

THE MINISTRY OF SCIENCE AND HIGHER EDUCATION OF THE RUSSIAN FEDERATION



ISSN 2687-0517

Computing, Telecommunications and Control

**Vol. 15, No. 2
2022**

Peter the Great St. Petersburg
Polytechnic University
2022

COMPUTING, TELECOMMUNICATIONS AND CONTROL

EDITORIAL COUNCIL

Prof. Dr. *Rafael M. Yusupov* corresponding member of RAS, St. Petersburg Institute for Informatics and Automation of the RAS, Russia,
Prof. Dr. *Sergey M. Abramov* corresponding member of RAS, full member of RAS, Ailamazyan Program Systems Institute of the RAS,
Prof. Dr. *Dmitry G. Arseniev* corresponding member of RAS, Peter the Great St. Petersburg Polytechnic University, Russia,
Prof. Dr. *Vladimir V. Voevodin* corresponding member of RAS, Lomonosov Moscow State University, Russia,
Prof. Dr. *Vladimir S. Zaborovsky*, Peter the Great St. Petersburg Polytechnic University, Russia,
Prof. Dr. *Vladimir N. Kozlov*, Peter the Great St. Petersburg Polytechnic University, Russia,
Prof. Dr. *Alexandr E. Fotiadi*, Peter the Great St. Petersburg Polytechnic University, Russia,
Prof. Dr. *Igor G. Chernorutsky*, Peter the Great St. Petersburg Polytechnic University, Russia.

EDITORIAL BOARD

Editor-in-chief

Prof. Dr. *Alexander S. Korotkov*, Peter the Great St. Petersburg Polytechnic University, Russia;

Members:

Assoc. Prof. Dr. *Vladimir M. Itsykson*, Peter the Great St. Petersburg Polytechnic University, Russia;
Prof. Dr. *Philippe Ferrari*, Grenoble Alpes University, France;
Prof. Dr. *Yevgeni Koucheryavy*, Tampere University of Technology, Finland;
Prof. Dr. *Wolfgang Krautschneider*, Hamburg University of Technology, Germany;
Prof. Dr. *Fa-Long Luo*, University of Washington, USA;
Prof. Dr. *Sergey B. Makarov*, Peter the Great St. Petersburg Polytechnic University, Russia;
Prof. Dr. *Emil Novakov*, Grenoble Alpes University, France;
Prof. Dr. *Nikolay N. Prokopenko*, Don State Technical University, Russia;
Prof. Dr. *Mikhail G. Putrya*, National Research University of Electronic Technology, Russia;
Sen. Assoc. Prof. Dr. *Evgeny Pyshkin*, University of Aizu, Japan;
Prof. Dr. *Viacheslav P. Shkodyrev*, Peter the Great St. Petersburg Polytechnic University, Russia;
Prof. Dr. *Peter V. Trifonov*, ITMO University, Russia;
Prof. Dr. *Igor A. Tsikin*, Peter the Great St. Petersburg Polytechnic University, Russia;
Prof. Dr. *Sergey M. Ustinov*, Peter the Great St. Petersburg Polytechnic University, Russia;
Prof. Dr. *Lev V. Utkin*, Peter the Great St. Petersburg Polytechnic University, Russia.

The journal is included in the List of Leading PeerReviewed Scientific Journals and other editions to publish major findings of PhD theses for the research degrees of Doctor of Sciences and Candidate of Sciences.

Open access journal is to publish articles of a high scientific level covering advanced experience, research results, theoretical and practical problems of informatics, electronics, telecommunications, and control.

The journal is indexed by Ulrich's Periodicals Directory, Google Scholar, EBSCO, ProQuest, Index Copernicus, VINITI RAS Abstract Journal (Referativnyi Zhurnal), VINITI RAS Scientific and Technical Literature Collection, Russian Science Citation Index (RSCI) database Scientific Electronic Library and Math-Net.ru databases.

The journal is registered with the Federal Service for Supervision in the Sphere of Telecom, Information Technologies and Mass Communications (ROSKOMNADZOR). Certificate ЭЛ No. ФС77-77378 issued 25.12.2019.

Editorial office

Dr. Sc., Professor A.S. Korotkov – Editor-in-Chief;

E.A. Kalinina – literary editor, proofreader; G.A. Pyshkina – editorial manager; A.A. Kononova – computer layout; D.Yu. Alekseeva – English translation.

Address: 195251 Polytekhnicheskaya Str. 29, St. Petersburg, Russia.

+7 (812) 552-6216, e-mail: infocom@spbstu.ru

Release date: 30.08.2022

© Peter the Great St. Petersburg Polytechnic University, 2022

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ



ISSN 2687-0517

Информатика, теле^{коммуникации} и управление

**Том 15, № 2
2022**

Санкт-Петербургский политехнический
университет Петра Великого
2022

ИНФОРМАТИКА, ТЕЛЕКОММУНИКАЦИИ И УПРАВЛЕНИЕ

РЕДАКЦИОННЫЙ СОВЕТ ЖУРНАЛА

Юсупов Р.М., чл.-кор. РАН, Санкт-Петербургский институт информатики и автоматизации РАН, Санкт-Петербург, Россия; *Абрамов С.М.*, чл.-кор. РАН, Институт программных систем им. А.К. Айламазяна РАН, Москва, Россия; *Арсеньев Д.Г.*, чл.-кор. РАН, д-р техн. наук, профессор, Санкт-Петербургский политехнический университет Петра Великого, Санкт-Петербург, Россия; *Воеводин В.В.*, чл.-кор. РАН, Московский государственный университет им. М.В. Ломоносова, Москва, Россия; *Зaborовский В.С.*, д-р техн. наук, профессор, Санкт-Петербургский политехнический университет Петра Великого, Санкт-Петербург, Россия; *Козлов В.Н.*, д-р техн. наук, профессор, Санкт-Петербургский политехнический университет Петра Великого, Санкт-Петербург, Россия; *Фотиади А.Э.*, д-р физ.-мат. наук, профессор, Санкт-Петербургский политехнический университет Петра Великого, Санкт-Петербург, Россия; *Черноруцкий И.Г.*, д-р техн. наук, профессор, Санкт-Петербургский политехнический университет Петра Великого, Санкт-Петербург, Россия.

РЕДАКЦИОННАЯ КОЛЛЕГИЯ ЖУРНАЛА

Главный редактор

Коротков А.С., д-р техн. наук, профессор, Санкт-Петербургский политехнический университет Петра Великого, Санкт-Петербург, Россия;

Редакционная коллегия:

Ицыксон В.М., канд. техн. наук, доцент, Санкт-Петербургский политехнический университет Петра Великого, Санкт-Петербург, Россия;

Феррари Ф., профессор, Университет Гренобль-Альпы, Гренобль, Франция;

Краутшайдер В., профессор, Гамбургский технический университет, Гамбург, Германия;

Кучерявый Е.А., канд. техн. наук, профессор, Университет Тампere, Финляндия.

Люо Ф.-Л., University of Washington, Washington, USA;

Макаров С.Б., д-р техн. наук, профессор, Санкт-Петербургский политехнический университет Петра Великого, Санкт-Петербург, Россия;

Новаков Э., профессор, Университет Гренобль-Альпы, Гренобль, Франция;

Прокопенко Н.Н., д-р техн. наук, профессор, Донской государственный технический университет, г. Ростов-на-Дону, Россия;

Путря М.Г., д-р техн. наук, профессор, Национальный исследовательский университет «Московский институт электронной техники», Москва, Россия;

Пышкин Е.В., профессор, Университет Айзу, Айзу-Вакаматсу, Япония;

Трифонов П.В., д-р техн. наук, доцент, Национальный исследовательский университет ИТМО, Санкт-Петербург, Россия;

Устинов С.М., д-р техн. наук, профессор, Санкт-Петербургский политехнический университет Петра Великого, Санкт-Петербург, Россия;

Уткин Л.В., д-р техн. наук, профессор, Санкт-Петербургский политехнический университет Петра Великого, Санкт-Петербург, Россия;

Цикин И.А., д-р техн. наук, профессор, Санкт-Петербургский политехнический университет Петра Великого, Санкт-Петербург, Россия;

Шкодырев В.П., д-р техн. наук, профессор, Санкт-Петербургский политехнический университет Петра Великого, Санкт-Петербург, Россия.

Журнал с 2002 года входит в Перечень ведущих рецензируемых научных журналов и изданий, в которых должны быть опубликованы основные результаты диссертаций на соискание ученой степени доктора и кандидата наук.

Сетевое издание открытого доступа публикует статьи высокого научного уровня, освещающие передовой опыт, результаты НИР, теоретические и практические проблемы информатики, электроники, телекоммуникаций, управления.

Сведения о публикациях представлены в Реферативном журнале ВИНИТИ РАН, в международной справочной системе «Ulrich's Periodical Directory», в Российской государственной библиотеке. В базах данных: Российский индекс научного цитирования (РИНЦ), Google Scholar, EBSCO, Math-Net.Ru, ProQuest, Index Copernicus.

Журнал зарегистрирован Федеральной службой по надзору в сфере информационных технологий и массовых коммуникаций (Роскомнадзор). Свидетельство о регистрации ЭЛ № ФС77-77378 от 25.12.2019.

Учредитель и издатель: Санкт-Петербургский политехнический университет Петра Великого, Санкт-Петербург, Российская Федерация.

Редакция журнала

д-р техн. наук, профессор А.С. Коротков – главный редактор;

Е.А. Калинина – литературный редактор, корректор; Г.А. Пышкина – ответственный секретарь, выпускающий редактор;

А.А. Кононова – компьютерная вёрстка; Д.Ю. Алексеева – перевод на английский язык.

Адрес редакции: Россия, 195251, Санкт-Петербург, ул. Политехническая, д. 29.

Тел. редакции +7(812) 552-62-16, e-mail: infocom@spbstu.ru

Дата выхода: 30.08.2022

© Санкт-Петербургский политехнический университет Петра Великого, 2022

Contents

Circuits and Systems for Receiving, Transmitting and Signal Processing

Shashikhin V.N., Savchuk S.A. Complete synchronization of chaotic systems at different parameter values 8

Satyshev V.I. Modern approaches to design of multi-channel delta-sigma ADCs 25

Kvashina N.V. Influence of algorithm parameters on static nonlinearity in switching-based calibrations for DACs 32

Simulations of Computer, Telecommunications, Control and Social Systems

Seliverstov Y.A., Komissarov A.A., Lesovodskay A.A., Bovykin P.G., Podtikhov A.V., Torsionov S.S., Tsyrkov D.A., Orlov S.A. Development and research of models of multi-class classifiers for a recommended system for preparing applications on the e-procurement 43

Timofeev S.V., Baenkhayeva A.V. Modeling of information confrontation: Research directions and mathematical tools 63



Содержание

Устройства и системы передачи, приёма и обработки сигналов

Шашихин В.Н., Савчук С.А. Полная синхронизация хаотических систем при различных значениях параметров	8
---	---

Сатышев В.И. Современные подходы к проектированию многоканальных дельта-сигма АЦП	25
--	----

Квашина Н.В. Влияние параметров алгоритма на статическую нелинейность в коммутационных калибровках ЦАП	32
---	----

Моделирование вычислительных, телекоммуникационных, управляющих и социально-экономических систем

Селиверстов Я.А., Комиссаров А.А., Лесоводская А.А., Бовыкин П.Г., Подтихов А.В., Торсионов С.С., Цирков Д.А., Орлов С.А. Разработка и исследование моделей многоклассовых классификаторов для рекомендательной системы подготовки заявок на портале единой информационной системы в сфере закупок	43
---	----

Тимофеев С.В., Баенхаева А.В. Моделирование информационного противоборства: направления исследований и математические инструменты	63
--	----



June is a hard month in St. Petersburg Universities. It is a time of exam sessions and diploma presentations. Peter the Great St. Petersburg Polytechnic University is not an exception. Nevertheless, the University students with high research skills are ready to publish their results and achievements. Thus, the Editorial Board emphasizes an attention of the Journal readers on the papers prepared by students of Electronics and Communications Institute Vladislav I. Satyshev “Modern Approaches to Design of Multi-Channel Delta-Sigma ADCs” and Natalya V. Kvashina “Influence of Algorithm Parameters on Static Nonlinearity in Switching-Based Calibrations for DACs”. The papers are based on results of the student diploma works.

We wish authors to continue the research and to present new articles in our Journal.

Chief in Editor

Prof. Dr. Alexander S. Korotkov

Director of the Electronics and Communications Institute

Circuits and Systems for Receiving, Transmitting and Signal Processing

Устройства и системы передачи, приёма и обработки сигналов

Research article

DOI: <https://doi.org/10.18721/JCSTCS.15201>

UDC 519.8



COMPLETE SYNCHRONIZATION OF CHAOTIC SYSTEMS AT DIFFERENT PARAMETER VALUES

V.N. Shashikhin¹✉, S.A. Savchuk²

^{1,2} Peter the Great St. Petersburg Polytechnic University,
St. Petersburg, Russian Federation

✉ shashikhin@bk.ru

Abstract. The article solves the problem of controlling the mode of complete synchronization of two chaotic systems at different values of their parameters. The control is based on the principle of linear feedback on the phase vector of chaotic generators. The introduction of feedback makes it possible to ensure the equality of the components of the phase vectors of the receiver and the transmitter due to the equality of their Lyapunov characteristic indicators. To change the Lyapunov spectrum, it is proposed to synthesize control by the modal control method based on the solution of the Sylvester linear matrix equation on the basis of the theorem on the topological equivalence of the nonlinear system and its linearized model. An example of using this technique to synchronize the chaotic oscillations of two Lorentz systems when transmitting information using chaotic masking is considered. Computational experiments confirm the operability of the proposed method of ensuring the synchronization of two chaotic systems.

Keywords: synchronization, chaotic generators, hidden data transmission system, chaotic masking, control with feedback

Citation: Shashikhin V.N., Savchuk S.A. Complete synchronization of chaotic systems at different parameter values. Computing, Telecommunications and Control, 2022, Vol. 15, No. 2, Pp. 8–24. DOI: 10.18721/JCSTCS.15201

Научная статья

DOI: <https://doi.org/10.18721/JCSTCS.15201>

УДК 519.8



ПОЛНАЯ СИНХРОНИЗАЦИЯ ХАОСТИЧЕСКИХ СИСТЕМ ПРИ РАЗЛИЧНЫХ ЗНАЧЕНИЯХ ПАРАМЕТРОВ

В.Н. Шашихин¹✉, С.А. Савчук²^{1,2} Санкт-Петербургский политехнический университет Петра Великого,
Санкт-Петербург, Российская Федерация

✉ shashihin@bk.ru

Аннотация. Изучена задача управления режимом полной синхронизации двух систем с хаотической динамикой и странным атрактором при различных значениях их параметров. Управление строится по принципу линейной обратной связи по фазовому вектору хаотических генераторов. Введение обратной связи позволяет обеспечить равенство компонент фазовых векторов приёмника и передатчика при равенстве их характеристических показателей Ляпунова. Для изменения спектра Ляпунова предложено на основе теоремы о топологической эквивалентности нелинейной системы и её линеаризованной модели в гиперболическом случае синтезировать управление методом модального управления на основе решения линейного матричного уравнения Сильвестра. Рассмотрен пример применения данной методики для синхронизации хаотических колебаний двух систем Лоренца при передаче информации с использованием хаотической маскировки. Вычислительные эксперименты подтверждают работоспособность предлагаемого метода на основе равенства одноименных компонент фазовых векторов передатчика и приёмника и соответствующих значений критериев полной синхронизации двух хаотических систем.

Ключевые слова: синхронизация, хаотические генераторы, система скрытой передачи данных, хаотическое маскирование, управление с обратной связью

Для цитирования: Shashikhin V.N., Savchuk S.A. Complete synchronization of chaotic systems at different parameter values // Computing, Telecommunications and Control. 2022. T. 15, № 2. C. 8–24. DOI: 10.18721/JCSTCS.15201

Introduction

At present, there is an increased interest of Russian and foreign authors in theoretical research and practical application of nonlinear systems with chaotic modes, which is caused by the widespread introduction of information technologies in various spheres of civil and military application [1–4]. One of the important problems in the creation of information systems is the problem of ensuring reliable and confidential communication [5].

One of the promising directions in the development of telecommunication systems is the use of deterministic chaos [6, 7], which allows providing a high level of protection and a wide band of the carrier with its large information capacity [8, 9]. The main issues in the construction of telecommunication systems based on chaotic generators are issues related to the synchronization of chaotic generators. To solve these problems, various types of synchronization are used: generalized synchronization [10–12], frequency synchronization [13], and phase synchronization [13–15].

Among the various systems of covert information transmission [16–19], systems with chaotic masking based on full synchronization have become widespread [20, 21]. To implement full synchronization, various approaches are used on the basis of rough chaotic systems [22, 23], hyperchaotic systems [24, 25], and the introduction of mutual influence between two chaotic generators [26, 27].

The paper proposes a method to ensure full synchronization of chaotic generators with different values of parameters caused by technological reasons. The considered method is based on the management of the spectrum of characteristic indicators of Lyapunov chaotic generators [28, 29].

The problem of complete synchronization of two generators with chaotic dynamics

Mathematical model of chaotic generators. Let the dynamics of a system of hidden information transmission, which consists of a transmitter, a receiver, and a communication line, be described by two autonomous families of ordinary differential equations

$$\dot{X}(t) = dX(t)/dt = F(X(t), \alpha, u^1(t)) + G(Z(t), X(t)), \quad X(0) = X_0, \quad (1)$$

$$\dot{Z}(t) = dZ(t)/dt = F(Z(t), \beta, u^2(t)) + G(X(t), Z(t)), \quad Z(0) = Z_0, \quad (2)$$

where $X(t) \in \mathbb{R}^n$ is the vector of the transmitter state, $Z(t) \in \mathbb{R}^n$ is the vector of the receiver state, $\alpha \in \mathbb{R}^s$ is the vector of the transmitter parameters, $\beta \in \mathbb{R}^s$ is the vector of the receiver parameters, $u^1(t), u^2(t) \in \mathbb{R}^m$ are the vectors of the transmitter and receiver control, $m \leq n$, $F: \mathbb{R}^n \otimes \mathbb{R}^s \otimes \mathbb{R}^m \rightarrow \mathbb{R}^n$ is a vector function that determines the dynamics of the transmitter and receiver behavior, $G: \mathbb{R}^n \otimes \mathbb{R}^n \rightarrow \mathbb{R}^n$ is a vector function that determines the nature of the communication between the transmitter and the receiver. The function G depends linearly on the state vector of the transmitter and receiver.

The vector function F has a given smoothness class by vector arguments $X(t)$, $Z(t)$, and $u^i(t)$, that is $F \in \mathbb{C}_X^r$, $F \in \mathbb{C}_Z^r$, and $F \in \mathbb{C}_{u^i}^r$, or satisfies the Lipschitz condition

$$\|F(X') - F(X'')\| \leq k \|X' - X''\|, \quad \|F(Z') - F(Z'')\| \leq k \|Z' - Z''\|, \quad k > 0.$$

The vector function F :

- 1) is unstable in relation to the setting of initial conditions – there is a value δ , such that for an arbitrary point $X \in \mathbb{R}^n$ and $\varepsilon > 0$, there is a point $Y \in \mathbb{R}^n$ for which the condition $\text{dist}[X(t_0), Y(t_0)] < \varepsilon$ follows $\text{dist}[X(t), Y(t)] \geq \delta$ for some $t > t_0$, where $\text{dist}[\ast, \ast]$ is the distance;
- 2) is topologically transitive – for any two open sets N, M , there is such l that $F^l(N) \cap M \neq \emptyset$;
- 3) has the element of regularity or otherwise density of periodic trajectories – in any vicinity of any point of phase space there is at least one and, therefore, infinitely many periodic trajectories.

These three conditions determine the presence of chaotic dynamics in systems (1) and (2) at certain values of their parameters.

Criteria for complete synchronization of chaotic systems. Complete synchronization of chaotic systems occurs when the evolution in time of the states of the interacting chaotic generators proceeds in the same way after the completion of the transient process. In this case, the equality of the similar components of the vectors of the transmitter and receiver state is ensured:

$$x_1(t) = z_1(t), \dots, x_i(t) = z_i(t), \dots, x_n(t) = z_n(t). \quad (3)$$

Various criteria are used to assess the full synchronization mode [30].

Synchronization error for each time point

$$x_1(t) - z_1(t), \dots, x_i(t) - z_i(t), \dots, x_n(t) - z_n(t). \quad (4)$$

The integral error



$$e = \int_0^{\infty} \|X(t) - Z(t)\| dt. \quad (5)$$

The similarity function

$$G^2(\tau) = \frac{\langle (z_i(t + \tau) - x_i(t))^2 \rangle}{\sqrt{(\langle x_i^2(t) \rangle \langle z_i^2(t) \rangle)}}, \quad (6)$$

where $x_i(t)$ and $z_i(t)$ are the same-name components of the phase vectors of the transmitter and receiver. In the case of full synchronization, the similarity function is zero.

Evaluation of the degree of synchronization

$$\nu_i = \frac{\sigma_{x_i+z_i}^2}{2(\sigma_{x_i}^2 + \sigma_{z_i}^2)}, \dots i = \overline{1, n}, \quad (7)$$

where $x_i(t)$ and $z_i(t)$ are the same-name state variables of the transmitter and receiver generators, $\sigma_{x_i}^2$ and $\sigma_{z_i}^2$ are the variance of the variables $x_i(t)$ and $z_i(t)$, $\sigma_{x_i+z_i}$ is the variance of their sum. The estimation of the degree of synchronization can take values $\nu_i \in [0.5; 1]$, while $\nu_i = 1$ corresponds to the case of complete synchronization, and with the independence of the processes $x_i(t)$ and $z_i(t)$, we have .

The cross-correlation coefficient

$$\rho(x_i, z_i) = \frac{\sum_{j=1}^K (x_{ij} - \bar{X})(z_{ij} - \bar{Z})}{\sqrt{\frac{\sum_{j=1}^K (x_{ij} - \bar{X})^2}{K}} \sqrt{\frac{\sum_{j=1}^K (z_{ij} - \bar{Z})^2}{K}}}, \quad (8)$$

$$\bar{X} = \frac{\sum_{j=1}^K x_{ij}}{K}, \quad \bar{Z} = \frac{\sum_{j=1}^K z_{ij}}{K},$$

where x_{ij}, z_{ij} are the values of the similar components of the phase vectors of the transmitter and receiver at time points $j = j_o, j_k$; $X = \{x_1, x_2, \dots, x_j, \dots\}$, $Z = \{z_1, z_2, \dots, z_j, \dots\}$ are a number of values of the components of the phase vectors of the transmitter and receiver.

The cross-correlation coefficient is equal to one when fully synchronized and tends to zero when there is no statistical relationship between the processes in the transmitter and receiver.

Full synchronization management task. At the same values of the parameters of chaotic generators, a full synchronization mode is observed. When changing the values of the parameters of one of the generators, desynchronization occurs and the identity of chaotic oscillations is lost, which disrupts the process of data transmission in the communication system built on the basis of these generators.

In this situation, the task arises of synthesizing control that restores the identity of chaotic oscillations $X(t) \equiv Z(t)$, $t > t_0$ at different values of transmitter and receiver parameters $\alpha \neq \beta$.

It is necessary to synthesize control in the form of linear stationary feedback on the phase vector of the transmitter and receiver when changing the parameters

$$u^1(t) = L^1 X(t); \quad u^2(t) = L^2 Z(t). \quad (9)$$

Feedback coefficients $L^i, i=1,2$ (9) are calculated based on the condition of equality of the spectrum of Lyapunov characteristic exponents of the receiver chaotic generator

$$\Omega(F, Z, \beta, u^2) = \left\{ \chi_1, \dots, \chi_i, \dots, \chi_n : \chi_1 > \chi_i \quad \forall i = \overline{2, n} \right\} \quad (10)$$

to the Lyapunov spectrum of the transmitter

$$\Omega(F, X, \alpha, u^1). \quad (11)$$

The equality of Lyapunov spectra makes it possible to ensure the identity of chaotic oscillations of generators and to ensure complete synchronization, which can be checked according to criteria (3)–(8).

Synthesis of the Lyapunov characteristic exponents spectrum control

Topological equivalence of nonlinear systems. The feedback control synthesis technique is based on the topological equivalence of the behavior of nonlinear systems to the behavior of a linearized system. The systems of nonlinear autonomous differential equations (1), (2) correspond to a vector field F . According to the Hartman–Grobman theorem [31, 32], a continuously differentiable vector field with a hyperbolic special point in some vicinity of this point is topologically equivalent to its linear part.

From this theorem, it follows that the qualitative behavior of the solutions to an autonomous system of nonlinear differential equations (1), (2) at a hyperbolic special point is completely determined by the behavior of the solutions to a system of linear differential equations with a constant operator (Jacobi matrix) of the linear part of the field at this point. At the hyperbolic special point of the nonlinear system, no Jacobi matrix eigenvalue lies on the imaginary axis.

The nature of the behavior of the solutions to linear differential equations corresponding to the nonlinear system is determined by the intrinsic values of the Jacobi matrix, the real parts of which are associated with its Lyapunov characteristic exponents. A characteristic exponent of a function $y(t)$ is a value that has a finite value or value $\pm\infty$ and is defined as

$$\lambda(y) \equiv \overline{\lim}_{t \rightarrow \infty} (t^{-1} \ln \|y(t)\|).$$

The characteristic exponent determines the change in the function according to the scale of indicative functions.

The nature of the behavior of solutions to nonlinear differential equations is determined by the Lyapunov spectrum. Negative characteristic exponents correspond to regular dynamics, in the presence of positive, negative and zero characteristic exponents in the Lyapunov spectrum, there is a chaotic dynamic.

Using one of the modal control methods, it is possible to provide the required spectrum of a linearized system, and by virtue of the theorem of topological equivalence – also the corresponding spectrum of a nonlinear system.

Linearization of a nonlinear system. Using the Taylor formula in the assumption of the smoothness of the vector function F by vector arguments $X(t), Z(t), u^i(t)$ ($F \in \mathbb{C}_X^r, F \in \mathbb{C}_{u^i}^r$, and $F \in \mathbb{C}_{u^i}^r$) in the vicinity of a special point and limited to the linear term of the Taylor series, we convert equation (1) and (2) to the quasi-linear form:



$$\dot{X}(t) = J(X^s)X(t) + f(X^s), \quad X(0) = X_0, \quad (12)$$

$$\dot{Z}(t) = J(Z^s)Z(t) + f(Z^s), \quad Z(0) = Z_0. \quad (13)$$

In equations (12), (13), the Jacobian matrices $J(X^s)$ and $J(Z^s)$ are calculated at singular points X^s , Z^s of nonlinear systems (1) and (2) according to the following formulas:

$$J(X^s) = \begin{bmatrix} \partial f_1 / \partial x_1 & \dots & \partial f_1 / \partial x_n \\ \dots & \dots & \dots \\ \partial f_n / \partial x_1 & \dots & \partial f_n / \partial x_n \end{bmatrix}_{X(t)=X^s}, \quad (14)$$

$$J(Z^s) = \begin{bmatrix} \partial f_1 / \partial z_1 & \dots & \partial f_1 / \partial z_n \\ \dots & \dots & \dots \\ \partial f_n / \partial z_1 & \dots & \partial f_n / \partial z_n \end{bmatrix}_{Z(t)=Z^s}. \quad (15)$$

If all singular points are evaluated

$$\|f(\xi^s)\| \leq q \|\xi\|,$$

and Jacobian matrices are calculated by formulas (14) and (15), then equations (12) and (13) will take the form of linearized systems (or equations in variations):

$$\dot{Y}(t) = J(X^s)Y(t), \quad (16)$$

$$\dot{W}(t) = J(Z^s)W(t). \quad (17)$$

Systems (16) and (17) can be used to synthesize control of nonlinear systems (1) and (2) in order to fully synchronize chaotic oscillations of receiver and transmitter generators.

Synthesis of control of complete synchronization of chaotic systems. Control of nonlinear systems (1) and (2) by introducing feedback consists in changing the spectrum of Lyapunov characteristic exponents to achieve the desired result – full synchronization of two chaotic systems at different values of parameters.

Using the modal control method based on the solution of the matrix algebraic Sylvester equation for a linearized system of a chaotic generator

$$\dot{W}(t) = J(Z^s)W(t) + Bu(t), \quad (18)$$

control in the form of static linear feedback

$$u(t) = LW(t) \quad (19)$$

is derived from the condition of equality of the spectrum of the closed system (18), (19)



$$\dot{W}(t) = J(Z^S)W(t) + BLW(t) = A_Z W(t), \quad (20)$$

$$A_Z = J(Z^S) + BL \quad (21)$$

to the required spectrum

$$\rho(A_Z) = \rho(\Phi).$$

Equation (18) is a linearized model of the nonlinear part of systems (1) and (2), according to which the decentralized feedback (19) is calculated from the state vector of the transmitter and receiver. Equations (18)–(20) are written in the coordinates of a linearized system and do not contain a mutual influence function G .

Let us represent the matrix (20) of a closed system by the expansion in the basis of its eigenvectors:

$$A_Z = S\Phi S^{-1}, \quad (22)$$

where Φ is the Jordan form of the matrix A_Z ; S is the matrix of eigenvectors of the matrix A_Z . Then substituting (22) into (21) and multiplying by S , we get:

$$J(Z^S)S - S\Phi = -BLS. \quad (23)$$

Let us eliminate the nonlinear component in the right-hand side of equation (23) by introducing a matrix factor $V = -LS$, then the equation (23) will take the form:

$$J(Z^S)S - S\Phi = BV. \quad (24)$$

Equation (24) is the Sylvester matrix equation. It is linear with respect to S and solvable if the following conditions are met [33]:

- 1) full rank matrix V ;
- 2) $\text{rank } D_u = n$, where $D_u = (B | J(Z^S)B | \dots | J(Z^S)^{n-1} B) \in R^{n \times mn}$;
- 3) $\text{rank } H_l = n$, where $H_l = (V^T | \Phi^T V^T | \dots | (\Phi^T)^{n-1} V^T) \in R^{n \times mn}$;
- 4) the spectra of matrices $J(Z^S)$ and Φ do not intersect;
- 5) the eigenvalues of the matrix Φ are pairwise different.

Having solved equation (24) with respect to the matrix S , one can calculate the matrix of feedback coefficients:

$$L = -VS^{-1}. \quad (25)$$

The matrix L found by formula (25) is used to control the Lyapunov spectrum of a chaotic generator on the transmitting and receiving sides:

$$\dot{X}(t) = dX(t)/dt = F(X(t), \alpha) + BL^1 X(t), \quad X(0) = X_0, \quad (26a)$$

$$\dot{Z}(t) = dZ(t)/dt = F(Z(t), \beta) + BL^2 Z(t), \quad Z(0) = Z_0, \quad (26b)$$



which ensures that receiver and transmitter performance is equal for complete synchronization.

Investigation of the information transmission system using chaotic masking

Examination of the system at the same values of transmitter and receiver parameters. The Lorentz system is considered as chaotic generators (1), (2) of the transmitter and receiver for the system with chaotic masking [34]:

$$\begin{cases} \dot{x}_1 = \sigma(x_2 - x_1), \\ \dot{x}_2 = x_1(r - x_3) - x_2, \\ \dot{x}_3 = x_1x_2 - bx_3 \end{cases} \quad (27)$$

with parameters $\alpha = (\alpha_1 = \sigma = 12; \alpha_2 = b = 2; \alpha_3 = r = 45)$.

System (27) has three singular points:

$$\begin{aligned} X^{s=1} &= (x_1 = 0; x_2 = 0; x_3 = 0)^T, \\ X^{s=2} &= (x_1 = -9.3808; x_2 = -9.3808; x_3 = 44.0000)^T, \\ X^{s=3} &= (x_1 = 9.3808; x_2 = 9.3808; x_3 = 44.000)^T \end{aligned}$$

and a Jacobian matrix,

$$J = \begin{bmatrix} -12 & 12 & 0 \\ 45 - x_3 & -1 & -x_1 \\ x_2 & x_1 & -2 \end{bmatrix},$$

the eigenvalues of which, calculated at the specified parameters at special points, are equal to

$$\rho\{\lambda_i J(X^{s=1}), i = \overline{1,3}\} = \begin{cases} \lambda_1 = -30.3799, \\ \lambda_2 = 7.3799, \\ \lambda_3 = -2.0000. \end{cases}$$

$$\rho\{\lambda_i J(X^{s=2}), i = \overline{1,3}\} = \begin{cases} \lambda_1 = -16.0791 + 0.0000j, \\ \lambda_2 = 0.5395 + 11.4481j, \\ \lambda_3 = 0.5395 - 11.4481j. \end{cases}$$

The Lyapunov spectrum of system (27)

$$\Omega(F, X, \alpha) = \{\chi_1 = 0.6283; \chi_2 = 0.4156; \chi_3 = -16.0441\}$$

contains a positive characteristic indicator, and the phase volume of the system is compressed, so a chaotic mode occurs in the Lorentz system at the specified parameters. The phase portrait of the system is shown in Fig. 1.

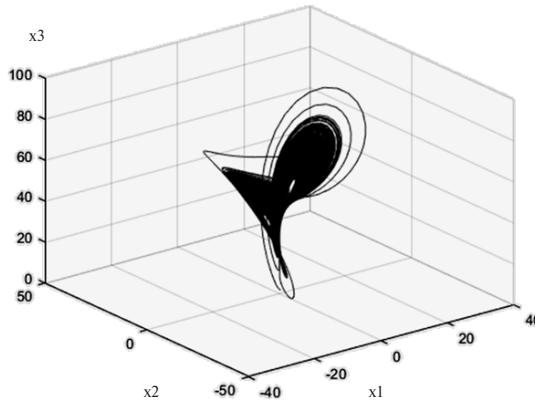


Fig. 1. Phase portrait of the Lorentz system

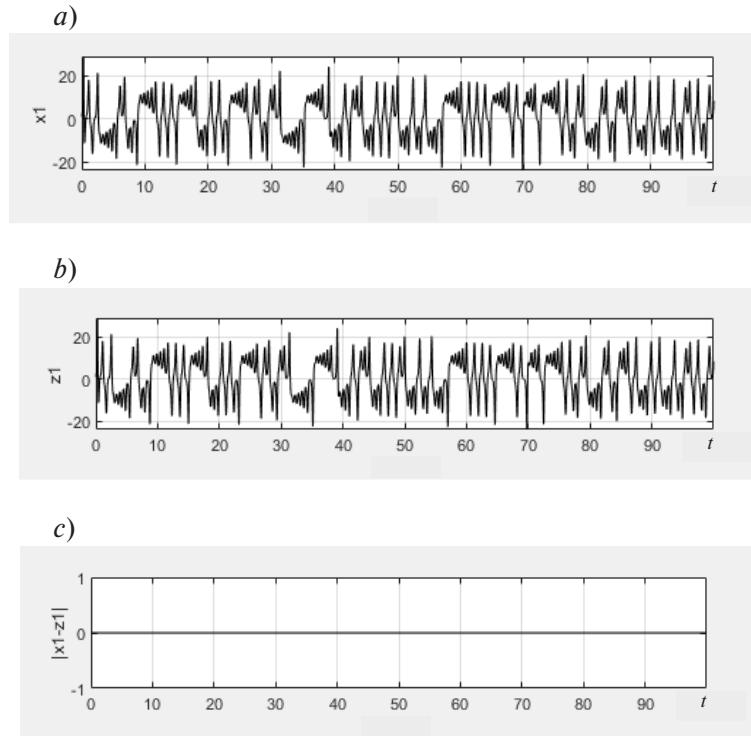


Fig. 2. Time diagrams with the same parameters of chaotic systems:
 a – transmitter component x_1 ; b – receiver component z_1 ; c – component difference ($x_1 - z_1$)

Fig. 2 shows time diagrams of the first components of the phase vector of the transmitter and receiver and their difference at the same values of parameters.

Since the parameters of the transmitter and receiver are equal ($\alpha = \beta$), all the given characteristics are equal. Therefore, there is a complete synchronization of chaotic oscillations in the system. The values of the complete synchronization criteria are given in Table 1.

Examination of the system at different values of transmitter and receiver parameters. At different parameter values, the identity of the phase vectors is violated, which leads to the desynchronization of the receiver and transmitter. This effect is considered when the receiver parameter values are changed:



$$\begin{aligned}\beta = & (\beta_1 = \sigma = 12.01; \beta_2 = b = 2.01; \beta_3 = r = 45) \neq \\ & \neq \alpha = (\alpha_1 = \sigma = 12.00; \alpha_2 = b = 2.00; \alpha_3 = r = 45).\end{aligned}$$

With the changed values of the parameters, the Lorentz system on the receiving side

$$\begin{aligned}\dot{z}_1 &= -z_2 - z_3 \\ \dot{z}_2 &= z_1 + az_2 \\ \dot{z}_3 &= b + z_3(z_1 - c),\end{aligned}$$

has the following characteristics:

- three singular points

$$\begin{aligned}X^{s=1} &= (x_1 = 0; x_2 = 0; x_3 = 0)^T, \\ X^{s=2} &= (x_1 = -9.4043; x_2 = -9.4043; x_3 = 44.0000)^T, \\ X^{s=3} &= (x_1 = 9.4043; x_2 = 9.4043; x_3 = 44.0000)^T;\end{aligned}$$

- Jacobian matrix eigenvalues

$$\rho\{\lambda_i J(X^{s=1}), i = \overline{1,3}\} = \begin{cases} \lambda_1 = -30.3955, \\ \lambda_2 = 17.3855, \\ \lambda_3 = -2.0100; \end{cases}$$

$$\rho\{\lambda_i J(X^{s=2}), i = \overline{1,3}\} = \begin{cases} \lambda_1 = -16.0987 + 0.0000j, \\ \lambda_2 = +0.5394 + 11.4746j, \\ \lambda_3 = +0.5394 - 11.4746j; \end{cases}$$

- Lyapunov spectrum

$$\Omega(F, X, \alpha) = \{\chi_1 = 0.6284; \chi_2 = 0.4156; \chi_3 = -16.0441\}.$$

Fig. 3 shows time diagrams of the first components of the phase vector of the transmitter and receiver and their difference at different values of parameters.

There is a significant difference between the components of the same phase vectors in the system, and data transmission becomes impossible. The values of the complete synchronization criteria (actually de-synchronization) are given in Table 1.

Synchronization when there is mutual influence. To restore synchronization between the transmitter and the receiver, the mutual influence between the subsystems is introduced [35]. The system of two Lorentz oscillators with different parameters and with the addition of a coupling coefficient is described by equations of the form:

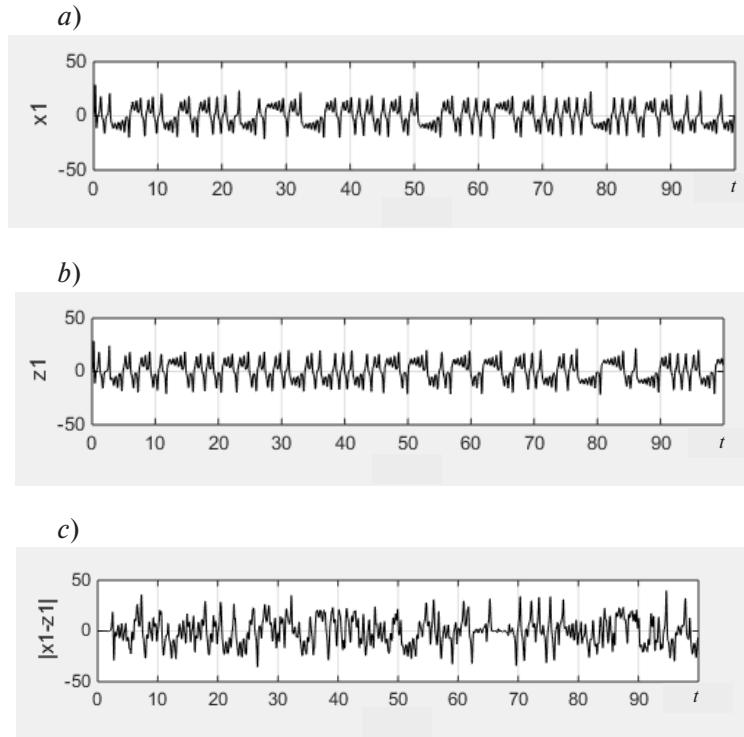


Fig. 3. Time diagrams with different parameters of chaotic systems:
 a – transmitter component x_1 ; b – receiver component z_1 ; c – component difference $(x_1 - z_1)$

$$\begin{cases} \dot{x}_1 = \sigma(x_2 - x_1) + \gamma(z_1 - x_1), \\ \dot{x}_2 = x_1(r - x_3) - x_2, \\ \dot{x}_3 = x_1x_2 - bx_3, \\ \dot{z}_1 = \sigma_2(z_2 - z_1) + \gamma(x_1 - z_1) \\ \dot{z}_2 = z_1(r_2 - z_3) - z_2, \\ \dot{z}_3 = z_1z_2 - b_2z_3. \end{cases} \quad (28)$$

The mutual influence function is determined by the difference of the first coordinates of the transmitter and receiver multiplied by a constant coefficient. With the parameter of interconnection $\gamma = 10$, the system (28) has the following Lyapunov characteristic exponents:

$$\Omega(F, X, Z, \alpha, \beta, \gamma) = \{\chi_1 = 0.8520; \chi_2 = 0.4764; \chi_3 = -0.2983; \chi_4 = -0.5547; \chi_5 = -15.8508; \chi_6 = -33.6646\}.$$

Fig. 4 shows time diagrams of the first components of the phase vector of the transmitter and receiver and their difference at different values of parameters and the presence of mutual connection between them.

The difference between the similar components of the phase vectors of the transmitter and the receiver is significantly reduced, which indicates the restoration of synchronization. The latter confirms the values of the synchronization criteria given in Table 1.

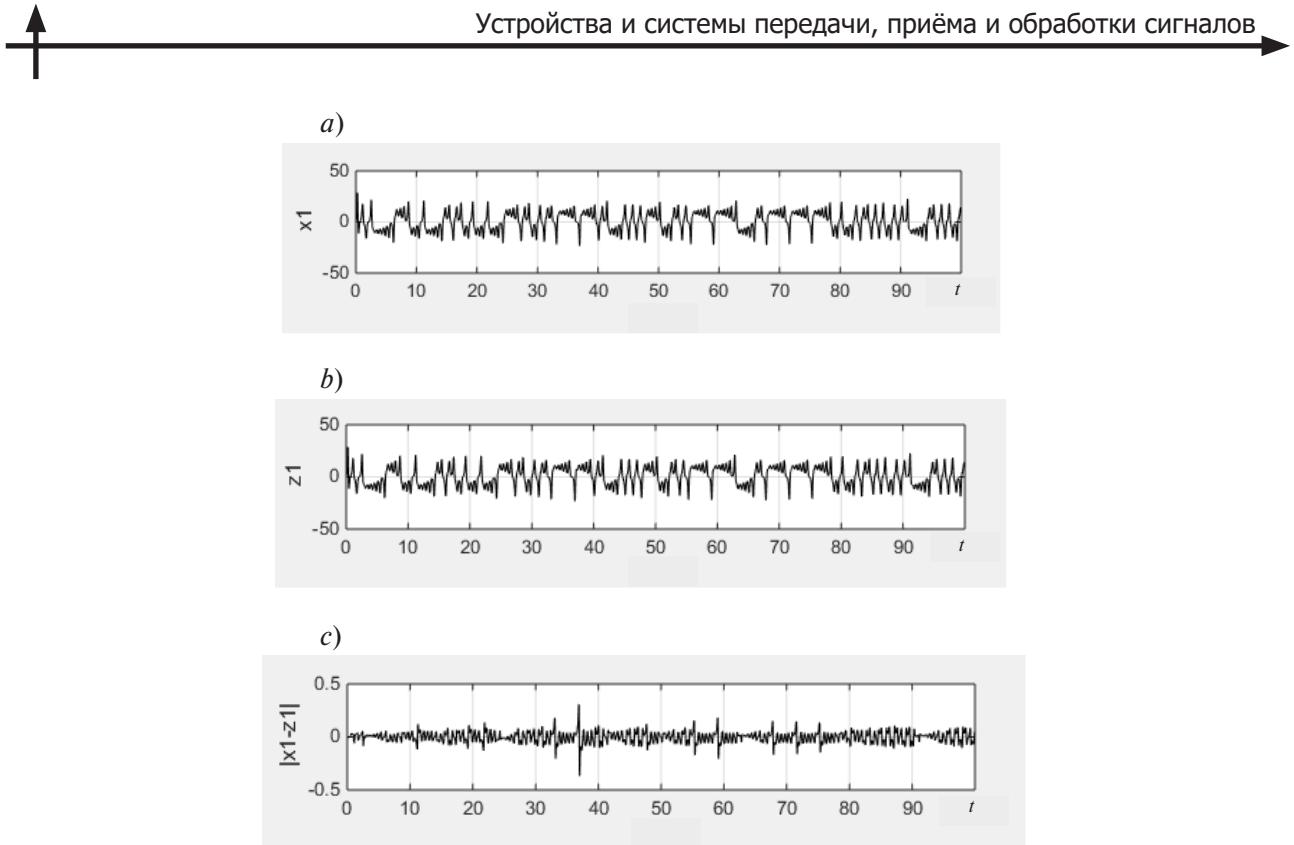


Fig. 4. Time diagrams with different parameters of chaotic systems and their interconnection:
 a – transmitter component x_1 ; b – receiver component z_1 ; c – component difference $(x_1 - z_1)$

The complete synchronization in the presence of mutual influence and management of the spectrum of characteristic exponents. To ensure the full synchronization mode of two chaotic generators at different values of parameters, feedback (19) is introduced into the Lorentz system of the transmitter and receiver and the system of autonomous differential equations (26) or (28) takes the form:

$$\begin{cases} \dot{x}_1 = \sigma(x_2 - x_1) - (l_{11}^1 x_1 + l_{12}^1 x_2 + l_{13}^1 x_3) + \gamma(z_1 - x_1), \\ \dot{x}_2 = x_1(r - x_3) - x_2 - (l_{21}^1 x_1 + l_{22}^1 x_2 + l_{23}^1 x_3), \\ \dot{x}_3 = x_1 x_2 - b x_3 - (l_{31}^1 x_1 + l_{32}^1 x_2 + l_{33}^1 x_3), \\ \dot{z}_1 = \sigma_2(z_2 - z_1) - (l_{11}^2 z_1 + l_{12}^2 z_2 + l_{13}^2 z_3) + \gamma(x_1 - z_1), \\ \dot{z}_2 = z_1(r_2 - z_3) - z_2 - (l_{21}^2 z_1 + l_{22}^2 z_2 + l_{23}^2 z_3), \\ \dot{z}_3 = z_1 z_2 - b_2 z_3 - (l_{31}^2 z_1 + l_{32}^2 z_2 + l_{33}^2 z_3). \end{cases} \quad (29)$$

Feedback coefficients L^i , $i = 1, 2$ are derived according to formula (25) and are equal to:

$$L^1 = \begin{bmatrix} 0.1203 & -0.0489 & 0 \\ -0.1835 & -0.1203 & 0 \\ 0 & 0 & 0.5 \end{bmatrix}; \quad L^2 = \begin{bmatrix} 0.1047 & -0.0426 & 0 \\ -0.1749 & -0.1147 & 0 \\ 0 & 0 & 0.49 \end{bmatrix}. \quad (30)$$

In formula (25), the matrix S is the solution to the Sylvester equation (24), $J^1(Z^S)$, $J^2(Z^S)$ is the solution to the Jacobian matrix of linearized models of the transmitter and receiver, $\Phi = \text{diag}\{\chi_1 = 0.0709; \chi_2 = 0.0184; \chi_3 = -5.4453\}$ is the matrix of the required eigenvalues of the closed linearized system of the form (20).

The feedback system (29), (30) has the following Lyapunov characteristic exponents:

$$\begin{aligned} \Omega(F, X, Z, \sigma^1, \sigma^2, b^1, b^2 r) = \\ = \{\chi_1 = 0.8704; \chi_2 = 0.3316; \chi_3 = -16.7021; \\ \chi_4 = 0.8157; \chi_5 = 0.3706; \chi_6 = -16.6863\}. \end{aligned}$$

Fig. 5 shows time diagrams of the first components of the phase vector of the transmitter and receiver and their difference with different parameters, the presence of mutual influence and control of the Lyapunov spectrum.

Thus, the control of the spectrum of Lyapunov characteristic exponents makes it possible to significantly reduce the difference between the similar components of the phase vectors of the transmitter and receiver, and the system is in the complete synchronization mode. The values of the full synchronization criteria, for this case, are given in Table 1.

Table 1
Criteria for complete synchronization of chaotic systems

Characteristics of the data transmission system	Criteria for complete synchronization								
	Dispersion coefficient			Correlation coefficient			Similarity function		
	$v_{x1,z1}$	$v_{x2,z2}$	$v_{x3,z3}$	$R_{x1,z1}$	$R_{x2,z2}$	$R_{x3,z3}$	$G_{x1,z1}$	$G_{x2,z2}$	$G_{x3,z3}$
Same transmitter and receiver parameters	1	1	1	1	1	1	0	0	0
Different transmitter and receiver parameters	0.4977	0.5020	0.4664	0.0047	0.0041	0.0673	0.0112	0.0106	0.0025
Existence of mutual influence	1	0.9999	0.9999	1	0.9999	0.9998	$1.5 \cdot 10^{-5}$	$3.5 \cdot 10^{-5}$	0.0012
Existence of mutual influence and control of Lyapunov spectrum	1	1	1	1	1	1	$3.3 \cdot 10^{-6}$	$1.3 \cdot 10^{-5}$	$6.1 \cdot 10^{-4}$

In data transmission systems with chaotic masking, the signal at the transmitter output $S(t)$ is formed as a sum of transmitted information $m(t)$ and the value of the first component of the phase vector of the chaotic oscillator $x_1(t)$, that is $S(t) = m(t) + x_1(t)$. The signal at the output of the system $m(t)$ represents the difference between the signal transmitted over the communication line $S(t)$ and the value of the first component of the phase vector of the chaotic generator of the receiver $z_1(t)$, i.e. $m(t) = S(t) - z_1(t)$. Fig. 6 shows time diagrams of the information transmission system with chaotic masking.

Thus, with equal values of transmitter and receiver parameters, there is a complete synchronization of chaotic generators in the system. With different parameters, the process becomes out of sync. The introduction of mutual influence and feedback into the system allows restoring complete synchronization of chaotic signals with different parameters of the transmitter and receiver.

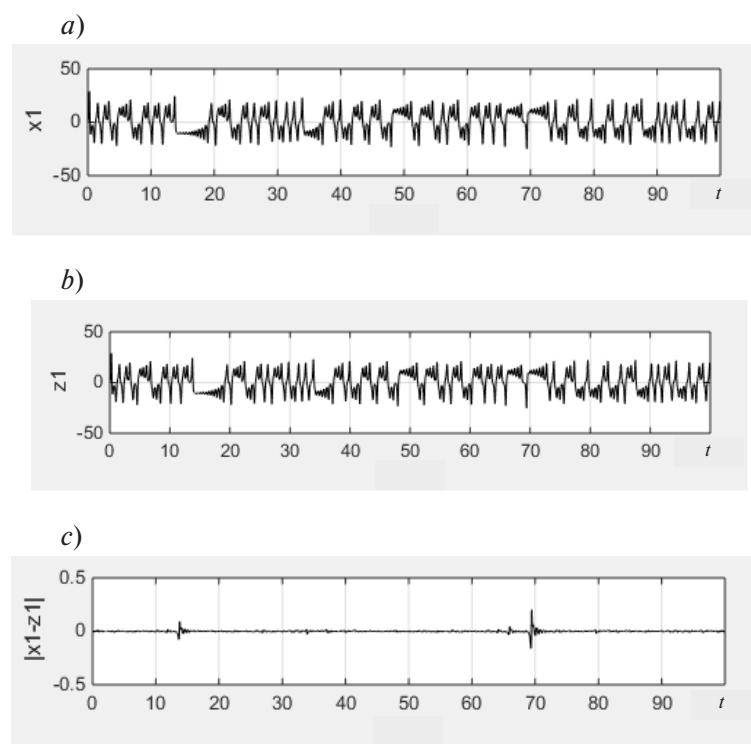


Fig. 5. Time diagrams with different parameters, mutual influence and control of Lyapunov spectrum:
a – transmitter component x_1 ; b – receiver component z_1 ; c – component difference ($x_1 - z_1$)

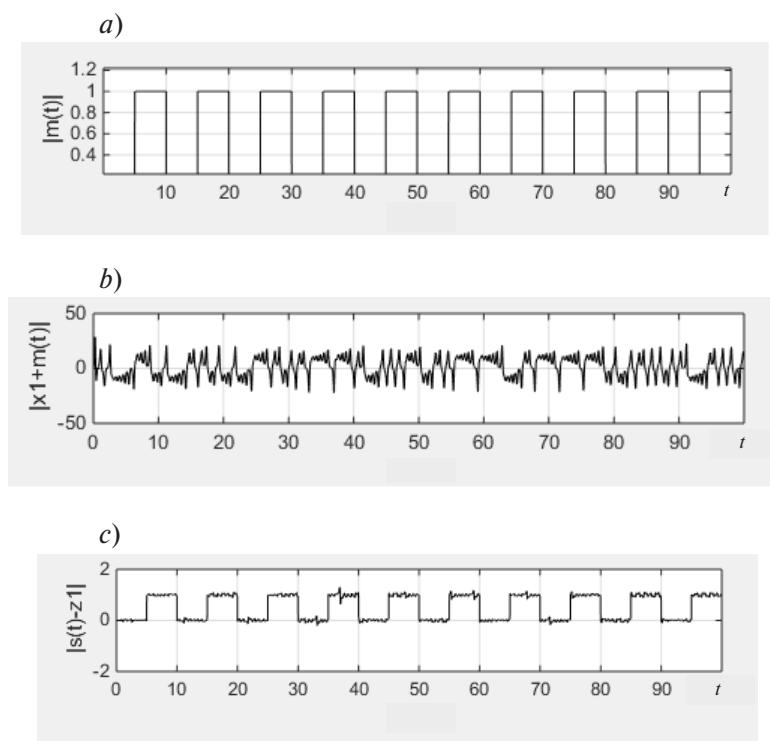


Fig. 6. Time diagrams in the transmission system: a – signal at the input of the system;
b – signal in the communication line; c – signal at the output of the system

Conclusion

A control synthesis technique for complete synchronization of chaotic oscillations in a nonlinear system using phase vector feedback is proposed. The feedback coefficient ensures the equality of the spectra of Lyapunov characteristic exponents at different values of the parameters of chaotic generators.

The possibility of using the proposed method of feedback synthesis by the example of a system consisting of two Lorentz systems has been investigated. The results of computational experiments confirmed the complete synchronization of chaotic oscillations in the information transmission system with chaotic masking.

The advantage of the proposed method of forming a complete synchronization of two chaotic systems, which combines the introduction of the mutual influence of subsystems and control of the Lyapunov spectrum, lies in the possibility of transmitting a useful signal of lower amplitude, which could make the communication line stealthier. In addition, a full synchronization mode is provided with different parameters of the transmitter and receiver and mutual influence in a wider range.

REFERENCES

1. Andriyevskiy B.R., Fradkov A.L. Control of chaos: Methods and applications. II. Applications. *Autom. Remote Control*, 2004, no. 4, Pp. 505–533. (rus)
2. Vahdati P.M., Kazemi A. Bifurcations and chaos in nonlinear dynamics power systems. *24th Iranian Conference on Electrical Engineering (ICEE)*, 2016, Pp. 1706–1711. DOI: 10.1109/IranianCEE.2016.7585796
3. Luongo A., Leamy M.J. Advances in stability, bifurcations and nonlinear vibrations in mechanical systems. *Nonlinear Dynamics*, 2021, Vol. 103, no. 4, Pp. 2993–2995.
4. Shashikhin V.N., Turulin A.V., Budnik S.V. Image encryption algorithm based on controlled chaotic maps. *Computing, Telecommunication and Control*, 2021, Vol. 14, no. 1, Pp. 7–21. DOI: 10.18721/JC-STCS.14101
5. Koronovskiy A.A., Moskalenko O.I. On the use of chaotic synchronization for secure communication. *Phys. Usp.*, 2009, Vol. 179, Pp. 1281–1310. (rus)
6. Dmitriev A.S., Panas A.I. *Dinamicheskiy khaos: novyye nositeli informatsii dlya sistem svyazi* [Dynamic chaos: New media for communication systems]. Moscow: Fizmatlit Publ., 2002. 252 p. (rus)
7. Abel A., Schwarz W. Chaos communication – principles, schemes, and system analysis *Proc. IEEE*, 2002, Vol. 90, no. 5, Pp. 691–710.
8. Baptista M.S., Macau E.E., Grebogi C. Conditions for efficient chaos-based communication. *Chaos*, 2003, Vol. 13, no. 1, Pp. 145–150.
9. Ivanyuk P.V., Politanskiy L.F., Eliyashiv O.M. Chaotic masking of information signals using generator based on the Liu system. *Tekhnologiya i Konstruirovaniye v Elektronnoi Apparature*, 2012, no. 3, Pp. 11–17. (rus)
10. Terry J.R., Van Wiggezen G.D. Chaotic communication using generalized synchronization. *Chaos, Soliton & Fractals*, 2001, Vol. 12, no. 1, Pp. 145–152.
11. Koronovskiy A.A., Moskalenko O.I., Popov P.V., Khramov A.Ye. Method for secure data transmission based on generalized synchronization. *Bulletin of the Russian Academy of Sciences: Physics*, 2008, Vol. 72, no. 1, Pp. 143–147. (rus)
12. Hramov A.E., Koronovskii A.A. An approach to chaotic synchronization. *Chaos*, 2004, Vol. 14, no. 3, Pp. 603–610.
13. Khodulin A.V., Kazakov L.N. Cascaded systems of phase-locked loop chaotic synchronization. *Model. Anal. Inform. Sist.*, 2009, Vol. 16, no. 4, Pp. 117–131. (rus)
14. Zakharova A.S., Vadivasova T.Ye., Anishchenko V.S. The interconnection of synchronization threshold with effective diffusion coefficient of instantaneous phase of chaotic self-sustained oscillations. *Nelin. Dinam.*, 2008, Vol. 4, no. 2, Pp. 169–180. (rus)



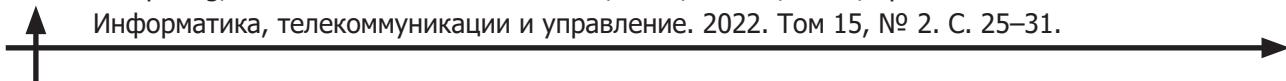
15. Matrosov V.V., Shalfeev V.D. Synchronization of chaotically modulated oscillations. *Vestnik of Lobachevsky University of Nizhni Novgorod*, 2011, Vol. 3, no. 5, Pp. 292–299. (rus)
16. Femat R., Jauregui-Ortiz R., Solis-Perales G. A chaos-based communication scheme via robust asymptotic feedback. *IEEE Trans. Circ. Syst.*, 2001, Vol. 48, no. 1, Pp. 1161–1169.
17. Feki M. An adaptive chaos synchronization scheme applied to secure communication. *Chaos, Solitons and Fractals*, 2003, Vol. 18, Pp. 141–148.
18. Abel A., Schwarz W. Chaos communication – principles, schemes, and system analysis *Proc. IEEE*, 2002, Vol. 90, no. 5, Pp. 691–710.
19. Agureyev K.I. Применение детерминированного хаоса для передачи информации [Application of deterministic chaos for information transmission]. *Izvestiya TulGU. Tekhnicheskiye Nauki*, 2014, Vol. 11, Pp. 197–212. (rus)
20. Belim S.V., Rakitskiy Yu.S. Using families of orthogonal functions for coding messages in a chaotic masking scheme. *Information and Control Systems*, 2019, no. 2, Pp. 76–82. DOI: 10.31799/1684-88-53-2019-2-76-82
21. Verichev N.N. Chaotic synchronization: 30 years later. *Vestnik Nauchno-Tekhnicheskogo Razvitiya [Bulletin of Scientific and Technical Development]*, 2018, Vol. 127, no. 3, Pp. 3–14. (rus). DOI: 10.18411/vntr2018-127-1
22. Demina N.V. The study of the unidirectionally coupled generators of robust chaos and wide band communication scheme based on its synchronization. *Izvestiya vuzov. Prikladnaya nelineynaya dinamika*, 2013, Vol. 21, no. 3, Pp. 18–28. (rus)
23. Kuznetsov S.P. From Anosov's dynamics on a surface of negative curvature to electronic generator of robust chaos. *Izv. Saratov Univ. (N.S.), Ser. Physics.*, 2016, Vol. 16, no. 3, Pp. 131–144. (rus). DOI: 10.18500/1817-3020-2016-16-3-131-144
24. Luo X. Circuitry implementation of novel four-dimensional nonautonomous hyperchaotic Liu system and its experimental studies on synchronization control. *Chinese Physics*, 2009, Vol. 18, no. 6, Pp. 2168–2175.
25. Kuptsov P.V., Kuznetsov S.P. Numerical test for hyperbolicity of chaotic dynamics in time delay systems. *Phys. Rev. E*, 2016, Vol. 94, no. 1, Pp. 1201–1212.
26. Yemelyanova Yu.P., Kuznetsov A.P. Synchronization of coupled Van der Pol and Kislov-Dmitriev self-oscillators. *Technical Physics*, 2011, Vol. 81, no. 4, Pp. 7–18. (rus)
27. Yemelyanov V.V., Yemelyanova Yu.P. Mutual synchronization of two coupled generators with delay. *Izvestiya VUZ. Applied Nonlinear Dynamics*, 2013, Vol. 21, no. 3, Pp. 52–60. (rus). DOI: 10.18500/0869-6632-2013-21-3-52-61
28. Shashihin V.N., Budnik S.V. Synthesis of control for nonlinear systems. *Automatic Control and Computer Sciences*, 2019, Vol. 53, no. 2, Pp. 97–106. DOI: 10.3103/S0146411619020068
29. Shashihin V.N., Budnik S.V. Stabilization of nonlinear systems with dynamic chaos. *Automatic Control and Computer Sciences*, 2021, Vol. 55, no. 3, Pp. 213–221. DOI: 10.3103/S0146411621030032
30. Shabunin A.B. Methods of measuring chaotic synchronization. *Izvestiya VUZ. Applied Nonlinear Dynamics*, 2009, Vol. 17, no. 5, Pp. 60–69. (rus). DOI: 10.18500/0869-6632-2009-17-5-60-69
31. Grobman D. Homeomorphy of dynamical systems. *Differ. Uravn.* (rus). 1959, Vol. 128, no. 5, Pp. 880–881.
32. Khartman F. *Obyknovennyye differenttsialnyye uravneniya* [Ordinary differential equations]. Moscow: Mir Publ., 1970. 720 p. (rus)
33. Kozlov V.N., Shashikhin V.N. Synthesis of decentralized robust stabilizing control for the systems with parametric perturbations. *Computing, Telecommunications and Control*, 2020, Vol. 13, no. 2, Pp. 49–60. DOI: 10.18721/JCSTCS.13205
34. Lorenz E.N. Deterministic nonperiodic flow. *J. Atmos. Sci.*, 1963, Vol. 20, Pp. 130–141.
35. Pankratova Ye.V., Belykh V.N. The peculiarities of transition to complete synchronization in networks of Hodgkin–Huxley elements. *Izvestiya VUZ. Applied Nonlinear Dynamics*, 2008, Vol. 16, no. 2, Pp. 3–19. (rus). DOI: 10.18500/0869-6632-2008-16-2-3-17

INFORMATION ABOUT AUTHORS / СВЕДЕНИЯ ОБ АВТОРАХ

Vladimir N. Shashikhin
Шашихин Владимир Николаевич
E-mail: shashihin@bk.ru

Sergey A. Savchuk
Савчук Сергей Андреевич
E-mail: ser_egiks@mail.ru

Submitted: 08.06.2022; Approved: 08.07.2022; Accepted: 25.08.2022.
Поступила: 08.06.2022; Одобрена: 08.07.2022; Принята: 25.08.2022.



Research article

DOI: <https://doi.org/10.18721/JCSTCS.15202>

UDC 621.3.087.92



MODERN APPROACHES TO DESIGN OF MULTI-CHANNEL DELTA-SIGMA ADCS

V.I. Satyshev¹ 

¹ Peter the Great St. Petersburg Polytechnic University,
St. Petersburg, Russian Federation

 sateshev5@yandex.ru

Abstract. Incremental delta-sigma ADC (IADC) and memoryless delta-sigma ADC are described. These two approaches allow to utilize a delta-sigma ADC, known for its increased resolution, in multi-channel systems due to the inter-sample interference suppression that two mentioned structures provide. In this paper, MATLAB/Simulink models of the mentioned structures are presented. In particular, limiting blocks are added to take into account nonlinearities due to finite power supply of integrators; coefficients of delta-sigma modulators were selected so as to maximize their signal-to-noise ratio; parameters of the raised cosine filter were selected to minimize crosstalk between channels. Results of simulations, namely power spectral density of the output signals and signal-to-noise ratio of the output signals, confirm operability of the described structures.

Keywords: analog-to-digital converter, delta-sigma modulation, inter-sample interference, incremental delta-sigma ADC, memoryless delta-sigma ADC

Citation: Satyshev V.I. Modern approaches to design of multi-channel delta-sigma ADCs. Computing, Telecommunications and Control, 2022, Vol. 15, No. 2, Pp. 25–31. DOI: 10.18721/
JCSTCS.15202

Научная статья

DOI: <https://doi.org/10.18721/JCSTCS.15202>

УДК 621.3.087.92



СОВРЕМЕННЫЕ ПОДХОДЫ К ПРОЕКТИРОВАНИЮ МНОГОКАНАЛЬНЫХ ДЕЛЬТА-СИГМА АЦП

В.И. Сатышев¹ ✉

¹ Санкт-Петербургский политехнический университет Петра Великого,
Санкт-Петербург, Российская Федерация

✉ sateshev5@yandex.ru

Аннотация. Рассмотрены инкрементальные дельта-сигма АЦП (IADC) и дельта-сигма АЦП без эффекта памяти. Данные подходы позволяют использовать дельта-сигма АЦП, известные высоким разрешением, в многоканальных системах, благодаря подавлению межвыборочной интерференции (ISI), которая достигается в двух рассматриваемых структурах. Приведены модели указанных структур в MATLAB/Simulink. В частности, введены ограничивающие блоки для учета нелинейных свойств реальных интеграторов, обусловленных конечностью напряжения питания; коэффициенты дельта-сигма модуляторов выбирались таким образом, чтобы максимизировать их отношение сигнал/шум; выбор параметров фильтра приподнятого косинуса проводился с целью минимизации межвыборочной интерференции в каналах. Результаты моделирований (спектральная плотность мощности выходных сигналов, отношение сигнал/шум выходных сигналов) подтверждают работоспособность рассматриваемых структур.

Ключевые слова: аналого-цифровой преобразователь, дельта-сигма модуляция, межвыборочная интерференция, инкрементальный дельта-сигма АЦП, дельта-сигма АЦП без эффекта памяти

Для цитирования: Satyshev V.I. Modern approaches to design of multi-channel delta-sigma ADCs // Computing, Telecommunications and Control. 2022. T. 15, № 2. C. 25–31. DOI: 10.18721/JCSTCS.15202

Introduction

Electronic signals can be divided into two distinct categories: analog signals, which are continuous in time and amplitude, and digital signals, which could be presented as a set of discrete values. Digital systems and devices for storing and processing information have become widespread in recent decades. However, all of the signals that can be collected via physical processes, such as human speech or temperature measurements, are analog. That rises a problem of converting such signals into digital form so that they could be processed in digital systems. The devices that carry out such a conversion are known as analog-to-digital converters or ADCs.

In practice the necessity of using one ADC to digitize several analog channels might arise. That process is known as multiplexing an ADC. There are many types of ADCs, and among them delta-sigma ADCs are known to achieve the highest resolution thanks to oversampling and the noise shaping effect [1, 2]. Unfortunately, it is not possible to use conventional delta-sigma ADCs in multi-channel devices due to the inter-sample interference (ISI). The problem of ISI also known as crosstalk is demonstrated in Fig. 1.

The aim of this work is to present two modern approaches to design a delta-sigma ADC in a way that it could be used in a multi-channel system, that is it is free of inter-sample interference.

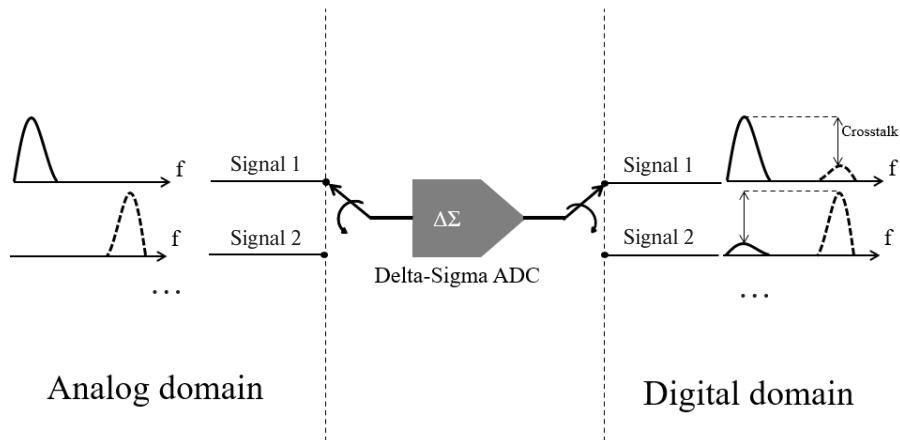


Fig. 1. Multiplexing a conventional delta-sigma ADC

Incremental delta-sigma ADCs

The first approach to design a multi-channel delta-sigma ADC is to use an incremental delta-sigma ADC (IADC). The main feature of IADC is the presence of a global reset pulse. The memories of both analog part and digital part of the ADC are reset at the beginning of the conversion of the subsequent analog signal sample. Thereby it is possible to completely remove ISI from the system. Moreover, since the delta-sigma modulators (DSMs) of such structures usually have finite impulse response, it is possible to use a finite impulse response decimation filter in contrast to much harder to implement infinite impulse response filters that are usually favored in conventional delta-sigma ADCs. The implementation of the decimation filter in this structure can be as straightforward as a cascade of integrators (CoIs). Other advantages of incremental structures include low latency and less tendency to idle tones.

The main disadvantage of incremental structures is increased thermal noise compared to conventional structures. To keep the same value of signal-to-noise ratio (SNR) as in conventional structures, it is necessary to increase the size of the input capacitor which results in higher power consumption of the amplifier that drives it [3–7]. There is a number of different approaches to improve characteristics

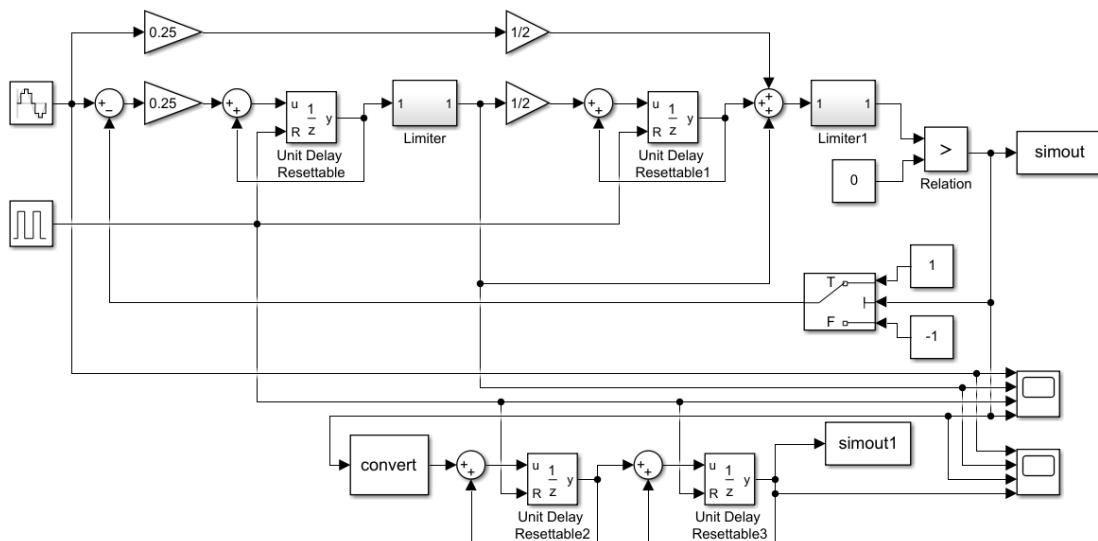


Fig. 2. Model of a second-order IADC in Simulink

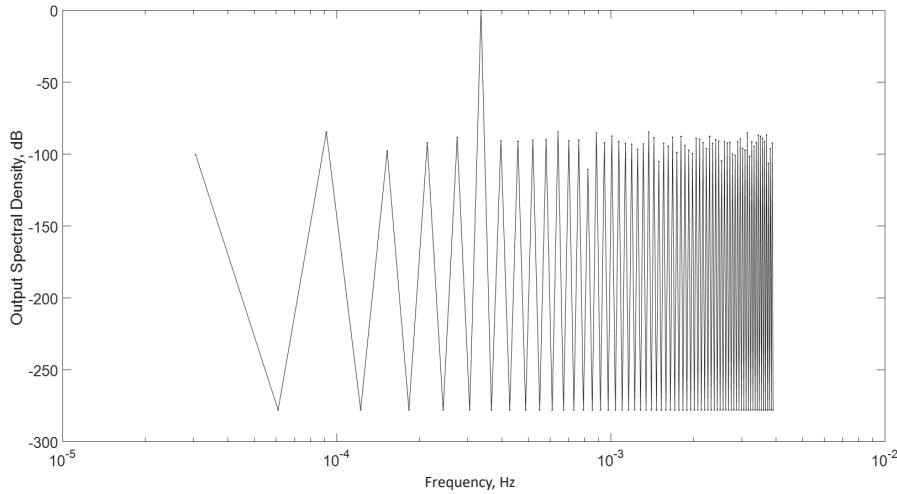


Fig. 3. Spectral density of the output signal after decimation

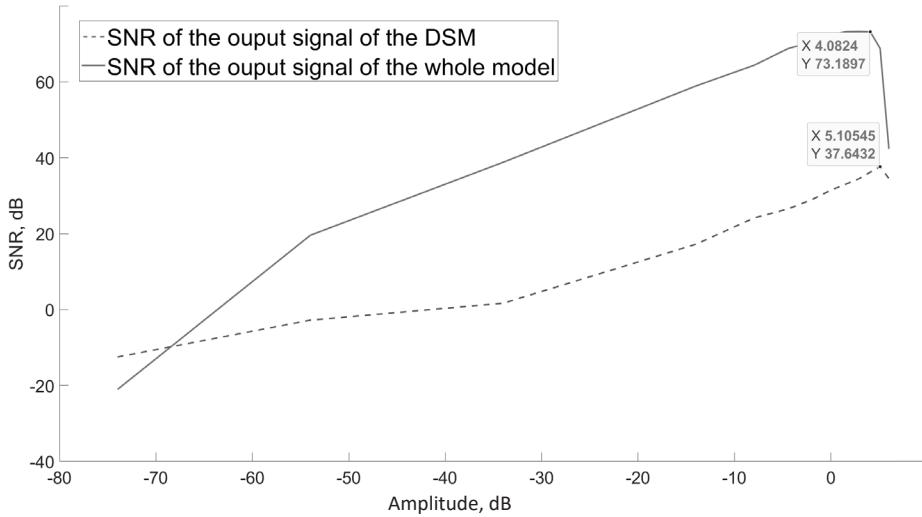


Fig. 4. Dependence of SNR from input amplitude

of incremental structures such as IADC with extended counting [8–11], hardware-sharing IADC [12], multi-step IADC [13–15], zoom ADC [16].

A second-order IADC was simulated in MATLAB/Simulink and is presented in Fig. 2. The model consists of a second-order delta-sigma ADC and two integrators that filter out the quantization noise. The decimation is carried out in a MATLAB script. In order to implement reset, functionally resettable unit delay blocks are used. Limiter blocks are added to take into account possible nonlinearities that could be created by real integrators due to limitations of power supply. Coefficients of delta-sigma modulator are selected so as to maximize SNR. Simulation results are presented in Fig. 3, 4.

Simulation results show that the reset pulses lead to a dramatic SNR decrease of the DSM output signal compared to conventional structures, but the SNR of the IADC as a whole remains good enough. As stated before, it is possible to add a multiplexer to the input of this model and configure the reset pulses in a way that would erase the memories of DSM and filter integrators at the beginning of each subsequent analog signal so as to achieve a zero ISI multi-channel conversion.

Memoryless delta-sigma ADCs

Another approach to design a multi-channel delta-sigma ADC is to use a memoryless delta-sigma ADC. This type of ADC relies heavily on implementing the decimation filter as a Nyquist filter, which is commonly used in communication systems. By utilizing zeros in its impulse response, it is possible to achieve a great ISI suppression. The main advantage of this structure comparing to IADC is that it achieves higher SNR because the impulse response of its decimation filter is not limited by the reset pulses which leads to narrower transition width and higher stop-band attenuation of the frequency response. That results in better overall noise suppression [17]. The principle of operation of the memoryless-type system is as follows. Let f_s be the oversampled DSM clock frequency, M – the oversampling ratio. Then if an analog signal is sampled at the frequency of f_s/M , up-sampled by a factor of M , passed through the DSM and Nyquist filter, the resulting signal would have zero ISI once in every M samples. These samples can be extracted by down-sampling the resulting signal by a factor of M (Fig. 5) [18].

A memoryless delta-sigma ADC utilizing a third-order DSM and a 4-bit quantizer was simulated in MATLAB/Simulink and is presented in Fig. 6. The model consists of a switch that acts as multiplexer, an upsample block, a third-order DSM, a 4-bit ADC-DAC structure (quantizer), a raised cosine decimation filter, which is a popular Nyquist filter, and two switches that act as a demultiplexer. Results of simulation are presented in Fig. 7.

Results of the simulations show that the memoryless delta-sigma ADC shown in Fig. 6 achieves a low noise level, which results in the SNR value of 81.9 dB, and great ISI suppression. The value of crosstalk between different channels is -94.5 dB.

Conclusion

Two modern approaches to design a multi-channel delta-sigma ADC are described. These structures are suitable for multi-channel systems as they aim to suppress the inter-sample interference, that limits

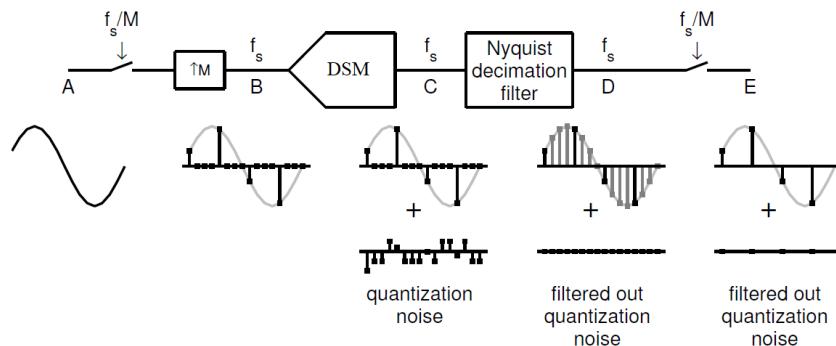


Fig. 5. Block diagram of memoryless delta-sigma ADC

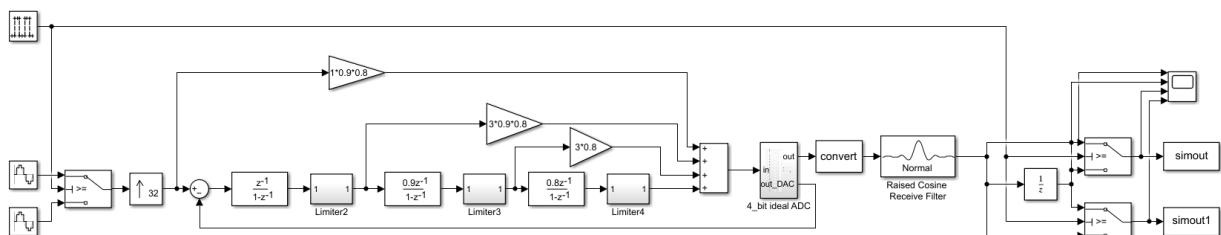


Fig. 6. Model of memoryless delta-sigma ADC in Simulink

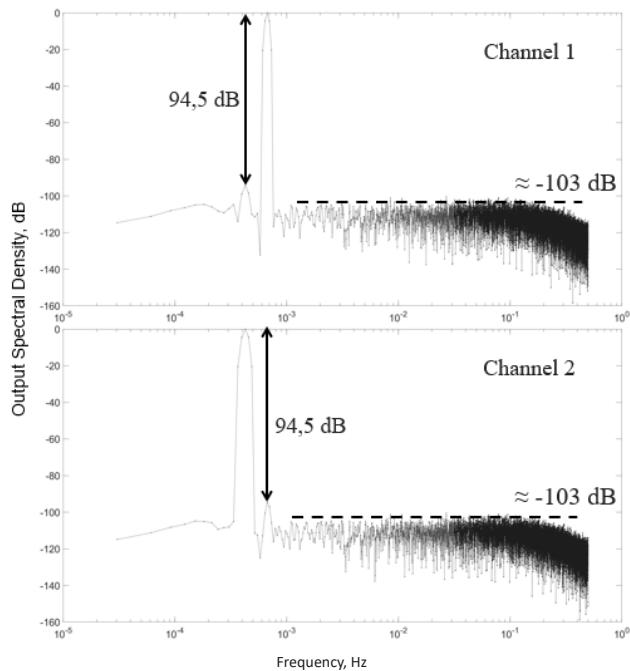


Fig. 7. Spectral density of the output signals

the usage of conventional delta-sigma ADCs in such systems. The MATLAB/Simulink simulation results of the second-order incremental delta-sigma ADC (IADC) and the third-order memoryless delta-sigma ADC are presented.

REFERENCES

1. Kozlov A.S., Pilipko M.M. A second-order sigma-delta modulator with a hybrid topology in 180nm CMOS. *2020 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering*, St. Petersburg and Moscow, Russia, 2020, Pp. 144–146. DOI: 10.1109/EIConRus49466.2020.9039246
2. Korotkov A.S., Morozov D.V., Pilipko M.M., Yenuchenko M.S. Sigma-delta ADC on SOI technology for working at high temperatures. *Radioelectronics and Communications Systems*, 2020, Vol. 63, no. 11, Pp. 586–595. DOI: 10.3103/S0735272720110035
3. Tan Z., Chen C.H., Chae Y., Temes G.C. Incremental delta-sigma ADCs: A tutorial review. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 2020, Vol. 67, no. 12, Pp. 4161–4173. DOI: 10.1109/TCSI.2020.3033458
4. Markus J., Silva J., Temes G.C. Theory and applications of incremental delta-sigma converters. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 2004, Vol. 51, no. 4, Pp. 678–690. DOI: 10.1109/TCSI.2004.826202
5. Kavusi S., Kakavand H., Gamal A.E. On incremental sigma-delta modulation with optimal filtering. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 2006, Vol. 53, no. 5, Pp. 1004–1015. DOI: 10.1109/TCSI.2006.870218
6. Chen C.H., He T., Zhang Y., Temes G.C. Incremental analog-to-digital converters for high-resolution energy-efficient sensor interfaces. *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, 2015, Vol. 5, no. 4, Pp. 612–623. DOI: 10.1109/JETCAS.2015.2502135



7. **Chae Y., et al.** A 2.1 M Pixels, 120 Frame/s CMOS image sensor with column-parallel delta-sigma ADC architecture. *IEEE Journal of Solid-State Circuits*, 2011, Vol. 46, no. 1, Pp. 236–247. DOI: 10.1109/JSSC.2010.2085910
8. **Agah A., Vleugels K., Griffin P.B., Ronaghi M., Plummer J.D., Wooley B.A.** A high-resolution low-power incremental sigma-delta ADC with extended range for biosensor arrays. *IEEE Journal of Solid-State Circuits*, 2010, Vol. 45, no. 6, Pp. 1099–1110. DOI: 10.1109/JSSC.2010.2048493
9. **Harjani R., Lee T.A.** FRC: A method for extending the resolution of Nyquist rate converters using oversampling. *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, 1998, Vol. 45, no. 4, Pp. 482–494. DOI: 10.1109/82.663805
10. **Maeyer J. De, Rombouts P., Weyten L.** A double-sampling extended-counting ADC. *IEEE Journal of Solid-State Circuits*, 2004, Vol. 39, no. 3, Pp. 411–418. DOI: 10.1109/JSSC.2003.822903
11. **Lee C.C., Flynn M.P.** A 14 b 23 MS/s 48 mW resetting sigma-delta ADC. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 2011, Vol. 58, no. 6, Pp. 1167–1177. DOI: 10.1109/TCSI.2010.2097716
12. **Kim J.H., et al.** A 14b extended counting ADC implemented in a 24Mpixel APS-C CMOS image sensor. *2012 IEEE International Solid-State Circuits Conference*, 2012, Pp. 390–392. DOI: 10.1109/ISSCC.2012.6177060
13. **Chen C.H., Zhang Y., He T., Chiang P.Y., Temes G.C.** A micro-power two-step incremental analog-to-digital converter. *IEEE Journal of Solid-State Circuits*, 2015, Vol. 50, no. 8, Pp. 1796–1808. DOI: 10.1109/JSSC.2015.2413842
14. **Zhang Y., Chen C.H., He T., Temes G.C.** Multi-step extended-counting analogue-to-digital converters. *Electronics Letters*, 2013, Vol. 49, no. 1, Pp. 30–31. DOI: 10.1049/el.2012.3655
15. **Zhang Y., Chen C.H., He T., Temes G.C.** A 16 b multi-step incremental analog-to-digital converter with single-opamp multi-slope extended counting. *IEEE Journal of Solid-State Circuits*, 2017, Vol. 52, no. 4, Pp. 1066–1076. DOI: 10.1109/JSSC.2016.2641466
16. **Chae Y., Souri K., K. Makinwa A.A.** A 6.3 μ W 20 bit incremental zoom-ADC with 6 ppm INL and 1 μ V Offset. *IEEE Journal of Solid-State Circuits*, 2013, Vol. 48, no. 12, Pp. 3019–3027. DOI: 10.1109/JSSC.2013.2278737
17. **Kumar R.S.A., Behera D., Krishnapura N.** Reset-free memoryless delta–sigma analog-to-digital conversion. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 2018, Vol. 65, no. 11, Pp. 3651–3661. DOI: 10.1109/TCSI.2018.2854707
18. **Behera D., Krishnapura N.** A 2-channel 1MHz BW, 80.5 dB DR ADC using a DS modulator and zero-ISI filter. *ESSCIRC 2014 – 40th European Solid State Circuits Conference (ESSCIRC)*, 2014, Pp. 415–418. DOI: 10.1109/ESSCIRC.2014.6942110

INFORMATION ABOUT AUTHOR / СВЕДЕНИЯ ОБ АВТОРЕ

Vladislav I. Satyshev
Сатышев Владислав Игоревич
E-mail: sateshev5@yandex.ru

Submitted: 24.06.2022; Approved: 28.07.2022; Accepted: 25.08.2022.
Поступила: 24.06.2022; Одобрена: 28.07.2022; Принята: 25.08.2022.

Research article

DOI: <https://doi.org/10.18721/JCSTCS.15203>

UDC 621.396.69



INFLUENCE OF ALGORITHM PARAMETERS ON STATIC NONLINEARITY IN SWITCHING-BASED CALIBRATIONS FOR DACS

N.V. Kvashina¹ 

¹ Peter the Great St. Petersburg Polytechnic University,
St. Petersburg, Russian Federation

 kvashina.nv@gmail.com

Abstract. This paper presents a comparative study of nonlinearity reduction by different algorithms of switching-based calibration method for DACs. Based on the known algorithms, one general parametric algorithm is proposed. The introduced parameters are a resolution of primary array and numbers of folding and decoupling operations. The intermediate options of such a general algorithm, defined by the parameters combination, are called “algorithm cases”. The algorithm cases are compared by the efficiency of static nonlinearity reduction in presence of random and systematic errors. For both errors, the folding operation improves DNL in contrast with the decoupling operation, which deteriorates it. The increment of folding operations number lowers the presence of systematic drift in DNL, which completely disappears at the maximum number. In the case of INL, the impact of both operations is determined by the elements order. If elements values depend on their indices, the folding operation deteriorates INL in contrast with the decoupling operation, which improves it. All fully unarized arrays provide the INL reduction, but do not decrease DNL.

Keywords: digital-to-analog converter, digital calibration, switching-based algorithm, mismatch, nonlinearity reduction, DNL, INL

Citation: Kvashina N.V. Influence of algorithm parameters on static nonlinearity in switching-based calibrations for DACs. Computing, Telecommunications and Control, 2022, Vol. 15, No. 2, Pp. 32–42. DOI: 10.18721/JCSTCS.15203

Научная статья

DOI: <https://doi.org/10.18721/JCSTCS.15203>

УДК 621.396.69



ВЛИЯНИЕ ПАРАМЕТРОВ АЛГОРИТМА НА СТАТИЧЕСКУЮ НЕЛИНЕЙНОСТЬ В КОММУТАЦИОННЫХ КАЛИБРОВКАХ ЦАП

Н.В. Квашина¹ ¹ Санкт-Петербургский политехнический университет Петра Великого,
Санкт-Петербург, Российская Федерация

kvashina.nv@gmail.com

Аннотация. Проведено сравнительное исследование снижения нелинейности ЦАП различными алгоритмами коммутационного метода калибровки. На основе известных алгоритмов представлен общий параметрический алгоритм. Предложенными параметрами алгоритма являются разрядность исходного массива и количество операций объединения и рассоединения. Промежуточные случаи общего алгоритма, определяющиеся комбинацией параметров, называются «случаями алгоритма». Проведено сравнение случаев алгоритма между собой по эффективности снижения статической нелинейности в присутствии случайной и систематической ошибок. Для обоих типов ошибок операция объединения улучшает DNL в отличие от операции рассоединения, ухудшающей её. Увеличение количества операций объединения снижает влияние систематического ухода в зависимости DNL, полностью исчезающего при максимальном количестве операций. В случае INL влияние обеих операций зависит от того, в каком порядке находятся элементы. Если наблюдается зависимость значений элементов от их индексов, то операция объединения ухудшает INL в отличие от операции рассоединения, которая её улучшает. Все массивы элементов, имеющие унарную архитектуру, демонстрируют снижение INL, однако не уменьшают DNL.

Ключевые слова: цифро-аналоговый преобразователь, цифровая калибровка, коммутационный алгоритм, отклонение, снижение нелинейности, дифференциальная нелинейность, интегральная нелинейность

Для цитирования: Kvashina N.V. Influence of algorithm parameters on static nonlinearity in switching-based calibrations for DACs // Computing, Telecommunications and Control. 2022. T. 15, № 2. С. 32–42. DOI: 10.18721/JCSTCS.15203

Introduction

A digital-to-analog converter (DAC) is a significant interface device, which is implemented in modern electronic communication systems for converting an input digital code into a corresponding output analog signal, that can be represented as current or voltage. The conversion process is realized due to DAC's weighting elements (or shortly "elements") forming an analog signal proportionally to their values, or weights. In this paper, as for the elements, current sources are considered.

According to the ratio of elements weights, three common DACs architectures are highlighted: unary, binary and segmented. In a unary architecture, all elements are identical in their weights (also called "unary-weighted elements") and their ratio equals 1. In a binary architecture, elements are binary-weighted, which means that their weights have a ratio as power of 2. As for the segmented architecture, it is constructed from both unary and binary architecture parts called "segments". In Fig. 1, there are all three common architectures of DACs with a representation of elements in each architecture.

For a correct conversion result, such characteristics of DAC transfer curve as linearity and monotonicity are substantial and require the elements weights ratio to be precise regarding the nominal ratio.

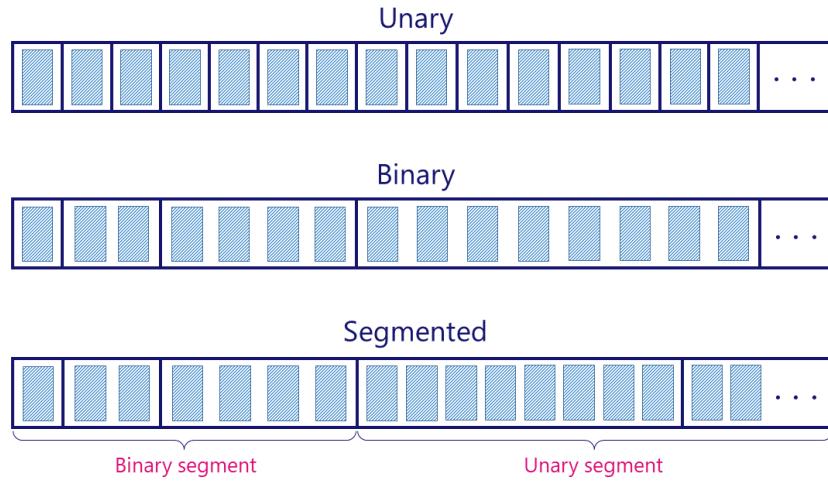


Fig. 1. Three common architectures of DACs

However, with the fabrication of integrated circuits, the elements mismatches unavoidably appear and deteriorate the accuracy of elements ratio, resulting in nonlinearity, or even nonmonotonicity, of transfer curve. Moreover, such unpleasant effects as aging, parasitic components and temperature deviations can also have an impact on further degradation of linearity. The mentioned issues are critical and make design of high-resolution (over 12 bits) DACs with a linear transfer curve quite challenging. In order to reduce the deteriorative impact of errors on DAC linearity and weaken the requirements for elements accuracy, a great number of electronic calibration methods was presented in [1–5] and classified in [6].

Among all calibration methods described in the classification, one particular method, i.e. switching-based calibration method, stands out due to minimization of the analog part. There is a great diversity of implementations of this calibration method (or just “algorithms”). These algorithms are also discussed thoroughly and classified in [7]. One of the introduced criteria for the classification is presence or absence of redundant elements, and the algorithms with the latter one are an object of this research. Some works did a comparison of such algorithms in terms of efficiency in nonlinearity reduction [8, 9], but provided poor coverage of intermediate implementations for the calibration method. Moreover, they do not study a static performance for the discussed algorithms in presence of systematic errors, as it was done for a dynamic performance in [8].

The goal of this research is to provide a comparative analysis of nonlinearity reduction by different implementations of switching-based calibration method in presence of both random and systematic errors. Additionally, this paper proposes a generalization of the known implementations providing intermediate options called “algorithm cases”.

The present paper is organized as follows. In Section 1, the basic operations of switching-based calibration method are introduced. Then, a description of a general algorithm for switching-based calibration method is presented in Section 2. In Section 3, the simulation results for algorithm cases in presence of random and systematic errors are discussed. Based on the results, the comparison of considered algorithm cases by their efficiency in nonlinearity reduction is provided. Finally, in Section 4, the conclusions of the present work are drawn.

1. Basic operations

The principle of switching-based calibration method without redundant elements lies in changing the switching pattern for elements in an initial non-calibrated array. The initial array consists of 2^N unary-weighted elements (N is a DAC resolution) and further is called the primary array. The switching pattern is changed by means of elements indices reordering. The reordering procedure is based on two

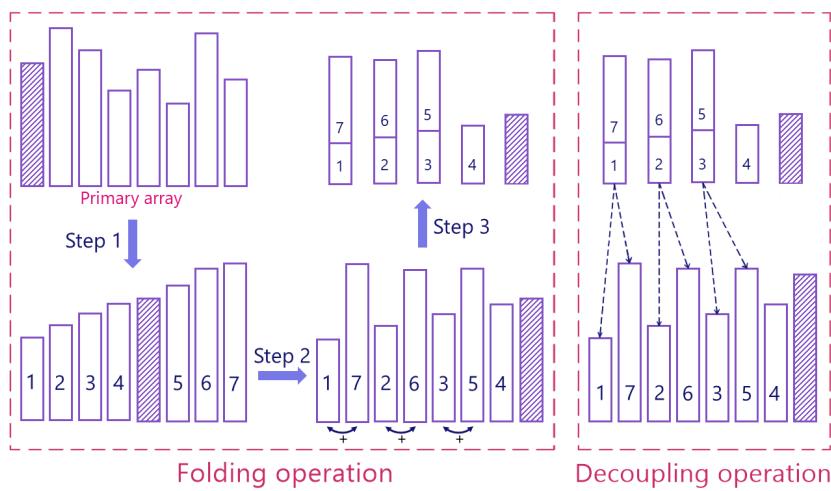


Fig. 2. The folding and decoupling operations

basic operations, folding and decoupling, which are applied to the primary array of elements as illustrated in Fig. 2. Here, rectangles heights correspond to the elements values.

To perform the folding operations, the 3 following steps must be done. At Step 1, the elements of primary array are compared with each one by their values and sorted in an ascending order. Then, the sorted elements change their places as it is shown in Step 2, so the couples of elements are created. Each couple contains the smallest elements preceding the greatest. At the final Step 3, the elements in couples are united. There is also a shaded dummy element in the array, which may be not involved in unification and can be absent. The result of folding operation is a segmented array with a decreased resolution of unary segment by one bit and with an increased resolution of binary segment by one bit.

The decoupling operation is performed with the application of only one step – the decomposition of previously united elements. It should be noted that after the decoupling all elements are placed in their order before unification. As the result, this operation increases the resolution of the unary array by one bit, which corresponds to the decrease in the resolution of the binary array. The number of decoupling operations cannot exceed the number of folding operations.

2. Description of general algorithm

Based on the previously known implementations [5, 9, 10], one general parametric algorithm is proposed. The algorithm depends on the following parameters: N is a primary array resolution, F and D are numbers of folding and decoupling operations respectively. A certain combination of N , F and D gives an individual algorithm case. In Fig. 3, a table of all algorithm cases is presented. The resolution of primary array N defines the maximum numbers of F and D . Colored areas in the table correspond to existent algorithm cases, while grayed out areas represent impossible cases of algorithm due to unrealizable combination of F and D .

It can be seen that with the increment of N the maximum numbers of folding and decoupling operations F and D also increase. Thus, the number of possible or existent algorithm cases grows with an arithmetic progression. The sum of arithmetic progression S_n is obtained by

$$S_n = \left(\frac{a_1 + a_n}{2} \right) \times n,$$

where a_1 and a_n are the first and last numbers in numerical sequence, n is a number of elements in the sequence. The a_1 is the number of algorithm cases for a row with $F = 1$ in Fig. 3, a_n is the number of

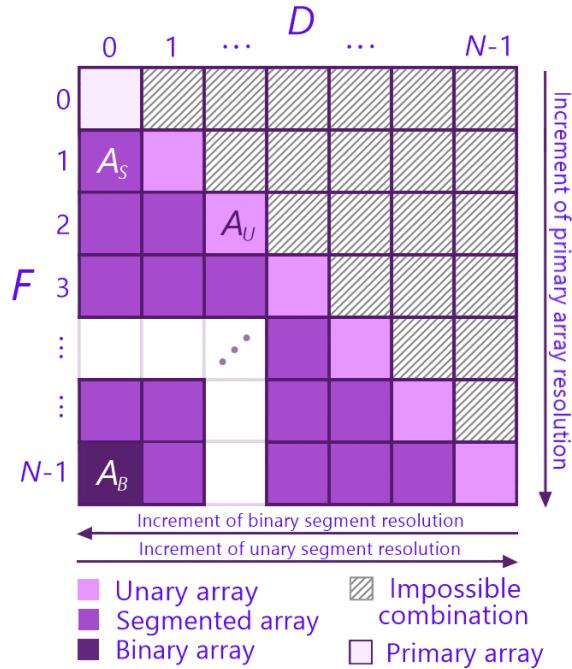


Fig. 3. The table of algorithm cases and corresponding architectures

algorithm cases for a row with $F = N - 1$ and n is the number of rows. So, the final result for the number of existent algorithm cases is:

$$M_{ex} = \frac{(N-1)(N+2)}{2}.$$

According to the combinations of F and D , there are three types of algorithm cases which allow to obtain either binary, or unary or segmented architecture of calibrated array. As it was mentioned before, the folding and decoupling operations change the resolutions of unary and binary segments of calibrated array, thus the resolutions can be obtained by the following expressions, where N_U and N_B are resolutions of unary and binary segments respectively:

$$\begin{aligned} N_B &= F - D, \\ N_U &= N - N_B = N - F + D. \end{aligned}$$

In order to construct the desired architecture of calibrated array, corresponding requirements for the numbers of folding and decoupling operations must be satisfied. So, a unary architecture is obtained when $F = D$, a segmented architecture requires $D < F < N - 1$ and a binary architecture is constructed when $F = N - 1$ and $D = 0$. The previously mentioned implementations are particular cases of the proposed general algorithm, that is:

- Switching sequence post adjustment – SSPA (further A_U – unary calibrated array) [5];
- Total 3-dimensional sort-and-combine – T3D-SC (further A_S – segmented calibrated array) [10];
- Complete-Folding (further A_B – binary calibrated array) [9].

In this work, they are called “boundary algorithm cases”. For A_U algorithm case $F = D = 2$, for A_S algorithm case $F = 1, D = 0$ and for A_B algorithm case $F = F_{\max} = N - 1, D = 0$. In Fig. 4, the calibrated array for every boundary algorithm case is presented.

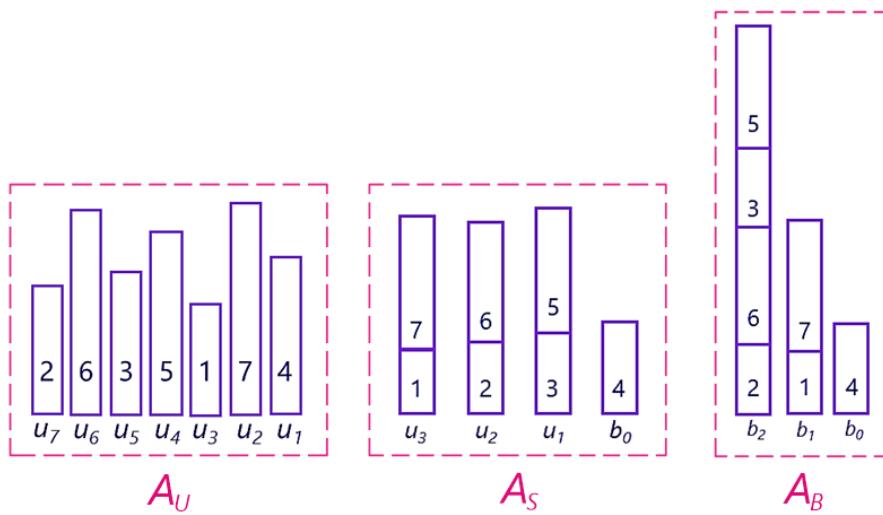


Fig. 4. Calibrated arrays for boundary algorithm cases

In order to study the efficiency of DAC static nonlinearity reduction by algorithm cases, a parametric algorithm was observed, and maximum averaged values of differential and integral nonlinearities (DNL and INL) were obtained in the research [11]. In the present work, the influence of algorithm parameters on DNL and INL exact values will be studied more thoroughly.

3. Simulation and comparison

In this section, the simulation and comparison of the results for random and systematic errors are presented. Since it is reasonable to calibrate high-resolution DACs (12 bits and more), here N resolution is 12. To perform a calibration of the primary array influenced by random errors, it is necessary to form an array of $m = 2^N$ unary elements. For each element of the primary array, a random deviation with a normal distribution is introduced. Such a normal distribution has the following parameters: mathematical expectation μ is 0 and the value of standard deviation σ is chosen to be 0.13 %. The value of σ is an example of deviation, which was derived while performing the Monte-Carlo simulation of an element (current source) for UMC 180 nm technology [12]. The number of generated primary arrays is 200. All mentioned parameters are highlighted in Table 1.

Table 1
Parameters of primary array with random errors

Resolution, N	Number of elements, m	Number of algorithm cases, M_{ex}	Mathematical expectation, μ	Standard deviation, σ	Number of arrays
12	4096	77	0	0.13 %	200

All possible algorithm cases were applied to a set of primary arrays with random deviations. As the result, transfer curves before and after calibration for each algorithm case were obtained and further used for DNL and INL calculations. In Fig. 5, the DNL and INL dependences on input code D and elements order are presented for boundary algorithm cases A_U , A_S and A_B . There is also a maximum normalized value defined for each nonlinearity dependence. As it was mentioned before, the elements

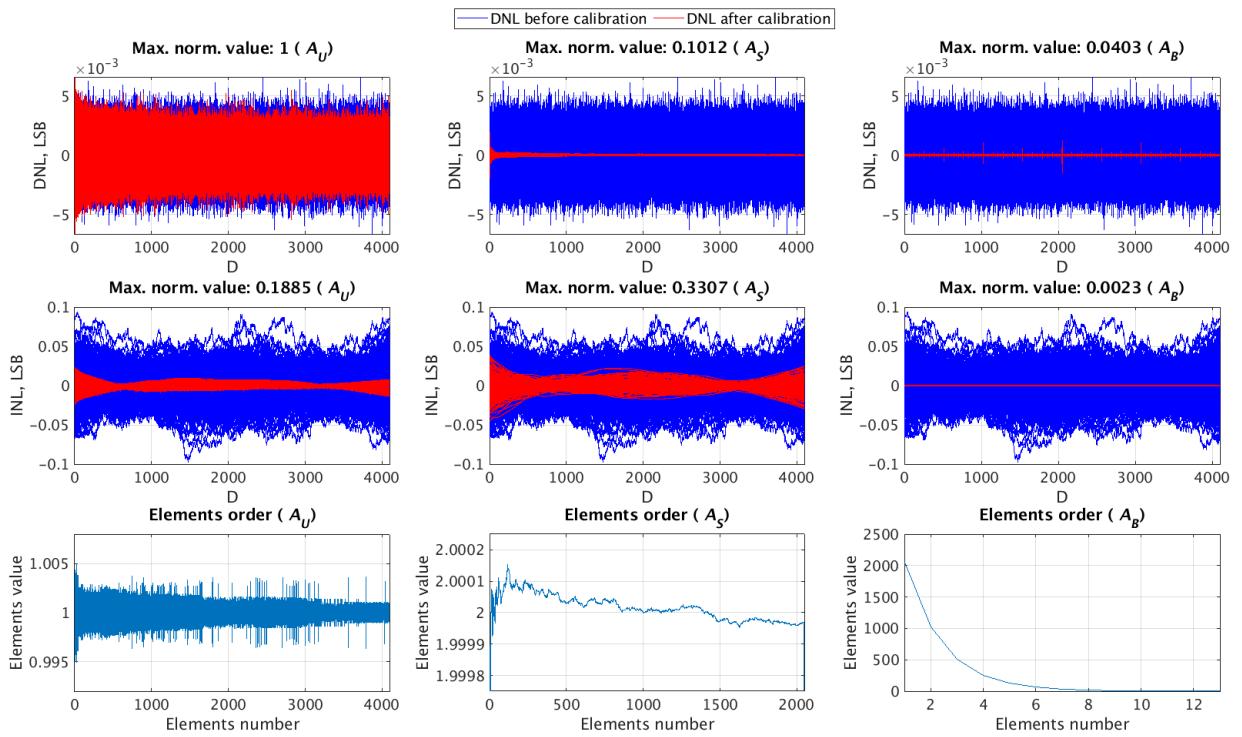


Fig. 5. Simulation results for DNL, INL and elements orders of boundary cases in presence of random errors

after application of A_U , A_S and A_B algorithm cases form unary, segmented and binary arrays respectively. The number of deviation cases illustrated in the graphs is 200.

The results show that the increment of binary segment resolution even only by one bit allows to greatly improve the DNL, which can be seen in the transition between A_U and A_S and between A_S and A_B algorithm cases. Such an improvement is possible due to presence of united elements after the folding operation. The unification reduces elements deviation and, consequently, DNL. For instance, transition from A_U to A_S reduces elements deviation and corresponding DNL by one order. However, the elements order obtained after unification deteriorated INL since element value decreases with its index. In contrast, splashes break the dependence of element value on its index in case of A_U . As a result, it leads to better INL reduction.

When comparing the obtained DNL dependences, a systematic drift of DNL can be noticed after the calibration for the A_U and A_S boundary cases, which is absent in the A_B case. To establish the reason for the systematic drift appearance, it is necessary to consider the results for DNL with fixed F from a pure unary, or unarized, array to a pure binary, or binarized, one. This will allow us to understand whether an increase in the resolution of a unary segment, or unarization, affects the appearance of a systematic drift.

In Fig. 6, the simulation results of DNL, INL and elements orders are presented. Here, algorithm cases from a bottom line of the table in Fig. 3 are considered, where $F = F_{\max} = 11$. The unarization power increases from left to right, so the leftmost case is a binarized array and the rightmost case is a unarized one. The number of deviation cases illustrated in the graphs is 200.

According to the obtained results, there is no systematic drift of DNL with an increase of unarization power. So, the unarization has no impact on the appearance of the systematic drift. Previously, the systematic drift in DNL dependences occurred for A_U and A_S algorithm cases when F was lesser than the maximum possible value F_{\max} . Meanwhile for A_B case, where $F = F_{\max}$, it was absent. Therefore, while studying algorithm cases with a fixed F_{\max} , no systematic drift was observed in the dependences.

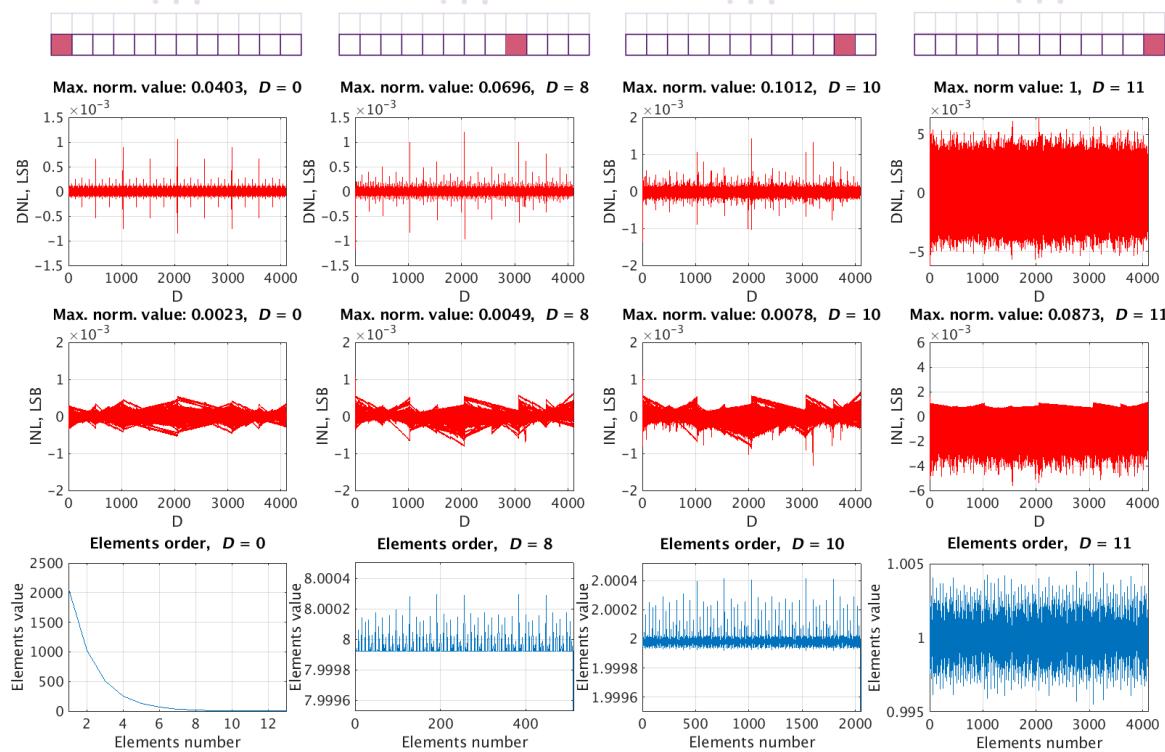


Fig. 6. Simulation results for DNL, INL and elements orders ($F = F_{\max} = 11$)

In order to establish the reason of systematic drift appearance, results of DNL, INL and elements orders for algorithm cases with different numbers of F are presented in Fig. 7. Here, algorithm cases that allow obtaining fully unarized arrays (i.e. cases on the diagonal in Fig. 3) are considered.

According to the results, the systematic drift can be seen. Moreover, with the decrement of F from the maximum possible value $F_{\max} = 11$ to the minimum possible value $F = 1$ the influence of systematic drift becomes worse. Hence, the systematic drift appears when F is lesser than the maximum possible value. Therefore, the more folding operations are done, the less is the systematic drift for DNL results in presence of random errors.

The forming of primary array influenced by systematic errors is similar to the one in presence of random errors except for the deviations introduction. A primary array of 2^N elements is formed ($N = 12$) for a further calibration. Then, according to the methodology for systematic errors forming with different profiles [13], a generation of profile parameters combinations and set of profiles systematic errors is held. Among the whole set of profiles, a joint approximation with anisotropic case is chosen to be applied to the primary array. The number of cases considered for various angles and shifts is 200.

Similarly to the case of random errors, boundary algorithm cases were applied to a primary array influenced by systematic errors. The DNL and INL were calculated using the obtained transfer curves before and after calibration. In Fig. 8, the DNL and INL dependences and elements orders are presented for boundary algorithm cases A_U , A_S and A_B .

According to the obtained results, the systematic drift, which was observed earlier for the DNL dependence in presence of random errors, provides an impact on DNL for the A_U and A_S algorithm cases in presence of systematic errors as well. However, the A_B algorithm case demonstrates the lack of systematic drift, which is also similar to the result for random errors. Therefore, the cause of such a behavior is the same for both types of errors, and one common statement can be concluded: the fewer folding operations are done, the worse is the systematic drift impact on the DNL results.

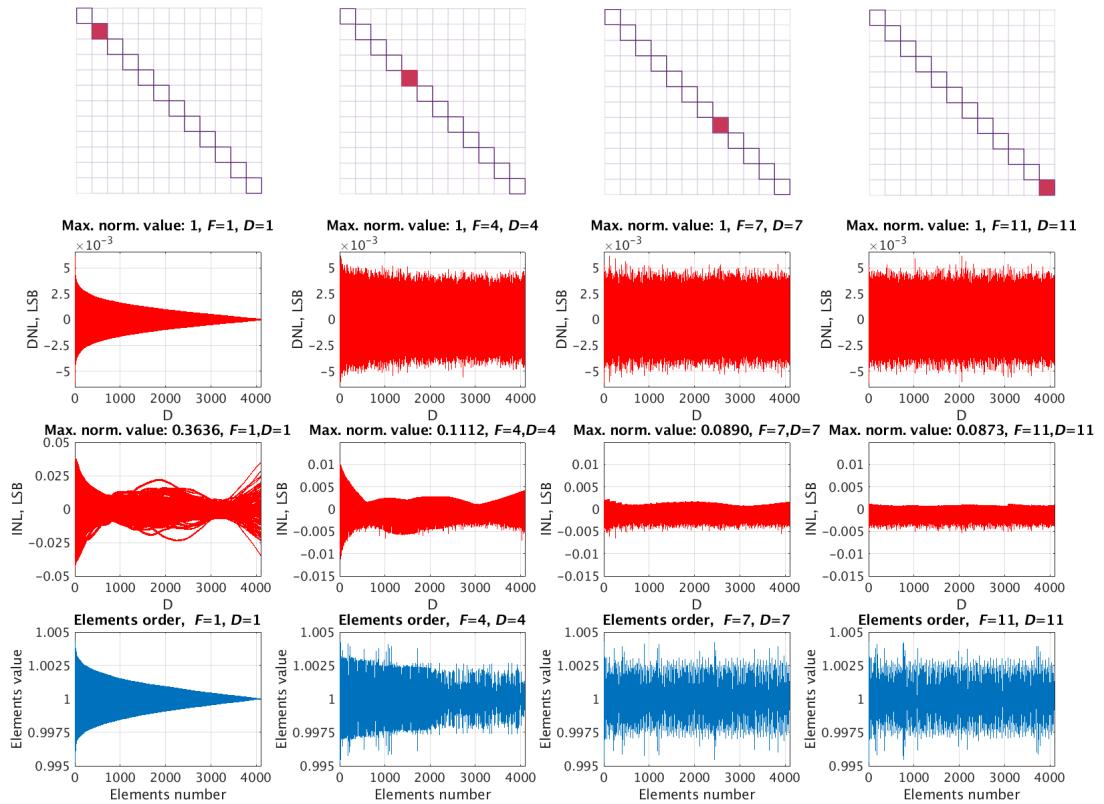


Fig. 7. Simulation results for DNL, INL and elements orders (diagonal cases)

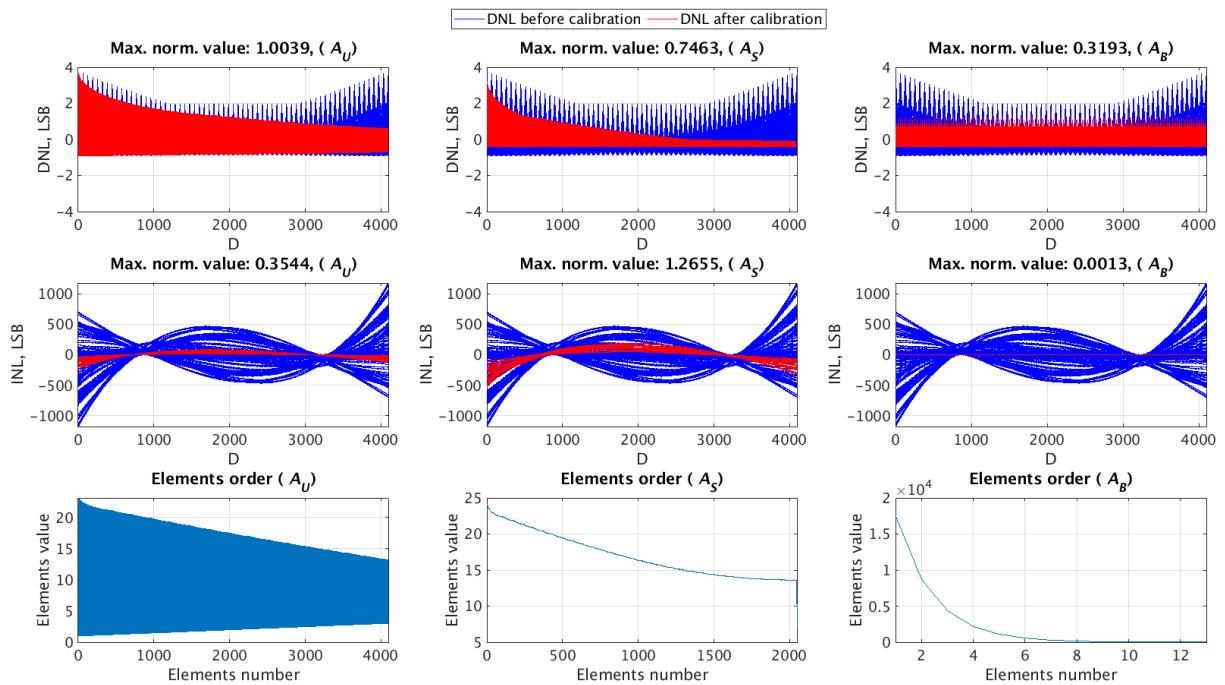


Fig. 8. Simulation results for DNL, INL and elements orders of boundary cases in presence of systematic errors

4. Conclusion

Digital-to-analog converters (DACs) are widespread devices for converting digital data into analog signals. The linearity of conversion depends on the elements ratio, which can deviate from the nominal value during integrated circuit fabrication. To overcome this issue, a calibration is one of the possible solutions. In this work, a non-redundant switching-based calibration method is considered and the general algorithm for such a method is proposed. The intermediate implementations of the algorithm are compared by efficiency of DNL and INL reduction.

The folding and decoupling operations are introduced. One folding operation ($F = 1$) decreases the resolution of a unary segment by one bit and increases the resolution of a binary segment by one bit. One decoupling operation ($D = 1$) decreases the resolution of a binary segment by one bit and increases the resolution of a unary segment by one bit. In terms of DNL, the folding operation improves nonlinearity due to reduction of elements deviations during the elements unification. The opposite result is demonstrated by the decoupling operation, which deteriorates nonlinearity during the elements splitting. In terms of INL, the impact of folding operation depends on the elements order after calibration, which also contributes to the nonlinearity. If the elements order represents a dependence of element value on its index, then the unification has a deteriorative impact on nonlinearity in contrast to the decoupling operation, which improves it. Research [11] concluded that a simultaneous variation of both F and D does not allow definite estimation of a more efficient algorithm case in nonlinearity reduction.

All fully unarized arrays ($F = D$) provide the INL reduction, however the DNL cannot be decreased. The similar results were demonstrated in [14, 15] for the switching schemes of unary DACs. The DNL reduction can be obtained by the increment of binary segment resolution – binarization. Binarization provides an error averaging for elements, which reduces elements deviations, as they will consist of two and more parts. Such an effect is demonstrated for switching schemes with element division in [13, 14, 16]. As a result, the greater is the resolution of the binary segment, the better is the nonlinearity reduction.

The number of applied folding operations F affects the presence of systematic drift in DNL dependences for both random and systematic errors. After the sorting step of the folding operation an explicit dependence of element value on its index occurs, which has an influence on the DNL dependence in the form of the systematic drift. With the increment of F , the reordering of elements is more thorough. Thus, the greater is F , the better elements are reordered and the lesser is an explicit dependence of element value on its index. At the maximum possible number of folding operations F_{\max} , the array of elements is completely reordered, which results in the absence of the explicit dependence. Hence, the systematic drift in case of F_{\max} is absent.

REFERENCES

1. Xu S., Lee J.W. A digital technique for calibrating and correcting nonlinearities in current-steering DACs. *2018 IEEE International Symposium on Circuits and Systems (ISCAS)*, 2018, Pp. 1–5. DOI: 10.1109/ISCAS.2018.8351366
2. Mukhopadhyay I., Mukadam M.Y., Narayanan R., O'Mahony F., Apsel A.B. Dual-calibration technique for improving static linearity of thermometer DACs for I/O. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, 2016, Vol. 24, no. 3, Pp. 1050–1058. DOI: 10.1109/TVLSI.2015.2424954
3. Cong Y., Geiger R.L. A 1.5-V 14-bit 100-MS/s self-calibrated DAC. *IEEE Journal of Solid-State Circuits*, 2003, Vol. 38, no. 12, Pp. 2051–2060. DOI: 10.1109/JSSC.2003.819163
4. Wu G., Wei Q., Chen X., Qiao F. A 14-bit 2.5-GS/s DAC using CAL-TRZ. *2016 5th International Conference on Modern Circuits and Systems Technologies (MOCAST)*, 2016, Pp. 1–4. DOI: 10.1109/MOCAST.2016.7495116

5. **Chen T., Gielen G.** A 14-bit 200-MHz current-steering DAC with switching sequence post-adjustment calibration. *2006 IEEE Asian Solid-State Circuits Conference*, 2006, Pp. 347–350. DOI: 10.1109/ASSCC.2006.357922
6. **Kvashina N., Yenuchenko M.** Overview of electronic calibrations for DACs. *2020 IEEE International Conference on Electrical Engineering and Photonics (EExPolytech)*, 2020, Pp. 42–46. DOI: 10.1109/EExPolytech50912.2020.9243988
7. **Kvashina N.V., Yenuchenko M.S.** Algorithms of switching-based calibration methods for DACs. *2021 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (ElConRus)*, 2021, Pp. 145–s149. DOI: 10.1109/ElConRus51938.2021.9396170
8. **Valet P., Schwingshakl D., Tonello A.M.** Comparative study on pre-distortion/calibration methods for current-steering digital-to-analog converters. *2021 IEEE 12th Latin America Symposium on Circuits and System (LASCAS)*, 2021, Pp. 1–4. DOI: 10.1109/LASCAS51355.2021.9459164
9. **Zeng T., Chen D.** New calibration technique for current steering DACs. *Proceedings of 2010 IEEE International Symposium on Circuits and Systems*, 2010, Pp. 573–576. DOI: 10.1109/ISCAS.2010.5537534
10. **Huang S., Li X., Li X.** A 14b 1GS/s DAC with SFDR > 80 dBc across the whole Nyquist band by mixed total 3-dimesional sort-and-combine and dynamic element matching. *2015 IEEE 11th International Conference on ASIC (ASICON)*, 2015, Pp. 1–4. DOI: 10.1109/ASICON.2015.7516911
11. **Kvashina N.V., Yenuchenko M.S.** Comparative analysis of switching-based calibration algorithms for DACs. *2022 Conference of Russian Young Researchers in Electrical and Electronic Engineering (ElConRus)*, 2022, Pp. 157–161. DOI: 10.1109/ElConRus54750.2022.9755635
12. **Kvashina N.V., Yenuchenko M.S.** A 12-bit current-steering segmented DAC with digital foreground calibration. *2021 International Conference on Electrical Engineering and Photonics (EExPolytech)*, 2021, Pp. 95–99. DOI: 10.1109/EExPolytech53083.2021.9614886
13. **Pilipko M.M., Yenuchenko M.S.** First- and second-order gradient compensation with only eight parts per element in unary DACs. *2021 International Conference of Russian Young Researchers in Electrical and Electronic Engineering (ElConRus)*, 2021, Pp. 158–163. DOI: 10.1109/ElConRus51938.2021.9396732
14. **Yenuchenko M.S., Pilipko M.M., Hauer J.** Analysis of nonlinearity reduction by binary and unary switching schemes in DACs. *2021 International Conference on Electrical Engineering and Photonics (EExPolytech)*, 2021, Pp. 112–115. DOI: 10.1109/EExPolytech53083.2021.9614711
15. **Konstantinov A.I., Yenuchenko M.S., Korotkov A.S.** Efficiency analysis of techniques for weighting elements arrangement on the chip of unary digital-to-analog converter. *2017 Radioelectronics and Communications Systems*, 2017, Vol. 60, no. 5, Pp. 225–232.
16. **Yenuchenko M.S., Pilipko M.M.** multi-flip technique compensating a gradient rotation in unary DACs. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 2021, Vol. 68, no. 3, Pp. 883–887. DOI: 10.1109/TCSII.2020.3021203

INFORMATION ABOUT AUTHOR / СВЕДЕНИЯ ОБ АВТОРЕ

Natalya V. Kvashina
Квашина Наталья Владимировна
E-mail: kvashina.nv@gmail.com

Submitted: 30.06.2022; Approved: 22.07.2022; Accepted: 25.08.2022.
Поступила: 30.06.2022; Одобрена: 22.07.2022; Принята: 25.08.2022.

Моделирование вычислительных, телекоммуникационных, управляющих и социально-экономических систем

Simulations of Computer, Telecommunications, Control and Social Systems

Научная статья

DOI: <https://doi.org/10.18721/JCSTCS.15204>

УДК 004.912, 004.85, 004.41



РАЗРАБОТКА И ИССЛЕДОВАНИЕ МОДЕЛЕЙ МНОГОКЛАССОВЫХ КЛАССИФИКАТОРОВ ДЛЯ РЕКОМЕНДАТЕЛЬНОЙ СИСТЕМЫ ПОДГОТОВКИ ЗАЯВОК НА ПОРТАЛЕ ЕДИНОЙ ИНФОРМАЦИОННОЙ СИСТЕМЫ В СФЕРЕ ЗАКУПОК

Я.А. Селиверстов¹✉, А.А. Комиссаров², А.А. Лесоводская³,
П.Г. Бобыкин⁴, А.В. Подтихоб⁵, С.С. Торсионов⁶,
Д.А. Цирков⁷, С.А. Орлов⁸

¹ Институт проблем транспорта имени Н.С. Соломенко РАН,
Санкт-Петербург, Российская Федерация;

^{1,2,3,4,5,6,7} Университет национальной технологической инициативы 2035,
Москва, Российская Федерация;

⁸ Томский государственный университет,
г. Томск, Российская Федерация

✉ Y.Seliverstov@2035.university

Аннотация. Обоснована актуальность разработки сервисов, способствующих подготовке тендерной документации, в части определения кода ОКПД 2 к формируемой заявке. Для решения задачи автоматической классификации заявок в соответствии с ОКПД 2 разработан алгоритм системы сравнительного анализа моделей классификаторов, осуществлена предобработка и запись в базу данных собранной информации в формате json. Разметка и подготовка данных для обучения моделей классификаторов осуществлена в среде PolyAnalyst. В качестве моделей многоклассовых классификаторов из библиотеки Scikit-Learn выбраны наивный байесовский классификатор, SVM-классификатор и классификатор на основе случайного леса. В качестве векторизаторов выбрана модель TFIDF и word-hashing. В качестве четвертого классификатора выбрана нейросетевая модель ruBert-base. Проведено обучение классификаторов и оценено качество их работы. По результатам валидации и тестирования лучшими оказались две модели: ruBert-base и модель наивного байесовского классификатора с векторизатором word-hashing. На основе результатов произведена тестовая классификация заявок.

Ключевые слова: рекомендательные системы, многоклассовая классификация, SVM, naive Bayes, ruBert-base, векторизаторы

Для цитирования: Селиверстов Я.А., Комиссаров А.А., Лесоводская А.А. и др. Разработка и исследование моделей многоклассовых классификаторов для рекомендательной системы подготовки заявок на портале единой информационной системы в сфере закупок // Computing, Telecommunications and Control. 2022. Т. 15, № 2. С. 43–62. DOI: 10.18721/JCSTCS.15204

Research article

DOI: <https://doi.org/10.18721/JCSTCS.15204>

UDC 004.912, 004.85, 004.41



DEVELOPMENT AND RESEARCH OF MODELS OF MULTI-CLASS CLASSIFIERS FOR A RECOMMENDED SYSTEM FOR PREPARING APPLICATIONS ON THE E-PROCUREMENT

*Y.A. Seliverstov¹✉, A.A. Komissarov², A.A. Lesovodskay³,
P.G. Bovykin⁴, A.V. Podtikhov⁵, S.S. Torsionov⁶,
D.A. Tsyrkov⁷, S.A. Orlov⁸*

¹ IPT RAS, St. Petersburg, Russian Federation;
^{1,2,3,4,5,6,7} University National Technology Initiative 2035,
Moscow, Russian Federation;

⁸ National Research Tomsk State University,
Tomsk, Russian Federation

✉ Y.Seliverstov@2035.university

Abstract. As a result of the analysis, the relevance of developing services that contribute to the preparation of tender documentation, in terms of determining the OKPD 2 code for the generated application, is indicated. To solve the problem of automatic classification of applications in accordance with OKPD 2, an algorithm for the system of comparative analysis of classifier models was developed. Further, preprocessing was carried out, and the collected information was written to the database in json format. Labeling and preparation of data for training classifier models was carried out in the PolyAnalyst environment. As a result of the analysis, a naive Bayes classifier, an SVM classifier, and a random forest classifier were selected as models of multiclass classifiers from the Scikit-Learn library. The TFIDF and word-hashing models were chosen as vectorizers. The ruBert-base neural network model was chosen as the fourth classifier. Classifiers were trained and the quality of their work was assessed. According to the results of validation and testing, two models turned out to be the best: ruBert-base and a model of a naive Bayes classifier with a word-hashing vectorizer. Based on the results, a test classification of applications was made.

Keywords: recommendation systems, multiclass classification, SVM, naive Bayes, ruBert-base, vectorizer

Citation: Seliverstov Y.A., Komissarov A.A., Lesovodskay A.A., et al. Development and research of models of multi-class classifiers for a recommended system for preparing applications on the e-procurement. Computing, Telecommunications and Control, 2022, Vol. 15, No. 2, Pp. 43–62.
DOI: 10.18721/JCSTCS.15204

Введение

Развитие современных информационных систем и технологий способствуют быстрому росту цифровизации во всех сферах государственного управления, в том числе и государственных закупок. Основную роль в этом процессе играют цифровые платформы, поскольку именно они формируют среду электронного взаимодействия потенциальных заказчиков и поставщиков, являясь провайдером сферы электронных закупок.

С 2016 года в России введена в эксплуатацию Единая информационная система (ЕИС) в сфере закупок [1]. ЕИС создана в целях информационного обеспечения контрактной системы в сфере закупок товаров, работ, услуг для обеспечения государственных и муниципальных нужд, закупок товаров, работ, услуг отдельными видами юридических лиц.

В настоящее время процедуры проведения тендеров занимают ключевое место в работе отдела закупок компаний. Большинство современных рекомендательных сервисов и систем [2–4] функ-



ционируют на веб-площадках. К их числу относятся рекомендательные сервисы проведения тендеров [5], выбора тендеров с вероятностью оценки выигрыша участников тендера [6], проверки корректности тендерной документации [7] и др.

Одним из важных этапов формирования электронной тендерной документации является процедура формирования заявки и выбора соответствующего кодификатора из перечня ОКПД 2 – Общероссийского классификатора продукции по видам экономической деятельности.

Стремительный рост различных секторов производства и расширение многообразия создаваемых товаров, продуктов и услуг, а также перепрофилирование предпринимательской деятельности самостоятельного субъекта усложняют задачу точной идентификации продуктовых границ отраслевого рынка в рамках Общероссийских классификаторов видов предпринимательской деятельности (ОКПД 2) [8].

По данным [9], ОКПД 2 содержит около 38 тысяч позиций. Столь большая величина затрудняет поиск и увеличивает временные затраты на его точное определение.

Указание неправильного кода ОКПД 2, его незнание или несоответствие реальной экономической деятельности коду ОКПД 2, указанному субъектом при регистрации, влечет множество нежелательных последствий, таких как определение ставки взносов при возникновении несчастных случаев; оформление патентов; потеря налоговых льгот, предоставляемых государством по некоторым видам экономической деятельности; блокирование платежей при внутренней банковской проверке; наложение высоких штрафных санкций при внесении ошибочного кода ОКПД 2 в Единую информационную систему в сфере закупок [10, 11].

Последнее подчеркивает актуальность разработки и внедрения рекомендательных сервисов проверки корректности заявки [12] в части выбора ОКПД 2 в системы электронного сопровождения корпоративного тендерного документооборота [13].

Анализ предметной области. В настоящее время рекомендательным системам посвящено множество работ российских и зарубежных ученых.

Теоретико-методологические особенности рекомендательных систем, а также различные характеристики и возможности методов прогнозирования в современных рекомендательных системах рассмотрены в [14, 15].

Подход совместной фильтрации, основанный на наивном байесовском классификаторе, изучен в [16]. Предложенная в [16] байесовская модель не только предоставляет рекомендации, но и дает объяснения этим прогнозам.

В [17] описана система, использующая кластеризацию и случайный лес в качестве многоуровневых стратегий для прогнозирования рекомендаций на основе оценок пользователей, ориентируясь при этом на мышление пользователей и текущие тенденции.

Решая предсказательную задачу моделирования динамических предпочтений пользователей на основе их исторического поведения, в [18] представлена глубокая двунаправленная последовательная модель под названием BERT4Rec для последовательной рекомендации. При обучении модели применяется задача Cloze, которая предсказывает элементы, используя как левый, так и правый контекст.

Влияние систем поиска и рекомендаций на продажи в электронной торговле изучено в [3]. В статье показано, как разные классы инструментов поиска и рекомендаций влияют на распределение продаж по продуктам, общий объем продаж и излишек потребителей.

В [19] авторы рассматривают тематическую диверсификацию, новый метод, разработанный для баланса и диверсификации персонализированных списков рекомендаций, чтобы отразить полный спектр интересов пользователя. Согласно [19] алгоритмы диверсификации использует Amazon для улучшения своих рекомендаций.

В [20] решается задача многоклассовой классификации слаботоксичного образовательного контента онлайн школ с использованием нейросетевой модели RoBERT.

В [21], опираясь на информационно-методическое обеспечение [22] на основе нейронной сети трансформера, разрабатывается метод многоклассовой классификации угроз.

В [23] и [24] рассмотрены системы на основе моделей классификации транспортных данных в части определения состояния качества дорог. В качестве моделей взяты классификаторы SVM и наивного байеса.

В [25] изучены методы машины опорных векторов и латентно-семантический анализ с различными векторизаторами для построения модели многоклассовой классификации слабоструктурированных текстовых документов.

В [26] выполнен подробный обзор подходов интеллектуального анализа текстов на русском языке.

Проведенный анализ свидетельствует о том, что модели классификаторов на основе наивного байесовского алгоритма, машины опорных векторов, случайного леса и нейросетевые алгоритмы, такие как BERT, широко используются в различных информационных системах, связанных с фильтрацией и классификацией веб-контента.

Постановка задачи. Цель настоящей работы – в исследовании современных программных методов классификации для определения наиболее точной модели классификации применительно к задаче классификации заявки тендерной документации к соответствующему классификатору из ОКПД 2. Данную модель классификации в последующем предполагается использовать в разработке рекомендательного сервиса, помогающего специалисту в сфере закупок по информации, содержащейся в подготавливаемой к тендеру заявке, быстро и точно определить соответствующий ей код первого уровня ОКПД 2.

Сервис планируется использовать совместно с Единой информационной системой в сфере закупок (портал <https://zakupki.gov.ru/>).

В рамках исследования планируется разработать краулер для сбора данных, осуществить сбор данных с портала ЕИС, произвести подготовку и разметку данных с использованием программных методов, построить и осуществить сравнительный анализ четырех моделей мультиклассовой классификации: наивного байесовского классификатора (Naive Bayes), классификатора на основе алгоритма случайного леса (Random Forest), классификатора на основе машины опорных векторов (SVM) и нейросетевой модели ruBERT (Bidirectional Encoder Representations from Transformers) [27]. Общее количество классов – 88. В качестве векторизаторов планируется рассмотреть два метода – TFIDF и word hashing [28].

Основная часть

Построение плана исследования целесообразно начать с разработки обобщенного алгоритма системы сравнительного анализа моделей классификаторов. Алгоритм в общем виде состоит из процедур, представленных на рис. 1.

На первом этапе осуществляется сбора данных заявок. Для этих целей разработан краулер на языке Python3 с использованием библиотеки Requests. С помощью разработанного краулера осуществлен сбор заявок закупок с портала <https://zakupki.gov.ru> (рис. 2).

Информация о закупках доступна по гостевому доступу (рис. 3) и хранится на FTP-сервере (<ftp://ftp.zakupki.gov.ru>) в виде дампов в формате xml в директории fcs_regions (рис. 4).

Сбор данных осуществляется по следующему алгоритму: 1) краулер заходит на ftp-сервер под гостевым доступом в директорию fcs_regions; 2) проходит по регионам и загружает RETR-запросом по очереди каждый xml-файл закупок в память; 3) «на лету» из памяти осуществляет парсинг метаданных, содержащихся в xml-файле, преобразуя их в json в соответствии с заданной структурой; 4) записывает json в базу данных; 5) удаляет xml-файл из памяти; 6) переходит к следующему по очереди xml-файлу закупок.

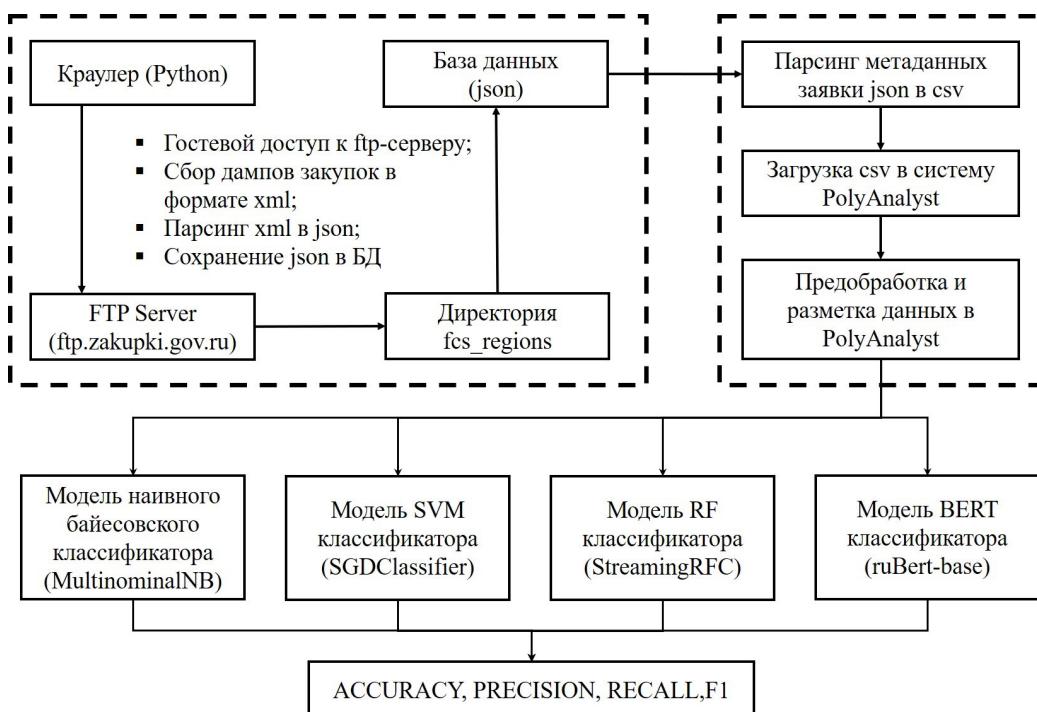


Рис. 1. Схема обобщенного алгоритма системы сравнительного анализа моделей классификаторов

Fig. 1. Scheme of the algorithm of the system for comparing classifier models

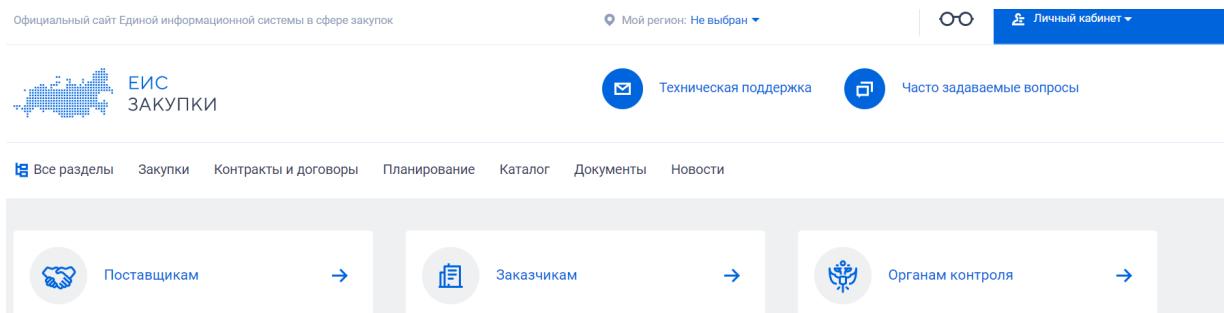


Рис. 2. Портал zakupki.gov.ru

Fig. 2. Portal zakupki.gov.ru

Цикл повторяется до тех пора, пока все метаданные заявок закупок не будут загружены в базу данных.

Структура json-файла закупок в базе данных представлена на рис. 5.

На втором этапе осуществляется парсинг лотов заявок в базе данных и перевод json в csv-формат с использованием скрипта, написанного на Python3, и загрузка csv в систему PolyAnalyst. В результате парсинга json формируется следующая структура csv файла метаданных закупок: id – номер заявки на ftp-сервере портала zakupki.gov.ru; number – реестровый номер извещения; reg – место подведения итогов; publish data – дата публикации заявки на портале ЕИС; purchaseResponsible – наименование организации заказчика; maxPrise – максимальная цена контракта закупки; purchaseObjectInfo – информация о закупках; name_purchaseObject – наимено-

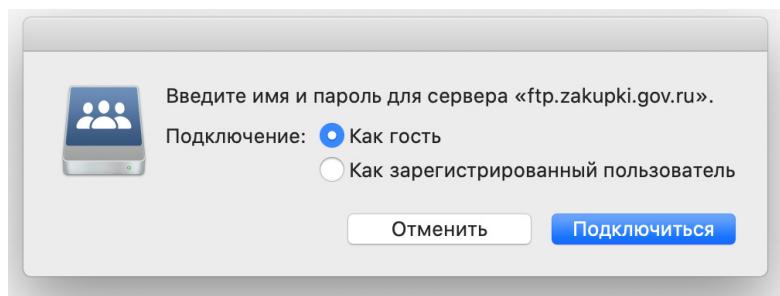


Рис. 3. Окно гостевого доступа к серверу <ftp://ftp.zakupki.gov.ru>
Fig. 3. Guest access window to the <ftp://ftp.zakupki.gov.ru> server

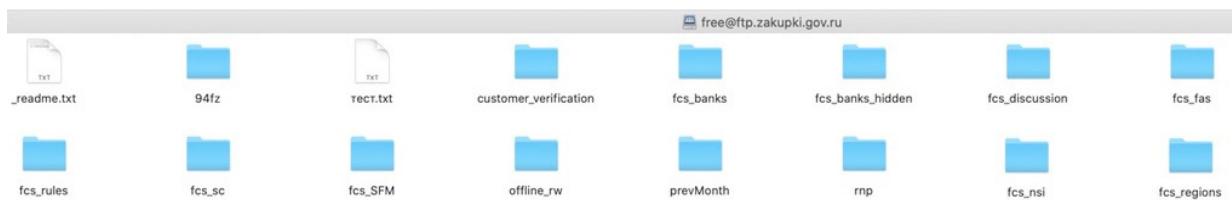


Рис. 4. Директории ЕИС на сервере <ftp://ftp.zakupki.gov.ru>
Fig. 4. Data directories on the ftp server <ftp://ftp.zakupki.gov.ru>

вание объекта закупок; OKPD2_codes – код ОКПД2; OKPD2_names – наименование ОКПД2; OKPD_codes – код ОКПД; OKPD_names – наименование ОКПД; KRTU_codes – код КРТУ; KRTU_names – наименование КРТУ.

Структура csv файла закупки `zakupki_all.csv` для загрузки в систему PolyAnalys для последующей обработки и разметки (рис. 6).

На третьем этапе в среде Polyanalyst [29] выполняется обработка и разметка данных. Графическое представление вычислительного скрипта представлено на рис. 7.

Вычислительный скрипт содержит следующие узлы:

1) узел производные колонки sum text – в данном узле формируется новое поле sum text путем объединения информации из смежных пяти полей:

```
sum text = [purchaseObjectInfo] + "\n" + [name_purchaseObject] + "\n" + [OKPD2_names] + "\n" + [OKPD_names] + "\n" + [KTRU_names];
```

2) узел модификации колонок – в данном узле идет переименование поля sum text в поле description;

3) узел CRC – в данном узле каждой уникальной строке присваивается уникальный номер;

4) узел обработки уникальных описаний – в данном узле по номерам CRC устраняются дубликаты записей, с помощью фильтра русского языка данные отфильтровываются от любых других иностранных слов, устраняются орфографические ошибки и «ё» заменяется на «е»;

5) узел объединения – в данном узле формируется исправленный текст с присвоением ему CRC;

6) узел фильтрации строк – в данном узле удаляются пустые поля description, т. е. те поля, в которых нет описания заявки;

7) узел `zakupki_clean.csv` – исправленный и размеченный классификаторами ОКПД и ОКПД2 текст сохраняется в csv.

На четвертом этапе осуществляется разработка и сравнение моделей классификаторов на языке Python 3. Количество классов ОКПД 2 – 88, т. к. классификация осуществляется только

```

{
    "id": 123,
    "lot": {
        "currency": {
            "code": "RUB",
            "name": "Российский рубль"
        },
        "maxPrice": 245099.1,
        "restrictInfo": "не установлено",
        "financeSource": "средства ТФ ОМС Курганской области",
        "purchaseObjects": [
            {
                "totalSum": 245099.1,
                "purchaseObject": {
                    "sum": 245099.1,
                    "OKPEI": {
                        "code": 236,
                        "nationalCode": "КАП/Ч"
                    },
                    "OKPD": {
                        "code": "40.30.10.111",
                        "name": "Тепловая энергия, отпущенная тепловыми электростанциями"
                    },
                    "name": "Теплоснабжение",
                    "price": 1287.15,
                    "quantity": {
                        "value": 190.42
                    }
                }
            },
            "quantityUndefined": false,
            "customerRequirements": {
                "customerRequirement": {
                    "customer": {
                        "regNum": 2431000024
                    },
                    "maxPrice": 245099.1,
                    "deliveryTerm": "с 01.01.2014г. по 31.12.2014г.",
                    "deliveryPlace": "Российская Федерация, 640018, Курганская область, Курган г, Советская, 81, -"
                }
            }
        ],
        "href": "http://zakupki.gov.ru/epz/order/notice/ep44/view/common-info.html?noticeId=123",
        "printForm": {
            "url": "http://zakupki.gov.ru/epz/order/notice/printForm/view.html?noticeId=123"
        },
        "placingWay": {
            "code": "EP44",
            "name": "Закупка у единственного поставщика (подрядчика, исполнителя)"
        },
        "attachments": {
            "attachment": {
                "url": "http://zakupki.gov.ru/44fz/filestore/public/1.0/download/priz/file.html?uid=EF3F84CF641E00A0E043AC110725D56E",
                "fileName": "обоснование.odt",
                "docDescription": "обоснование"
            }
        },
        "docPublishDate": "2014-01-09T09:48:28.348+04:00",
        "purchaseNumber": 243100002414000420,
        "purchaseObjectInfo": "Теплоснабжение здания ТФ ОМС Курганской области советская, 81",
        "purchaseResponsible": {
            "responsibleOrg": {
                "regNum": 2431000024,
                "fullName": "Территориальный фонд обязательного медицинского страхования Курганской области",
                "factAddress": "Российская Федерация, 640018, Курганская обл, Курган г, Советская, д.81, -",
                "postAddress": "Российская Федерация, 640018, Курганская обл, Курган г, Советская, д.81, -"
            },
            "responsibleInfo": {
                "contactFax": "+7-3522-463142",
                "contactMail": "ur@ktfoms.orbitel.ru",
                "contactPhone": "+7-3522-413637",
                "contactPerson": {
                    "lastName": "Белов",
                    "firstName": "Николай",
                    "middleName": "Григорьевич"
                },
                "orgFactAddress": "Российская Федерация, 640018, Курганская обл, Курган г, Советская, д.81, -",
                "orgPostAddress": "Российская Федерация, 640018, Курганская обл, Курган г, Советская, д.81, -"
            },
            "responsibleRole": "СУ"
        }
    }
}

```

Рис. 5. Структура метаданных заявки в базе данных

Fig. 5. Json file of purchases in the database

по первому уровню. Предобработка табличных размеченных данных `zakupki_clean.csv` осуществляется с использованием библиотеки Pandas, лемматизация текста – с использованием библиотеки `rutystem3`, сериализация и десериализация объектов (запись, чтение и загрузка дампов) – с использованием модуля `pickle`, фильтрация стоп-слов с помощью библиотеки `nltk`, очистка текста от лишних пробелов и знаков препинания осуществляется с помощью регулярных выражений. Модели многоклассовых классификаторов `MultinomialNB` (наивный байесовский классификатор), `SGDClassifier` (машина опорных векторов), `StreamingRFC` (случайный лес), а также `TFIDF` векторизатор импортируются из библиотеки `sklearn`. Векторизатор `word hashing` создается на основе триграмм с использованием функции `permutations` из библиотеки `itertools`. Модель нейросетевого классификатора на основе `ruBert-base` от Сбербанка [29] с токинезато-

Моделирование вычислительных, телекоммуникационных, управляющих и социально-экономических систем

#	id	number	reg	publish_date	purchaseResponsible	maxPrice	purchaseObjectInfo	name_purchaseObject	OKPD2_names	OKPO_code	OKPD_name	KTRU_code	KTRU_name	
4481	25,912,894	00,000,000,000	Республика Карелия	3/20/2021	УПРАВЛЕНИЕ ФЕДЕРАЛЬНОЙ СЛУЖБЫ ПО ТЕХНИЧЕСКОМУ РЕГУЛИРОВАНИЮ И МЕТРОЛОГИИ	104,553.20	Поставка картриджей для принтеров	Картридж Q3651A для принт	28.23.25.000 28.23	Части и принадлежности	32.50.13.120-00327	Линза интраокулярная		
4481	25,912,902	00,000,000,000	Республика Карелия	3/12/2021	ГОСУДАРСТВЕННЫЙ БЮДЖЕТНОЙ ОРГАН	941,078.00	Поставка линз интраокулярных	Линза интраокулярная для з	68.10.11.0000	Услуги по покупке и продаже				
4481	25,912,903	00,000,000,000	Республика Карелия	3/12/2021	МУНИЦИПАЛЬНОЕ КАЗЕННОЕ УЧРЕЖДЕНИЕ	1,335,081.60	Приобретение благоустройств	Линзы интраокулярные однокомнатные	68.10.11.0000	Услуги по покупке и продаже				
4481	25,912,957	00,000,000,000	Республика Карелия	3/12/2021	АДМИНИСТРАЦИЯ ОЛОНЦЕВСКОГО МУНИЦИПАЛЬНОГО КОММУНИКАЦИОННОГО АДМИНИСТРАТИВНОГО ОРГАНА	837,000.00	Услуги по покупке и продаже	Линзы интраокулярные однокомнатные	13.92.29.190	Изделия текстильные	19.20.21.125-00001	Изделия текстильные		
4481	25,913,087	00,000,000,000	Республика Карелия	3/12/2021	ГОСУДАРСТВЕННОЕ КАЗЕННОЕ УЧРЕЖДЕНИЕ	34,077.00	Поставка хозяйственных т	Линзы интраокулярные однокомнатные	26.40.41.000 26.40	Бензин автомобилей	26.40.41.000 26.40	Бензин автомобилей		
4481	25,913,098	00,000,000,000	Республика Карелия	3/12/2021	ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ УЧРЕЖДЕНИЕ	612,040.00	поставка автомобилного	Бензин автомобильный АИ-95	68.10.11.0000	Услуги по покупке и продаже				
4481	25,913,116	00,000,000,000	Республика Карелия	3/12/2021	МУНИЦИПАЛЬНОЕ КАЗЕННОЕ УЧРЕЖДЕНИЕ	1,339,502.40	Приобретение благоустро	Бензин автомобильный АИ-95	21.20.23.111	Препараты диагност	26.40.31.190-00000	Цифровой микшер		
4481	25,913,123	00,000,000,000	Республика Карелия	3/12/2021	ГОСУДАРСТВЕННОЕ КАЗЕННОЕ УЧРЕЖДЕНИЕ	1,880,000.00	поставка светового и звуко	Бензин автомобильный АИ-95	68.10.11.0000	Услуги по покупке и продаже				
4481	25,913,244	00,000,000,000	Республика Карелия	3/12/2021	ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ УЧРЕЖДЕНИЕ	1,511,973.60	поставка экспресс-тестов	Бензин автомобильный АИ-95	68.10.11.0000	Услуги по покупке и продаже				
4481	25,913,252	00,000,000,000	Республика Карелия	3/12/2021	МУНИЦИПАЛЬНОЕ КАЗЕННОЕ УЧРЕЖДЕНИЕ	1,719,691.20	поставка светового и звуко	Бензин автомобильный АИ-95	68.10.11.0000	Услуги по покупке и продаже				
4481	25,913,321	00,000,000,000	Республика Карелия	3/12/2021	МУНИЦИПАЛЬНОЕ КАЗЕННОЕ УЧРЕЖДЕНИЕ	1,976,097.60	Приобретение благоустро	Бензин автомобильный АИ-95	68.10.11.0000	Услуги по покупке и продаже				
4481	25,913,425	00,000,000,000	Республика Карелия	3/12/2021	МУНИЦИПАЛЬНОЕ КАЗЕННОЕ УЧРЕЖДЕНИЕ	1,989,360.00	Приобретение благоустро	Бензин автомобильный АИ-95	68.10.11.0000	Услуги по покупке и продаже				
4481	25,913,491	00,000,000,000	Республика Карелия	3/12/2021	МУНИЦИПАЛЬНОЕ КАЗЕННОЕ УЧРЕЖДЕНИЕ	2,528,697.60	Приобретение благоустро	Бензин автомобильный АИ-95	68.10.11.0000	Услуги по покупке и продаже				
4481	1,407,288	00,000,000,000	Орловская область	7/1/2014	Администрация Орловского района	1,133,543.48	Приобретение жилого пом	Бензин автомобильный АИ-95	70.12.11.0000	Услуги по покупке и продаже				

Рис. 6. Структура csv файла закупок (zakupki_all.csv) в PolyAnalyst

Fig. 6. csv-file (zakupki all.csv) in PolyAnalyst

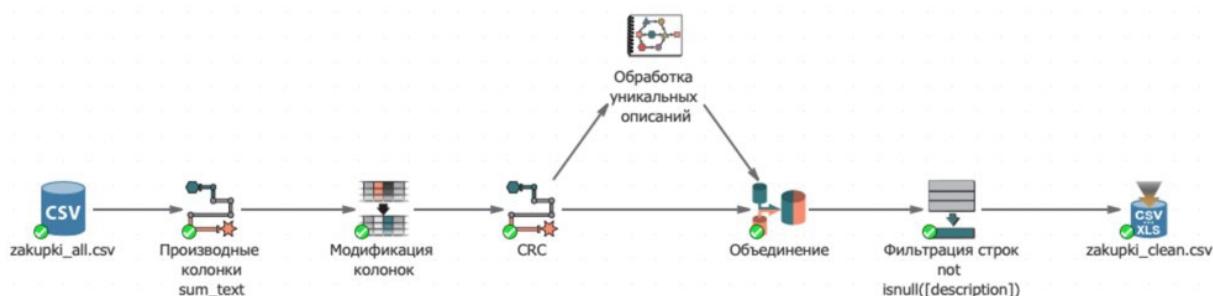


Рис. 7. Вычислительный скрипт обработки данных в PolyAnalyst

Fig. 7. Data processing script in PolyAnalyst

```

classifier nb tfidf
Train [1 / 5] Accuracy = 54.99%, Precision = 46.71%, Recall = 41.61%, F1 = 39.64%:
100% [REDACTED] 459/459 [00:30<00:00, 15.00it/s]
Val [1 / 5] Accuracy = 42.21%, Precision = 40.54%, Recall = 35.26%, F1 = 33.24%:
100% [REDACTED] 51/51 [00:01<00:00, 38.40it/s]
Train [2 / 5] Accuracy = 54.34%, Precision = 47.06%, Recall = 48.03%, F1 = 42.33%:
100% [REDACTED] 459/459 [00:30<00:00, 15.22it/s]
Val [2 / 5] Accuracy = 40.58%, Precision = 40.53%, Recall = 35.63%, F1 = 32.76%:
100% [REDACTED] 51/51 [00:01<00:00, 39.04it/s]
Train [3 / 5] Accuracy = 53.68%, Precision = 46.12%, Recall = 49.75%, F1 = 42.08%:
100% [REDACTED] 459/459 [00:30<00:00, 15.26it/s]
Val [3 / 5] Accuracy = 39.94%, Precision = 40.07%, Recall = 35.77%, F1 = 32.42%:
100% [REDACTED] 51/51 [00:01<00:00, 39.16it/s]
Train [4 / 5] Accuracy = 53.28%, Precision = 45.68%, Recall = 50.49%, F1 = 41.75%:
100% [REDACTED] 459/459 [00:30<00:00, 15.28it/s]
Val [4 / 5] Accuracy = 39.63%, Precision = 40.11%, Recall = 35.95%, F1 = 32.25%:
100% [REDACTED] 51/51 [00:01<00:00, 38.87it/s]
Train [5 / 5] Accuracy = 53.03%, Precision = 45.45%, Recall = 50.88%, F1 = 41.55%:
100% [REDACTED] 459/459 [00:29<00:00, 15.34it/s]
Val [5 / 5] Accuracy = 39.53%, Precision = 39.80%, Recall = 35.99%, F1 = 32.00%:
100% [REDACTED] 51/51 [00:01<00:00, 38.83it/s]
  
```

Рис. 8. Характеристики модели MultinomialNB с TFIDF

Fig. 8. Characteristics of the MultinomialNB model with TFIDF

classifier nb hashing	
Train [1 / 5]	Accuracy = 55.88%, Precision = 45.33%, Recall = 45.08%, F1 = 42.44%:
100%	[██████████] 459/459 [00:59<00:00, 7.78it/s]
Val [1 / 5]	Accuracy = 46.62%, Precision = 41.17%, Recall = 40.32%, F1 = 38.17%:
100%	[██████████] 51/51 [00:02<00:00, 20.86it/s]
Train [2 / 5]	Accuracy = 54.65%, Precision = 43.96%, Recall = 51.24%, F1 = 44.29%:
100%	[██████████] 459/459 [00:59<00:00, 7.74it/s]
Val [2 / 5]	Accuracy = 45.76%, Precision = 40.29%, Recall = 41.37%, F1 = 37.84%:
100%	[██████████] 51/51 [00:02<00:00, 20.99it/s]
Train [3 / 5]	Accuracy = 54.14%, Precision = 42.88%, Recall = 52.18%, F1 = 43.55%:
100%	[██████████] 459/459 [00:59<00:00, 7.77it/s]
Val [3 / 5]	Accuracy = 45.36%, Precision = 39.75%, Recall = 41.69%, F1 = 37.44%:
100%	[██████████] 51/51 [00:02<00:00, 21.01it/s]
Train [4 / 5]	Accuracy = 53.87%, Precision = 42.34%, Recall = 52.58%, F1 = 43.07%:
100%	[██████████] 459/459 [00:59<00:00, 7.74it/s]
Val [4 / 5]	Accuracy = 45.15%, Precision = 39.51%, Recall = 41.84%, F1 = 37.23%:
100%	[██████████] 51/51 [00:02<00:00, 20.96it/s]
Train [5 / 5]	Accuracy = 53.71%, Precision = 42.03%, Recall = 52.59%, F1 = 42.72%:
100%	[██████████] 459/459 [00:59<00:00, 7.72it/s]
Val [5 / 5]	Accuracy = 45.04%, Precision = 39.58%, Recall = 42.08%, F1 = 37.21%:
100%	[██████████] 51/51 [00:02<00:00, 20.95it/s]

Рис. 9. Характеристики модели MultinomialNB с word hashing

Fig. 9. Characteristics of the MultinomialNB model with word hashing

classifier svm tfidf	
Train [1 / 5]	Accuracy = 12.04%, Precision = 15.04%, Recall = 4.35%, F1 = 4.62%:
100%	[██████████] 459/459 [00:58<00:00, 7.85it/s]
Val [1 / 5]	Accuracy = 8.40%, Precision = 10.94%, Recall = 3.44%, F1 = 3.20%:
100%	[██████████] 51/51 [00:01<00:00, 40.53it/s]
Train [2 / 5]	Accuracy = 12.07%, Precision = 15.88%, Recall = 4.35%, F1 = 4.64%:
100%	[██████████] 459/459 [00:45<00:00, 10.12it/s]
Val [2 / 5]	Accuracy = 10.16%, Precision = 12.78%, Recall = 4.06%, F1 = 4.33%:
100%	[██████████] 51/51 [00:01<00:00, 40.88it/s]
Train [3 / 5]	Accuracy = 11.95%, Precision = 14.97%, Recall = 4.31%, F1 = 4.59%:
100%	[██████████] 459/459 [00:44<00:00, 10.26it/s]
Val [3 / 5]	Accuracy = 9.41%, Precision = 12.26%, Recall = 3.84%, F1 = 3.64%:
100%	[██████████] 51/51 [00:01<00:00, 41.76it/s]
Train [4 / 5]	Accuracy = 11.99%, Precision = 15.00%, Recall = 4.30%, F1 = 4.58%:
100%	[██████████] 459/459 [00:45<00:00, 10.10it/s]
Val [4 / 5]	Accuracy = 8.24%, Precision = 11.10%, Recall = 3.20%, F1 = 2.90%:
100%	[██████████] 51/51 [00:01<00:00, 41.74it/s]
Train [5 / 5]	Accuracy = 12.13%, Precision = 15.13%, Recall = 4.40%, F1 = 4.69%:
100%	[██████████] 459/459 [00:44<00:00, 10.37it/s]
Val [5 / 5]	Accuracy = 9.69%, Precision = 13.07%, Recall = 3.72%, F1 = 4.04%:
100%	[██████████] 51/51 [00:01<00:00, 41.85it/s]

Рис. 10. Характеристики модели SGDClassifier с TFIDF

Fig. 10. Characteristics of the SGDClassifier model with TFIDF

classifier svm hashing	
Train [1 / 5]	Accuracy = 34.95%, Precision = 23.10%, Recall = 14.24%, F1 = 13.95%: 100% [██████████] 459/459 [01:55<00:00, 3.97it/s]
Val [1 / 5]	Accuracy = 31.21%, Precision = 22.82%, Recall = 12.13%, F1 = 12.02%: 100% [██████████] 51/51 [00:03<00:00, 15.82it/s]
Train [2 / 5]	Accuracy = 40.38%, Precision = 26.75%, Recall = 14.49%, F1 = 14.98%: 100% [██████████] 459/459 [01:46<00:00, 4.30it/s]
Val [2 / 5]	Accuracy = 31.37%, Precision = 22.19%, Recall = 11.33%, F1 = 10.92%: 100% [██████████] 51/51 [00:02<00:00, 20.84it/s]
Train [3 / 5]	Accuracy = 39.87%, Precision = 27.21%, Recall = 13.46%, F1 = 14.08%: 100% [██████████] 459/459 [01:49<00:00, 4.21it/s]
Val [3 / 5]	Accuracy = 34.79%, Precision = 22.79%, Recall = 11.75%, F1 = 11.36%: 100% [██████████] 51/51 [00:02<00:00, 20.95it/s]
Train [4 / 5]	Accuracy = 39.02%, Precision = 27.41%, Recall = 12.74%, F1 = 13.44%: 100% [██████████] 459/459 [01:51<00:00, 4.13it/s]
Val [4 / 5]	Accuracy = 30.07%, Precision = 23.37%, Recall = 10.48%, F1 = 10.55%: 100% [██████████] 51/51 [00:02<00:00, 20.85it/s]
Train [5 / 5]	Accuracy = 38.12%, Precision = 27.31%, Recall = 12.17%, F1 = 12.91%: 100% [██████████] 459/459 [01:50<00:00, 4.15it/s]
Val [5 / 5]	Accuracy = 30.44%, Precision = 22.98%, Recall = 10.32%, F1 = 10.31%: 100% [██████████] 51/51 [00:02<00:00, 21.12it/s]

Рис. 11. Характеристики модели SGDClassifier с word hashing

Fig. 11. Characteristics of the SGDClassifier model with word hashing

classifier rf_tfidf	
Train [1 / 5]	Accuracy = 1.09%, Precision = 0.15%, Recall = 1.49%, F1 = 0.19%: 100% [██████████] 459/459 [2:56:29<00:00, 23.07s/it]
Val [1 / 5]	Accuracy = 1.22%, Precision = 0.23%, Recall = 1.48%, F1 = 0.24%: 100% [██████████] 51/51 [38:14<00:00, 44.99s/it]
Train [2 / 5]	Accuracy = 1.08%, Precision = 0.15%, Recall = 1.49%, F1 = 0.19%: 100% [██████████] 459/459 [8:34:57<00:00, 67.31s/it]
Val [2 / 5]	Accuracy = 1.10%, Precision = 0.21%, Recall = 1.45%, F1 = 0.22%: 100% [██████████] 51/51 [1:17:48<00:00, 91.54s/it]
Train [3 / 5]	Accuracy = 1.10%, Precision = 0.18%, Recall = 1.50%, F1 = 0.21%: 100% [██████████] 459/459 [10:34:42<00:00, 43.07s/it]
Val [3 / 5]	Accuracy = 1.23%, Precision = 0.25%, Recall = 1.50%, F1 = 0.26%: 100% [██████████] 51/51 [2:58:33<00:00, 34.99s/it]
Train [4 / 5]	Accuracy = 1.09%, Precision = 0.15%, Recall = 1.49%, F1 = 0.19%: 100% [██████████] 459/459 [14:21:37<00:00, 27.31s/it]
Val [4 / 5]	Accuracy = 1.11%, Precision = 0.24%, Recall = 1.48%, F1 = 0.24%: 100% [██████████] 51/51 [4:32:11<00:00, 61.54s/it]
Train 5 / 5]	Accuracy = 1.11%, Precision = 0.21%, Recall = 1.54%, F1 = 0.23%: 100% [██████████] 459/459 [18:55:12<00:00, 47.31s/it]
Val [5 / 5]	Accuracy = 1.13%, Precision = 0.28%, Recall = 1.51%, F1 = 0.27%: 100% [██████████] 51/51 [6:32:11<00:00, 44.32s/it]

Рис. 12. Характеристики модели StreamingRFC с TFIDF

Fig. 12. Characteristics of the StreamingRFC model with TFIDF

<u>classifier_rf hashing</u>	
Train [1 / 5]	Accuracy = 1.11%, Precision = 0.12%, Recall = 1.47%, F1 = 0.17%: 100% [██████████] 459/459 [4:01:31<00:00, 31.57s/it]
Val [1 / 5]	Accuracy = 1.15%, Precision = 0.19%, Recall = 1.53%, F1 = 0.23%: 100% [██████████] 51/51 [53:37<00:00, 63.09s/it]
Train [2 / 5]	Accuracy = 1.12%, Precision = 0.12%, Recall = 1.48%, F1 = 0.18%: 100% [██████████] 459/459 [12:02:32<00:00, 94.45s/it]
Val [2 / 5]	Accuracy = 1.12%, Precision = 0.18%, Recall = 1.50%, F1 = 0.22%: 100% [██████████] 51/51 [1:48:05<00:00, 127.18s/it]
Train [3 / 5]	Accuracy = 1.12%, Precision = 0.14%, Recall = 1.50%, F1 = 0.21%: 100% [██████████] 459/459 [23:32:26<00:00, 27.44s/it]
Val [3 / 5]	Accuracy = 1.17%, Precision = 0.22%, Recall = 1.64%, F1 = 0.33%: 100% [██████████] 51/51 [2:35:13<00:00, 73.22s/it]
Train [4 / 5]	Accuracy = 1.12%, Precision = 0.12%, Recall = 1.48%, F1 = 0.18%: 100% [██████████] 459/459 [1:16:29:32<00:00, 65.48s/it]
Val [4 / 5]	Accuracy = 1.16%, Precision = 0.24%, Recall = 1.58%, F1 = 0.24%: 100% [██████████] 51/51 [1:48:05<00:00, 127.18s/it]
Train [5 / 5]	Accuracy = 1.14%, Precision = 0.18%, Recall = 1.61%, F1 = 0.31%: 100% [██████████] 459/459 [2:11:55:18<00:00, 72.48s/it]
Val [5 / 5]	Accuracy = 1.18%, Precision = 0.32%, Recall = 1.88%, F1 = 0.30%: 100% [██████████] 51/51 [3:21:43<00:00, 87.28s/it]

Рис. 13. Характеристики модели StreamingRFC с word hashing

Fig. 13. Characteristics of the StreamingRFC model with word hashing

ром BertTokenizer импортируется из библиотеки transformers. В качестве метрик качества классификаторов используются Accuracy, Recall, Precision и мера F1. Количество эпох обучения для каждого классификатора – 10.

Результаты обучения моделей классификаторов MultinomialNB, SGDClassifier, StreamingRFC, с векторизаторами TFIDF и word hashing на пяти эпохах с указанными метриками качества представлены на рис. 8–13. Маркером и шрифтом выделены лучшие значения на трейне и валидации.

Лучшие метрики обучения нейросетевой модели классификатора на основе ruBert-base соответствуют седьмой эпохе, они представлены на рис. 14–17.

Лучшей среди четырёх представленных выше моделей классификаторов оказалась модель классификатора ruBert-base. В сравнении с моделью классификатора на основе наивного байеса с векторизатором word hashing на валидационной выборке (см. табл. 1), она показала результаты в среднем на 15 % выше.

Таблица 1

Сравнение MultinomialNB + word hashing с ruBert-base

Table 1

Comparison of MultinomialNB + word hashing with ruBert-base

		Accuracy	Precision	Recall	F1
MultinomialNB c word hashing	Train	0.5588	0.4553	0.4508	0.4244
	Val	0.4662	0.4117	0.4032	0.3817
ruBert-base	Train	1	1	1	1
	Val	0.7141	0.5802	0.5815	0.5760

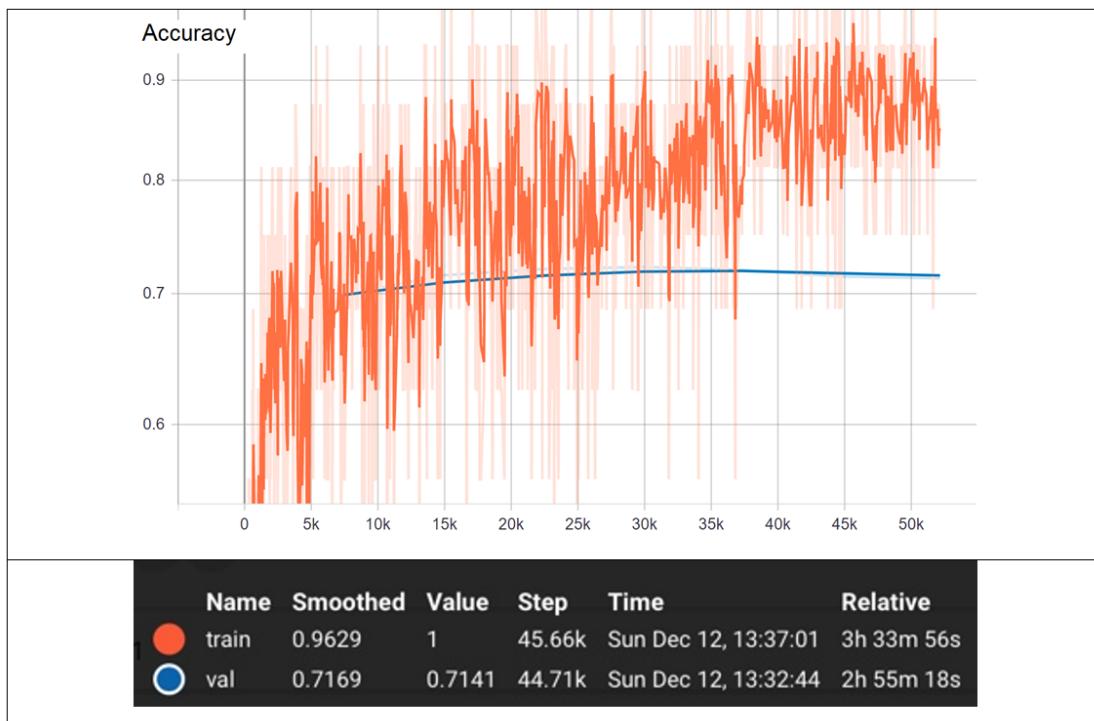


Рис. 14. Метрика Accuracy модели классификатора ruBert-base

Fig. 14. Accuracy metric of ruBert-base classifier model

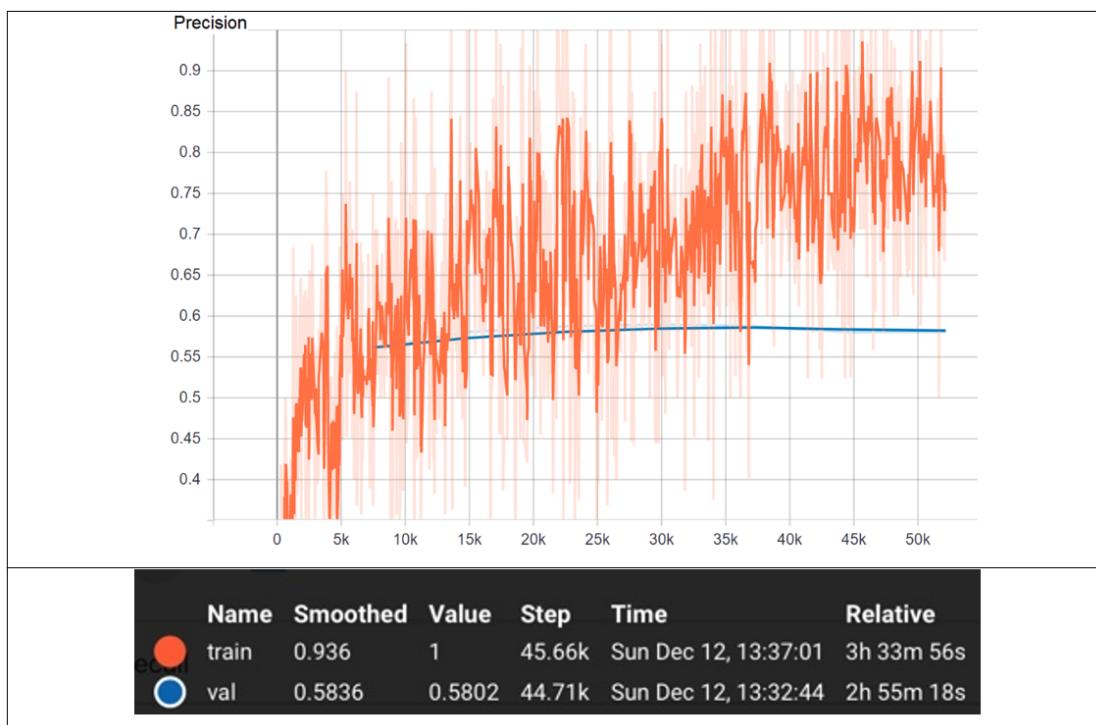


Рис. 15. Метрика Precision модели классификатора ruBert-base

Fig. 15. The Precision metric of the ruBert-base classifier model

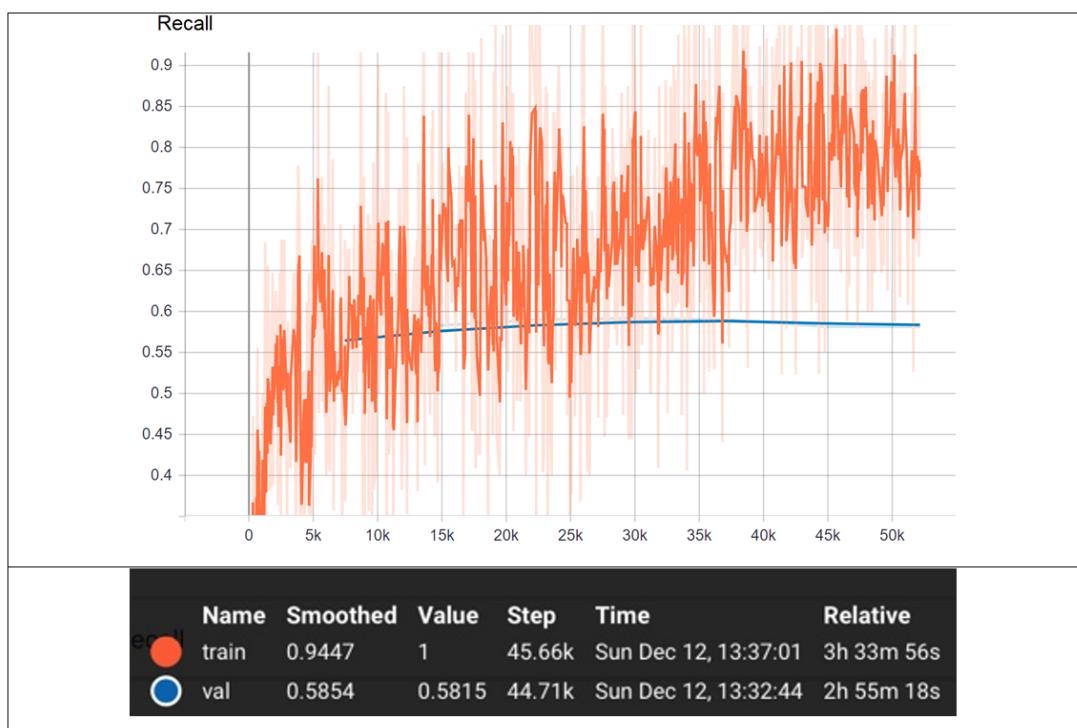


Рис. 16. Метрика Recall модели классификатора ruBert-base

Fig. 16. Recall metric of ruBert-base classifier model

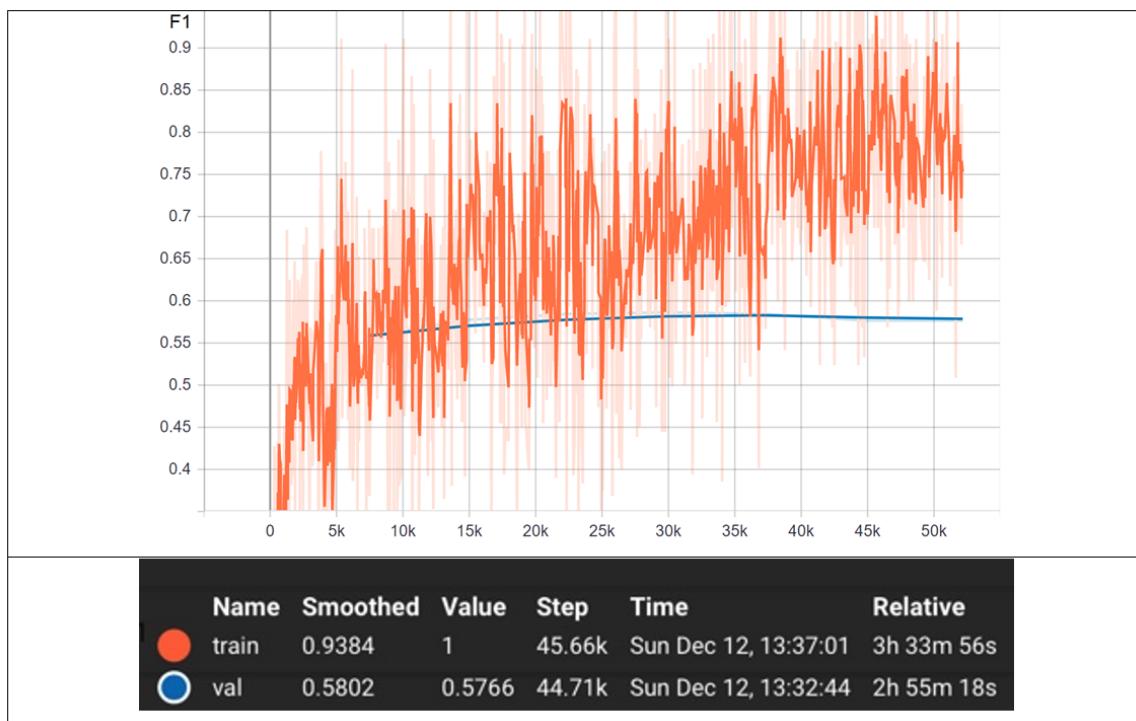


Рис. 17. Метрика F1 модели классификатора ruBert-base

Fig. 17. F1 metric of ruBert-base classifier model

Худшие результаты классификации показал алгоритм случайного леса, основные метрики качества не превысили 1 %. Это можно объяснить тем, что при решении задач многоклассовой классификации с большим количеством классов и обучающей выборкой случайный лес использует множество деревьев: каждому дереву в лесу передаются одни и те же входные данные, на основании которых оно должно вернуть свое предсказание. После чего также происходит голосование на полученных прогнозах. Весь этот процесс занимает много времени. Процесс разделения классов повторяется с использованием «жадной», рекурсивной процедуры, пока дерево не достигнет максимальной глубины, или в каждом узле не останутся только образцы одного класса. Выращивание деревьев такой размерности затратная процедура как по времени, так и по вычислительной мощности. Что подтвердилось в ходе обучения: уже на второй эпохе алгоритм случайного леса с векторизатором word hashing затратил на обучение более 12 часов.

Тестирование классификаторов в режиме предсказывания класса заявки в соответствии с первым уровнем классификатора ОКПД2 (см. табл. 2) показало, что лучшие результаты продемонстрировали наивный байесовский классификатор и ruBert-base.

Таблица 2
Результаты тестирования классификаторов в режиме предсказывания кода заявки из ОКПД2
Table 2
Results of testing classifiers in the order code prediction mode from OKPD2

Текст	SVM		MultinomialNB		StreamingRFC		ruBert-base	ОКПД 2
	TFIDF	W-hash	TFIDF	W-hash	TFIDF	W-hash		
Текущий ремонт помещений	27	45	41	43	49	49	43	43
Поставка песка и гравия природного	28	14	8	8	49	62	8	8
Закупка персональных компьютеров	27	14	26	26	49	49	26	26
Закупка медицинских масок	32	13	14	32	49	80	14	32,21, (14)
На оказание услуг автомойки для автомобилей Федерального бюджетного учреждения здравоохране- ния «Центр гигиены и эпидемиологии в Камчатском крае»	74	71	20	45	62	80	45	45

Таким образом, на основе сравнительного анализа определено, что для рекомендательной системы по предсказанию кодов классификатора ОКПД 2 по тексту заявки целесообразнее использовать модели многоклассовой классификации на основе наивного байесовского алгоритма с векторизатором word hashing и нейросетевую модель ruBert-base.

Выводы

Осуществлена разработка и проведено сравнение моделей многоклассовых классификаторов для построения рекомендательной системы, используемой при подготовке заявок на портале ЕИС в сфере закупок <https://zakupki.gov.ru/>. Модели классификаторов обучались на данных, собранных с портала ЕИС.

В ходе сравнительного анализа хорошие результаты продемонстрировали модель наивного байесовского классификатора с векторизатором word hashing и модель ruBert-base. Метрики ка-

чества на модели наивного байесовского классификатора на валидационной выборке составили: accuracy = 0.4662, precision = 0.4117, recall = 0.4032, F1 = 0.3817. Метрики на модели ruBert-base составили: accuracy = 0.7141, precision = 0.5802, recall = 0.5817, F1 = 0.5760. В результате тестирования классификаторов в режиме предсказывания кода заявки из ОКПД 2 на наборе из пяти примеров классификатор на основе наивного байеса с векторизатором word hashing отработал без ошибок, классификатор ruBert-base на четвертом примере определил класс неточно. Исходя из результатов сравнительного анализа, данные две модели классификаторов предлагается использовать для построения рекомендательной системы.

В дальнейшем планируется рассмотреть широкий класс нейросетевых моделей трансформерной архитектуры, такие как gpt [30], gpt-2 [31], gpt-3 [32], roBERT [33], megatronBert [34] и др. Также планируется реализовать все уровни классификатора ОКПД2.

Подобные подходы позволяют расширить функции существующих электронных торговых площадок в части развития широкого класса рекомендательных сервисов [35, 36] для интеллектуального анализа тендерной документации в сфере закупок.

СПИСОК ЛИТЕРАТУРЫ

1. **Мамедова Н.А., Александрова К.С., Позднякова Н.А.** Анализ информационного обеспечения контрактной системы на примере официального сайта ЕИС в сфере закупок // Internat. J. of Professional Science. 2018. No. 4. Pp. 24–34.
2. **Rashid A.M., Albert I., Cosley D., Lam S.K., McNee S.M., Konstan J.A., et al.** Getting to know you: Learning new user preferences in recommender systems // Proc. of the Internat. Conf. on Intelligent User Interfaces. 2002. Pp. 127–34.
3. **Hinz O., Eckert J.** The impact of search and recommendation systems on sales in electronic commerce // Bus Inf Syst Eng. 2010. no. 2. Pp. 67–77. DOI: 10.1007/s12599-010-0092-x
4. **García Rodríguez M.J., Rodríguez Montequín V., Ortega Fernandez F., Villanueva Balsera J.M.** Bidders recommender for public procurement auctions using machine learning: Data analysis, algorithm, and case study with tenders from Spain // Complexity. 2020. DOI: 10.1155/2020/8858258
5. **Sunmola Funlade, Shehu Yusuf.** A case study on performance features of electronic tendering systems // Procedia Manufacturing. 2020. Vol. 51. Pp. 1586–1591. DOI: 10.1016/j.promfg.2020.10.221
6. **Mehrabd A., Grilo A.** Advanced engineering informatics tender calls search using a procurement product named entity recognizer // Advanced Engineering Informatics. 2018. Vol. 36.
7. **García Rodríguez M.J., Rodríguez Montequín V., Ortega Fernandez F., Villanueva Balsera J.M.** Public procurement announcements in Spain: Regulations, data analysis, and award price estimator using machine learning // Complexity. 2019.
8. **Мигунов В.В.** Проблема идентификации отраслевого рынка в ОКВЭД: методический подход // Конкурентоспособность территорий. Матер. XXIV Всерос. экономического форума молодых ученых и студентов. В 4-х ч. Екатеринбург, 2021. С. 64–66.
9. **Молчанов А.Н., Рейт Д.В., Смирнов С.О.** Анализ проблемного поля основных общероссийских классификаторов технико-экономической информации (ОКВЭД 2, ОКПД 2, ОКОФ) // Управление развитием крупномасштабных систем MLSD'2017. Матер. X междунар. конф. в 2-х т. Институт проблем управления им. В.А. Трапезникова; РАН, 2017. С. 358–361.
10. **Игнатова Г.В., Иноzemцева В.В.** Информационное обеспечение закупочной деятельности в цифровой экономике // Вестник Саратовского государственного социально-экономического университета. 2019. № 3 (77). С. 15–19.

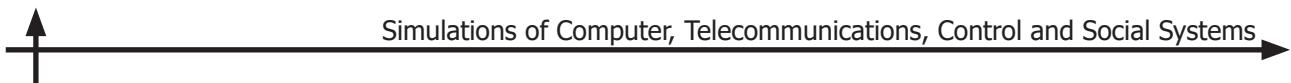
11. Горюнова Н.Д., Громов Д.Н., Кузьмин А.Е. Организация внутреннего финансового контроля государственными заказчиками при планировании закупок товаров, работ и услуг // Управленческое консультирование. 2019. № 5 (125). С. 42–52. DOI: 10.22394/1726-1139-2019-5-42-52
12. Schoenherr T., Tummala V. Electronic procurement: A structured literature review and directions for future research // Internat. J. of Procurement Management. 2007. no. 1. Pp. 8–37. DOI: 10.1504/IJPM.20-07.015353
13. Masudin I., Aprilia G.D., NugrahaA., Restuputri D.P. Impact of e-procurement adoption on company performance: Evidence from Indonesian manufacturing industry // Logistics. 2021. no. 5. P. 16. <https://doi.org/10.3390/logistics5010016>
14. Milano S., Taddeo M., Floridi L. Recommender systems and their ethical challenges // AI & Soc. 2020. no. 35. Pp. 957–967. <https://doi.org/10.1007/s00146-020-00950-y>
15. Isinkaye F.O., Folajimi Y.O., Ojokoh B.A. Recommendation systems: Principles, methods and evaluation // Egyptian Informatics J. 2015. Vol. 16(3). Pp. 261–273. DOI: 10.1016/j.eij.2015.06.005
16. Valdiviezo-Diaz P., Ortega F., Cobos E., Lara-Cabrera R. A collaborative filtering approach based on naive Bayes classifier // IEEE Access. 2019. Vol. 7. Pp. 108581–108592.
17. Ajesh A., Nair J., Jijin P.S. A random forest approach for rating-based recommender system // 2016 Internat. Conf. on Advances in Computing, Communications and Informatics (ICACCI). DOI: 10.1109/icacci.2016.7732225
18. Sun F., Liu J., Wu J., Pei C., Lin X., Ou W., Jiang P. BERT4Rec // Proc. of the 28th ACM Internat. Conf. on Information and Knowledge Management – CIKM’19. DOI: 10.1145/3357384.3357895
19. Ziegler C.N., McNee S.M., Konstan J.A., Lausen G. Improving recommendation lists through topic diversification // Proc. of the 14th Internat. Conf. on World Wide Web. 2005. Pp. 22–32.
20. Seliverstov Y.A., Komissarov A.A., Poslovskaia E.D., Lesovodskaya A.A., Podtikhov A.V. Detection of low-toxic texts in similar sets using a modified XLM-RoBERTa neural network and toxicity confidence parameters // XXIV Internat. Conf. on Soft Computing and Measurements (SCM). 2021. Pp. 161–164. DOI: 10.1109/SCM52931.2021.9507117
21. Zueva N., Kabirova M., Kalaidin P. Reducing unintended identity bias in Russian hate speech detection // Proc. of the 4th Workshop on Online Abuse and Harms. 2010. Pp. 65–69. DOI: 10.18653/v1/P17, arXiv preprint arXiv:2010.11666
22. Smetanin S. Toxic comments detection in Russian // Proc. of the Internat. Conf. on Computational Linguistics and Intellectual Technologies. Dialogue 2020. DOI: 10.28995/NNNN-NNNN/2020-19-1-11
23. Seliverstov Y.A., Seliverstov S.A., Malygin I.G., Korolev O. Traffic safety evaluation in Northwestern Federal District using sentiment analysis of Internet users' reviews // Transportation Research Procedia. 2020. no. 50. Pp. 626–635. DOI: 10.1016/j.trpro.2020.10.074
24. Селиверстов Я.А., Чигур В.И., Сазанов А.М., Селиверстов С.А., Свищунова А.С. Разработка системы для тонового анализа отзывов пользователей портала “AUTOSTRADA.INFO/RU” // Труды СПИИРАН. 2019.18:2. С. 354–389.
25. Wang Qing, Peng Rongqun, Wang Jiaqiang, Xie Yushu, Zhou Yangfan. Research on text classification method of LDA- SVM based on PSO optimization. 2019. Pp. 1974–1978. DOI: 10.1109/CAC48633.2019.8996952
26. Двойникова А.А., Карпов А.А. Аналитический обзор подходов к распознаванию тональности русскоязычных текстовых данных // Информационно-управляющие системы. 2020. № 4. С. 20–30. DOI: 10.31799/1684-8853-2020-4-20-30
27. Kuratov Y., Arkhipov M. Adaptation of deep bidirectional multilingual transformers for Russian language // Computational Linguistics and Intellectual Technologies: Papers from the Annual Internat. Conf. Dialogue. RSUH, Moscow, 2019. Pp. 333–340.
28. Huang Po-Sen, He Xiaodong, Gao Jianfeng, Deng Li, Acero Alex, Heck Larry. Learning deep structured semantic models for web search using clickthrough data. 2013. Pp. 2333–2338. DOI: 10.1145/2505515.2505665

29. **Usama Khalid, Mirza Omer Beg, Muhammad Umair Arshad.** RUBERT: A bilingual Roman Urdu BERT using cross lingual transfer learning // arXiv:2102.11278v1 [cs.CL] 22 Feb 2021.
30. **Radford A., Narasimhan K., Salimans T., Sutskever I.** Improving language understanding by generative pre-training // URL: https://cdn.openai.com/research-covers/language-unsupervised/language_understanding_paper.pdf, 2018.
31. **Vig J., Belinkov Y.** Analyzing the structure of attention in a transformer language model. 2019. Pp. 63–76. DOI: 10.18653/v1/W19-4808
32. **Dehouche Nassim.** Plagiarism in the age of massive Generative Pre-trained Transformers (GPT-3): The best time to act was yesterday. The next best time is now // Ethics in Science and Environmental Politics. 2021. P. 21. DOI: 10.3354/esep00195
33. **Liu Yinhan, Ott Myle, Goyal Naman, Du Jingfei, Joshi Mandar, Chen Danqi, Levy Omer, Lewis Mike, Zettlemoyer Luke, Stoyanov Veselin.** RoBERTa: A robustly optimized BERT pretraining approach. 2019.
34. **Shoeybi Mohammad, Patwary Md. Mostofa Ali, Puri Raul, Legresley Patrick, Casper Jared, Catan-zaro Bryan.** Megatron-LM: Training multi-billion parameter language models using model parallelism. 2019 // URL: <https://deepai.org/publication/megatron-lm-training-multi-billion-parameter-language-models-using-model-parallelism>
35. **Allal-Chérif Oihab, Virginia Simón-Moya, Ballester Antonio C.** Intelligent purchasing: How artificial intelligence can redefine the purchasing function // J. of Business Research. 2021. no. 124. Pp. 69–76. DOI: 10.1016/j.jbusres.2020.11.050
36. **Milana Carlo, Ashta Arvind.** Artificial intelligence techniques in finance and financial markets: A survey of the literature // Strategic Change. 2021. no. 30. Pp. 189–209. DOI: 10.1002/jsc.2403

REFERENCES

- Mamedova N.A., Aleksandrova K.S., Pozdnyakova N.A.** Analiz informatsionnogo obespecheniya kontraktnoy sistemy na primere Ofitsialnogo sayta YeIS v sfere zakupok. *International Journal of Professional Science*, 2018, no. 4, Pp. 24–34. (rus)
- Rashid A.M., Albert I., Cosley D., Lam S.K., McNee S.M., Konstan J.A., et al.** Getting to know you: Learning new user preferences in recommender systems. *Proceedings of the International Conference on Intelligent User Interfaces*, 2002, Pp. 127–34.
- Hinz O., Eckert J.** The impact of search and recommendation systems on sales in electronic commerce. *Bus Inf Syst Eng*, 2010, no. 2, Pp. 67–77. DOI: 10.1007/s12599-010-0092-x
- Garcia Rodriguez M.J., Rodriguez Montequin V., Ortega Fernandez F., Villanueva Balsara J.M.** Bidders recommender for public procurement auctions using machine learning: Data analysis, algorithm, and case study with tenders from Spain. *Complexity*, 2020. DOI: 10.1155/2020/8858258
- Sunmola Funlade, Shehu, Yusuf.** A case study on performance features of electronic tendering systems. *Procedia Manufacturing*, 2020, Vol. 51, Pp. 1586–1591. DOI: 10.1016/j.promfg.2020.10.221
- Mehrbod A., Grilo A.** Advanced engineering informatics tender calls search using a procurement product named entity recognizer. *Advanced Engineering Informatics*, 2018, Vol. 36.
- Garcia Rodriguez M.J., Rodriguez Montequin V., Ortega Fernandez F., Villanueva Balsara J.M.** Public procurement announcements in Spain: Regulations, data analysis, and award price estimator using machine learning. *Complexity*, 2019.
- Migunov V.V.** Problema identifikatsii otrslevogo rynka v OKVED: metodicheskiy podkhod. *Konkuren-tosposobnost territoriy. Materialy XXIV Vserossiyskogo ekonomicheskogo foruma molodykh uchenykh i studentov*. Yekaterinburg, 2021. Pp. 64–66. (rus)
- Molchanov A.N., Reut D.V., Smirnov S.O.** Analiz problemnogo polya osnovnykh obshcherossiyskikh klassifikatorov tekhniko-ekonomiceskoy inorfomatsii (OKVED 2, OKPD 2, OKOF). *Upravleniye razvitiyem*

- krupnomasshtabnykh sistem MLSD'2017. Materialy X mezhdunarodnoy konferentsii.* Institut problem upravleniya im. V.A.Trapeznikova; RAS, 2017. Pp. 358–361. (rus)
10. **Ignatova G.V., Inozemtseva V.V.** Informatsionnoye obespecheniye zakupochnoy deyatelnosti v tsifrovoy ekonomike. *Vestnik Saratovskogo Gosudarstvennogo Sotsialno-Ekonomicheskogo Universiteta*, 2019, no. 3 (77), Pp. 15–19. (rus)
11. **Goryunova N.D., Gromov D.N., Kuzmin A.Ye.** Organizatsiya vnutrennego finansovogo kontrolya gosudarstvennymi zakazchikami pri planirovaniy zakupok tovarov, rabot i uslug. *Upravlencheskoye Konsultirovaniye*, 2019, no. 5 (125), Pp. 42–52. (rus). DOI: 10.22394/1726-1139-2019-5-42-52
12. **Schoenherr T., Tummala V.** Electronic procurement: A structured literature review and directions for future research. *International Journal of Procurement Management*, 2007, no. 1, Pp. 8–37. DOI: 10.1504/IJPM.2007.015353
13. **Masudin I., Aprilia G.D., Nugraha A., Restuputri D.P.** Impact of e-procurement adoption on company performance: Evidence from Indonesian manufacturing industry. *Logistics*, 2021, no. 5, P. 16. <https://doi.org/10.3390/logistics5010016>
14. **Milano S., Taddeo M., Floridi L.** Recommender systems and their ethical challenges. *AI & Soc.*, 2020, no. 35, Pp. 957–967. <https://doi.org/10.1007/s00146-020-00950-y>
15. **Isinkaye F.O., Folajimi Y.O., Ojokoh B.A.** Recommendation systems: Principles, methods and evaluation. *Egyptian Informatics Journal*, 2015, Vol. 16(3), Pp. 261–273. DOI: 10.1016/j.eij.2015.06.005
16. **Valdiviezo-Diaz P., Ortega F., Cobos E., Lara-Cabrera R.** A collaborative filtering approach based on naive Bayes classifier. *IEEE Access*, 2019, Vol. 7, Pp. 108581–108592.
17. **Ajesh A., Nair J., Jijin P.S.** A random forest approach for rating-based recommender system. *2016 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*. DOI: 10.1109/icacci.2016.7732225
18. **Sun F., Liu J., Wu J., Pei C., Lin X., Ou W., Jiang P.** BERT4Rec. *Proceedings of the 28th ACM International Conference on Information and Knowledge Management – CIKM'19*. DOI: 10.1145/3357384.3357895
19. **Ziegler C.N., McNee S.M., Konstan J.A., Lausen G.** Improving recommendation lists through topic diversification. *Proceedings of the 14th International Conference on World Wide Web*, 2005, Pp. 22–32.
20. **Seliverstov Y.A., Komissarov A.A., Poslovskaya E.D., Lesovodskaya A.A., Podtikhov A.V.** Detection of low-toxic texts in similar sets using a modified XLM-RoBERTa neural network and toxicity confidence parameters. *XXIV International Conference on Soft Computing and Measurements (SCM)*, 2021, Pp. 161–164. DOI: 10.1109/SCM52931.2021.9507117
21. **Zueva N., Kabirova M., Kalaidin P.** Reducing unintended identity bias in Russian hate speech detection. *Proceedings of the 4th Workshop on Online Abuse and Harms*, 2010, Pp. 65–69. DOI: 10.18653/v1/P17, arXiv preprint arXiv:2010.11666
22. **Smetanin S.** Toxic comments detection in Russian. *Proceedings of the International Conference on Computational Linguistics and Intellectual Technologies*, Dialogue 2020. DOI: 10.28995/NNNNNNNN/2020-19-1-11
23. **Seliverstov Y.A., Seliverstov S.A., Malygin I.G., Korolev O.** Traffic safety evaluation in Northwestern Federal District using sentiment analysis of Internet users' reviews. *Transportation Research Procedia*, 2020, no. 50, Pp. 626–635. DOI: 10.1016/j.trpro.2020.10.074
24. **Seliverstov Y.A., Chigur V.I., Sazanov A.M., Seliverstov S.A., Svistunova A.S.** Razrabotka sistemy dlya tonovogo analiza otzygov polzovateley portala "AUTOSTRADA.INFO/RU". *Tr. SPIIRAN*, 2019, 18:2, Pp. 354–389. (rus)
25. **Wang Qing, Peng Rongqun, Wang Jiaqiang, Xie Yushu, Zhou Yangfan.** *Research on text classification method of LDA- SVM based on PSO optimization*, 2019, Pp. 1974–1978. DOI: 10.1109/CAC48633.2019.8996952
26. **Dvoynikova A.A., Karpov A.A.** Analiticheskiy obzor podkhodov k raspoznavaniyu tonalnosti russkoyazychnykh tekstovyykh dannykh. *Informatsionno-Upravlyayushchiye Sistemy*, 2020, no. 4, Pp. 20–30. (rus). DOI: 10.31799/1684-8853-2020-4-20-30



27. **Kuratov Y., Arkhipov M.** Adaptation of deep bidirectional multilingual transformers for Russian language. *Computational Linguistics and Intellectual Technologies: Papers from the Annual International Conference “Dialogue”*. RSUH, Moscow, 2019, Pp. 333–340.
28. **Huang Po-Sen, He Xiaodong, Gao Jianfeng, Deng Li, Acero Alex, Heck Larry.** *Learning deep structured semantic models for web search using clickthrough data*, 2013, Pp. 2333–2338. DOI: 10.1145/2505515.2505665
29. **Usama Khalid, Mirza Omer Beg, Muhammad Umair Arshad.** RUBERT: A bilingual Roman Urdu BERT using cross lingual transfer learning. *arXiv:2102.11278v1 [cs.CL]* 22 Feb 2021.
30. **Radford A., Narasimhan K., Salimans T., Sutskever I.** Improving language understanding by generative pre-training. Available: https://cdn.openai.com/research-covers/language-unsupervised/language_understanding_paper.pdf, 2018
31. **Vig J., Belinkov Y.** *Analyzing the structure of attention in a transformer language model*, 2019, Pp. 63–76. DOI: 10.18653/v1/W19-4808
32. **Dehouche Nassim.** Plagiarism in the age of massive generative pre-trained transformers (GPT-3): The best time to act was yesterday. The next best time is now. *Ethics in Science and Environmental Politics*, 2021, P. 21. DOI: 10.3354/esep00195
33. **Liu Yinhan, Ott Myle, Goyal Naman, Du Jingfei, Joshi Mandar, Chen Danqi, Levy Omer, Lewis Mike, Zettlemoyer Luke, Stoyanov Veselin.** *RoBERTa: A robustly optimized BERT pretraining approach*, 2019.
34. **Shoeybi Mohammad, Patwary Md. Mostofa Ali, Puri Raul, Legresley Patrick, Casper Jared, Catanzaro Bryan.** Megatron-LM: Training multi-billion parameter language models using model parallelism. 2019. Available: <https://deepai.org/publication/megatron-lm-training-multi-billion-parameter-language-models-using-model-parallelism>
35. **Allal-Chérif Oihab, Virginia Simón-Moya, Ballester Antonio C.** Intelligent purchasing: How artificial intelligence can redefine the purchasing function. *Journal of Business Research*, 2021, no. 124, Pp. 69–76. DOI: 10.1016/j.jbusres.2020.11.050
36. **Milana Carlo, Ashta Arvind.** Artificial intelligence techniques in finance and financial markets: A survey of the literature. *Strategic Change*, 2021, no. 30, Pp. 189–209. DOI: 10.1002/jsc.2403

INFORMATION ABOUT AUTHORS / СВЕДЕНИЯ ОБ АВТОРАХ

Селиверстов Ярослав Александрович
Yaroslav A. Seliverstov
E-mail: Y.Seliverstov@2035.university

Комиссаров Андрей Александрович
Andrey A. Komissarov
E-mail: Andrew.Komissarov@gmail.com

Лесоводская Алина Алексеевна
Alina A. Lesovodskaya
E-mail: a.lesovodskaya@2035.university

Бовыкин Павел Геннадьевич
Pavel G. Bovykin
E-mail: p.bovykin@2035.university

Подтихов Артур Владимирович
Artur V. Podtikhov
E-mail: a.podtikhov@2035.university

Торсионов Станислав Сергеевич
Stanislav S. Torsionov
E-mail: s.torsionov@2035.university

Цирков Дмитрий Алексеевич
Dmitriy A. Tsyrkov
E-mail: tsyrkov@2035.university

Орлов Сергей Александрович
Sergei A. Orlov
E-mail: orlov@ftf.tsu.ru

*Поступила: 08.04.2022; Одобрена: 23.08.2022; Принята: 25.08.2022.
Submitted: 08.04.2022; Approved: 23.08.2022; Accepted: 25.08.2022.*



Научная статья

DOI: <https://doi.org/10.18721/JCSTCS.15205>

УДК 517.938:070



МОДЕЛИРОВАНИЕ ИНФОРМАЦИОННОГО ПРОТИВОБОРСТВА: НАПРАВЛЕНИЯ ИССЛЕДОВАНИЙ И МАТЕМАТИЧЕСКИЕ ИНСТРУМЕНТЫ

С.В. Тимофеев¹✉, А.В. Баенхаева²

^{1,2} Байкальский государственный университет,
г. Иркутск, Российская Федерация

✉ timofeevsv12@gmail.com

Аннотация. Представлен обзор результатов первого этапа исследования, цель которого – освоение новых возможностей при изучении такой сложной по структуре системы, как средства массовой информации (СМИ). Успехи в развитии фундаментальной математики и, как неизбежное следствие, прорыв в области цифровых технологий, позволяют активно участвовать в совершенствовании методов получения, обработки, хранения и распространения информации – то, чем традиционно занимаются СМИ. Поэтому логичной кажется идея использовать математический аппарат для моделирования некоторых звеньев коммуникации с аудиторией. С применением методов теории динамических систем описана стадия распространения через СМИ информации, направленной на продвижение в общество новой системы взглядов, и сопровождающее этот процесс информационное противоборство. С этой целью выделен ряд параметров, с помощью которых можно оценить реакцию аудитории на появление новостей. В зависимости от соотношения этих параметров представлены сценарии дальнейшего распространения появившейся информации и сделаны заключения о готовности общества к смене имеющихся концепций.

Ключевые слова: математическая модель, дифференциальные уравнения, продвижение информации, информационное противоборство, альтернативные точки зрения

Для цитирования: Тимофеев С.В., Баенхаева А.В. Моделирование информационного противоборства: направления исследований и математические инструменты // Computing, Telecommunications and Control. 2022. Т. 15, № 2. С. 63–75. DOI: 10.18721/JCSTCS.15205



Research article

DOI: <https://doi.org/10.18721/JCSTCS.15205>

UDC 517.938:070



MODELING OF INFORMATION CONFRONTATION: RESEARCH DIRECTIONS AND MATHEMATICAL TOOLS

S.V. Timofeev¹✉, A.V. Baenkhaeva²

^{1,2} Baikal State University, Irkutsk, Russian Federation

✉ timofeevsv12@gmail.com

Abstract. The article provides an overview of the first stage study results, the purpose of which is to master new opportunities in the study of such a complex structure of a system as the mass media (mass media). Success in the fundamental mathematics development and, as an inevitable consequence, a breakthrough in the field of digital technologies, make it possible to participate actively in the methods' improving of obtaining, processing, storing and distributing information to its consumers – something that the media traditionally do. Therefore, the idea of mathematical apparatus using for modeling some communication links with the audience seems logical. Using the methods of the dynamic systems theory, the stage of dissemination through the information media aimed at promoting a new system of views into society, and the informational confrontation accompanying this process, are described. For this purpose, a number of parameters have been identified, that makes possible to assess the audience reaction to the news appearance. Depending on these parameters' ratio, scenarios for the further information dissemination that has appeared are presented and conclusions about the society's readiness to existing concepts change are drawn.

Keywords: mathematical model, differential equations, dissemination of new information, information confrontation, alternative view

Citation: Timofeev S.V., Baenkhaeva A.V. Modeling of information confrontation: Research directions and mathematical tools. Computing, Telecommunications and Control, 2022, Vol. 15, No. 2, Pp. 63–75. DOI: 10.18721/JCSTCS.15205

Введение

В современном мире стремительный рост цифровых технологий привел к мощной трансформации системы массмедиа [1]. Развитые изменения таких характеристик, как доступность, скорость, сила влияния привела к тому, что СМИ стали мощнейшим инструментом воздействия на аудиторию. Помимо этого, существенно расширились возможности и инструменты проведения разностороннего анализа происходящих изменений в функционировании самих СМИ. В работе [2], например, изложена мысль о том, что «...средства массовой информации есть не просто сложная структурированная система. По своему исторически пройденному пути и современному состоянию она есть одна из самых динамично развивающихся систем...». Как важный вывод – сформулированная идея о назревшей задаче «разработать модель развития системы СМИ в историческом плане и в плане прогноза ее будущего». Задача описать развитие СМИ как динамической системы показалась любопытной, поскольку это предполагает построение соответствующей математической модели. А это, в свою очередь, предвещает возможное появление новых выводов и точек зрения.

В качестве пробной попытки авторами было предложено рассмотреть СМИ как структуру, развивающуюся эволюционно. С использованием аппарата дифференциальных уравнений удалось описать эволюцию развития основных медиаканалов [3]. Под медиаканалами, опираясь на определения и характеристики СМИ, данные Л.М. Земляновой [4], понималась совокупность



взаимодополняющих однородных источников информации¹. Такого подхода к описанию СМИ при анализе литературных источников найдено не было. Полученная математическая модель оказалась весьма содержательной. В ходе её анализа и наблюдения за современными тенденциями в СМИ было замечено, что определенная доля технологического и рекламного ресурса стала использоваться в сфере, о которой до недавнего времени можно было рассуждать лишь в формате science fiction. Речь шла о бурно развивающихся в настоящее время технологиях виртуальной реальности. Это навело на мысль, что, быть может, уже создаются предпосылки появления качественно нового вида СМИ. И действительно, в междисциплинарном ключе удалось обосновать [5], что такие технологии предоставляют эффективную возможность получать информацию на совершенно новом уровне. Тем самым на основе проведенного анализа был сделан вывод, что мы находимся на начальной стадии формирования новой информационно-коммуникационной среды.

При дальнейшем исследовании пришло понимание того, что для повышения эффективности следует делать акцент на изучении какого-то определенного звена в структуре СМИ. И особенно привлекательной для нас представлялась возможность изучить, как средства массовой информации взаимодействуют с аудиторией.

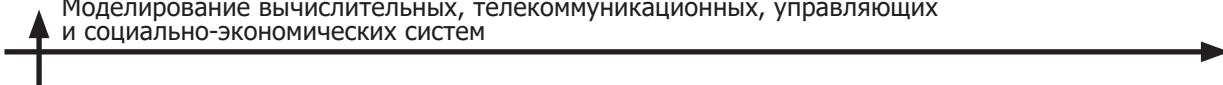
В качестве первого этапа исследования принимается решение смоделировать процесс продвижения новой информации через СМИ в каком-либо сегменте общества или обществе в целом.

Интерес к этому определялся тем, что информация с равным успехом может применяться как для объединения и стабилизации общества, так и для разъединения и дестабилизации его. Поскольку все зависит от целевых установок инициатора информационного воздействия и от потенциала объекта воздействия, который либо желает принять эти установки, либо намерен защитить себя от внешнего информационного «давления» [6]. Успех в продвижении принципиально новых идей в общество во многом зависит от позиций основных действующих сил. С одной стороны, – влиятельных СМИ, обладающих способностью формировать общественное мнение, а, с другой, – таких субъектов общества, как экспертные сообщества, органы исполнительной власти, политические партии, общественные организации, которые имеют возможность задействовать другую часть СМИ для освещения альтернативной точки зрения и «раскручивания» своих концепций в социуме [7]. Здесь мы имеем дело с неким информационным противоборством².

Моделированию этого процесса посвящен ряд заслуживающих внимания работ. Нами в статье [8] проведен детальный обзор некоторых из них. Так, например, в [9] с использованием методики постановки нелинейных задач [10, 11] предложена математическая модель информационного обмена среди граждан в различных общественных структурах. При полученных связях между переменными модели поставлена задача максимального увеличения численности группы заинтересованных в определенной информации людей. В монографии [12] подробно рассмотрены теоретико-игровые модели информационного управления и информационного противоборства в социальных сетях. Системное исследование моделей информационного противоборства в социуме проведено в работах [13–16]. Опираясь на базовую модель [13], исследования развивались по нескольким направлениям. Например, в работах [14, 15] данная модель информационной борьбы рассмотрена в случае, когда одна из противоборствующих сторон периодически увеличивает интенсивность пропаганды на некоторое время, понижая её затем до прежнего уровня. Исследуя в другом направлении, авторы постепенно вводили дополнительные факторы, детализируя описание информационной борьбы. Расширенная модель представляет собой систему из восьми нелинейных обыкновенных дифференциальных уравнений [16]. В каждой из представленных в

¹ В этом смысле все печатные источники – газеты, журналы и прочие печатные издания – могут служить примером медиаканала.

² Под информационным противоборством в данной работе подразумевается взаимодействие нескольких субъектов, обладающих несогласующими интересами и осуществляющих информационное воздействие на один и тот же управляемый субъект.



кратком обзоре работ получены интересные результаты. При этом было замечено, что их объединяет общий подход при моделировании процесса информационного противоборства. Все модели описываются через характеристики численности различных групп реципиентов в качестве фазовых переменных для динамических систем либо – в теоретико-игровых моделях – через количественные показатели игроков и агентов, которые, так или иначе, относятся к одной из конфликтующих в информационном пространстве сторон. В дальнейшем, вероятно, для достоверности каждой из полученных теоретических моделей потребуется подтверждение её адекватности эмпирическими данными, для чего необходимо применение стандартных социологических инструментов в виде опроса выборочной совокупности с последующим использованием традиционных статистических методов для идентификации параметров и анализа данных.

Однако в арсенале ученых появились новые инструменты для исследовательской работы, такие, например, как Data Science – наука о данных, Text Mining – интеллектуальный анализ текста. Открылись и новые возможности. Нами предложен принципиально другой подход для анализа распространения новой информации через СМИ, который не опирается на теорию выборки для изучения общественного мнения, а предполагает работу с большими данными. Социологи, изучающие общественное мнение, заявляют, что присущие Big Data характеристики дают более эффективные возможности для построения прогнозов, чем традиционные методы [17, 18]. Отсутствие выборок ($n = All$) и непосредственная работа с генеральными совокупностями или очень крупными их частями, масштабируемость данных, постоянный автоматизированный сбор данных в архивы и возможность их быстрой обработки приводят в конечном итоге к высокой достоверности и востребованности прогнозирования «в реальном времени». Поэтому, на наш взгляд, более актуально в современных условиях проводить исследование не измерением количества людей, принимающих позицию «за» или «против» определенной точки зрения, а анализом объема и интенсивности поступающей в СМИ информации, направленной на достижение несоппадающих по интересам задач. В данной статье описаны результаты первого этапа системных исследований в этом направлении.

В рамках нового подхода при моделировании процесса движения информационных потоков и сопровождающего его информационного противоборства было решено выделить в качестве ключевых факторов следующие величины, изменяющиеся в течение времени t .

$N(t)$ (от англ. news) – количественная характеристика объема новостной информации, соответствующая продвижению новых взглядов в информационном пространстве. Если рассмотреть СМИ как агрегированный канал, обеспечивающий передачу информации обществу как потребителю, то очевидным становится интерес любых экономически или политически активных сил иметь возможность целенаправленно воздействовать на этот канал, контролируя информационное воздействие на социальную систему.

$C(t)$ (от англ. censorship) – число органов в структуре общества, обладающих властью и информационными ресурсами, цель которых – сохранение ранее принятых в обществе концепций (например, идеологических или технологических). Учитывая всеобъемлющую и неограниченную информатизацию всех сфер общества, становится очевидным, что контроль над системой СМИ в некоторой оперативной перспективе способен оказывать решающее влияние на настроения в отдельных социальных группах, а длительный контроль на основе продуманной стратегии управления способен менять систему ценностей общества в целом. В мировой практике органы, которым поручено осуществлять надзор за процессом регулирования, могут быть самыми разными: в одних странах это правительственные департаменты, в других – регулятивные органы, в третьих – независимые организации [19].

$A(t)$ (от англ. alternative view) – количественная характеристика объема информационного потока (возможно, генерируемая по инициативе органов цензуры), противопоставленного распространению новой концепции в информационном пространстве.

$i(t)$ (от англ. index) – показатель доли населения, лояльно относящейся к новым идеям, появляющимся в СМИ на момент времени t . Данный показатель характеризуется формулой $i = 1 - \frac{I^*}{I}$. Здесь I (%) соответствует доли аудитории, которая полностью принимает сложившиеся в обществе положения до начала наблюдений; I^* (%) – соответствующая характеристика принятия этих положений при распространении в СМИ новых идей.

Основой формализации послужила идея моделирования иммунного ответа организма на проявление вирусной атаки [20]. Опираясь на эту идею, в работе [21] предложена для изучения математическая модель распространения новой информации в обществе, представляющая собой систему обыкновенных дифференциальных уравнений:

$$\begin{aligned} \frac{dN}{dt} &= \beta N - \gamma AN, \\ \frac{dC}{dt} &= \alpha AN - \mu(C - C_*), \\ \frac{dA}{dt} &= \rho C - \eta \gamma AN - \lambda A, \\ \frac{di}{dt} &= \sigma N - \omega i. \end{aligned} \tag{1}$$

Здесь параметр $\beta \geq 0$ показывает, с какой интенсивностью распространяется новая информация через СМИ; параметр $\gamma \geq 0$ характеризует возможность нейтрализации эффекта от появившейся информации после изложения альтернативного мнения. Коэффициент $\alpha \geq 0$ описывает интенсивность реакции на противоборство противоположных точек зрения; параметр $\mu > 0$ обратно пропорционален времени работы дополнительно созданных органов (будем предполагать, что в обществе всегда есть специальный ресурс в количестве C_* для защиты прежней концепции). Средняя скорость появления новостей из одного органа информации C будет характеризовать параметр $\rho \geq 0$, а $\eta \geq 0$ – количество информации A , направленное на нейтрализацию влияния сообщений N ; коэффициент $\lambda > 0$ обратно пропорционален времени забывания информации A . Параметры $\sigma \geq 0$, $\omega \geq 0$ характеризуют соответственно темп принятия новой идеи и возврат в силу инерции мышления к прежней концепции.

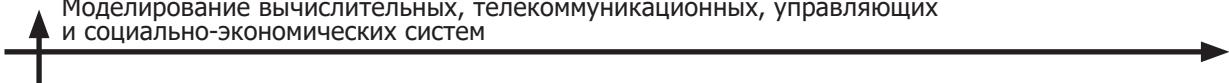
Безусловно, предложенная математическая модель не учитывает абсолютно все тонкости и детали при описании процесса распространения новой информации в обществе посредством СМИ. Однако этот обобщенный вид модели позволяет связать в систему основные факторы и помогает глубже понять процесс информационного противоборства.

Результаты исследования системы (1)

Цель данной работы – избегая, по возможности, строгих математических формулировок, соединить воедино и систематизировать все новые результаты исследования системы (1), а также дать четкую интерпретацию влияющих на поведение системы соотношений параметров.

При изучении предложенной модели предполагалось, что в обществе (или его сегменте) до некоторого момента преобладает определенная концепция (например, идеологическая или технологическая). Для системы (1) эта ситуация определена как точка покоя:

$$X_{1st} = (N_{1st}, C_{1st}, A_{1st}, i_{1st}) = \left(0, C_*, \frac{\rho C_*}{\lambda}, 0 \right).$$



При трактовке в [21] этого стационарного решения подразумевалось, что для поддержки сформированной системы ценностей административный ресурс в количестве $C = C_*$ задействует в СМИ достаточное с его точки зрения количество информации $A = \frac{\rho C_*}{\lambda}$.

Неожиданное появление в средствах массовой информации новостей N , расходящихся с системой устоявшихся взглядов, способно вызвать в обществе разную реакцию в зависимости от его готовности воспринимать или не воспринимать новую информацию. Для модели эти случаи определяются разными значениями параметров системы (1). Используя метод функции Ляпунова и другие качественные методы исследования дифференциальных уравнений, показано и математически обосновано [21–23], что решения системы в зависимости от соотношения параметров имеют существенные различия в поведении. При этом, принимая во внимание содержательный смысл решений системы (1), выделен ряд основных сценариев поведения общества.

Поведение решений системы (1) в случае наличия двух стационарных решений

В начале исследования из всего пространства параметров данной системы определились несколько важных соотношений [21], которые были сгруппированы в две области –

$\Omega_1 : \begin{cases} \gamma\rho C_* > \lambda\beta \\ \mu\gamma > \alpha\rho \end{cases}$ и $\Omega_2 : \begin{cases} \gamma\rho C_* < \lambda\beta \\ \mu\gamma < \alpha\rho \end{cases}$. При этих значениях параметров у системы (1) появляется

ещё одно стационарное решение или точка покоя $X_{2st} = (N_{2st}, C_{2st}, A_{2st}, i_{2st})$, охарактеризованное в [21] как толерантное состояние общества, где уживаются различные системы взглядов, поддерживаемые своей частью аудитории. Таким образом, определена потенциальная возможность компромисса между различными мнениями и идеями. Будет он найден или нет, зависит от реакции общества на публикуемые в СМИ новости, и, соответственно, от соотношений параметров, которыми описывается эта реакция. В [24] дана содержательная интерпретация решений системы для параметров из каждой области Ω_1 и Ω_2 в зависимости от количества появившейся в СМИ информации N и информации A , противоположной по смыслу. Геометрическая интерпретация для каждого случая представлена на рис. 1 и 2.

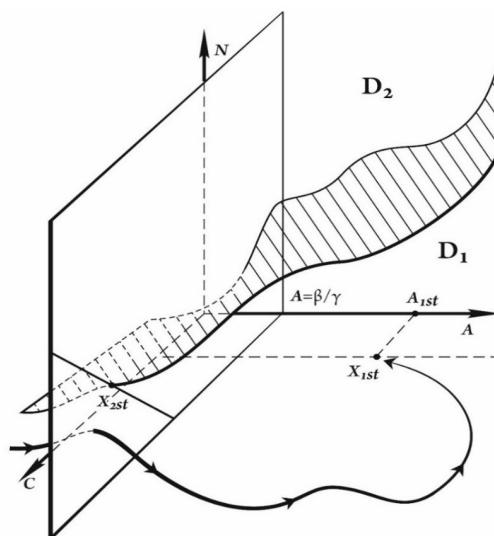


Рис. 1. Поведение решения системы (1) в области параметров Ω_1
Fig. 1. Behavior of the solution of the system (1) in the parameter domain Ω_1

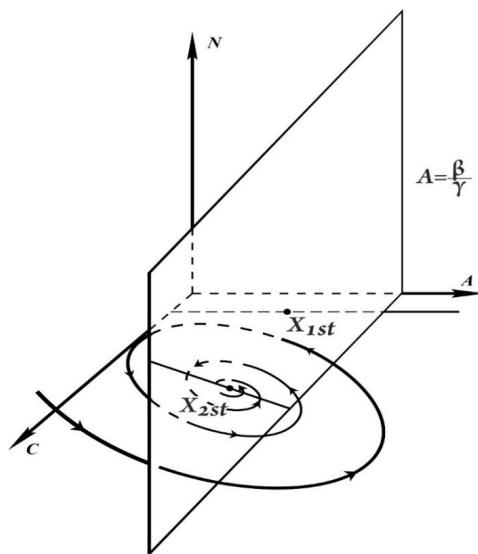


Рис. 2. Поведение решения системы (1) в области параметров Ω_2
Fig. 2. Behavior of the solution of the system (1) in the parameter domain Ω_2

Сценарий 1. Рис. 1 иллюстрирует ситуацию, когда реакция общества на «вброс» новых положений описывается параметрами из области Ω_1 .

На фазовом портрете для переменных C, A, N , имеющих смысл основных факторов при моделировании, изображена область D_1 , где точно известно поведение системы. Если в некоторый момент времени траектория решения будет находиться в области D_1 , то она со временем стремится к точке покоя X_{1st} . Смысл заключается в том, что, если показатели факторов C, A и N численно находятся в области D_1 , то новая идея, внедряемая в общество через СМИ, не имеет поддержки. Постепенно количество информации N , способствующей распространению новой идеи, убывает, и традиционная система взглядов вновь становится доминирующей.

По фактической динамике противоборства можно оценить, в какой области – D_1 или D_2 – находятся значения A, N и C . Каждая соперничающая сторона может определить необходимость и размер изменения количества информации в СМИ для получения нужного ей эффекта, повысив вероятность достижения поставленной цели. Полученный результат показывает, что при грамотном управлении числа и активности органов информационной защиты C теоретически возможно отразить любую нежелательную для социума информационную атаку. При этом со временем традиционные идеи и положения возвращаются на свои позиции.

Сценарий 2. На рис. 2 проиллюстрирована ситуация, когда общество готово к компромиссу.

Пусть информационное противоборство, моделируемое системой (1), характеризуется параметрами из Ω_2 с более сильным условием:

$$\mu\eta\gamma + \beta\eta\gamma < \rho\alpha. \quad (2)$$

Представленные соотношения описывают полную готовность общества воспринимать новые идеи и положения. Декларируемые через СМИ новшества успешно внедряются в устоявшееся мировоззрение и находят отклик. Со временем старые и новые представления приходят к совместному сосуществованию со своими долями приятия в обществе. Геометрически фазовый портрет означает, что все траектории системы (1), со временем уменьшая колебания, стремятся к точке покоя X_{2st} .

При этом в [22] показано, что без условия (2) достичь толерантности к обеим системам взглядов может и не удастся. Та часть аудитории, которая принимает новые взгляды, со временем убедит другую часть в неизбежности смены традиционных устоев, и доля консервативной части населения станет мала.

Поведение решений системы (1) в случае только одного стационарного решения

Продолжая изучать построенную модель, удалось выделить ещё две области параметров – $\Lambda_1 : \begin{cases} \gamma\rho C_* < \lambda\beta \\ \mu\eta\gamma > \alpha\rho \end{cases}$ и $\Lambda_2 : \begin{cases} \gamma\rho C_* > \lambda\beta \\ \mu\eta\gamma < \alpha\rho \end{cases}$. При данных соотношениях вторая точка покоя системы

$X_{2st} = (N_{2st}, C_{2st}, A_{2st}, i_{2st})$ отсутствует. Эти области характеризуют состояние общества, в котором невозможны какие-либо компромиссы. В статье [23] математически обосновано, что в этих случаях реализуются два противоположных сценария.

Сценарий 3. Если появление в СМИ новой неординарной информации вызывает реакцию, описанную параметрами в виде неравенств из Λ_1 , то можно сделать вывод об абсолютной готовности общества к полной смене традиционной системы взглядов. Любое появление в СМИ новых идей и мнений, не совпадающих с общепринятыми, найдет поддержку в обществе. В этом случае осуществляется полная смена адресатами ранее доминирующей концепции. При этом реализация данного сценария может быть осуществлена разными способами, поэтому из-за их разнообразия геометрическая интерпретация для этого случая не дана.

Сценарий 4. Пусть теперь реакция на незнакомую информацию описывается соотношениями, обозначенными в Λ_2 и усиленным условием (2). В этом случае правильным будет заключение о том, что в обществе традиционная концепция занимает прочное положение. Причиной этому могут быть либо полное одобрение происходящих в обществе процессов, либо неготовность к смене устоявшихся взглядов. Также это может происходить из-за высокой эффективности органов цензуры, которые не дают возможности новой информации заполнить информационное пространство. Геометрическая интерпретация этому случаю представлена на рис. 3.

Здесь любое решение системы (1) всегда достаточно быстро приближается к точке покоя X_{1st} .

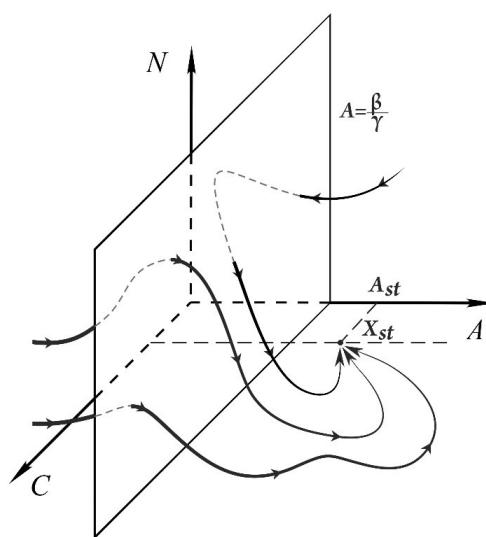


Рис. 3. Поведение решения системы (1) в области параметров Λ_2 при выполнении условия (2)
Fig. 3. Behavior of the solution of the system (1) in the parameter domain Λ_2 when the condition (2) is met

Интерпретация соотношений параметров

Качественно различное поведение решений системы в областях параметров Ω_1 , Ω_2 , Λ_1 и Λ_2 требует глубокого понимания смысла соответствующих соотношений. Правильная трактовка неравенств дает возможность понять причины того или иного сценария распространения новостной информации в обществе.

Рассмотрим неравенство $\gamma\rho C_* < \lambda\beta$, входящее в Ω_2 и Λ_1 . После деления обеих частей данного соотношения на произведение $\lambda\gamma$ будем иметь $\frac{\rho C_*}{\lambda} < \frac{\beta}{\gamma}$. Левая часть, как описано выше при характеристике X_{1st} , имеет смысл информационной поддержки распространенной в обществе системы взглядов. Правая часть – отношение интенсивности распространения новой поступающей информации через СМИ (параметр β) к характеристике нейтрализации её эффекта от изложения противоположной точки зрения (параметр γ). Правая часть тем больше, чем больше интенсивность распространения или меньше вероятность нейтрализации эффекта. Следовательно, неравенство можно трактовать двумя способами. Первый – слабая информационная поддержка традиционных взглядов ввиду неадекватной оценки реальной опасности или ограниченных возможностей (например, финансовых, политических) органов цензуры и СМИ. Второй – достаточно агрессивная и убедительная реклама новых позиций с малой вероятностью эффективных противодействий. Соответственно, обратное неравенство $\gamma\rho C_* > \lambda\beta$ говорит о серьезной информационной поддержке общественных традиций или грамотном противодействии через СМИ появившейся опасности информационной атаки.

Сами же возможности органов, которые заинтересованы в защите традиционных взглядов, характеризуются неравенствами $\mu\gamma < \alpha\rho$ ($\mu\gamma > \alpha\rho$) или более сильным неравенством (2). Как видно из уравнений системы (1), составляющие этих неравенств (рис. 4) влияют на динамику переменных C и A .

Рассмотрим неравенство $\mu\gamma > \alpha\rho$, входящее в Ω_1 и Λ_1 . Для верной трактовки его лучше разбить в соответствии с уравнениями системы на смысловые группы $\mu > \alpha$ & $\gamma > \rho$.

Для С. Органы цензуры, стоящие на страже интересов старой концепции, медленно реагируют на надвигающиеся изменения (α мало). Но причина может заключаться и в том, что дополнительно созданный ресурс для информационной защиты рассчитан на непродолжительное время (μ обратно пропорционально времени работы). Поэтому он не использует весь свой потенциал для нейтрализации информационного воздействия: $\mu > \alpha$.

Для А. Информации, направленной на противодействие информационной атаки, появляется меньше, чем её генерируют оппозиционные издания: $\gamma > \rho$. Возможно, эта информация не отличается разнообразием или убедительностью. Скорее всего, позиции цензурирующих органов так слабы, что они не могут дать достойный ответ агрессивному потоку новой информации. Аппарат управления органами не обладает компетенцией улавливать надвигающиеся угрозы, что

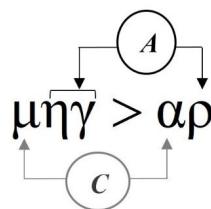
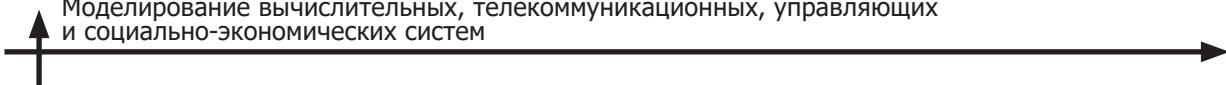


Рис. 4. Соответствие между переменными и параметрами

Fig. 4. Correspondence between variables and parameters



приводит к тому, что те не успевают увеличивать мощность ответа или для этого просто нет возможностей (финансовых, политических или административных).

Таким образом, система (1) с соотношениями параметров, например, из Λ_1 описывает ситуацию, которую можно кратко описать как системный кризис: нехватка финансирования, слабый менеджмент, слабая эффективность. Также можно предположить, что завершился жизненный цикл старой системы взглядов, и аппарат управления намеренно поддается экспансии новой концепции.

Рассмотрим теперь важное неравенство (2), которое можно записать как $\alpha\rho > (\mu + \beta)\eta\gamma$. Более сильное, чем просто неравенство $\alpha\rho > \mu\eta\gamma$, оно обеспечивает безальтернативность поведения системы (1) как в области параметров Ω_2 , так и в Λ_2 . Изучим его подробней, также разбив на смысловые группы $\alpha > (\mu + \beta)$ & $\rho > \eta\gamma$.

Для С. Неравенство $\alpha > \mu$ могло бы значить, что при появлении реальной информационной опасности количество ответственных за информационную защиту органов увеличено на длительный период времени или даже на постоянной основе. Это может подразумевать экстенсивное воздействие на проникновение неугодной информации. Но неравенство $\alpha > (\mu + \beta)$ предполагает интенсивность воздействия. Вновь появившиеся органы эффективно нейтрализуют путем серьёзной аргументации любую появившуюся информацию на нежелательную тему.

Для А. Неравенство $\rho > \eta\gamma$ также говорит о повышенной эффективности информационных сообщений органов защиты. Каждое из них увеличивает вклад в нейтрализацию угрожающей традиционным основам информации.

Например, для параметров из Λ_2 :
$$\begin{cases} \gamma\rho C_* > \lambda\beta \\ \mu\eta\gamma < \alpha\rho \end{cases}$$
, с учетом соотношения (2),

система (1) характеризует общество с четко отлаженным механизмом цензуры: хорошее финансирование, эффективный менеджмент, высокая эффективность. В этом случае попытка информационной атаки обречена на провал, и позиция традиционной концепции в обществе останется прочной.

Заключение

Обзор приведенных в данной статье результатов можно рассматривать как итог первого этапа системного исследования, в основе которого лежит новый подход к моделированию процесса информационного противоборства в обществе. Для модели, представленной в обзоре, проведены полная систематизация и подробная интерпретация результатов, полученных математическими методами в работах [21–24].

Изложенный во второй части введения материал определяет новизну и преимущества предложенного подхода к моделированию по сравнению с подходами других исследователей.

Проведена группировка по различным соотношениям параметров математической модели, количественные характеристики которых позволяют оценить реакцию общества на появление в СМИ новой, претенциозной информации. С учетом появления технологий больших данных и интеллектуального анализа текста такую оценку можно осуществить без теории выборки для изучения общественного мнения.

Для каждого из полученных соотношений подробно описана динамика факторов, использованных при моделировании. Это дает возможность осуществить прогноз вероятного сценария информационного противоборства, а также позволяет управлять системой в зависимости от реакции аудитории.

Содержательная трактовка параметрических соотношений помогает понять причины того или иного варианта распространения новостей в обществе.

Следующий этап исследования предполагает решение очередных важных задач. Предстоит информационная стадия, на которой необходимо определить, какая статистическая информация будет соответствовать по смыслу переменным рассмотренной модели, а также её сбор и систематизация. Это наиболее сложная часть исследования, поскольку необходимо уйти от абстрактных понятий «традиционная», «новая» информация и рассмотреть детали на конкретных информационных объектах. В современных условиях, когда уже существуют системы мониторинга и анализа социальных сетей и онлайн-медиа, такая задача вполне выполнима. Далее предстоит не менее ответственная фаза исследования – идентификационная. Она подразумевает статистический анализ модели и количественную оценку её параметров. Завершить этот трудоёмкий этап предполагается верификацией изучаемой модели.

Оценивая предстоящий объём работ, мы готовы к сотрудничеству со всеми заинтересованными лицами и организациями.

СПИСОК ЛИТЕРАТУРЫ

1. Суходолов А.П., Бычкова А.М., Ованесян С.С. Журналистика с искусственным интеллектом // Вопросы теории и практики журналистики. 2019. Т. 8. № 4. С. 647–667.
2. Суходолов А.П., Рачков М.П. К созданию теории средств массовой информации: постановка задачи // Вопросы теории и практики журналистики. 2016. Т. 5. № 1. С. 6–13.
3. Баенхаева А.В., Тимофеев С.В. Эволюционный подход к развитию средств массовой информации: построение математической модели // Известия Байкальского государственного университета. 2016. Т. 26. № 5. С. 825–833.
4. Землянова Л.М. Коммуникативистика и средства информации: англо-рус. толковый словарь концепций и терминов. М.: Изд-во Моск. ун-та, 2004.
5. Суходолов А.П., Тимофеев С.В. СМИ и виртуальная реальность: новые возможности и перспективы // Вопросы теории и практики журналистики. 2018. Т. 7. № 4. С. 567–580.
6. Информационное право: актуальные проблемы теории и практики. Под общ. ред. И.Л. Бачило. М.: Юрайт, 2009.
7. Марушак А.В. Политико-социальный образ России в американском медиапространстве // Журналистский ежегодник. 2012. № 1. С. 93–96.
8. Суходолов А.П., Кузнецова И.А., Тимофеев С.В. Анализ подходов в моделировании средств массовой информации // Вопросы теории и практики журналистики. 2017. Т. 6. № 3. С. 287–305. DOI: 10.17150/2308-6203.2017.6(3).287-305
9. Джашитов В.Э., Панкратов В.М., Резчиков А.Ф., Джашитов А.Э. Математическое моделирование и управление в системах информирования и информационного обмена в обществе // Проблемы управления. 2009. № 6. С. 2–8.
10. Леонов Г.А. Динамические принципы прогнозирования и управления // Проблемы управления. 2008. № 5. С. 31–35.
11. Вышкунд С.А., Деветьярова А.А. О динамике некоторых математических моделей в социологии // Известия вузов. Сер. Прикладная нелинейная динамика. 1994. № 2. С. 17–26.
12. Губанов Д.А., Новиков Д.А., Чхартишвили А.Г. Социальные сети: модели информационного влияния, управления и противоборства. М.: Физматлит, 2010. 228 с.
13. Михайлов А.П., Маревцева Н.А. Модели информационной борьбы // Математическое моделирование. 2011. Т. 23. № 10. С. 19–32.
14. Михайлов А.П., Петров А.П., Прончева О.Г., Маревцева Н.А. Модель информационного противоборства в социуме при периодическом дестабилизирующем воздействии // Математическое моделирование. 2017. Т. 29. № 2. С. 23–32.

15. Михайлов А.П., Петров А.П., Прончева О.Г. Модель информационного противоборства в социуме с кусочно-постоянной функцией дестабилизирующего воздействия // Математическое моделирование. 2018. Т. 30. № 7. С. 47–60.
16. Михайлов А.П., Петров А.П., Маревцева Н.А., Третьякова И.В. Развитие модели распространения информации в социуме // Математическое моделирование. 2014. Т. 26. № 3. С. 65–74.
17. Докторов Б.З. От соломенных опросов к постгэллаповским опросным методам. М.: Радуга, 2013.
18. Одинцов А.В. Социология общественного мнения и вызов Big Data // Мониторинг общественного мнения: Экономические и социальные перемены. 2017. № 3. С. 30—43.
19. Хосейн Г. Ограничение и сдерживание глобальных потоков данных. М.: МЦБС, 2008. 68 с.
20. Марчук Г.И. Математические модели в иммунологии. М.: Наука, 1980.
21. Тимофеев С.В., Суходолов А.П. Модель распространения новой информации в обществе // Научно-технические ведомости СПбГПУ. Физико-математические науки. 2019. Т. 12. № 4. С. 119–134.
22. Тимофеев С.В., Баенхаева А.В. Модель информационного противоборства в СМИ: важный случай в пространстве параметров // System Analysis & Mathematical Modeling. 2020. Т. 2. № 4. С. 44–52.
23. Тимофеев С.В., Баенхаева А.В. Математическое моделирование информационного противоборства // Научно-технические ведомости СПбГПУ. Физико-математические науки. 2021. Т. 14. № 1. С. 164–176.
24. Тимофеев С.В. Математическая модель распространения новой информации в обществе // Вопросы теории и практики журналистики. 2020. Т. 9. № 1. С. 5–17.

REFERENCES

1. Sukhodolov A.P., Bychkova A.M., Ovanesyan S.S. Zhurnalistika s iskusstvennym intellektom. *Voprosy Teorii i Praktiki Zhurnalistiki*, 2019, Vol. 8, no. 4, Pp. 647–667. (rus)
2. Sukhodolov A.P., Rachkov M.P. K sozdaniyu teorii sredstv massovoy informatsii: postanovka zadachi. *Voprosy Teorii i Praktiki Zhurnalistiki*, 2016, Vol. 5, no. 1, Pp. 6–13. (rus)
3. Bayenkhayeva A.V., Timofeyev S.V. Evolyutsionnyy podkhod k razvitiyu sredstv massovoy informatsii: postroyeniye matematicheskoy modeli. *Izvestiya Baykalskogo Gosudarstvennogo Universiteta*, 2016, Vol. 26, no. 5, Pp. 825–833. (rus)
4. Zemlyanova L.M. *Kommunikativistika i sredstva informatsii: Anglo-rus. tolkovyy slovar kontseptsiy i terminov*. Moscow: Mosc. University Publ., 2004. (rus)
5. Sukhodolov A.P., Timofeyev S.V. SMI i virtualnaya realnost: novyye vozmozhnosti i perspektivy. *Voprosy Teorii i Praktiki Zhurnalistiki*, 2018, Vol. 7, no. 4, Pp. 567–580. (rus)
6. *Informatsionnoye pravo: aktualnyye problemy teorii i praktiki*. Moscow: Yurayt Publ., 2009. (rus)
7. Marushchak A.V. Politiko-sotsialnyy obraz Rossii v amerikanskom mediaprostranstve. *Zhurnalistskiy Yezhegodnik*, 2012, no. 1, Pp. 93–96. (rus)
8. Sukhodolov A.P., Kuznetsova I.A., Timofeyev S.V. Analiz podkhodov v modelirovaniyu sredstv massovoy informatsii. *Voprosy Teorii i Praktiki Zhurnalistiki*, 2017, Vol. 6, no. 3, Pp. 287–305. (rus). DOI: 10.17150/2308-6203.2017.6(3).287-305
9. Dzhashitov V.E., Pankratov V.M., Rezchikov A.F., Dzhashitov A.E. Matematicheskoye modelirovaniye i upravleniye v sistemakh informirovaniya i informatsionnogo obmena v obshchestve. *Problemy Upravleniya*, 2009, no. 6, Pp. 2–8. (rus)
10. Leonov G.A. Dinamicheskiye printsipy prognozirovaniya i upravleniya. *Problemy Upravleniya*, 2008, no. 5, Pp. 31–35. (rus)
11. Vyshkind S.A., Devetyarova A.A. O dinamike nekotorykh matematicheskikh modeley v sotsiologii. *Izvestiya Vuzov. Prikladnaya Nelineynaya Dinamika*, 1994, no. 2, Pp. 17–26. (rus)



12. **Gubanov D.A., Novikov D.A., Chkhartishvili A.G.** *Sotsialnyye seti: modeli informatsionnogo vliyaniya, upravleniya i protivoborstva*. Moscow: Fizmatlit Publ., 2010. 228 p. (rus)
13. **Mikhaylov A.P., Marevtseva N.A.** Modeli informatsionnoy borby. *Matematicheskoye Modelirovaniye*, 2011, Vol. 23, no. 10, Pp. 19–32. (rus)
14. **Mikhaylov A.P., Petrov A.P., Proncheva O.G., Marevtseva N.A.** Modeli informatsionnogo protivoborstva v sotsiume pri periodicheskem destabiliziruyushchem vozdeystvi. *Matematicheskoye Modelirovaniye*, 2017, Vol. 29, no. 2, Pp. 23–32. (rus)
15. **Mikhaylov A.P., Petrov A.P., Proncheva O.G.** Modeli informatsionnogo protivoborstva v sotsiume s kusochno-postoyannoy funktsiyey destabiliziruyushchego vozdeystviya. *Matematicheskoye Modelirovaniye*, 2018, Vol. 30, no. 7, Pp. 47–60. (rus)
16. **Mikhaylov A.P., Petrov A.P., Marevtseva N.A., Tretyakova I.V.** Razvitiye modeli rasprostraneniya informatsii v sotsiume. *Matematicheskoye Modelirovaniye*, 2014, Vol. 26, no. 3, Pp. 65–74. (rus)
17. **Doktorov B.Z.** *Ot solomennykh oprosov k postgellapovskim oprosnym metodam*. Moscow: Raduga Publ., 2013. (rus)
18. **Odintsov A.V.** Sotsiologiya obshchestvennogo mneniya i vyzov Big Data. *Monitoring Obshchestvennogo Mneniya: Ekonomicheskiye i Sotsialnyye Peremeny*, 2017, no. 3, Pp. 30–43. (rus)
19. **Khoseyn G.** *Ogranicheniya i sderzhivaniye globalnykh potokov dannykh*. Moscow: MTsBS, 2008. 68 p. (rus)
20. **Marchuk G.I.** *Matematicheskiye modeli v immunologii*. Moscow: Nauka Publ., 1980. (rus)
21. **Timofeyev S.V., Sukhodolov A.P.** Modeli rasprostraneniya novoy informatsii v obshchestve. *Nauchno-Tekhnicheskiye Vedomosti SPbGPU. Fiziko-Matematicheskiye Nauki*, 2019, Vol. 12, no. 4, Pp. 119–134. (rus)
22. **Timofeyev S.V., Bayenkhayeva A.V.** Modeli informatsionnogo protivoborstva v SMI: vazhnyy sluchay v prostranstve parametrov. *System Analysis & Mathematical Modeling*, 2020, Vol. 2, no. 4, Pp. 44–52. (rus)
23. **Timofeyev S.V., Bayenkhayeva A.V.** Matematicheskoye modelirovaniye informatsionnogo protivoborstva. *Nauchno-Tekhnicheskiye Vedomosti SPbGPU. Fiziko-Matematicheskiye Nauki*, 2021, Vol. 14, no. 1, Pp. 164–176. (rus)
24. **Timofeyev S.V.** Matematicheskaya modeli rasprostraneniya novoy informatsii v obshchestve. *Voprosy Teorii i Praktiki Zhurnalistiki*, 2020, Vol. 9, no. 1, Pp. 5–17. (rus)

INFORMATION ABOUT AUTHORS / СВЕДЕНИЯ ОБ АВТОРАХ

Тимофеев Сергей Викторович
Sergey V. Timofeev
E-mail: timofeevsv12@gmail.com

Баенхаева Аюна Валерьевна
Ayuna V. Baenkhaeva
E-mail: ayunab2000@mail.ru

Поступила: 16.02.2022; Одобрена: 27.06.2022; Принята: 25.08.2022.
Submitted: 16.02.2022; Approved: 27.06.2022; Accepted: 25.08.2022.