, John Wiley & Sons, 2013. 432 p.
 16. Kish L. B. Moore’s law and the energy requirement of computing versus performance. *IEEE Proceedings: Circuits, Devices and Systems*, 2004, vol. 151, no. 2, pp. 190–194.
 17. Kohonen T. *Self-Organizing Maps*. New York, 2001. 501 p.
 18. Kutuzov O. I., Tatarnikova T. M. Model of a self-similar traffic generator and evaluation of buffer storage for classical and fractal queuing system. *Moscow Workshop on Electronic and Networking Technologies, MWENT*, 2018, pp. 1–3.
 19. Morville P., Callender J. *Search Patterns: Design for Discovery*. O’Reilly, 2010. 192 p.

УВАЖАЕМЫЕ АВТОРЫ!

Научная электронная библиотека (НЭБ) продолжает работу по реализации проекта SCIENCE INDEX. После того как Вы регистрируетесь на сайте НЭБ (<http://elibrary.ru/defaultx.asp>), будет создана Ваша личная страничка, содержание которой составят не только Ваши персональные данные, но и перечень всех Ваших печатных трудов, имеющих в базе данных НЭБ, включая диссертации, патенты и тезисы к конференциям, а также сравнительные индексы цитирования: РИНЦ (Российский индекс научного цитирования), h (индекс Хирша) от Web of Science и h от Scopus. После создания базового варианта Вашей персональной страницы Вы получите код доступа, который позволит Вам редактировать информацию, помогая создавать максимально объективную картину Вашей научной активности и цитирования Ваших трудов.

УДК 004.93

doi:10.31799/1684-8853-2019-2-76-82

Использование семейств ортогональных функций для кодирования данных в схеме хаотической маскировки

С. В. Белим^а, доктор физ.-мат. наук, профессор, orcid.org/0000-0002-5785-5160, sbelim@mail.ruЮ. С. Ракицкий^а, канд. техн. наук, доцент, orcid.org/0000-0001-9419-5424^аОмский государственный университет им. Ф. М. Достоевского, Мира пр., 55а, Омск, 644077, РФ

Введение: хаотическая маскировка сообщений применяется для скрытой передачи данных по каналам связи. Существующие схемы хаотической маскировки требуют сложной системы согласования хаотических генераторов и не устойчивы к передаче сигнала по зашумленному каналу связи. Одним из возможных путей решения данной проблемы является использование алгоритмов кодирования сообщений, позволяющих извлекать их из шумового контейнера без дополнительной информации о генераторе шума. **Цель:** разработка схемы хаотической маскировки, не требующей согласования генераторов шума и устойчивой к передаче по зашумленным каналам. **Методы:** представление передаваемого сигнала в виде суперпозиции ортогональных функций с весовыми коэффициентами, определяемыми битами передаваемого сообщения. Извлечение сообщения из хаотического сигнала путем вычисления проекции полученного сигнала на семейство ортогональных функций. **Результаты:** на базе метода кодирования сообщения с помощью семейства ортогональных функций разработана схема хаотической маскировки для скрытой передачи данных. Отличительной особенностью предложенной схемы является отсутствие необходимости синхронизации генераторов шума передающей и принимающей сторонами. Скрытие сообщения происходит простым подмешиванием полезного сигнала к передаваемому хаотическому. Извлечение полезного сигнала осуществляется благодаря свойству ортогональности семейства функций, участвующих в кодировании сообщения. Предполагается, что проекция шума на ортогональные функции имеет малое значение. Для вычисления проекции сигнала на ортогональные функции выполнено численное интегрирование. Проведен компьютерный эксперимент для двух семейств простых тригонометрических функций. Определено, что при отношении уровня шума к полезному сигналу 38 дБ вероятность передачи одного байта сообщения без потерь превышает 0,95, что удовлетворяет требованиям к современным системам хаотической маскировки. Доказано, что результаты извлечения сообщения существенно зависят от точности численного интегрирования. Для повышения допустимого уровня шумового сигнала необходимо применять более точные методы интегрирования. **Практическая значимость:** на основе результатов исследования могут быть спроектированы и реализованы системы со скрытой передачей сообщений. Предложенный подход позволяет повысить устойчивость схемы хаотической маскировки к шумам в каналах связи.

Ключевые слова — хаотическая маскировка, ортогональные функции, скрытая передача сообщений, модуляция дифференциального хаоса.

Для цитирования: Белим С. В., Ракицкий Ю. С. Использование семейств ортогональных функций для кодирования данных в схеме хаотической маскировки. *Информационно-управляющие системы*, 2019, № 2, с. 76–82. doi:10.31799/1684-8853-2019-2-76-82

For citation: Belim S. V., Rakitskiy Yu. S. Using families of orthogonal functions for coding messages in a chaotic masking scheme. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2019, no. 2, pp. 76–82 (In Russian). doi:10.31799/1684-8853-2019-2-76-82

Введение

В последнее время получили развитие схемы использования динамического хаоса для сокрытия самого факта передачи содержательного сообщения. Полезный сигнал каким-либо образом подмешивается к хаотическому сигналу, для выработки которого применяются хаотические генераторы. Причем интенсивность хаотического сигнала значительно превышает интенсивность полезного сигнала. Основная проблема состоит в извлечении принимающей стороной полезной составляющей принятого сигнала. Для этого по каналу связи либо какими-то другими способами передается дополнительная информация о хаотическом генераторе. На основе дополнительной информации принимающая сторона самостоя-

тельно генерирует хаотическую составляющую и вычитает ее из полученного сигнала. Данную ситуацию принято характеризовать как формирование двух связанных идентичных хаотических генераторов. На сегодня получили развитие несколько различных подходов: хаотическая маскировка [1], переключение хаотических режимов [2], нелинейное подмешивание передаваемого сообщения к хаотическому сигналу [3], модулирование управляющих параметров хаотического генератора [4].

Исторически одним из первых и наиболее простых в реализации является метод хаотической маскировки сообщения [1]. Передаваемое сообщение $m(t)$ складывается с хаотическим сигналом $x(t)$. Полученный смешанный сигнал $m'(t) = m(t) + x(t)$ передается принимающей стороне.

Основная задача, которая стоит перед принимающей стороной, заключается в синхронизации своего хаотического генератора $u(t)$ с хаотическим генератором передающей стороны: $u(t) = x(t)$. После этого передаваемое сообщение может быть получено путем простого вычитания $m(t) = m'(t) - u(t)$. В разработанных в настоящее время схемах хаотической маскировки уровень шума по сравнению с уровнем полезного сигнала составляет 35–65 дБ [5]. Схема хаотической маскировки эффективна, если уровень шума в канале связи достаточно низок, в противном случае резко снижается качество передачи информации. Также к значительному снижению качества полученного сигнала приводит рассинхронизация управляющих параметров генераторов шума [6–8].

Одним из возможных подходов к вычитанию шума из передаваемого сигнала является ортогонализация хаотических сигналов. Такие схемы нашли развитие в рамках общего подхода к модуляции сигналов, получившего название Differential Chaos Shift Keying (DCSK) [9]. Одной из проблем подобных систем является построение ортогональных хаотических сигналов. Для ее решения используются обычные генераторы шума с последующим применением к ним преобразования Гильберта [10–12], преобразования Грамма — Шмидта [13–15], кодов Уолша [16–18]. Также следует отметить, что для схем на основе ортогональных хаотических сигналов характерна низкая пропускная способность. Каждый ортогональный хаотический фрейм несет один бит информации.

В данной статье предложена схема хаотической маскировки сигнала, не требующая синхронизации генераторов шума. Основная идея состоит в кодировании сообщения на основе семейства ортогональных функций, которое позволяет в дальнейшем извлекать его без синхронизации хаотических генераторов передающей и принимающей сторонами. Данное кодирование было ранее применено для шифрования сообщений [19] и формирования устойчивых цифровых водяных знаков [20].

Кодирование сообщения

Пусть сообщение представлено в виде конечной последовательности битов $C = c_1 c_2 \dots c_N$.

Разобьем сообщение на подпоследовательности C_j длиной n ($j = 1, \dots, M, M = \lfloor N/n \rfloor$) и представим каждую такую подпоследовательность в виде вектора

$$C_j = (c_{jn}, \dots, c_{(j+1)n-1}).$$

Выберем систему ортогональных на отрезке $[0, b]$ функций $f_i(t), i = 1, \dots, n$, где t — временная

ось. Пусть для данного семейства функций условие ортогональности имеет вид

$$\int_0^b w(t) f_i(t) f_j(t) dt = \delta_{ij},$$

$$\text{где } \delta_{ij} = \begin{cases} 0, & i \neq j; \\ 1, & i = j. \end{cases}$$

Построим вектор-функцию в n -мерном пространстве, координатами которой служат ортогональные функции из выбранного семейства:

$$f(t) = (f_1(t), f_2(t), \dots, f_n(t)).$$

Сопоставим каждой подпоследовательности исходного сообщения функцию $F_j(x)$, вычисляемую как скалярное произведение:

$$F_j(t) = c_j \cdot f(t).$$

Тогда все сообщение представимо в виде вектор-функции

$$F(t) = (F_1(t), F_2(t), \dots, F_M(t)).$$

Если исходное сообщение представить в виде матрицы, строками которой служат векторы подпоследовательностей:

$$MC = \begin{pmatrix} C_1 \\ C_2 \\ \vdots \\ C_M \end{pmatrix},$$

то вектор $F(t)$ может быть получен как

$$F(t) = MC \cdot f(t).$$

Сформируем функцию $F_c(t)$ на отрезке $[0, Mb]$, соответствующую сообщению, для передачи по каналу связи:

$$F_c(t) = \begin{cases} F_1(t), & t \in [0, b]; \\ F_2(t-b), & t \in [b, 2b]; \\ \vdots & \vdots \\ F_M(t-(M-1)b), & t \in [(M-1)b, Mb]. \end{cases}$$

Функция $F_c(x)$ несет в себе полную информацию об исходном сообщении.

После получения функции $F_c(t)$, определенной на отрезке $[0, Mb]$, для извлечения сообщения необходимо выполнить следующие действия.

1. Разделить отрезок, на котором определена функция $F_c(t)$, на M частей. Данный отрезок определяется длительностью передачи сообще-

ния. Число M является параметром системы обмена сообщениями и известно обеим сторонам.

2. Определить координаты вектор-функции

$$\mathbf{F}(t) = (F_1(t), F_2(t), \dots, F_M(t)).$$

Все функции $F_j(t)$, $j = 1, \dots, M$ определяются на отрезке $[0, b]$ с помощью соотношения

$$F_j(t) = F_c(t + (j - 1)b).$$

3. Для каждой функции $F_j(t)$ вычислить набор величин

$$a_{ij} = \int_0^b w(t) f_i(t) F_j(t) dt, \quad i = 1, \dots, n.$$

4. Сформировать последовательность

$$C_j = c_1^{(j)} c_2^{(j)} \dots c_n^{(j)}, \quad c_i^{(j)} = \begin{cases} 1, & a_{ij} \geq 0,5; \\ 0, & a_{ij} < 0,5. \end{cases}$$

5. Объединить подпоследовательности C_i в одну последовательность с помощью конкатенации:

$$C = C_1 C_2 \dots C_M.$$

Полученная последовательность C будет совпадать с исходной, переданной по каналу связи. В идеальной системе величины a_{ij} должны принимать только единичные и нулевые значения. Однако даже в отсутствие помех в канале связи на третьем шаге необходимо вычислять интегралы, значение которых в общем случае может быть определено только численно. Для вычисления величин a_{ij} необходимо выполнить дискретизацию функции $F_c(x)$ с некоторым шагом x :

$$F_{ci} = F_c(ix), \quad i = 0, \dots, Mb/x - 1.$$

Дискретизация функции может использоваться не только на этапе вычисления интегралов, но и при отправке сообщения. Тогда правильный выбор параметра b влияет на работоспособность схемы в целом.

Схема скрытой передачи сообщения

Общая схема системы скрытой передачи сообщений на основе хаотической маскировки представлена на рис. 1.

Для скрытой передачи сообщения используем схему хаотической маскировки, сложив передаваемое сообщение $F_c(t)$ с хаотическим шумом. Пусть генератор шума вырабатывает сигнал $H(t)$. Тогда по каналу связи будет передаваться сигнал

$$S(t) = F_c(t) + H(t),$$

причем интенсивность хаотического сигнала I_H превышает интенсивность полезного сигнала I_F . Соотношение сигнал/шум принято измерять в децибелах, т. е. вычислять величину $k = 10 \log_{10}(I_H/I_F)$.

Для извлечения передаваемого сообщения из сигнала $S(t)$ необходимо выполнить те же преобразования, что и для извлечения сообщения из функции $F_c(t)$:

1) разделить отрезок, на котором определена функция $S(t)$, на M частей;

2) определить координаты вектор-функции

$$\mathbf{S}(t) = (S_1(t), S_2(t), \dots, S_M(t));$$

все функции $S_j(t)$, $j = 1, \dots, M$ определяются на отрезке $[0, b]$ с помощью соотношения

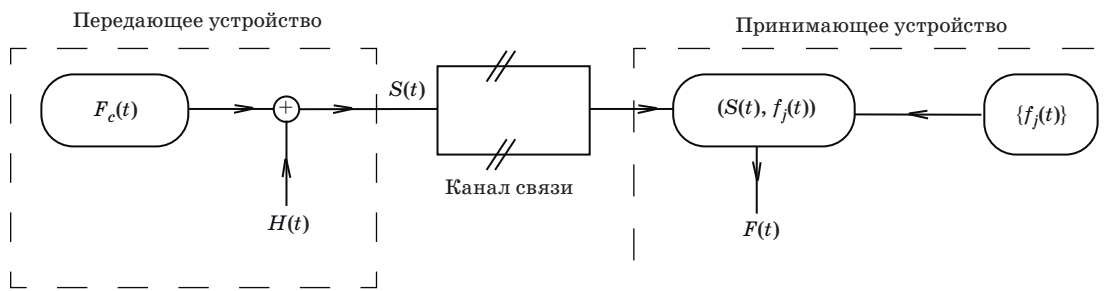
$$S_j(t) = S(t + (j - 1)b);$$

3) для каждой функции $S_j(x)$ вычислить набор величин

$$d_{ij} = \int_0^b w(t) f_i(t) S_j(t) dt, \quad i = 1, \dots, n;$$

4) сформировать последовательность

$$E_j = e_1^{(j)} e_2^{(j)} \dots e_n^{(j)}, \quad e_i^{(j)} = \begin{cases} 1, & d_{ij} \geq 0,5; \\ 0, & d_{ij} < 0,5; \end{cases}$$



■ Рис. 1. Структурная схема системы скрытой передачи сообщений

■ Fig. 1. The structure scheme of a system for the hidden messages transmission

5) объединить подпоследовательности E_i в одну последовательность с помощью конкатенации:

$$E = E_1 E_2 \dots E_M.$$

В силу аддитивного подхода к формированию сигнала $S(x)$ над хаотическим сигналом $H(x)$ выполняются те же самые преобразования. Хаотический сигнал $H(t)$ будет представлен в виде вектор-функции

$$\mathbf{H}(t) = (H_1(t), H_2(t), \dots, H_n(t)).$$

Каждая функция $H_j(t)$, $j = 1, \dots, n$ определена на отрезке $[0, b]$. Для величин d_{ij} можем записать

$$d_{ij} = \int_0^b w(t) f_i(t) (F_j(t) + H_j(t)) dt = a_{ij} + h_{ij},$$

где $h_{ij} = \int_0^b w(t) f_i(t) H_j(t) dt$.

Для правильного декодирования сообщения, в идеальном случае, необходимо выполнение требования $d_{ij} = a_{ij}$, т. е. $h_{ij} = 0$. Однако в силу того, что при вычислении $e_i^{(j)}$ используется пороговая схема, достаточно выполнения условия $|h_{ij}| < 0,5$. Данное ограничение накладывает условия как на выбор семейства ортогональных функций, так и на соотношение интенсивностей сигнала и шума. В идеальном случае при равномерном распределении случайных величин $H_j(t)$ и интегрировании на бесконечном интервале ($b \rightarrow \infty$) все $h_{ij} = 0$. В реальных системах приходится работать с конечными интервалами ($b < \infty$). В результате чего статистические характеристики случайных функций $H_j(t)$ отличаются от идеальных, и интегралы принимают ненулевые значения ($|h_{ij}| > 0$).

Сформулируем основные требования к семействам ортогональных функций, которые могут быть использованы в предложенной схеме.

1. Интервал ортогональности должен быть конечным. В противном случае невозможно сформировать кадры закодированного сообщения.

2. Ни одна из функций, в том числе весовая, не должна иметь особенностей на интервале ортогональности и ограниченную область допустимых значений. Это требование необходимо для эффективной реализации численного интегрирования.

3. Вычисление значений функции должно иметь низкую трудоемкость.

4. Определенный интеграл от функций в интервале ортогональности с учетом весовой функции должен быть равен нулю. Это требование позволит минимизировать значения величин h_{ij} .

Следующие факторы могут приводить к ненулевому значению величин h_{ij} :

1) выбор семейства ортогональных функций $f_i(t)$, $i = 1, \dots, n$;

2) ограниченная длина отрезка b ;

3) погрешности численного интегрирования, связанные с выбором численного метода и шага дискретизации.

Компьютерный эксперимент

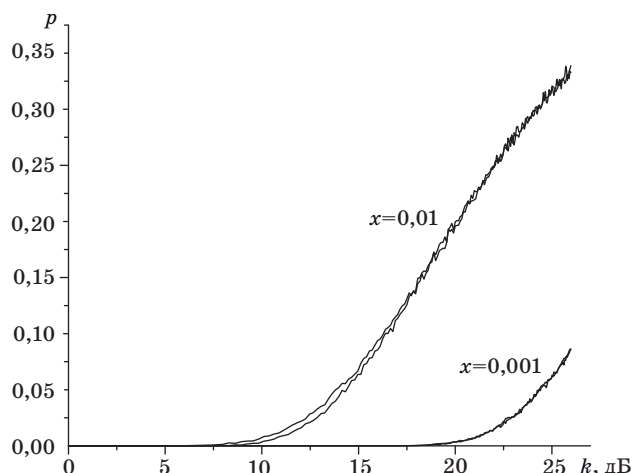
В рамках компьютерного эксперимента протестируем предложенную схему хаотической маскировки. Для тестирования было использовано два семейства простых тригонометрических функций, ортогональных на отрезке $[0, 1]$:

$$f_j(t) = \sqrt{2} \cos(\pi j t), \quad j = 1, \dots, 8;$$

$$g_j(t) = \sqrt{2} \sin(\pi j t), \quad j = 1, \dots, 8.$$

Весовая функция для этих семейств ортогональных функций $w(t) = 1$. Интегралы вычислялись методом трапеций. Для выявления влияния точности численного интегрирования на результаты извлечения полезного сигнала рассматривалось два случая с шагом дискретизации $x = 0,01$ и $x = 0,001$. Интенсивность полезного сигнала $I_F = 1$. Интенсивность хаотического сигнала I_H менялась от 0 до 20 с шагом 0,1. В качестве генератора шума использовался линейный конгруэнтный генератор с равномерным распределением. В качестве скрытого сообщения использована битовая последовательность длиной 2048 символов, которая формировалась из последовательно записанных чисел от 0 до 255 в однобайтовом представлении. Для каждого семейства ортогональных функций производилось кодирование сообщения, зашумление и извлечение 100 раз. В каждом эксперименте генератор псевдослучайной последовательности инициализировался различными зернами, что обеспечивало различные сигналы $H(t)$. В каждом эксперименте определялось количество неверно извлеченных битов. После этого значения, полученные в различных экспериментах, усреднялись, и вычислялась вероятность потери отдельного бита при передаче сообщения по предложенной схеме. На рис. 2 приведены графики зависимости вероятности неверного извлечения бита от соотношения сигнал/шум k для семейств ортогональных функций $f_j(t)$ и $g_j(t)$ при различных шагах численного интегрирования.

Из рисунка можно сделать несколько выводов. Во-первых, оба семейства ортогональных функций $f_j(t)$ и $g_j(t)$ обеспечивают мало отличимые друг от друга результаты. Во-вторых, точность вычисления интегралов, выражающаяся в выборе шага дискретизации, существенно влияет



■ **Рис. 2.** Зависимость вероятности p неверного определения одного бита передаваемого сообщения от соотношения сигнал/шум

■ **Fig. 2.** Dependence of insecure determination probability for one bit of the transferred message p on a ratio signal/noise

на результаты извлечения сообщения. При шаге дискретизации $x = 0,001$ для уровня шума 25 дБ вероятность неверного определения значения одного бита составляет $p = 0,08$. Вероятность передачи одного байта сообщения без потерь может быть определена по формуле $P = 1 - (1 - p)^8$.

Отсюда следует, что вероятность передачи одного байта сообщения без потерь при $k = 28$ дБ для $x = 0,01$ составляет $P = 0,96$, а для $x = 0,001$ получаем значение $P = 0,49$. Введем верхнюю границу пороговой вероятности передачи одного байта сообщения без потерь P_0 . Согласно работе [5], рекомендуемым значением является $P_0 = 0,95$. В этом случае совместное использование методов хаотической маскировки и алгоритмов помехоустойчивого кодирования позволяет передавать сообщение без потерь [5]. Значение верхней границы вероятности передачи одного байта сообщения без потерь достигается при вероятности неверного определения одного бита $p_0 = 0,31$. При шаге дискретизации $x = 0,01$ верхняя граница передачи одного байта сообщения без потерь достигается при отношении уровня шума к уровню сигнала $k = 25,2$ дБ, что заметно ниже уровня, приемлемого для современных систем хаотической маскировки. Для шага дискретизации $x = 0,001$ аналогичный уровень шума составляет $k = 38,3$ дБ, что сопоставимо с результатами, использующими схему синхронизации генераторов шума [5]. Из этого можно сделать вывод, что шаг дискретизации $x = 0,001$ позволяет построить схему скрытой передачи сообщений, обеспечивающую устойчивость передачи сообщений при более высоких значениях интенсивности шума. Устойчивость схемы может быть повыше-

на с помощью уменьшения шага дискретизации и использования более точных методов интегрирования.

Заключение

Предложенный в данной статье метод кодирования сообщений позволяет реализовать схему хаотической маскировки сигнала, не требующую согласования генераторов хаоса отправляющей и принимающей сторон. В результате существенно повышается устойчивость схемы к случайным шумам, присутствующим в каналах передачи информации. Основной проблемой построения предложенной схемы является обеспечение высокой точности численного интегрирования, влияющей на правильность извлечения передаваемого сообщения. Величина ошибки при вычислении интегралов может варьироваться с помощью шага дискретизации функции, кодирующей сообщение. Уменьшение шага дискретизации приводит к снижению количества ошибочно извлекаемых битов. Как следствие, появляется возможность увеличить интенсивность маскирующего хаотического сигнала. При значениях шага дискретизации 0,001 удастся построить систему, сопоставимую по уровню хаотического сигнала с актуальными на данный момент системами хаотической маскировки.

Также следует отметить более высокую пропускную способность предложенной схемы по сравнению с результатами, полученными в работах [10–18], в которых свойства ортогональности используются для формирования хаотической составляющей сигнала. Схемы, предложенные в этих статьях, позволяют в одном фрейме кодировать не более двух бит исходного сообщения. В разработанном в настоящей статье подходе в одном фрейме передается восемь бит исходного сообщения.

Конфиденциальность передачи информации в предложенной схеме может быть обеспечена с помощью введения дополнительных параметров в ортогональные функции. Например, в рассмотренной схеме на основе простых тригонометрических функций может быть выбрано семейство ортогональных функций

$$f_j(t) = \sqrt{2} \cos(\omega \pi j t), \quad j = 1, \dots, N.$$

Два параметра, ω и N , известны только передающей и принимающей сторонам. До начала сеанса принимающая и передающая стороны по защищенному каналу обмениваются информацией об используемой частоте ортогональных функций ω и их количестве N . Без знания этих параметров извлечение сообщения невозможно.

В данной статье в качестве маскировки использован хаотический сигнал. Однако может быть использован и любой белый шум, что не

скажется на работоспособности схемы в целом. Выбор той или иной маскировки может осуществляться исходя из конкретной реализации.

Литература

1. **Cuomo K. M., Oppenheim A. V., Strogatz S. H.** Synchronization of Lorenz-based chaotic circuits with applications to communications. *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, 1993, vol. 40, iss. 10, pp. 626–633. doi:10.1109/82.246163
2. **Dedieu H., Kennedy M. P., Hasler M.** Chaos shift keying: modulation and demodulation of a chaotic carrier using self-synchronizing Chua's circuits. *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, 1993, vol. 40, iss. 10, pp. 634–642. doi:10.1109/82.246164
3. **Dmitriev A. S., Panas A. I., Starkov S. O.** Experiments on speech and music signals transmission using chaos. *International Journal of Bifurcation and Chaos*, 1995, vol. 5, no. 4, pp. 1249–1254. doi:10.1142/S0218127495000910
4. **Yang T., Chua L. O.** Secure communication via chaotic parameter modulation. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, 1996, vol. 43, iss. 9, pp. 817–819. doi:10.1109/81.536758
5. **Downes P. T.** Secure communication using chaotic synchronization. *SPIE*, 1993, vol. 2038, pp. 227–234.
6. **Perez G., Cerderia H. A.** Extracting messages masked by chaos. *Phys Rev Lett*, 1995, vol. 74, pp. 1970–1973. doi:10.1103/PhysRevLett.74.1970
7. **Short K. M.** Unmasking a modulated chaotic communication scheme. *International Journal of Bifurcation and Chaos*, 1996, vol. 6, no. 2, pp. 367–375. doi:10.1142/S0218127496000114
8. **Ponomarenko V. I., Prokhorov M. D.** Extracting information masked by the chaotic signal of a time-delay system. *Phys Rev E*, 2002, vol. 66, pp. 026215. doi: 10.1103/PhysRevE.66.026215
9. **Kolumban G., Vizvari G. K., Schwarz W., Abel A.** Differential chaos shift keying: a robust coding for chaos communication. *Proc. Intern. Workshop on Non-linear Dynamics of Electronic Systems (NDES)*, 1996, pp. 92–97.
10. **Galias Z., Maggio G. M.** Quadrature chaos-shift keying: theory and performance analysis. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, 2001, vol. 48, no. 12, pp. 1510–1519. doi:10.1109/TCSI.2001.972858
11. **Kaddoum G.** Design and performance analysis of a multiuser OFDM based differential chaos shift keying communication system. *IEEE Trans. Commun.*, 2016, vol. 64, no. 1, pp. 249–260. doi:10.1109/TCOMM.2015.2502259
12. **Yang H., Tang W. K. S., Chen G.** System design and performance analysis of orthogonal multi-level differential chaos shift keying modulation scheme. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, 2016, vol. 63, no. 1, pp. 146–156. doi:10.1109/TCSI.2015.2510622
13. **Wren T. J., Yang T. C.** Orthogonal chaotic vector shift keying in digital communications. *IET Communications*, 2010, vol. 4, no. 6, pp. 739–753. doi:10.1049/iet-com.2009.0122
14. **Wang L., Cai G., Chen G.** Design and performance analysis of a new multiresolution M-ary differential chaos shift keying communication system. *IEEE Transaction on Wireless Communications*, 2015, vol. 14, no. 9, pp. 5197–5208. doi:10.1109/TWC.2015.2434820
15. **Kaddoum G., Soujeri E., Arcila C., Eshteiwi K.** I-DCSK: An improved noncoherent communication system architecture. *IEEE Transactions on Circuits and Systems II: Exp. Briefs*, 2015, vol. 62, no. 9, pp. 901–905. doi:10.1109/TCSII.2015.2435831
16. **Xu W. K., Wang L., Kolumban G.** A novel differential chaos shift keying modulation scheme. *International Journal of Bifurcation and Chaos*, 2011, vol. 21, no. 3, pp. 799–814. doi:10.1142/S0218127411028829
17. **Huang T., Wang L., Xu W., Lau F. C.** A multilevel code-shifted differential chaos-shift-keying system. *IET Communications*, 2016, vol. 10, no. 10, pp. 1189–1195. doi:10.1109/TCSII.2017.2764916
18. **Escribano F. J., Kaddoum G., Wagemakers A., Giard P.** Design of a new differential chaos-shift-keying system for continuous mobility. *IEEE Trans. Commun.*, 2016, vol. 64, no. 5, pp. 2066–2078. doi:10.1109/TCOMM.2016.2538236
19. **Белим С. В., Белим С. Ю.** Шифрование сообщений на основе собственных функций операторов. *Математические структуры и моделирование*, 2008, вып. 18, с. 95–97.
20. **Белим С. В., Илюшечкин Е. А.** Применение семейств ортогональных функций для построения устойчивых цифровых водяных знаков. *Математические структуры и моделирование*, 2014, вып. 32, с. 225–231.

UDC 004.93

doi:10.31799/1684-8853-2019-2-76-82

Using families of orthogonal functions for coding messages in a chaotic masking scheme

S. V. Belim^a, Dr. Sc., Phys.-Math., Professor, orcid.org/0000-0002-5785-5160, sbelim@mail.ruYu. S. Rakitskiy^a, PhD, Tech., Associate Professor, orcid.org/0000-0001-9419-5424^aF. M. Dostoevskiy Omsk State University, 55a, Mira Pr., 644077, Omsk, Russian Federation

Introduction: Chaotic masking of messages is used for hidden data transfer over a communication channel. The existing chaotic masking schemes require a sophisticated system to coordinate the chaotic generators. These systems are not stable in the case of a noisy communication channel. One of the ways to solve this problem is using coding algorithms which allow you to extract messages from a noisy container without any extra information about the chaotic generator. **Purpose:** Development of a chaotic masking scheme which does not need the noise generators to be coordinated and which is immune to transfer over noisy channels. **Methods:** The signal is represented as a superposition of orthogonal functions with weighting coefficients determined by the bits in the message being transferred. A message is extracted from the chaotic signal by calculating its projection on a family of orthogonal functions. **Results:** Based on the proposed method, a chaotic masking scheme is constructed for hidden data transfer. It does not need the chaotic generators of the transmitter and receiver to be synchronized. Messages are concealed by mixing the useful and chaotic signals. The useful signal can be extracted because the functions used for the coding are orthogonal. It is assumed that noise projection on orthogonal functions is very small. To calculate this projection, digital integration was used. A computer experiment was performed for two families of simple trigonometric functions. It showed that the proposed scheme allows you to recover the signal with a high precision for noise-to-signal ratio 38 dB (the probability of a correct byte transfer is more than 0.95). The success of recovering strongly depends on the digital integration precision. To increase the allowable noise level, you need to use more precise integration methods. **Practical relevance:** The results of the study can be used for developing systems with hidden message transmission. The proposed approach allows you to improve the stability of a chaotic masking scheme to the noise in communication channels.

Keywords — chaotic masking, orthogonal functions, hidden transmission of messages, differential chaos shift keying.

For citation: Belim S. V., Rakitskiy Yu. S. Using families of orthogonal functions for coding messages in a chaotic masking scheme. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2019, no. 2, pp. 76–82 (In Russian). doi:10.31799/1684-8853-2019-2-76-82

References

1. Cuomo K. M., Oppenheim A. V., Strogatz S. H. Synchronization of Lorenz-based chaotic circuits with applications to communications. *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, 1993, vol. 40, iss. 10, pp. 626–633. doi:10.1109/82.246163
2. Dedieu H., Kennedy M. P., Hasler M. Chaos shift keying: modulation and demodulation of a chaotic carrier using self-synchronizing Chua's circuits. *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, 1993, vol. 40, iss. 10, pp. 634–642. doi:10.1109/82.246164
3. Dmitriev A. S., Panas A. I., Starkov S. O. Experiments on speech and music signals transmission using chaos. *International Journal of Bifurcation and Chaos*, 1995, vol. 5, no. 4, pp. 1249–1254. doi:10.1142/S0218127495000910
4. Yang T., Chua L. O. Secure communication via chaotic parameter modulation. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, 1996, vol. 43, iss. 9, pp. 817–819. doi:10.1109/81.536758
5. Downes P. T. Secure communication using chaotic synchronization. *SPIE*, 1993, vol. 2038, pp. 227–234.
6. Perez G., Cerderia H. A. Extracting messages masked by chaos. *Phys Rev Lett*, 1995, vol. 74, pp. 1970–1973. doi:10.1103/PhysRevLett.74.1970
7. Short K. M. Unmasking a modulated chaotic communication scheme. *International Journal of Bifurcation and Chaos*, 1996, vol. 6, no. 2, pp. 367–375. doi:10.1142/S0218127496000114
8. Ponomarenko V. I., Prokhorov M. D. Extracting information masked by the chaotic signal of a time-delay system. *Phys Rev E*, 2002, vol. 66, pp. 026215. doi: 10.1103/PhysRevE.66.026215
9. Kolumban G., Vizvari G. K., Schwarz W., Abel A. Differential chaos shift keying: a robust coding for chaos communication. *Proc. Intern. Workshop on Non-linear Dynamics of Electronic Systems (NDES)*, 1996, pp. 92–97.
10. Galias Z., Maggio G. M. Quadrature chaos-shift keying: theory and performance analysis. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, 2001, vol. 48, no. 12, pp. 1510–1519. doi:10.1109/TCSI.2001.972858
11. Kaddoum G. Design and performance analysis of a multiuser OFDM based differential chaos shift keying communication system. *IEEE Trans. Commun.*, 2016, vol. 64, no. 1, pp. 249–260. doi:10.1109/TCOMM.2015.2502259
12. Yang H., Tang W. K. S., Chen G. System design and performance analysis of orthogonal multi-level differential chaos shift keying modulation scheme. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, 2016, vol. 63, no. 1, pp. 146–156. doi:10.1109/TCSI.2015.2510622
13. Wren T. J., Yang T. C. Orthogonal chaotic vector shift keying in digital communications. *IET Communications*, 2010, vol. 4, no. 6, pp. 739–753. doi:10.1049/iet-com.2009.0122
14. Wang L., Cai G., Chen G. Design and performance analysis of a new multiresolution M-ary differential chaos shift keying communication system. *IEEE Transaction on Wireless Communications*, 2015, vol. 14, no. 9, pp. 5197–5208. doi:10.1109/TWC.2015.2434820
15. Kaddoum G., Soujeri E., Arcila C., Eshteiwi K. I-DCSK: An improved noncoherent communication system architecture. *IEEE Transactions on Circuits and Systems II: Exp. Briefs*, 2015, vol. 62, no. 9, pp. 901–905. doi:10.1109/TCSII.2015.2435831
16. Xu W. K., Wang L., Kolumban G. A novel differential chaos shift keying modulation scheme. *International Journal of Bifurcation and Chaos*, 2011, vol. 21, no. 3, pp. 799–814. doi:10.1142/S0218127411028829
17. Huang T., Wang L., Xu W., Lau F. C. A multilevel code-shifted differential chaos-shift-keying system. *IET Communications*, 2016, vol. 10, no. 10, pp. 1189–1195. doi:10.1109/TC-SII.2017.2764916
18. Escribano F. J., Kaddoum G., Wagemakers A., Giard P. Design of a new differential chaos-shift-keying system for continuous mobility. *IEEE Trans. Commun.*, 2016, vol. 64, no. 5, pp. 2066–2078. doi:10.1109/TCOMM.2016.2538236
19. Belim S. V., Belim S. Yu. Encrypting message on based eigenfunction of operators. *Matematicheskie struktury i modelirovanie* [Mathematical Structures and Modeling], 2008, iss. 18, pp. 95–97 (In Russian).
20. Belim S. V., Ilushechkin E. A. Applying families of orthogonal function to building resistant digital watermarks. *Matematicheskie struktury i modelirovanie* [Mathematical Structures and Modeling], 2014, iss. 32, pp. 225–231 (In Russian).

Многопользовательские распределенные измерительно-управляющие системы как элемент современной цифровой образовательной среды

В. А. Комаров^а, канд. техн. наук, доцент, orcid.org/0000-0001-9210-9908

А. В. Сарафанов^б, доктор техн. наук, профессор, orcid.org/0000-0003-4264-9388

С. Р. Тумковский^в, доктор техн. наук, профессор, orcid.org/0000-0002-1647-2171, STumkovskiy@hse.ru

^аАО «Информационные спутниковые системы им. акад. М. Ф. Решетнёва», Ленина ул., 52, Железнодорожск, 662972, РФ

^бООО Витте Консалтинг (ГК «АЙ-ТЕКО»), Кедрова ул., 15, Москва, 117036, РФ

^вНациональный исследовательский университет «Высшая школа экономики», Мясницкая ул., 20, Москва, 101000, РФ

Постановка проблемы: одним из ключевых сдерживающих факторов качественной подготовки инженерных кадров для цифровой экономики является отсутствие в региональных университетах России доступа к современному лабораторному оборудованию. Среди эффективных подходов к решению указанной проблемы находится организация удаленного доступа к экспериментальному оборудованию на основе концепции мультиарендности. **Цель:** разработка способа организации высокотехнологичной цифровой образовательной среды, обеспечивающей выполнение лабораторных исследований в удаленном многопользовательском режиме при подготовке инженерных кадров с применением технологии контекстного обучения. **Методы:** создание унифицированной структуры сегмента цифровой образовательной среды, основой которой являются многопользовательские распределенные измерительно-управляющие системы с унифицированной структурой и специальной методикой адаптации таких систем под различные условия их эксплуатации. **Результаты:** разработана и апробирована унифицированная структура сегмента цифровой образовательной среды. Ее основными элементами являются web-лаборатория, обеспечивающая регламентированный доступ обучаемых к лабораторным исследованиям, и многопользовательские распределенные измерительно-управляющие системы, реализующие эти исследования посредством параллельного обращения обучаемых к разделяемым ресурсам – автоматизированным лабораторным макетам/стандам/установкам. Созданный программный комплекс для сбора и обработки статистической информации о функционировании многопользовательских распределенных измерительно-управляющих систем как систем массового обслуживания лежит в основе предложенной методики их адаптации к различным условиям функционирования. Предложенный сегмент цифровой образовательной среды позволяет реализовывать ряд направлений контекстного обучения, в том числе применение в экспериментальных исследованиях современных компьютерных измерительных технологий; исследование характеристик цифрового двойника лабораторного макета/установки с применением типовых методик; идентификацию и верификацию математических моделей как исследуемых объектов, устройств, процессов и явлений, так и образцов многопользовательских систем в виде соответствующих систем массового обслуживания; апробацию различных алгоритмов диспетчеризации разделяемого ресурса. **Практическая значимость:** предложенная структура сегмента образовательной среды обеспечивает подготовку инженеров с применением инновационных аппаратно-программных решений; эксплуатацию экспериментального оборудования на основе концепции мультиарендности, в десятки раз сокращая себестоимость оснащения одного рабочего места обучающегося; применение на системном уровне технологии контекстного обучения.

Ключевые слова – автоматизированный лабораторный практикум с удаленным доступом, автоматизированный лабораторный макет, удаленный доступ, web-лаборатория, мультиарендность, виртуальный лабораторный стенд, многопользовательский доступ, многопользовательская распределенная измерительно-управляющая система, цифровая образовательная среда.

Для цитирования: Комаров В. А., Сарафанов А. В., Тумковский С. Р. Многопользовательские распределенные измерительно-управляющие системы как элемент современной цифровой образовательной среды. *Информационно-управляющие системы*, 2019, № 2, с. 83–94. doi:10.31799/1684-8853-2019-2-83-94

For citation: Komarov V. A., Sarafanov A. V., Tumkovskiy S. R. Multi-user distributed information-control systems as an element of modern digital educational environment. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2019, no. 2, pp. 83–94 (In Russian). doi:10.31799/1684-8853-2019-2-83-94

Введение

В настоящее время сдерживающими факторами качественной подготовки высококвалифицированных специалистов, востребованных в рамках цифровой экономики, является отсутствие в организациях высшего и среднего профессионального образования специалистов необходимой квалификации для разработки и преподава-

ния современных учебных программ в области техники и технологий; методик получения обучающимися практических навыков в процессе обучения; доступа к высокотехнологичному лабораторному оборудованию; доступа к современным программному обеспечению и технологиям.

Неравномерное региональное распределение ресурсов в РФ приводит к тому, что качество подготовки специалистов в регионах существенно

отличается от качества подготовки в центральных вузах и тем более в ведущих западных образовательных центрах [1].

Выровнять качество оказываемых образовательных услуг в любой локации РФ можно за счет:

- предоставления удаленного доступа к высококачественным учебным материалам в цифровых форматах [2];

- организации удаленного доступа к лабораторному оборудованию, виртуальным лабораторным практикумам, цифровым двойникам высокотехнологичного оборудования и технологических процессов [3];

- методической поддержки образовательного процесса за счет применения создаваемых лучшими учебными заведениями учебно-методических разработок;

- привития практических навыков в рамках проектной модели обучения;

- создания базы знаний атомарных объектов учебного контента («квантов знаний», снабженных контрольно-измерительными материалами), что позволит автоматически формировать учебные курсы и образовательные программы в зависимости от требуемых компетенций;

- формирования индивидуального трека обучения и построения вектора профессионального развития на основе фиксации и анализа цифрового следа обучающегося и т. д.

Наибольшее развитие сегодня в РФ получило создание и предоставление удаленного доступа к высококачественному цифровому контенту массовых открытых онлайн-курсов, представленных на следующих интернет-ресурсах: <https://openedu.ru/>; [coursera https://www.coursera.org/](https://www.coursera.org/); <https://online.edu.ru/ru/about> и др.

Существенным недостатком данных платформ является либо отсутствие, либо предоставление в очень ограниченном масштабе удаленного доступа к современному лабораторному оборудованию, что в свою очередь резко ограничивает их использование в инженерном образовании. При этом процесс подготовки специалистов в области техники и технологий базируется на различных лабораторных исследованиях, затраты на организацию и проведение которых при традиционном подходе составляют до 60–70 % общей стоимости обучения. Организация лабораторных исследований в системе e-learning с применением современных цифровых технологий позволяет, во-первых, реализовывать полноценную подготовку специалистов в области техники и технологий; во-вторых, кратно сократить затраты на аппаратную часть лабораторных исследований; в-третьих, сделать современную цифровую образовательную среду (ЦОС) высокотехнологичной, что обеспечит ее гибкость для применения педагогических технологий контекстного обучения.

Такой подход также является актуальным в процессе развития экосистемы цифровой экономики РФ в рамках хаба «Образование и кадры» [4] в целях ее вхождения и дальнейшего развития в составе мирового экономического пространства [4].

В мировой практике в качестве перспективного направления к форме эксплуатации специализированного экспериментального оборудования в удаленном режиме позиционируется метод, основанный на концепции мультиарендности [5–11]. Основным подходом реализации данного метода являются многопользовательские распределенные измерительно-управляющие системы (МРИУС) [5, 6, 8].

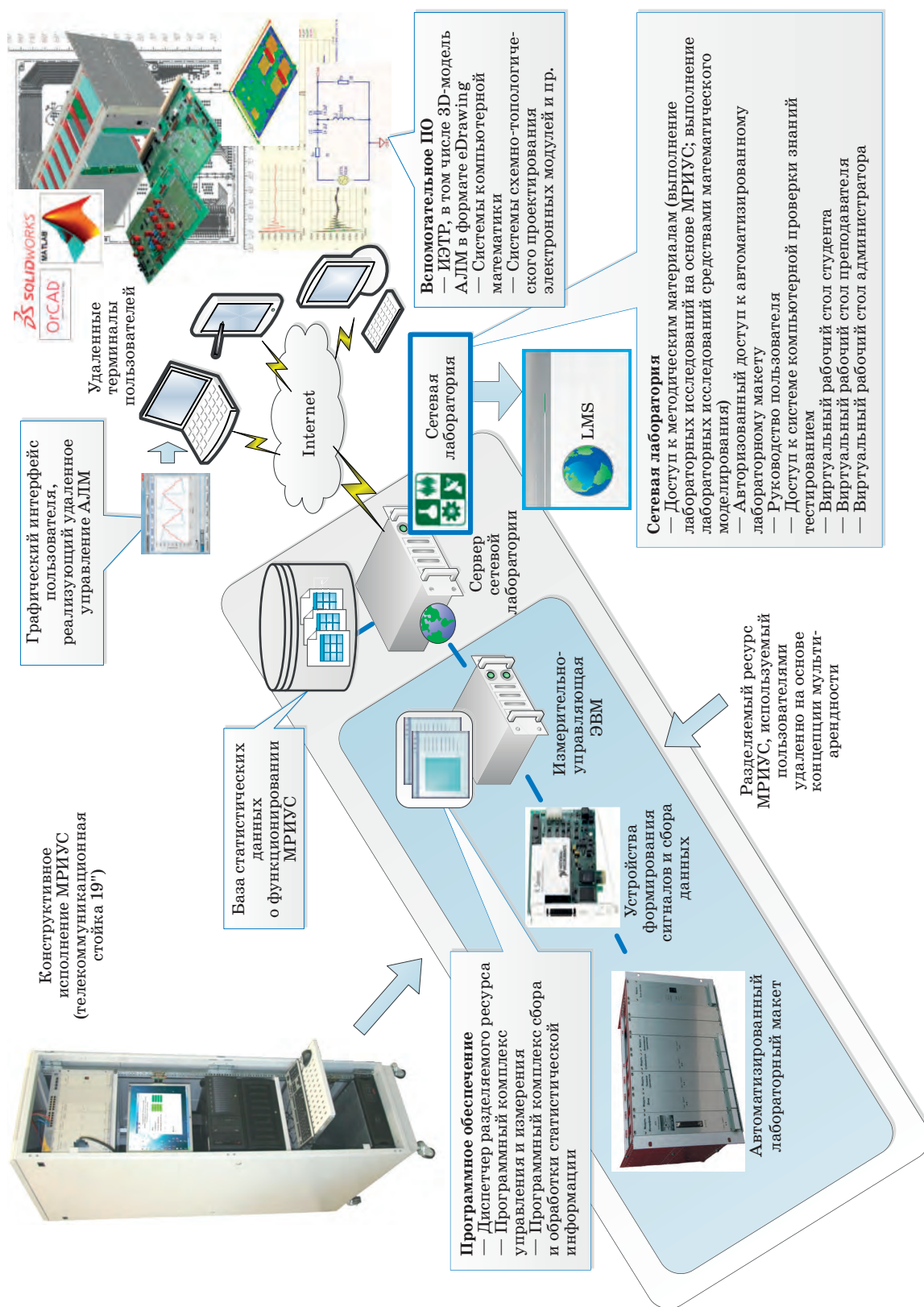
Таким образом, разработка методологических подходов применения МРИУС в качестве инструментария информационно-коммуникационных технологий (ИКТ) в составе ЦОС для организации дистанционных лабораторных исследований на базе современного экспериментального оборудования применительно к подготовке специалистов в области техники и технологий с использованием методов электронной педагогики и контекстного обучения является достаточно актуальной задачей [12].

Организация цифровой образовательной среды с применением МРИУС

Опыт авторов в области разработки, внедрения и эксплуатации МРИУС в различных предметных областях [6, 13–15] на протяжении более 10 лет привел к созданию унифицированной структуры сегмента ЦОС на основе МРИУС (рис. 1) и специальной методики адаптации МРИУС под различные условия их эксплуатации. Предложенная структура придает ЦОС характер высокотехнологичной среды и обеспечивает возможность совмещения натурального и вычислительного экспериментов. При этом соответствующее экспериментальное оборудование для выполнения натурального эксперимента используется на основе концепции мультиарендности.

Согласно представленному построению сегмента современной ЦОС, для каждого изучаемых в рамках лабораторных исследований объектов, устройств, процессов или явлений формируется соответствующая инфраструктура ЦОС, включающая в свой состав все или несколько из приведенных на рис. 1 компонентов. В общем случае в составе сегмента современной ЦОС в соответствии с предложенной структурой построения выделяются следующие ее основные компоненты.

1. Сетевая лаборатория [6, 13, 15], размещаемая на специализированном сервере, представляющая собой комплекс специализированных интерфейсов и интерактивных форм, на основе которых обеспечивается регламентированный



■ Рис. 1. Унифицированная структура сегмента современной ЦОС на основе МРИУС
 ■ Fig. 1. The unified structure of the modern digital educational environment based on distributed measurement and control system

авторизованный доступ как к методическому обеспечению лабораторных исследований, так и к ресурсам МРИУС, а также другим компонентам ЦОС (см. рис. 1).

2. МРИУС, включающая в свой состав:

- автоматизированный лабораторный макет/стенд/установку (АЛМ), определяющий специфику исследуемых объектов, устройств, процессов или явлений;

- устройства сбора данных с АЛМ, а также устройства формирования тестовых и (или) управляющих сигналов;

- измерительно-управляющую ЭВМ, обеспечивающую под управлением программного комплекса управления и измерения выполнение необходимых функциональных операций на основе АЛМ, устройств формирования сигналов и сбора данных в соответствии с поступающими с терминалов пользователей запросами;

- графический интерфейс пользователя (обеспечивает возможность выполнения пользователем экспериментальных исследований на основе АЛМ в интерактивном диалоговом режиме с удаленных терминалов посредством формирования в цифровом виде соответствующих запросов. Интерфейс реализуется, как правило, на основе web-технологий или в виде отдельных исполняемых приложений);

- диспетчер разделяемого ресурса (обеспечивает разрешение конфликтов при совместном обращении пользователей к разделяемому ресурсу МРИУС, а также осуществляет управление процедурой обслуживания поступивших запросов в соответствии с реализуемым алгоритмом диспетчеризации);

- программный комплекс сбора и обработки статистической информации (фиксирует с определенной дискретностью в базе данных основные данные о процессе взаимодействия каждого пользователя с разделяемым ресурсом МРИУС);

- удаленные терминалы пользователей (стационарные и мобильные ПЭВМ, планшеты — входят функционально).

3. База статистических данных о функционировании МРИУС (реализует накопление и хранение данных о параметрах процесса функционирования МРИУС, зафиксированных программным комплексом сбора и обработки статистической информации).

4. Комплекс вспомогательного программного обеспечения, включающий в свой состав:

- интерактивное электронное техническое руководство (комплект конструкторской документации, включая 3D-модель конструкции АЛМ, руководство пользователя, комплект технических отчетов по результатам исследований физических характеристик АЛМ [электрических, тепловых, показателей надежности и т. п.]

методами математического моделирования на стадии проектирования МРИУС) [16];

- комплекс базовых математических моделей (3D-геометрических, электрических, тепловых, надежности), представляющих собой формализованное описание цифровых двойников автоматизированных установок/стендов/макетов и (или) исследуемых объектов, устройств, процессов и явлений с точки зрения различных аспектов моделирования [17];

- набор специализированных САЕ/CAD/САМ-систем (используются для исследования характеристик цифровых двойников);

- интерфейс доступа к базе данных о функционировании МРИУС (реализуется в составе сетевой лаборатории или в виде автономного приложения).

Аппаратно-программные компоненты разделяемого ресурса МРИУС (см. рис. 1), решающие задачи автоматизации процесса формирования сигналов, сбора и обработки данных, измерения физических величин и их зависимостей, реализуются на основе типовых решений в области компьютерных измерительных технологий (National Instruments, Keysight, Adlink, «Актаком» и др.). Доступ к вспомогательному программному обеспечению в рамках ЦОС может быть предоставлен как непосредственно через сетевую лабораторию на основе концепции Software as a Service, так и путем установки на соответствующие терминалы пользователей в рамках специально оборудованных локаций доступа (компьютерных классов или отдельных рабочих мест) [11].

Организация выполнения научных и учебных экспериментальных исследований на базе рассмотренной ЦОС с применением МРИУС реализуется через обращение пользователей к специализированной сетевой web-лаборатории. Лаборатория может функционировать автономно, как элемент ЦОС, или инкорпорироваться в LMS-систему.

В общем случае в состав ЦОС входит несколько МРИУС, комплексированных на уровне сервера сетевой лаборатории.

Дистанционное выполнение соответствующих экспериментальных исследований осуществляется пользователем в интерактивном диалоговом режиме на основе доступного через сетевую лабораторию графического интерфейса пользователя. Посредством взаимодействия с элементами графического интерфейса, функционирующего на удаленном терминале (см. рис. 1), пользователь формирует необходимые параметры эксперимента (настройки конфигурации исследуемого узла или устройства, параметры исследуемого процесса, значения управляющих воздействий и тестовых сигналов и пр.), которые в дальнейшем в виде запроса передаются по сети на измерительно-управляющую ЭВМ. Запросы с термина-

лов пользователей, поступающие на измерительно-управляющую ЭВМ в порядке, определяемом алгоритмом работы диспетчера разделяемого ресурса, выполняются на основе АЛМ и соответствующих устройств формирования сигналов и сбора данных (см. рис. 1) посредством их «кратковременного захвата». Результаты выполнения запросов пользователей в виде измеренных параметров, графиков, зависимостей и массивов данных передаются обратно на терминал пользователя для их последующей обработки и отображения на графическом интерфейсе [5–11].

В целях исследования загрузки экспериментального оборудования, используемого на основе концепции мультиарендности, исследования функционирования МРИУС как системы массового обслуживания, оценки длительности выполняемых в соответствии с поступающими запросами функциональных операций, проведения оценки времени наработки оборудования, а также оценки удобства работы пользователей в состав аппаратно-программного обеспечения МРИУС дополнительно введен программный комплекс сбора и обработки статистической информации (см. рис. 1). В процессе функционирования МРИУС данный программный комплекс обеспечивает фиксацию моментов времени отправки заданий и получения результатов их выполнения на терминалах пользователей; времени поступления, начала и завер-

шения их выполнения на измерительно-управляющей ЭВМ; длительности соответствующих функциональных операций, выполняемых аппаратно-программным обеспечением МРИУС в соответствии с поступающим потоком запросов; служебной информации о параметрах сеанса работы пользователя (IP-адрес, идентификатор сессии и пр.), а также содержания всех обрабатываемых запросов персонализированно для каждого пользователя. Это позволяет централизованно формировать базу статистических данных о функционировании конкретной МРИУС посредством обновления ее содержимого с заданным интервалом времени (например, 1 час, 12 часов, 24 часа и т. д.). Фрагмент экспериментально полученных статистических данных для промышленного образца МРИУС — многопользовательского аппаратно-программного комплекса с удаленным доступом (АПК УД) «Электроника» [14] — показан на рис. 2.

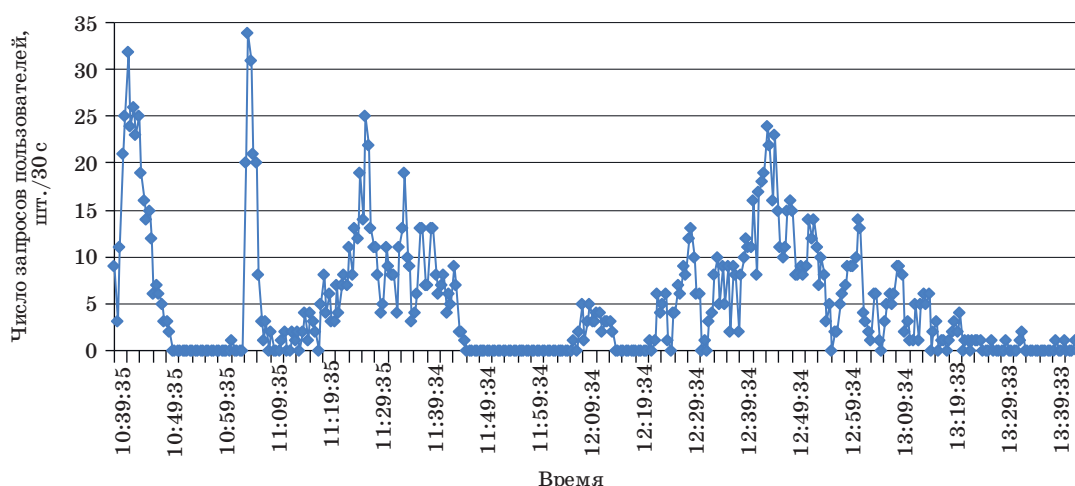
Примеры основных целевых функций, отражающих различные варианты обработки и представления непрерывно накапливаемой статистической информации в процессе функционирования МРИУС (рис. 3), приведены в таблице.

Диаграмма, построенная по результатам обработки статистической информации о моментах времени поступления запросов пользователей на измерительно-управляющую ЭВМ на примере АПК УД «Электроника» (см. рис. 3), характеризу-

IP RTF8(10.2.73.96)	IP RTF8(10.2.73.96)	IP RTF8(10.2.73.96)	IP RTF8(10.2.73.96)	IP RTF8(10.2.73.96)	IP RTF8(10.2.73.96)
ID 2044319	ID 2044319	ID 2044319	ID 2044319	ID 2044319	ID 2044319
N модуля 3	N модуля 3	N модуля 3	N модуля 3	N модуля 3	N модуля 3
N submodule 0	N submodule 0	N submodule 0	N submodule 0	N submodule 0	N submodule 0
Задание Задание 1	Задание Задание 1	Задание Задание 1	Задание Задание 1	Задание Задание 1	Задание Задание 1
Время отправки 9:10:36,375 30.11.2017	Время отправки 9:11:05,796 30.11.2017	Время отправки 9:11:33,046 30.11.2017	Время отправки 9:11:47,109 30.11.2017	Время отправки 9:12:00,078 30.11.2017	Время отправки 9:12:10,578 30.11.2017
Время начала выполнения 9:10:37,484 30.11.2017	Время начала выполнения 9:11:06,906 30.11.2017	Время начала выполнения 9:11:34,171 30.11.2017	Время начала выполнения 9:11:48,234 30.11.2017	Время начала выполнения 9:12:01,187 30.11.2017	Время начала выполнения 9:12:11,687 30.11.2017
Тобсл,с 1,109	Тобсл,с 1,418	Тобсл,с 1,419	Тобсл,с 1,43	Тобсл,с 1,351	Тобсл,с 1,413
Тож,с 1,161	Тож,с 1,471	Тож,с 1,47	Тож,с 1,482	Тож,с 1,402	Тож,с 1,465

■ **Рис. 2.** Фрагмент статистических данных процесса функционирования многопользовательского АПК УД «Электроника»: *IP* — сетевой адрес удаленного терминала пользователя; *ID* — уникальный идентификатор сеанса работы; *N модуля* — номер объектного модуля АПК УД, определяющий тип исследуемого электрорадиоэлемента; *N submodule* — номер типонаминала электрорадиоэлемента в составе соответствующего объектного модуля

■ **Fig. 2.** Fragment of statistical data of the functioning process the multi-user hardware and software complex with remote access “Electronics”: *IP* — the network address of the remote user terminal; *ID* — a unique identifier for the user session; *N module* — the number of the object module, which determines the type of the radio electronic element under study; *N submodule* — the number of the type label of an electrical radio element as part of the corresponding object module



■ **Рис. 3.** Фрагмент временной диаграммы числа поступивших запросов с терминалов пользователей на измерительно-управляющую ЭВМ для сеанса одновременной работы 22 терминалов пользователей в течение двух академических часов

■ **Fig. 3.** A fragment of the timing diagram of requests from user terminals to the measuring and controlling computer for a simultaneous work of 22 user terminals during two academic hours

■ Основные варианты обработки и представления статистической информации о функционировании МРИУС

■ The main variants of processing and presentation of statistical information on the functioning of distributed measurement and control system

Характеристика зависимости, вычисляемой на основе обработки экспериментальных данных	Целевая функция вычисляемой зависимости	Направление изменений, вносимых в МРИУС или в процесс ее эксплуатации
Гистограмма распределения числа запросов в течение анализируемого периода (дня/ недели/семестра и пр.)	Анализ модели поведения пользователей в МРИУС как в эргатической системе	Пересмотр организационных аспектов практических занятий на базе МРИУС — доступ к МРИУС с дополнительных локаций для определенных периодов времени эксплуатации. Верификация параметров математической модели МРИУС
Гистограмма распределения фактического числа пользователей в течение анализируемого временного интервала	Количественная оценка фактической загрузки МРИУС по числу обслуживаемых пользователей	Пересмотр организационных аспектов практических занятий на базе МРИУС — доступ к МРИУС с дополнительных локаций для определенных периодов времени эксплуатации
Процентное соотношение количества запросов по их принадлежности к отдельным компонентам АЛМ МРИУС (например, объектным модулям, исследуемым устройствам, узлам и пр.)	Анализ загрузки отдельных компонентов АЛМ МРИУС	Управление нагрузкой МРИУС: 1) реорганизация расписания доступа к МРИУС 2) подключение к МРИУС дополнительных (востребованных пользователями) экземпляров АЛМ или его компонентов 3) оперативное изменение алгоритма диспетчеризации разделяемого ресурса МРИУС с учетом специфики выполняемых исследований (измерений)
Число запросов с каждого терминала пользователей к МРИУС в пределах одной локации	Анализ удельной нагрузки на один персональный компьютер при выполнении экспериментальных исследований на базе МРИУС	1. Возможное использование мощностей компьютерного учебного класса общего назначения для доступа к МРИУС 2. Организация дополнительных локаций для доступа к МРИУС
Гистограмма распределения объема циркулирующих данных в вычислительной сети в соответствии с полученной статистикой выполнения запросов за анализируемый период времени	Выявление «пиков» загрузки вычислительной сети при эксплуатации МРИУС	Выработка требований к пропускной способности сети для конкретной локации

ет входящий поток запросов пользователей. На основе дальнейшей статистической обработки определяются параметры и вид распределения времени размышления пользователей между повторными запросами, которые учитываются в процессе верификации математической модели динамики функционирования МРИУС как системы массового обслуживания [18].

Методика адаптации МРИУС под различные условия эксплуатации

Анализ опыта внедрения ряда образцов МРИУС [6, 13–15, 19, 20] в учебный процесс и результатов исследования процесса их функционирования (времени реакции МРИУС на запросы пользователей, ее функциональных зависимостей от числа одновременно работающих пользователей и др.) на основе массива накопленных статистических данных выявили комплекс аспектов, требующих адаптации МРИУС к изменяющимся условиям их эксплуатации.

1. Значительное отклонение (более 40 %) в сторону увеличения количества пользователей от заложенных при проектировании МРИУС значений. Данный факт связан, например, с появлением дополнительных групп обучающихся и необходимостью введения для них дополнительных локаций доступа к МРИУС.

2. Значительное увеличение загрузки отдельных компонентов АЛМ. Проявляется, например, при формировании дополнительных групп обучающихся и локаций доступа к определенным составным компонентам АЛМ МРИУС. Примером может служить выполнение отдельных лабораторных исследований, в рамках которых дополнительно изучаются вопросы верификации и (или) идентификации параметров математических моделей отдельных исследуемых объектов, устройств, процессов и явлений на основе АЛМ или установки/стенда/макета в целом. Также данный аспект проявляется вследствие возникновения пиковой нагрузки, связанной, как правило, с предсессионной (аномальной) активностью обучающихся.

3. Необходимость задействования отдельных сочетаний исследовательских возможностей МРИУС с разработкой нового методического обеспечения в рамках реализации междисциплинарного (сквозного) использования разделяемого ресурса МРИУС (курсовое проектирование, курсы повышения квалификации, НИРС).

4. Результаты вариативности поведения пользователей с точки зрения влияния на время реакции МРИУС, выражающиеся:

— в возможном изменении в несколько раз времени «размышления» пользователей, например,

в зависимости от сложности и специфики выполняемых экспериментальных исследований;

— в наличии интервалов преимущественного формирования «однотипных» запросов, например, имеющих принадлежность к одному виду выполняемых измерений (осциллограммы, вольт-амперные характеристики, частотные характеристики) или определенному компоненту АЛМ, в рамках сеансов практических или лабораторных занятий.

Предложенная методика адаптации МРИУС к изменяющимся условиям эксплуатации в рамках выделенных аспектов в виде модели бизнес-процесса фрагмента жизненного цикла МРИУС в нотации BPMN (Business Process Model and Notation) приведена на рис. 4.

Методика отражает основные аспекты адаптации МРИУС к условиям ее эксплуатации. Данные аспекты реализуются посредством выполнения последовательности следующих основных шагов.

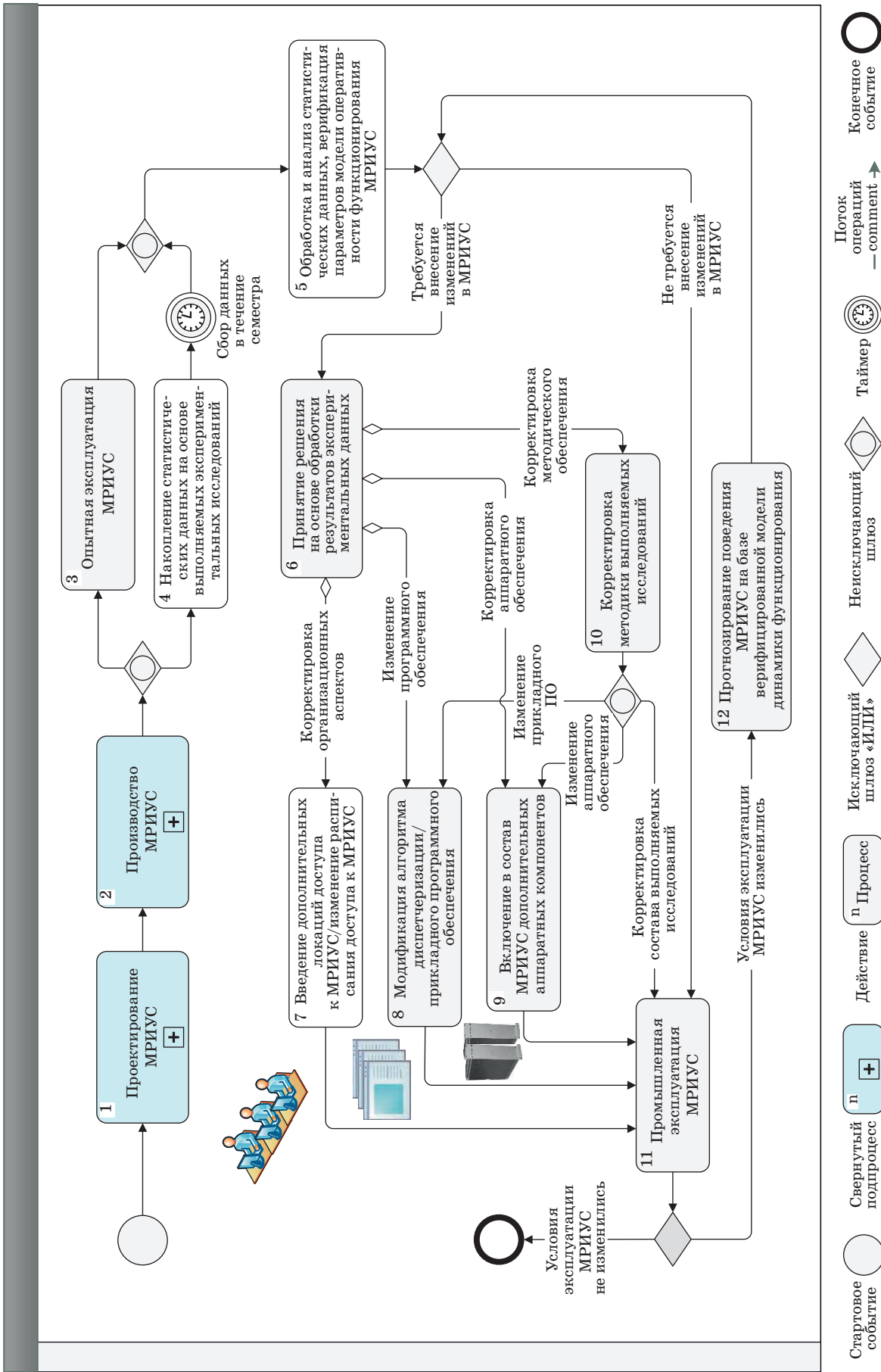
1. На основе анализа накапливаемых в процессе опытной эксплуатации МРИУС (блоки 3, 4) статистических данных (см. таблицу) выполняются идентификация и верификация соответствующей модели динамики функционирования МРИУС как системы массового обслуживания (блок 5), а также принятие решения о реализации соответствующих мероприятий (блок 6).

2. Верифицированная модель динамики функционирования МРИУС в дальнейшем используется для прогнозирования ее поведения при изменении ее параметров и (или) условий эксплуатации (блок 12).

3. Накопленные в процессе опытной эксплуатации статистические данные анализируются в целях принятия решения о необходимости внесения соответствующих технических и организационных изменений в МРИУС (блоки 7–10).

4. После внесения соответствующих изменений в МРИУС осуществляется ее дальнейшая эксплуатация (блок 11). При возникновении изменений условий эксплуатации МРИУС прогнозирование ее поведения и принятие решений о необходимости реализации комплекса мероприятий в дальнейшем осуществляется на основе верифицированной модели динамики ее функционирования (блок 12).

На основе накапливаемых статистических данных о работе каждого пользователя с МРИУС, комплексированных с данными о его работе в сетевой лаборатории (журнала работы соответствующего сервера), в дальнейшем формируется его «цифровая тень» пребывания в ЦОС. Анализ «цифровой тени» дает возможность выполнять соответствующую персонализацию ЦОС в части графического интерфейса пользователя МРИУС или декомпозиции состава выполняемых исследований, например, в соответствии с выделяе-



■ Рис. 4. Методика адаптации МРИУС к изменяющимся условиям эксплуатации
 ■ Fig. 4. Methods of adaptation distributed measurement and control system to changing operating conditions

мыми профилями пользователей (уровнем подготовки, особенностями субъективного восприятия контента и др.).

Направления контекстного обучения, реализуемые в цифровой образовательной среде с применением МРИУС

Реализация технологий контекстного обучения на основе рассмотренных выше сегмента ЦОС и принципов функционирования МРИУС выполняется в следующих направлениях.

1. Проведение лабораторных исследований в предметной области, определяемой конкретной МРИУС с применением современных компьютерных измерительных технологий (приобретаемый навык: работа с современными ИКТ-инструментами для решения прикладных задач).

2. Выполнение лабораторных исследований в предметной области, определяемой конкретной МРИУС с применением математической модели лабораторного макета/установки (приобретаемый навык: исследование характеристик технического объекта с применением математических моделей).

3. Исследование характеристик цифрового двойника лабораторного макета/установки с применением типовых методик сквозного автоматизированного проектирования современных приборов/мехатронных устройств (приобретаемый навык: применение цифрового двойника и интернет-технологий в процессе разработки новых образцов электронных/мехатронных устройств).

4. Анализ структуры интерактивного электронного технического руководства и его применение на различных стадиях жизненного цикла изделия (приобретаемый навык: применение компонентов технологии непрерывной информационной поддержки жизненного цикла изделия).

5. Идентификация и верификация как математических моделей исследуемых объектов, устройств, процессов и явлений, так и математических моделей образцов МРИУС в виде соответствующих систем массового обслуживания для формирования инструментария предсказательной (предиктивной) аналитики для этапов проектирования и эксплуатации (приобретаемый навык: применение методов математического моделирования для снижения объемов натурных испытаний технических объектов).

6. Анализ и выявление статистических закономерностей; апробация различных алгоритмов диспетчеризации; сравнение экспериментальных исследований с результатами математического моделирования режимов работы МРИУС на основе накопленного массива статистических данных о функционировании как АЛМ в целом, так и

его отдельных компонентов (объектных модулей/устройств/узлов, входящих в его состав) (приобретаемый навык: исследование принципов работы и основных характеристик современного ИКТ-инструментария).

7. Изучение и демонстрация принципов построения и функционирования МРИУС автоматизации учебного и научного эксперимента как сегмента современной ЦОС и как современного средства профессиональной практико-ориентированной подготовки специалистов в области техники и технологий (приобретаемый навык: владение методами электронной педагогики, повышающими качество преподавания технических дисциплин).

Реализация первого направления актуальна как для системы среднего образования, так и для систем среднего профессионального и высшего образования [1]. Реализация направлений 2–4 достаточно актуальна для направлений подготовки студентов в области техники и технологий системы высшего образования. Приобретаемые при этом студентами навыки позволяют им в дальнейшем легко адаптироваться в соответствующей профессиональной среде, которой в настоящее время и в перспективе будет соответствовать высокий уровень качества проведения научно-исследовательских и опытно-конструкторских работ (НИР и ОКР), выполняемых с широким применением ИКТ. Пятое и шестое направления при реализации в системе высшего образования позволяют организовать НИР бакалавров и магистров, прививая при этом им элементы ИТ-культуры при проведении НИР и ОКР. Реализация седьмого направления позволяет по линии университетских структур повышения квалификации и переподготовки кадров привить школьным учителям и преподавателям вузов профессиональные и пользовательские навыки в части электронной педагогики и концепции построения современного ИКТ-инструментария, а также компьютерных измерительных технологий для экспериментальных исследований в рамках ЦОС. В разрезе электронной педагогики данное направление позволяет раскрыть новые возможности реализации на базе сегмента ЦОС таких методов и форм обучения, как электронное, мобильное, сетевое и смешанное.

Приобретаемые пользователями сегмента ЦОС навыки в характеризуемых выше направлениях относятся непосредственно к профессиональной деятельности будущих специалистов в сфере цифровой экономики.

Заключение

Разработанная авторами, апробированная и описанная выше структура сегмента ЦОС обла-

дает отличительными особенностями, которые позволяют:

— обеспечить ЦОС возможностью подготовки современных специалистов с применением вышеперечисленных методов электронной педагогики и инновационного ИКТ-инструментария;

— применять специализированное экспериментальное оборудование (лабораторные макеты, установки, стенды) на основе концепции мультиарендности, обеспечивающей в десятки раз сокращение себестоимости оснащения одного рабочего места для проведения лабораторных исследований [13];

— оперативно модифицировать аппаратно-программное обеспечение МРИУС в целях эффективного их применения в соответствии с изменяющимися условиями эксплуатации;

— на системном уровне применять технологии контекстного обучения для ряда ключевых направлений цифровой экономики;

— значительно расширить возможности и качество современных образовательных ресурсов — электронных учебно-методических комплексов (компонента «Электронный лабораторный практикум»/ «Лабораторный практикум удаленного доступа») [21].

Дальнейшее совершенствование рассмотренного сегмента ЦОС планируется в части развития направлений и методов персонализированной адаптации форм представления контента в ЦОС к индивидуальному уровню подготовки и особенностям восприятия контента обучаемым, которые определяются его цифровым следом.

Литература

1. Подлесный С. А. Электронное обучение и обеспечение его качества. *Инженерное образование*, 2013, № 12, с. 104–111.
2. Ivannikov A. D., Tumkovskiy S. R. Internet portals as integration means of access to educational resources. *Proc. of 10th Intern. Technology, Education and Development Conf., INTED2016*, Valencia, IATED Academy, 2016, pp. 7337–7343. doi:10.21125/inted.2016.0731
3. Мазурицкий М. И., Солдатов А. В. Интерактивные сетевые научно-образовательные ресурсы для естественно-научного образования. *Высшее образование в России*, 2014, № 1, с. 80–87.
4. Программа «Цифровая экономика РФ» (утверждена распоряжением Правительства РФ от 28 июня 2017 г. № 1632-р).
5. *Cyber-physical laboratories in engineering and science education*/ Eds. M. E. Auer, A. K. M. Azad, A. Edwards, T. De Jong. Springer International Publishing AG, part of Springer Nature, 2018. 434 p. <https://doi.org/10.1007/978-3-319-76935-6>
6. Глинченко А. С., Дектерев М. Л., Комаров В. А., Сарафанов А. В. Сетевой учебно-исследовательский центр коллективного пользования уникальным лабораторным оборудованием на базе веб-портала как элемент системы дистанционного образования. *Открытое образование*, 2009, № 5, с. 18–29.
7. Евдокимов Ю. К., Кирсанов А. Ю., Салахова А. Ш. Дистанционные автоматизированные учебные лаборатории и технологии дистанционного учебного эксперимента в техническом вузе. *Открытое образование*, 2009, № 5, с. 101–116.
8. Липай Б. Р., Маслов С. И. Интернет-лаборатория Основы электротехники и электроники как пример современного учебного комплекса с удаленным доступом для открытого инженерного образования. *Вестник Московского энергетического института*, 2017, № 2, с. 71–76. doi:10.24160/1993-6982-2017-2-71-76
9. Markan C. M., Kumar G., Mittal S., Gupta P., Gupta S., Satsangi A., Gupta A., Kapur G. Remote triggered analog communication laboratory for e-learning. *International Journal of Online Engineering (iJOE)*, 2013, vol. 9, no. 5, pp. 48–55. doi:10.3991/ijoe.v9i5.2773
10. Ёжин М. Н., Степанов М. М. Организация многопользовательского удаленного доступа к распределенной гетерогенной системе лабораторного оборудования на основе схем программируемой логики для дистанционных практикумов по цифровой схемотехнике. *Современные информационные технологии и ИТ-образование*, 2017, т. 13, № 4, с. 65–81. doi:10.25559/SITITO.2017.4.425
11. Koike N. Cyber laboratory: Migration to the hybrid cloud solution for device dependent hardware experiments. *Information Technology Based Higher Education and Training (ITHET)*, York, 2014, pp. 1–5. doi:10.1109/ITHET.2014.7155681
12. Носкова Т. Н., Павлова Т. Б., Яковлева О. В. ИКТ-инструменты профессиональной деятельности педагога: сравнительный анализ российского и европейского опыта. *Интеграция образования*, 2018, т. 22, № 1, с. 25–45. doi: 10.15507/1991-9468.090.022.201801.025-045
13. Сарафанов А. В., Комаров В. А., Худоногов Д. Ю., Суковатый А. Г. Изучение физических эффектов с использованием дистанционных технологий. *Информатизация образования и науки*, 2012, № 4(16), с. 49–63.
14. Глинченко А. С., Егоров Н. М., Комаров В. А., Сарафанов А. В. *Исследование параметров и характеристик полупроводниковых приборов с применением интернет-технологий*. М., ДМК Пресс, 2008. 352 с.
15. *Сетевая лаборатория «Полупроводниковые приборы»*. <http://semicond.vlab.sfedu.ru/> (дата обращения: 18.12.2018).

16. ГОСТ 2.611–2011. Единая система конструкторской документации. Электронный каталог изделий. Общие положения. М., Стандартинформ, 2014. 28 с.
17. ГОСТ Р 57412–2017. Компьютерные модели в процессах разработки, производства и эксплуатации изделий. Общие положения. М., Стандартинформ, 2017. 10 с.
18. Комаров В. А., Сарафанов А. В. Имитационное моделирование процесса функционирования многопользовательских распределенных измерительно-управляющих систем. *Измерительная техника*, 2011, № 2, с. 16–19.
19. Пат. 2481621 РФ. МПК G06F 15/16, G06F 17/00. Способ функционирования распределенной измерительно-управляющей системы, А. С. Глинченко, В. А. Комаров, А. В. Сарафанов (РФ). № 2012111637; завл. 26.03.2012; опублик. 10.05.2013, Бюл. № 13. 15 с.
20. Володина Д. Н., Дектерев М. Л., Комаров В. А., Преснякова Г. О., Сарафанов А. В., Суковатый А. Г., Трухин А. А., Худоногов Д. Ю. *Исследования физических явлений в электрических цепях с применением интернет-технологий*. М.: ДМК Пресс, 2015. 432 с.
21. ГОСТ Р 55751–2013. Информационно-коммуникационные технологии в образовании. Электронные учебно-методические комплексы. Требования и характеристики. М., Стандартинформ, 2014. 11 с.

UDC 62.519

doi:10.31799/1684-8853-2019-2-83-94

Multi-user distributed information-control systems as an element of modern digital educational environmentV. A. Komarov^a, PhD, Tech., Associate Professor, orcid.org/0000-0001-9210-9908A. V. Sarafanov^b, Dr. Sc., Tech., Professor, orcid.org/0000-0003-4264-9388S. R. Tumkovskiy^c, Dr. Sc., Tech., Professor, orcid.org/0000-0002-1647-2171, STumkovskiy@hse.ru^aSC «Academician M. F. Reshetnev «Information Satellite System», 52, Lenin St., 662972, Zheleznogorsk, Russian Federation^bООО Vitte Consulting (GK «I-Teco»), 15, Kedrov St., 117036, Moscow, Russian Federation^cNational Research University «Higher School of Economics», 20, Myasnitskaya St., 101000, Moscow, Russian Federation

Introduction: One of the key deterrents to the quality training of engineers for digital economy is the lack of access to modern laboratory equipment in Russian regional universities. An effective approach to solving this problem is organizing remote access to experimental equipment based on the concept of multi-tenancy. **Purpose:** Developing a method to organize a high-tech digital educational environment, providing laboratory research in a remote multi-user mode for the training of engineers with contextual learning technology. **Methods:** Creating a unified structure for a digital educational environment segment, based on multi-user distributed measuring-control systems with a unified structure and special methods of adapting such systems to various conditions of their operation. **Results:** A unified structure for a digital educational environment segment has been developed and tested. Its main elements are a web laboratory providing regulated access of trainees to the laboratory studies, and multi-user distributed measuring-control systems implementing these studies by providing the trainees with parallel access to their shared resources: automated laboratory layouts/stands/installations. The created software complex for the collection and processing of statistical information about the functioning of multi-user distributed measuring-control systems as mass service systems underlies the proposed method of their adaptation to various operating conditions. The proposed segment of digital educational environment allows you to implement a number of contextual training directions, including: application in experimental research of modern computer measuring technologies; studying the characteristics of a digital twin of a laboratory layout/installation using standard techniques; identification and verification of mathematical models of objects, devices, processes and phenomena under study, or samples of multi-user systems in the form of appropriate mass service systems; approbation of various shared resource dispatch algorithms. **Practical relevance:** The proposed structure of an educational environment segment provides training of engineers using innovative hardware and software solutions; operation of experimental equipment on the basis of multi-tenancy concept, reducing the cost of equipping a trainee workplace in dozens of times; system-level application of contextual learning technology.

Keywords — automated laboratory practical work with remote access, automated laboratory layout, remote access, web laboratory, multi-tenancy, virtual laboratory stand, multi-user access, distributed measurement and control system, digital educational environment.

For citation: Komarov V. A., Sarafanov A. V., Tumkovskiy S. R. Multi-user distributed information-control systems as an element of modern digital educational environment. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2019, no. 2, pp. 83–94 (In Russian). doi:10.31799/1684-8853-2019-2-83-94

References

- Podlesnyi S. A. Quality assurance and quality enhancement in E-Learning. *Inzhenernoe obrazovanie*, 2013, no. 12, pp. 104–111 (In Russian).
- Ivannikov A. D., Tumkovskiy S. R. Internet portals as integration means of access to educational resources. *Proc. of 10th Intern. Technology, Education and Development Conf., INTED2016*, Valencia, IATED Academy, 2016, pp. 7337–7343. doi:10.21125/inted.2016.0731
- Mazuritsky M. I., Soldatov A. V. Interactive network resources for natural science education and research. *Vysshhee obrazovanie v Rossii*, 2014, no. 1, pp. 80–87 (In Russian).
- The program “Digital Economy of the Russian Federation” (approved by the Government of the Russian Federation in its resolution No. 1632-r dated July 28, 2017) (In Russian).
- Cyber-physical laboratories in engineering and science education*. Eds. M. E. Auer, A. K. M. Azad, A. Edwards, T. De

- Jong. Springer International Publishing AG, part of Springer Nature, 2018. 434 p. <https://doi.org/10.1007/978-3-319-76935-6>
6. Glinchenko A. S., Dekterev M. L., Zaharin K. N., Komarov V. A., Sarafanov A. V. The network uchebno-research centre of collective using the unique labware on the basis of a web-portal as an element of system of remote formation. *Otkrytoe obrazovanie*, 2009, no. 5, pp. 18–29 (In Russian).
 7. Evdokimov Y. K., Kirsanov A. Y., Salahova A. Sh. The remote automated educational laboratories and technologies of remote educational experiment in the technical university. *Otkrytoe obrazovanie*, 2009, no. 5, pp. 101–116 (In Russian).
 8. Lipay B. R., Maslov S. I. The Internet-based laboratory “Fundamentals of electrical engineering and electronics” as an example of a modern remotely accessible educational package for obtaining open engineering education. *Vestnik Moskovskogo energeticheskogo instituta*, 2017, no. 2, pp. 71–76 (In Russian). doi:10.24160/1993-6982-2017-2-71-76
 9. Markan C. M., Kumar G., Mittal S., Gupta P., Gupta S., Satsangi A., Gupta A., Kapur G. Remote triggered analog communication laboratory for e-learning. *International Journal of Online Engineering (iJOE)*, 2013, vol. 9, no. 5, pp. 48–55. doi: 10.3991/ijoe.v9i5.2773
 10. Yokhin M. N., Stepanov M. M. Multiuser remote access to distributed heterogeneous system of programmable logic based laboratory equipment for remote digital circuits design labs. *Sovremennye informatsionnye tekhnologii i IT-obrazovanie*, 2017, vol. 13, no. 4, pp. 65–81 (In Russian). doi:10.25559/SITITO.2017.4.425
 11. Koike N. Cyber laboratory: Migration to the hybrid cloud solution for device dependent hardware experiments. *Information Technology Based Higher Education and Training (IThET)*, York, 2014, pp. 1–5. doi:10.1109/IThET.2014.7155681
 12. Noskova T. N., Pavlova T. B., Yakovleva O. V. ICT tools of professional teacher activity: a comparative analysis of Russian and European experience. *Integratsiya obrazovaniya*, 2018, vol. 22, no. 1, pp. 25–45 (In Russian). doi: 10.15507/1991-9468.090.022.201801.025-045
 13. Sarafanov A. V., Komarov V. A., Sukovaty A. G., Khudonogov D. U. Investigation of physical effects using remote technologies. *Informatizatsiya obrazovaniya i nauki*, 2012, no. 4(16), pp. 49–63 (In Russian).
 14. Glinchenko A. S., Egorov N. M., Komarov V. A., Sarafanov A. V. *Issledovanie parametrov i kharakteristik poluprovodnikovyykh priborov s primeneniem internet-tekhnologiy* [Research of parameters and characteristics of semiconductor devices using Internet technologies]. Moscow, DMK Press Publ., 2008. 352 p. (In Russian).
 15. *Setevaya laboratoriya “Poluprovodnikovyye pribory”* [Network laboratory “Semiconductor devices”]. Available at: <http://semicond.vlab.sfedu.ru/> (accessed 18 December 2018).
 16. State Standard 2.611–2011. Unified system for design documentation. Electronic catalog of products. General provisions. Moscow, Standartinform Publ., 2012. 28 p. (In Russian).
 17. State Standard R 57412–2017. Computer models in processes, development of production and operation of products. General provisions. Moscow, Standartinform Publ., 2017. 10 p. (In Russian).
 18. Komarov V. A., Sarafanov A. V. Simulation of the operation of multiuser distributed measurement and control systems. *Izmeritel'naya tekhnika* [Measurement Techniques], 2011, vol. 54, no. 2, pp. 129–134 (In Russian). doi:10.1007/s11018-011-9695-y
 19. Glinchenko A. S., Komarov V. A., Sarafanov A. V. *Sposob funktsionirovaniya raspredelennoy izmeritel'no-upravlyayushchey sistemy* [The method of functioning of a distributed measuring and control system]. Patent RF, no. 2481621, 2013.
 20. Volodina D. N., Dekterev M. L., Komarov V. A., Presnyakova G. O., Sarafanov A. V., Sukovaty A. G., Trukhin A. A., Khudonogov D. U. *Issledovaniya fizicheskikh yavleniy v elektricheskikh tsepyakh s primeneniem Internet-tekhnologiy* [Research of physical effects in electrical circuits using Internet technologies]. Moscow, DMK Press Publ., 2015. 432 p. (In Russian).
 21. State Standard R 55751–2013. Information and communication technologies in education. Electronic teaching materials. Requirements and specifications. Moscow, Standartinform Publ., 2014. 11 p. (In Russian).

МЕЖДУНАРОДНАЯ КОНФЕРЕНЦИЯ ПО КОМПЬЮТЕРНОЙ МАТЕМАТИКЕ – CAM-2019

22–24 июля 2019 г.

Место проведения: Санкт-Петербургский государственный электротехнический университет «ЛЭТИ», зал видеоконференций, 5-й корпус

Адрес: ул. Профессора Попова, 5, Санкт-Петербург, 197376, РФ

Конференция CAM-2019 — это форум для ученых и преподавателей, которые развивают направление алгоритмической математики и заинтересованы в продвижении важных идей дискретной математики и теоретической информатики в образовании, в технических университетах и школах, в том числе посредством разработки компьютерных инструментов для поддержки исследований и продуктивного мышления.

Организатор

Санкт-Петербургский государственный электротехнический университет «ЛЭТИ», Россия

Оргкомитет

Михаил Куприянов, Санкт-Петербургский государственный электротехнический университет «ЛЭТИ», Россия

Сергей Поздняков, Санкт-Петербургский государственный электротехнический университет «ЛЭТИ», Россия

Программный комитет

Олег Иванов (председатель), Санкт-Петербургский государственный университет, Россия

Сергей Поздняков (сопредседатель), Санкт-Петербургский государственный электротехнический университет «ЛЭТИ», Россия

Николай Васильев (сопредседатель), ПОМИ РАН, Россия

Сергей Абрамович, Государственный университет Нью-Йорка в Потсдаме, США

Семён Адлай, ФИЦ ИУ РАН, Россия

Алкивиадис Акритас, Университет Фессалии, Греция

Владимир Дубровский, МГУ, Россия

Виктор Фрейман, Университет Монктона, Канада

Дима Григорьев, CNRS, Лилль, Франция

Юрай Хромкович, ЕТН Цюрих, Швейцария

Мирыяна Иванович, Университет Нови-Сад, Сербия

Эугениус Куриловас, Вильнюсский технический университет им. Гедиминаса, Литва

Юрий Матиясевиц, ПОМИ РАН, Россия

Кристиан Меркат, Университет Клода Бернара, Лион, Франция

Дмитрий Павлов, ИПА РАН, Россия

Михаэль Вейгенд, Университет Мюнстера, Германия

Татьяна Мюллари, Георгиевский университет, Гренада, Вест-Индия

Сергей Соловьев, Тулузский институт компьютерных исследований, Франция

Секции

Алгоритмическая математика и математическое моделирование

Обучение алгоритмической математике и математическому моделированию

Компьютерные инструменты для поддержки преподавания и изучения математики и информатики

Направления работы

Алгоритмическая математика

Математическое моделирование

Дискретный анализ

Теория управления

Теоретическая информатика

Символьные вычисления

Нерешенные проблемы теоретической и прикладной информатики

Разработка программного обеспечения

Компьютерное обучение

Программное обеспечение в области математического образования и исследований

Обучение алгоритмической математике и математическому моделированию

Компьютерные инструменты и технологии обучения для поддержки продуктивного мышления в математике и информатике

Информационная поддержка

Журнал «Компьютерные инструменты в науке и образовании» (КИО) и журнал «Компьютерная математика», который выйдет в 2019 году в качестве преемника журнала КИО и будет выпускаться полностью на английском языке.

Журнал «Информационно-управляющие системы».

Рабочий язык конференции

Английский

Контрольные даты

Прием материалов до 01 мая 2019 г.

Стоимость участия

Для граждан РФ и стран СНГ:

полный пакет участника — 7500 руб.

для студентов — бесплатно

Для иностранных участников:

полный пакет участника — 250 EURO

для аспирантов — 100 EURO

для студентов — 75 EURO

Дополнительная информация и справки

cam19@ipo.spb.ru

Сайт: <https://www.easychair.org/cfp/CAM-2019>

Для загрузки тезисов: <https://easychair.org/conferences/?conf=cam20190>

Тел.: +7-921-743-92-34

АНДРЕЕВ
Андрей
Васильевич



Аспирант кафедры вычислительных систем и сетей Санкт-Петербургского государственного университета аэрокосмического приборостроения. В 2011 году окончил магистратуру Санкт-Петербургского государственного университета аэрокосмического приборостроения по специальности «Информатика и вычислительная техника». Является автором четырех научных публикаций. Область научных интересов — службы каталогов, свободное программное обеспечение, теория множеств, теория графов, операционные системы. Эл. адрес: aathexf@gmail.com

ВАСИЛЬЕВ
Николай
Николаевич



Старший научный сотрудник лаборатории теории представлений и динамических систем Санкт-Петербургского отделения Математического института им. В. А. Стеклова РАН. В 1975 году окончил математико-механический факультет Ленинградского государственного университета по специальности «Алгебра и теория чисел». В 1990 году защитил диссертацию на соискание ученой степени кандидата физико-математических наук. Является автором более 120 научных публикаций. Область научных интересов — компьютерная алгебра, вычислительная алгебраическая геометрия и коммутативная алгебра, теория динамических систем, небесная механика. Эл. адрес: vasiliev@pdmi.ras.ru

ЖЕЛЕЗНЫ
Милош



Доцент, заместитель декана по концепции обучения и педагогической деятельности Западночешского университета, Пльзень, Чехия. В 1994 году окончил магистратуру Западночешского университета, Пльзень, Чехия, по специальности «Кибернетика и техника управления». В 2002 году защитил диссертацию на соискание доктора наук (PhD). Является автором 70 научных публикаций. Область научных интересов — человеко-машинное взаимодействие, обработка аудиовизуальных сигналов, распознавание образов. Эл. адрес: zelezny@kky.zcu.cz

БЕЛИМ
Сергей
Викторович



Профессор, заведующий кафедрой информационной безопасности Омского государственного университета им. Ф. М. Достоевского. В 1996 году окончил Омский государственный университет по специальности «Физика». В 2009 году защитил диссертацию на соискание ученой степени доктора физико-математических наук. Является автором 230 научных публикаций. Область научных интересов — защита информации, модели безопасности компьютерных систем, интеллектуальные системы обработки изображений. Эл. адрес: sbelim@mail.ru

ГОРДЕЕВ
Александр
Владимирович



Профессор кафедры вычислительных систем и сетей Санкт-Петербургского государственного университета аэрокосмического приборостроения. В 1975 году окончил Ленинградский институт авиационного приборостроения по специальности «Электронные вычислительные машины». В 1998 году защитил диссертацию на соискание ученой степени доктора технических наук. Является автором более 90 научных публикаций. Область научных интересов — моделирование параллельных вычислительных систем, распределенные вычисления, вычислительные сети. Эл. адрес: avg@aanet.ru

ИВАНЬКО
Денис
Викторович



Младший научный сотрудник лаборатории речевых и мультимедийных интерфейсов Санкт-Петербургского института информатики и автоматизации РАН, аспирант по программе двойной аспирантуры в Университете ИТМО (РФ) и Университете Ульма (Германия). В 2015 году окончил магистратуру Университета ИТМО по специальности «Речевые информационные системы». Является автором 12 научных публикаций. Область научных интересов — аудиовизуальное распознавание речи, автоматическое чтение речи по губам диктора, машинное обучение. Эл. адрес: denis.ivanko11@gmail.com

КАРПОВ
Алексей
Анатольевич



Главный научный сотрудник, руководитель лаборатории речевых и многомодальных интерфейсов Санкт-Петербургского института информатики и автоматизации РАН.

В 2002 году окончил Санкт-Петербургский государственный университет аэрокосмического приборостроения по специальности «Вычислительные машины, комплексы, системы и сети». В 2013 году защитил диссертацию на соискание ученой степени доктора технических наук. Является автором более 300 научных публикаций, включая три монографии и три патента. Область научных интересов — многомодальные интерфейсы и системы, речевые технологии, автоматическое распознавание и синтез речи, компьютерная паралингвистика.

Эл. адрес: karпов@iias.spb.su

КЛЮЧАРЁВ
Александр
Анатольевич



Доцент кафедры компьютерных технологий и программной инженерии Санкт-Петербургского государственного университета аэрокосмического приборостроения, почетный работник высшего профессионального образования РФ.

В 1971 году окончил Севастопольский приборостроительный институт по специальности «Общая радиотехника».

В 1982 году защитил диссертацию на соискание ученой степени кандидата технических наук. Является автором более 50 научных публикаций и 10 изобретений.

Область научных интересов — обработка экспериментальных данных, системы автоматизации натурального эксперимента, информационно-измерительные системы.

Эл. адрес: ak@aanet.ru

КОМАРОВ
Владимир
Александрович



Начальник группы комплексного моделирования космических систем АО «Информационные спутниковые системы имени академика М. Ф. Решетнёва», доцент Сибирского федерального университета, Железногорск.

В 2005 году окончил Красноярский государственный технический университет по специальности «Техническая эксплуатация транспортного радиооборудования».

В 2009 году защитил диссертацию на соискание ученой степени кандидата технических наук. Является автором более 90 научных публикаций.

Область научных интересов — компьютерные измерительные технологии, системы дистанционного управления объектами и др.

Эл. адрес: VKomarov@iss-reshetnev.ru

КОНОВАЛОВ
Петр
Алексеевич



Студент 3-го курса кафедры теоретической кибернетики математико-механического факультета Санкт-Петербургского государственного университета.

Область научных интересов — мультиагентные системы, алгоритмы навигации мобильных роботов.

Эл. адрес: petrkon98@gmail.com

КОРНИЛОВА
Анастасия
Валерьевна



Инженер-программист компании ООО «КиберТехЛабс», Санкт-Петербург.

В 2018 году окончила бакалавриат Санкт-Петербургского государственного университета по специальности «Математическое обеспечение и администрирование информационных систем».

Является автором пяти научных публикаций.

Область научных интересов — техническое зрение, цифровая обработка сигналов, автономные системы, мультиагентные системы, образовательная робототехника.

Эл. адрес: kornilova.anastasiia@gmail.com

КОРОЛЕВ
Дмитрий
Михайлович



Аспирант, начальник лаборатории «Комплексных систем управления мобильными роботами» Центрального научно-исследовательского и опытно-конструкторского института робототехники и технической кибернетики, Санкт-Петербург.

В 2010 году окончил Балтийский государственный технический университет им. Д. Ф. Устинова по специальности «Мехатроника».

Область научных интересов — методика построения унифицированных систем управления, диагностики и контроля узлов роботов специального назначения.

Эл. адрес: kalter.eden@gmail.com

**КОТЕНКО
Игорь
Витальевич**



Профессор, главный научный сотрудник, заведующий лабораторией проблем компьютерной безопасности Санкт-Петербургского института информатики и автоматизации РАН.
В 1983 году окончил Военно-космическую академию им. А. Ф. Можайского по специальности «Математическое обеспечение автоматизированных систем управления», в 1987 году — Военную академию связи по специальности «Инженерная автоматизированных систем управления».
В 1999 году защитил диссертацию на соискание ученой степени доктора технических наук.
Является автором более 500 научных публикаций.
Область научных интересов — безопасность компьютерных сетей, обнаружение компьютерных атак, межсетевые экраны и др.
Эл. адрес: ivkote@comsec.spb.ru

**ПАРАСИДИС
Иоаннис**



Доцент кафедры электротехники Технологического института Фессалии, Греция.
В 1980 году окончил математический факультет Казахского государственного университета им. С. М. Кирова (Алма-Ата) по специальности «Математика».
В 1989 году защитил диссертацию на соискание ученой степени кандидата физико-математических наук.
Является автором 30 научных публикаций.
Область научных интересов — расширения операторов, краевые задачи для дифференциальных и интегродифференциальных уравнений, спектральные задачи, обратные задачи, разностные уравнения.
Эл. адрес: paras@teilar.gr

**ПРОВИДАС
Ефимиос**



Доцент кафедры машиностроения Технологического института Фессалии, Греция.
В 1984 году окончил бакалавриат математического факультета Аристотельского университета города Салоники, Греция.
В 1991 году защитил диссертацию (PhD) на факультете математики и статистики Брунельского университета, Великобритания.
Является автором 32 научных публикаций.
Область научных интересов — прикладная математика, дифференциальные, интегральные и разностные уравнения, численные методы, метод конечных элементов.
Эл. адрес: providas@teilar.gr

**НОВИКОВА
Евгения
Сергеевна**



Доцент кафедры информационных систем Санкт-Петербургского государственного электротехнического университета «ЛЭТИ», старший научный сотрудник лаборатории проблем компьютерной безопасности Санкт-Петербургского института информатики и автоматизации РАН.
В 2007 году окончила университет «ЛЭТИ» по специальности «Компьютерная безопасность».
В 2010 году защитила диссертацию на соискание ученой степени кандидата технических наук.
Является автором 80 научных публикаций и трех патентов на изобретения.
Область научных интересов — информационная безопасность, безопасность промышленного интернета вещей, визуальная аналитика, протоколы аутентификации.
Эл. адрес: novikova@comsec.spb.ru

**ПОЙМАНОВА
Екатерина
Дмитриевна**



Старший преподаватель кафедры безопасности информационных систем Санкт-Петербургского государственного университета аэрокосмического приборостроения.
В 2005 году окончила Санкт-Петербургский государственный университет телекоммуникаций им. проф. М. А. Бонч-Бруевича по специальности «Информационные системы в области связи».
Является автором более 15 научных публикаций.
Область научных интересов — системы и технологии хранения данных.
Эл. адрес: e.d.poymanova@gmail.com

**РАКИЦКИЙ
Юрий
Сергеевич**



Доцент кафедры информационной безопасности Омского государственного университета им. Ф. М. Достоевского.
В 2008 году окончил Омский государственный университет по специальности «Компьютерная безопасность».
В 2011 году защитил диссертацию на соискание ученой степени кандидата технических наук.
Является автором 26 научных публикаций.
Область научных интересов — защита информации, информационная безопасность, модели безопасности компьютерных систем.
Эл. адрес: yrakitskiy@gmail.com

РЕНЕВА
Галина
Вадимовна



Студент 3-го курса кафедры теоретической кибернетики математико-механического факультета Санкт-Петербургского государственного университета. Область научных интересов — мультиагентные системы управления.
Эл. адрес: renevag@gmail.com

РУКАВИЦЫН
Андрей
Николаевич



Аспирант, программист АО «Научно-инженерный центр Санкт-Петербургского электротехнического университета». В 2016 году окончил магистратуру Санкт-Петербургского государственного электротехнического университета «ЛЭТИ» по специальности «Информатика и вычислительная техника». Является автором 11 научных публикаций и пяти патентов на изобретения. Область научных интересов — интеллектуальный анализ данных, распределенный анализ данных, нейронные сети.
Эл. адрес: rkvtsn@gmail.com

РЮМИН
Дмитрий
Александрович



Младший научный сотрудник лаборатории речевых и многомодальных интерфейсов Санкт-Петербургского института информатики и автоматизации РАН, аспирант 3-го года обучения Университета ИТМО. В 2015 году окончил магистратуру Университета ИТМО по специальности «Речевые информационные системы». Является автором 10 научных публикаций. Область научных интересов — многомодальные интерфейсы, машинное зрение, автоматическое распознавание жестов, распознавание образов.
Эл. адрес: dl_03.03.1991@mail.ru

САРАФАНОВ
Альберт
Викторович



Директор по развитию бизнеса департамента управления предприятиями ООО Вите Консалтинг (ГК «Ай-теко»), Москва. В 1986 году окончил Красноярский политехнический институт Универсальности «Конструирование и производство радиоэлектронной аппаратуры». В 2001 году защитил диссертацию на соискание ученой степени доктора технических наук. Является автором более 260 научных публикаций. Область научных интересов — методы комплексного математического моделирования физических процессов, интегрированные компьютерные технологии проектирования бортовых радиоэлектронных средств, автоматизация физических экспериментов, компьютерные измерительные технологии и др.
Эл. адрес: sarafanov@i-teco.ru

СОВЕТОВ
Борис
Яковлевич



Профессор кафедры информационных систем Санкт-Петербургского государственного электротехнического университета «ЛЭТИ», заслуженный деятель науки и техники РФ, действительный член Российской академии образования. В 1960 году окончил Ленинградский электротехнический институт им. В. И. Ульянова (Ленина) по специальности «Автоматика и телемеханика». В 1974 году защитил диссертацию на соискание ученой степени доктора технических наук. Является автором более 400 научных публикаций. Область научных интересов — теория и практика информационных систем и технологий.
Эл. адрес: bysovetov@mail.ru

ТАТАРНИКОВА
Татьяна
Михайловна



Профессор кафедры безопасности информационных систем Санкт-Петербургского государственного университета аэрокосмического приборостроения. В 1993 году окончила Восточно-Сибирский технологический институт по специальности «Электронно-вычислительные машины, комплексы, системы и сети». В 2007 году защитила диссертацию на соискание ученой степени доктора технических наук. Является автором более 100 научных публикаций. Область научных интересов — инфокоммуникации, взаимодействие неоднородных сетей.
Эл. адрес: tm-tatarn@yandex.ru

**ТУМКОВСКИЙ
Сергей
Ростиславович**



Профессор департамента компьютерной инженерии Московского института электроники и математики им. А. Н. Тихонова Национального исследовательского университета «Высшая школа экономики», лауреат премии Правительства РФ в области науки и техники.

В 1981 году окончил Московский институт электронного машиностроения по специальности «Автоматика и телемеханика».

В 2007 году защитил диссертацию на соискание ученой степени доктора технических наук. Является автором более 140 научных и методических публикаций, 11 свидетельств о регистрации ПО.

Область научных интересов — информационные системы и технологии в различных сферах деятельности.

Эл. адрес: STumkovskiy@hse.ru

**ФРАДКОВ
Александр
Львович**



Заведующий лабораторией управления сложными системами Института проблем машиноведения РАН, профессор Санкт-Петербургского государственного университета и Университета ИТМО.

В 1971 году окончил математико-механический факультет Ленинградского государственного университета.

В 1986 году защитил диссертацию на соискание ученой степени доктора технических наук.

Является автором более 600 научных публикаций, в том числе десяти патентов и 17 монографий. Область научных интересов — нелинейное и адаптивное управление в физико-технических системах, управление колебательными и хаотическими системами, математическое моделирование и др.

Эл. адрес: fradkov@mail.ru

**ШИРОКОЛОБОВ
Илья
Юрьевич**



Инженер-программист компании ООО «КиберТехЛабс», Санкт-Петербург.

В 2011 году окончил Санкт-Петербургский государственный университет по специальности «Математик, системный программист».

Область научных интересов — мультиагентные системы управления, навигация, техническое зрение, робототехника.

Эл. адрес: shirokolobov.ilya@gmail.com

**ФОМЕНКОВА
Анастасия
Алексеевна**



Ассистент кафедры компьютерных технологий и программной инженерии Санкт-Петербургского государственного университета аэрокосмического приборостроения.

В 2012 году окончила Киевский политехнический институт по специальности «Оборудование фармацевтических и микробиологических производств».

Область научных интересов — проектирование, математическое и имитационное моделирование биореакторов, систем биологической очистки сточных вод.

Эл. адрес: a.a.fomenkova@mail.ru

**ХИМЕНКО
Виталий
Иванович**



Профессор, директор департамента развития НИО ЦИТ «Петрокомета», Санкт-Петербург, заслуженный деятель науки РФ, лауреат премии Правительства РФ в области науки и техники.

В 1969 году окончил Куйбышевский политехнический институт по специальности «Приборные устройства».

В 1992 году защитил диссертацию на соискание ученой степени доктора технических наук.

Является автором более 150 научных публикаций.

Область научных интересов — статистическая радиофизика, обработка информации в системах живой и неживой природы.

Эл. адрес: vih.13@yandex.ru

**ЯРОШ
Дмитрий
Сергеевич**



Студент 2-го курса Санкт-Петербургского государственного университета по специальности «Программная инженерия».

Область научных интересов — встраиваемые системы, оптимизация приложений, мультиагентные системы, алгоритмы навигации мобильных роботов.

Эл. адрес: dimajarosh@gmail.com