

АСТРАХАНСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

ПРИКАСПИЙСКИЙ ЖУРНАЛ: управление и высокие технологии

НАУЧНО-ТЕХНИЧЕСКИЙ ЖУРНАЛ

2020

№ 1 (49)

Журнал включен в перечень рецензируемых научных изданий, рекомендованных ВАК России для публикации основных научных результатов диссертаций на соискание ученой степени кандидата наук, на соискание ученой степени доктора наук по следующим научным специальностям.

Группа специальностей 05.11.00 «Приборостроение, метрология и информационно-измерительные приборы и системы»:

- 05.11.01 – Приборы и методы измерения (по видам измерений) (технические науки);
- 05.11.16 – Информационно-измерительные и управляющие системы (по отраслям) (технические науки);
- 05.11.17 – Приборы, системы и изделия медицинского назначения (технические науки).

Группа специальностей 05.13.00 «Информатика, вычислительная техника и управление»:

- 05.13.01 – Системный анализ, управление и обработка информации (по отраслям) (технические науки);
- 05.13.05 – Элементы и устройства вычислительной техники и систем управления (технические науки);
- 05.13.10 – Управление в социальных и экономических системах (технические науки);
- 05.13.11 – Математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей (технические науки);
- 05.13.18 – Математическое моделирование, численные методы и комплексы программ (технические науки);
- 05.13.19 – Методы и системы защиты информации, информационная безопасность (технические науки).

Журнал входит в базу данных Ulrich's Periodicals Directory.

Астрахань
Издательский дом «Астраханский университет»
2020

Рекомендовано к печати редакционно-издательским советом
Астраханского государственного университета

ПРИКАСПИЙСКИЙ ЖУРНАЛ:
управление и высокие технологии
НАУЧНО-ТЕХНИЧЕСКИЙ ЖУРНАЛ

2020
№ 1 (49)

Редакционная коллегия

И.М. Ажмухамедов, доктор технических наук, профессор, зав. кафедрой «Информационная безопасность» Астраханского государственного университета (**главный редактор**)

И.В. Аникин, доктор технических наук, доцент, заведующий кафедрой «Системы информационной безопасности» Казанского национального исследовательского технического университета им. А.Н. Туполева-КАИ

А.А. Большаков, доктор технических наук, профессор, профессор кафедры «Системы автоматизированного проектирования и управления» Санкт-Петербургского государственного технологического института (технического университета)

Ю.М. Брумитейн, кандидат технических наук, доцент, доцент кафедры «Информационные технологии» Астраханского государственного университета (**ответственный секретарь**)

Ж.И. Батырканов, доктор технических наук, профессор, профессор Кыргызского государственного технического университета им. И. Разакова (Кыргызская Республика, г. Бишкек)

С.Н. Гончаренко, доктор технических наук, профессор, профессор Национального исследовательского технологического университета «МИСиС» (г. Москва)

Л.А. Демидова, доктор технических наук, профессор, профессор кафедры «Вычислительной и прикладной математики» Рязанского государственного радиотехнического университета (г. Рязань)

И.Ю. Квятковская, доктор технических наук, профессор, директор Института информационных технологий и коммуникаций Астраханского государственного технического университета

А.Г. Кравец, доктор технических наук, профессор, профессор кафедры «Системы автоматизированного проектирования и поискового конструирования» Волгоградского государственного технического университета

Ю.В. Литовка, доктор технических наук, профессор, профессор кафедры «Системы автоматизированной поддержки принятия решений» Тамбовского государственного технического университета

А.А. Лобатый, доктор технических наук, профессор, заведующий кафедрой «Информационные системы и технологии» Белорусского национального технического университета (Республика Беларусь, г. Минск)

В.В. Морозов, заслуженный деятель науки РФ, доктор технических наук, профессор, заведующий кафедрой «Технология машиностроения» Владимирского государственного университета им. А.Г. и Н.Г. Столетовых (г. Владимир)

Е.В. Никульчев, доктор технических наук, профессор, профессор кафедры «Управление и моделирование систем» Московского технологического университета (МИРЭА) (г. Москва)

В.О. Осипян, доктор физико-математических наук, доцент, профессор кафедры «Информационные технологии» Кубанского государственного университета (г. Краснодар)

И.Ю. Петрова, доктор технических наук, профессор, первый проректор Астраханского государственного архитектурно-строительного университета, заведующая кафедрой САИР Астраханского государственного архитектурно-строительного университета

А.В. Рыбаков, кандидат физико-математических наук, директор «Физико-математического института» Астраханского государственного университета; доцент кафедры электротехники, электроники и автоматики Астраханского государственного университета

А.В. Скрипаль, доктор физико-математических наук, профессор, заведующий кафедрой «Медицинская физика» Саратовского национального исследовательского государственного университета им. Н.Г. Чернышевского

И.Б. Старченко, доктор технических наук, профессор, ООО «Параметрика», научный руководитель (г. Таганрог Ростовской области)

Ю.Ю. Тарасевич, доктор физико-математических наук, профессор, профессор Астраханского государственного университета, заведующий лабораторией «Математическое моделирование и информационные технологии в науке и образовании»

Т.Л. Тен, доктор физико-математических наук, профессор кафедры «Информационно-вычислительные системы» Карагандинского экономического университета (Республика Казанстан, г. Караганда)

Е.Н. Тищенко, доктор экономических наук, профессор, заведующий кафедрой «Информационные технологии и защита информации» Ростовского государственного экономического университета (РИНХ) – г. Ростов-на-Дону

М.А. Ураксеев, доктор технических наук, профессор, профессор кафедры «Информационно-измерительная техника» Уфимского государственного авиационного технического университета

Д.А. Усанов, заслуженный деятель науки РФ, доктор физико-математических наук, профессор, заведующий кафедрой «Физика твердого тела» Саратовского национального исследовательского государственного университета им. Н.Г. Чернышевского

С.А. Филист, доктор технических наук, профессор, профессор кафедры «Биомедицинская инженерия» Юго-Западного государственного университета (г. Курск)

Л.Р. Фиопова, доктор технических наук, профессор, декан факультета Вычислительной техники, заведующая кафедрой «Информационное обеспечение управления и производства» Пензенского государственного университета

В.А. Цимбал, заслуженный деятель науки РФ, доктор технических наук, профессор, профессор кафедры «Автоматизированные системы управления» (Филиал Военной академии РВСН им. Петра Великого МО в г. Серпухов Московской области)

Н.К. Юрков, заслуженный деятель науки РФ, доктор технических наук, профессор, заведующий кафедрой «Конструирование и производство радиоаппаратуры» Пензенского государственного университета

N.A. Kolesova, PhD, Check Point Software Technologies LTD, Tel-Aviv, Israel

Serg Miranda, PhD (Toulouse University, France), – Master thesis at UCLA (University of California, Los Angeles with an INRIA Scholarship), Professor of Computer Science, University of Nice – Sophia Antipolis (Nice, France), Director of the CS dept. and MBDS innovation lab (www.mbd-s.fr.org)

Журнал выходит 4 раза в год
Все материалы, поступающие в редколлегию журнала,
проходят независимое рецензирование

© Астраханский государственный университет,
Издательский дом «Астраханский университет», 2020
© Свиридов В. Б., дизайн обложки, 2020

ASTRAKHAN STATE UNIVERSITY

**PRIKASPIYSKIY ZHURNAL:
Upravlenie i Vysokie Tekhnologii**

**CASPIAN JOURNAL:
Control and High Technologies**

A SCIENTIFIC AND TECHNICAL JOURNAL

2020

No. 1 (49)

The journal is included in the list of the reviewed scientific journals recommended by VAK of Russia for the publication of the main scientific results of theses for the candidate of science degree, for the doctor of science degree on the following scientific specialties.

Group of specialties 05.11.00 “Instrument engineering, measurement science, information and measuring devices and systems”:

05.11.01 – Instruments and measurement methods (by measurement types) (technical sciences);

05.11.16 – Information-measuring and control systems (by branches) (technical sciences);

05.11.17 – Medical instruments, systems and items (technical sciences).

Group of specialties 05.13.00 “Informatics, computer technique and control”:

05.13.01 – System analysis, information control and processing (by branches) (technical sciences);

05.13.05 – Components and devices of computational tools and control systems (technical sciences);

05.13.10 – Management in social and economic systems (technical sciences);

05.13.11 – Mathematical software and software for computing machines, computer systems and networks (technical sciences);

05.13.18 – Mathematical modelling, numerical methods and complexes of programmes (technical sciences);

05.13.19 – Information security methods and systems, information security (technical sciences).

The journal is included into the database Ulrich’s Periodicals Directory.

Astrakhan
Publishing House “Astrakhan University”
2020

Recommended by the Editorial and Publishing Board
of Astrakhan State University

**CASPIAN JOURNAL:
Control and High Technologies**

A SCIENTIFIC AND TECHNICAL JOURNAL

**2020
No. 1 (49)**

Editorial Board

I.M. Azhmukhamedov, Doct. Sci. (Engineering), Professor, Head of Information Security Department, Astrakhan State University (Editor-in-Chief)

I.V. Anikin, Doct. Sci. (Engineering), Associate Professor, Head of Information Security System Department, Federal State Budgetary Educational Institution of Higher Education «Kazan National Research Technical University named after A.N. Tupolev – KA»

A.A. Bolshakov, Doct. Sci. (Engineering), Professor of «Systems of Automated Design Engineering and Control» department, St. Petersburg State Technological Institute (Technical University)

Yu.M. Brumsteyn, Cand. Sci. (Engineering), Associate Professor, Associate Professor of Astrakhan State University (Executive Editor)

Zh.I. Batyrkanov, Doct. Sci. (Engineering), Professor, Professor of the Kyrgyz State Technical University named after I. Razza-kov (Kyrgyz Republic, Bishkek)

S.N. Goncharenko, Doct. Sci. (Engineering), Professor, Professor of the National University of Science and Technology «MISIS» (Moscow)

L.A. Demidova, Doct. Sci. (Engineering), Professor, Professor of the Computational and Applied Mathematics Department, Ryazan State Radio Engineering University (Ryazan)

I.Yu. Kvyatkovskaya, Doct. Sci. (Engineering), Professor, Head of “Information Technologies and Communications” Institute of the Astrakhan State Technical University

A.G. Kravets, Doct. Sci. (Engineering), Professor, Professor of the Automated Design Engineering Systems and Search Constructing Department, Volgograd State Technical University

Yu.V. Litovka, Doct. Sci. (Engineering), Professor, Professor of the Department of Automated Support System for Decision-Making, Tambov State Technical University

A.A. Lobatyy, Doct. Sci. (Engineering), Professor, Head of Information Systems and Technologies Department, Belarusian National Technical University (Belarus, Minsk)

V.V. Morozov, Honored Worker of Science of the Russian Federation, Doct. Sci. (Engineering), Professor of the Vladimir State University named after A.G. and N.G. Stoletov (Vladimir)

E.V. Nikulchev, Doct. Sci. (Engineering), Professor, Professor of the System Management and Modeling Department, Moscow Technological University (Moscow)

V.O. Osipyann, Doct. Sci. (Physics and Mathematics), Professor of the Kuban State University (Krasnodar)

I.Yu. Petrova, Doct. Sci. (Engineering), Professor, First Vice-Rector of the Astrakhan State Architectural and Construction University, Head of the CAD department of Astrakhan State Architectural and Construction University

A.V. Rybakov, Cand. Sci. (Physics and Mathematics), Director of the Institute of Physics and Mathematics, Astrakhan State University

A.V. Skripal, Doct. Sci. (Physics and Mathematics), Professor, Head of Medical Physics Department of the Saratov national research State University named after N.G. Chernyshevsky

I.B. Starchenko, Doct. Sci. (Engineering), Professor, OOO «Parametrica» (Taganrog, Rostov Oblast), Research Supervisor

Yu.Yu. Tarasevich, Doct. Sci. (Physics and Mathematics), Professor, Professor of the Astrakhan State University, head of the laboratory «Mathematical modeling and information technologies in science and education»

T.L. Ten, Doct. Sci. (Engineering), Professor, Karaganda Economic University (Republic of Kazakhstan, Karaganda)

E.N. Tishchenko, Doct. Sci. (Economics), Professor, Head of the Information Technologies & Information Security Department, Rostov State University of Economics, Rostov-on-Don

M.A. Urakseev, Doct. Sci. (Engineering), Professor, Professor of the Information and Measuring Equipment department of Ufa State Aviation Technical University

D.A. Usanov, Honored Worker of Science of the Russian Federation, Doct. Sci. (Physics and Mathematics), head of the department «Solid State Physics», Saratov National Research State University named after N.G. Chernyshevsky

S.A. Filist, Doct. Sci. (Engineering), Professor, Professor of Biomedical Engineering Department, Southwest State University (Kursk)

L.R. Fionova, Doct. Sci. (Engineering), Professor, Dean of the Computer Technology Faculty, Head of the Department «Information Support of Management and Production, Penza State University

V.A. Tsimbal, Doct. Sci. (Engineering), Honored Worker of Science of the Russian Federation, Professor, Professor of the Automated Control Systems Department (Branch of the Military Academy of the Russian Strategic Missile Forces named after Peter the Great of the Moscow Oblast, Serpukhov, Moscow Oblast)

N.K. Yurkov, Honored worker of science of the Russian Federation, Doct. Sci. (Engineering), Professor, Head of the department «Designing and production of the radio equipment», Penza State University

N.A. Kolesova, PhD, Check Point Software Technologies LTD, Tel-Aviv, Israel

Serg Miranda, PhD (Toulouse University, France), – Master thesis at UCLA (University of California, Los Angeles with an INRIA Scholarship), Professor of Computer Science dept., University of Nice – Sophia Antipolis (Nice, France), Director of the CS department and MBDS innovation lab (www.mbd-fr.org)

The journal is published four times a year
All materials that come to the Editorial Board of the journal
are subject to independent peer-review

© Astrakhan State University,
Publishing House “Astrakhan University”, 2020
© V. B. Sviridov, cover design, 2020

АВТОРСКИЙ УКАЗАТЕЛЬ

<i>Ажмухамедов И. М.</i>	111–121	<i>Коновалова Д. И.</i>	61–76
<i>Бондарев А. В.</i>	61–76		77–93
<i>Брумштейн Ю. М.</i>	61–76	<i>Крюкова Е. С.</i>	121–131
	77–93	<i>Литовка Ю. В.</i>	131–143
<i>Васильев Н. В.</i>	61–76	<i>Миловидова А. А.</i>	33–46
<i>Власенко А. В.</i>	144–155	<i>Новосельцева А. В.</i>	155–161
	162–169	<i>Нуриев Н. К.</i>	10–20
	169–178	<i>Овчаренко С. С.</i>	94–111
<i>Воля Я. И.</i>	20–32	<i>Паращук И. Б.</i>	121–131
<i>Врублевская К. А.</i>	77–93	<i>Печеный Е. А.</i>	10–20
<i>Гордеев И. И.</i>	94–111	<i>Роткин В. М.</i>	111–121
<i>Дзьобан П. И.</i>	144–155	<i>Сай Ван Квонг</i>	49–60
	162–169	<i>Сизова А. А.</i>	94–111
	169–178	<i>Соколовский Ю. Б.</i>	111–121
<i>Жерлицын С. А.</i>	20–32	<i>Соловьев Д. С.</i>	131–143
<i>Жук Р. В.</i>	144–155	<i>Соловьева И. А.</i>	131–143
	162–169	<i>Сотников В. В.</i>	20–32
	169–178	<i>Старыгина С. Д.</i>	10–20
<i>Иглина Н. А.</i>	77–93	<i>Частикова В. А.</i>	20–32
<i>Клюев С. Г.</i>	155–161	<i>Щербаков М. В.</i>	49–60

СОДЕРЖАНИЕ

ИНФОРМАТИКА, ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА И УПРАВЛЕНИЕ

СИСТЕМНЫЙ АНАЛИЗ, УПРАВЛЕНИЕ И ОБРАБОТКА ИНФОРМАЦИИ

- Е. А. Печеньй, Н. К. Нуриев, С. Д. Старыгина**
Самоорганизующаяся кластеризация потока больших данных 10–20
- В. А. Частикова, С. А. Жерлицын, Я. И. Воля, В. В. Сотников**
Нейросетевая технология обнаружения аномального сетевого трафика 20–32
- А. А. Миловидова**
Формирование нечеткой системы оценки для системы поддержки
принятия решений при управлении процессом переработки
в условиях неопределенности информации о качестве материала 33–48
- Сай Ван Кеонг, М. В. Щербаков**
Прогнозирование отказов сложных многообъектных систем
на основе комбинации нейросетей:
пути повышения точности прогнозирования 49–60

УПРАВЛЕНИЕ В СОЦИАЛЬНЫХ И ЭКОНОМИЧЕСКИХ СИСТЕМАХ

- А. В. Бондарев, Н. В. Васильев,
Д. И. Коновалова, Ю. М. Брумштейн**
Разработка, введение в эксплуатацию и использование «сайтов-визиток»:
анализ целей, возможных подходов, трудоемкости, структуры рисков 61–76
- Н. А. Иглина, Ю. М. Брумштейн,
К. А. Врублевская, Д. И. Коновалова**
Нелояльность персонала организаций в условиях развития
информационно-телекоммуникационных технологий:
анализ возможных источников информации, подходов
к выявлению, моделей и методов управления 77–93

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ, ЧИСЛЕННЫЕ МЕТОДЫ И КОМПЛЕКСЫ ПРОГРАММ

- И. И. Гордеев, С. С. Овчаренко, А. А. Сизова**
Нахождение проводящего остова
в двумерной решетке методом заливки 94–111
- В. М. Роткин, Ю. Б. Соколовский, И. М. Ажмухамедов**
Ротор Дарье и оптимизированные ветротурбины:
сравнительный анализ 111–121
- Е. С. Крюкова, И. Б. Паращук**
Математическая модель, предназначенная
для оценки качества электронной библиотеки:
синтез числа градаций пространства состояний 121–131
- Д. С. Соловьев, И. А. Соловьева, Ю. В. Литовка**
Применение решения обратной задачи математического моделирования
гальванического процесса для оптимизации неравномерности толщины покрытия 131–143

ИНФОРМАЦИОННАЯ БЕЗОПАСНОСТЬ И ЗАЩИТА ИНФОРМАЦИИ**А. В. Власенко, П. И. Дзьобан, Р. В. Жук**Обзор инструментов машинного обучения
и их применения в области кибербезопасности..... 144–155**А. В. Новосельцева, С. Г. Ключев**

Современные методы атак деанонимизации на сеть TOR..... 155–161

Р. В. Жук, П. И. Дзьобан, А. В. ВласенкоПостроение взаимосвязи между нарушителем
информационной безопасности и уязвимостями информационных активов
в информационных системах обработки персональных данных..... 162–169**Р. В. Жук, П. И. Дзьобан, А. В. Власенко**Определение актуальности угроз информационной безопасности
в информационных системах обработки персональных данных
с использованием математического аппарата нейронных сетей 169–178**ИНФОРМАЦИЯ О ПРЕДСТОЯЩИХ НАУЧНЫХ МЕРОПРИЯТИЯХ..... 179****ПРАВИЛА ДЛЯ АВТОРОВ 180**

CONTENTS

INFORMATICS, COMPUTER TECHNIQUE AND CONTROL

SYSTEM ANALYSIS, CONTROL AND INFORMATION PROCESSING

- E. A. Pechenny, N. K. Nuriev, S. D. Starygina**
Self-organizing clustering of great data flow 10–20
- V. A. Chastikova, S. A. Zherlitsyn,
Ya. I. Volya, V. V. Sotnikov**
Neural network technology for detecting
anomalous network traffic 20–32
- A. A. Milovidova**
Creation of the fuzzy measurement system for decision support system
when managing the production process in an uncertain information
about the quality of the material being processed 33–48
- Sai Van Cuong, M. V. Shcherbakov**
Failure prediction of complex multiple-component systems
based on a combination of neural networks:
ways to improve the accuracy of forecasting 49–60

CONTROL IN SOCIAL AND ECONOMIC SYSTEMS

- A. A. Bondarev, N. V. Vasilyev,
D. I. Konovalova, Yu. M. Brumshteyn**
Development, commissioning and use of “business card sites”:
analysis of targets, possible approaches, labor consumption, risk structure 61–76
- N. A. Iglina, Yu. M. Brumshteyn,
K. A. Vrublevskaya, D. I. Konovalova**
Non-liability of organization personnel in the conditions
of information-telecommunication technologies development:
analysis of possible information sources,
identification approaches, models and management methods 77–93

MATHEMATICAL MODELLING, NUMERICAL METHODS AND PROGRAM SYSTEMS

- I. I. Gordeev, S. S. Ovcharenro, A. A. Sizova**
The determination of the conducting backbone
in a two-dimensional lattice by the flooding method 94–111
- V. M. Rotkin, Yu. B. Sokolovsky, I. M. Azhmukhamedov**
Darier rotor and optimized wind turbines:
comparative analysis 111–121
- E. S. Kryukova, I. B. Parashchuk**
Mathematical mode for estimating
the quality of electronic library:
synthesis of the number of gradations of space states 121–131
- D. S. Solovjev, I. A. Solovjeva, Yu. V. Litovka**
Inverse problem solution of the mathematical modeling
for galvanic process to optimize the non-uniformity
of the coating thickness 131–143

INFORMATION SAFETY AND INFORMATION PROTECTION***A. V. Vlasenko, P. I. Dzoban, R. V. Zhuk***

Analytical review of machine learning tools
and their applications in the field of cyber security 144–155

A. V. Novoseltseva, S. G. Klyuev

Modern methods of attacks of deanonymization
on the Tor network..... 155–161

R. V. Zhuk, P. I. Dzoban, A. V. Vlasenko

Building a relationship between an information security
intruder and vulnerabilities of information assets
in information systems for processing personal data 162–169

R. V. Zhuk, P. I. Dzoban, A. V. Vlasenko

Determining the relevance of information security threats
in information systems for processing personal data
using the mathematical apparatus of neural networks..... 169–178

INFORMATION ON FORTHCOMING SCIENTIFIC EVENTS 179

RULES FOR THE AUTHORS 180

ИНФОРМАТИКА, ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА И УПРАВЛЕНИЕ

СИСТЕМНЫЙ АНАЛИЗ, УПРАВЛЕНИЕ И ОБРАБОТКА ИНФОРМАЦИИ

УДК 004.9

САМООРГАНИЗУЮЩАЯСЯ КЛАСТЕРИЗАЦИЯ ПОТОКА БОЛЬШИХ ДАННЫХ

Статья поступила в редакцию 06.02.2020, в окончательном варианте – 26.02.2020.

Печеный Евгений Абрамович, Казанский национальный исследовательский технологический университет, 420015, Российская Федерация, Республика Татарстан, г. Казань, ул. К. Маркса, 68, кандидат технических наук, доцент, РИНЦ AuthorID 408103, e-mail: platova51@mail.ru

Нуриев Наиль Кашипович, Казанский национальный исследовательский технологический университет, 420015, Российская Федерация, Республика Татарстан, г. Казань, ул. К. Маркса, 68, доктор педагогических наук, кандидат технических наук, профессор, заведующий кафедрой информатики и прикладной математики, ORCID <https://orcid.org/0000-0002-9557-5493>, РИНЦ AuthorID 527783, e-mail: nurievnk@mail.ru

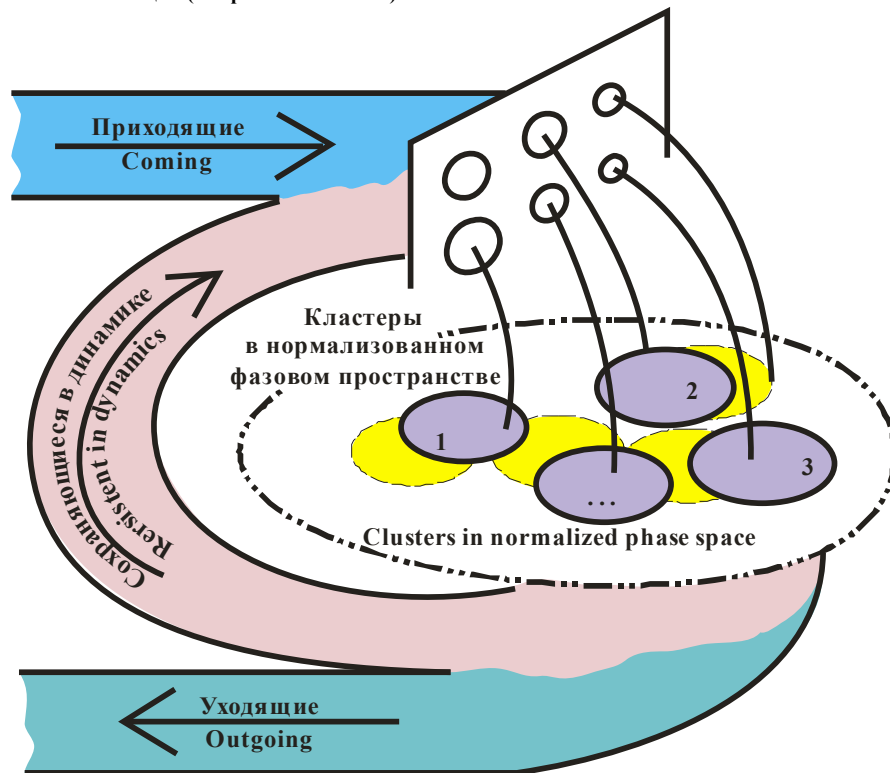
Старыгина Светлана Дмитриевна, Казанский национальный исследовательский технологический университет, 420015, Российская Федерация, Республика Татарстан, г. Казань, ул. К. Маркса, 68,

кандидат педагогических наук, доцент, ORCID <https://orcid.org/0000-0002-3401-6452>, РИНЦ AuthorID 237000, e-mail: svetacd_kazan@mail.ru

В работе представлена математическая модель и дано описание алгоритма на основе аппарата кластерного анализа, ориентированного на проведение процедур классификации «больших данных». В качестве кластеров предложено использовать сфероиды, для построения которых предварительно производится нормирование переменных и преобразование их в безразмерную форму. Простота аналитического описания формы кластеров служит эффективной защитой алгоритма от «проклятия размерности», обеспечивает сохранение его работоспособности при большом числе классифицируемых признаков. Отличительной особенностью разработанного алгоритма является его способность функционировать в динамическом режиме, т.е. в условиях изменений свойств объектов, присутствующих в кластерах; пополнения кластеров потоком новых объектов; удаления некоторых объектов из числа классифицируемых. Для обеспечения однозначности выделяемых классификационных категорий в алгоритме предусмотрена защита от пересечений кластеров. Важным и полезным эксплуатационным качеством алгоритма является его самоорганизуемость. Он может обрабатывать данные в потоке без участия оператора, выполняя, по мере необходимости, коррекцию положения и размеров кластеров. Процедура коррекции представляет собой последовательность итераций, в ходе которой осуществляется сближение геометрических центров кластеров с центрами группировок объектов, имеющих в их составе. В статье приведена блок-схема алгоритма, который был реализован программно. Работа этого алгоритма продемонстрирована и графически проиллюстрирована на примере сравнительно небольшого массива данных, элементы которого описываются двумя классифицирующими признаками.

Ключевые слова: большие данные, классификация объектов, кластеры, динамическая кластеризация, самоорганизация, самообучение

Графическая аннотация (Graphical annotation)



SELF-ORGANIZING CLUSTERING OF GREAT DATA FLOW

The article was received by the editorial board on 06.02.2020, in the final version – 26.02.2020.

Pechenyu Eugene A., Kazan National Research Technological University, 68 Karl Marks St., Kazan, 420015, Republic of Tatarstan, Russian Federation,

Cand. Sci. (Engineering), Associate Professor, RISC AuthorID 408103, e-mail: platova51@mail.ru

Nuriev Nail K., Kazan National Research Technological University, 68 Karl Marks St., Kazan, 420015, Republic of Tatarstan, Russian Federation,

Doct. Sci. (Pedagogy), Cand. Sci. (Engineering), Professor, Head of the Department of Informatics and Applied Mathematics, ORCID <https://orcid.org/0000-0002-9557-5493>, RISC AuthorID 527783, e-mail: nurievnk@mail.ru

Starygina Svetlana D., Kazan National Research Technological University, 68 Karl Marks St., Kazan, 420015, Republic of Tatarstan, Russian Federation,

Cand. Sci. (Pedagogy), Associate Professor, ORCID <https://orcid.org/0000-0002-3401-6452>, RISC AuthorID 237000, e-mail: svetacd_kazan@mail.ru

This paper presents a mathematical model and provides a description of an algorithm, based on clustering apparatus focused on carrying out big data classification procedures. Spheroids are proposed to be used as clusters, for which to be constructed the variables are preliminarily normalized and transformed to be nondimensional. Simplicity of analytically defining the forms of clusters serves for the algorithm as an efficient protection from the curse of dimensionality and makes it efficient for a great number of criteria to be classified. The distinctive feature of the algorithm developed is its ability to function dynamically, i. e., in the conditions of changing properties of elements available in clusters and refilling the clusters with the stream of new elements. To ensure the unambiguity of the classification categories distinguished, the algorithm provides protection from cluster intersections. An important and useful operating characteristic of the algorithm is its self-organizability. It can process stream data without the intervention of an operator, correcting the locations and sizes of clusters, where necessary. Correcting procedure represents a sequence of iterations, in which the geometric centers of clusters approach to the centers of groups of objects available in the clusters. The paper presents a control-flow chart that was implemented in software. The algorithm operation is demonstrated and illustrated graphically, exemplified by a comparatively small data array, the elements of which are defined by two classification criteria.

Key words: big data, object classification, clusters, dynamic clustering, self-organization, self-study

Стремительное развитие информационных технологий и интернета, которым ознаменовался конец XX – начало XXI в., сделало возможным формирование и передачу очень больших объемов информации. Это, в свою очередь, привело к необходимости создания специальных методов, позволяющих быстро обрабатывать массивы так называемых «больших данных» [1, 2, 6, 9, 16, 17, 20]. Особый интерес и актуальность имеют задачи, связанные с обработкой массивов больших данных в процессе их динамических изменений. Действительно, полностью сложившиеся, неизменяемые большие объемы информации сейчас можно отыскать только в архивных хранилищах и каталогах. Массивы больших данных, являясь, по сути дела, выборками каких-то генеральных совокупностей, находятся в процессе постоянного обмена элементами друг с другом. Иначе говоря, существует поток (потоки), пополняющий массив новыми элементами, и поток уноса (гибели) части элементов массива, что влечет за собой изменения его численного состава. Изменения могут происходить также с признаками, характеризующими состояние элементов. Следовательно, «большие данные» представляют собой сложные динамические образования, подверженные как качественным, так и количественным изменениям. При разработке алгоритмов, предназначенных для работы с большими данными, это обстоятельство должно быть обязательно учтено.

Остановимся более подробно на понятии «обработка данных», которое является весьма широким и нуждается, по мнению авторов, в некоторой конкретизации. Этот термин может быть с достаточным основанием применен к ряду разнородных процедур: устранению недостоверных результатов и исправлению неточностей, пакетированию данных, классификации элементной базы больших данных по определенным признакам и т.п. Здесь мы будем понимать под обработкой только процедуру классификации по совокупности количественно измеримых признаков.

Обратим внимание на довольно существенные различия классификационных техник для качественных и количественных признаков. Например, классификация социума по признакам пола, расовой или национальной принадлежности, группе крови, как правило, дает однозначный результат и не вызывает затруднений. Решение же задачи классификации по фактору возраста далеко не так очевидно и зависит не только от состава данных, но и от целей исследования. Где следует установить границы молодости, зрелости и старости; как будет происходить изменение этих границ в течение 10, 20, 50 лет, в том числе при изменении среднего возраста населения стран; достаточно ли для вынесения обоснованных суждений этих трех классификационных областей, или целесообразно введение дополнительных – эти вопросы не имеют готовых ответов, пригодных для любых ситуаций. Скорее всего, мнение специалистов-медиков будет отлично от мнения социологов, поскольку физиологическое развитие человеческой личности не совпадает с изменениями ее социального статуса. Перечисленные особенности процедур классификации при наличии количественно измеримых признаков делает подобные задачи достаточно сложными для формализации, а трудоемкость их решения многократно возрастает с ростом размерности пространства классифицирующих признаков. В настоящей работе для решения задачи классификации на массиве больших данных в условиях динамических изменений последних предлагается использовать алгоритм, в основе которого лежит аппарат кластерного анализа.

Кластерный анализ как инструмент группировки объектов внутри исследуемого массива данных известен давно [4, 5, 7, 10–15, 18] и достаточно широко применяется для решения практических задач. Основная его идея состоит в разбиении множества всех элементов массива на конечное число непересекающихся подмножеств (кластеров), элементы которых близки друг к другу в смысле введенной метрики. Это позволяет в ряде случаев отождествить свойства отдельных объектов со свойствами кластера, которому они принадлежат, и заменить решение нескольких частных задач решением одной общей задачи по единому для всех элементов кластера критерию. Такой подход часто и успешно используется при решении проблем логистики, вопросов энергоснабжения и водоснабжения, управления городским хозяйством и пр. Практическое применение кластерного анализа ограничивается отсутствием универсальных рекомендаций по выбору формы кластеров, их размеров и метрики. Эти вопросы решаются, как правило, с использованием эвристических методов, а результаты в значительной мере зависят от интуиции и опыта лица, принимающего решение.

Рассмотрим массив данных, элементы которого характеризуются n количественно измеримыми признаками. Мощность массива Q не является постоянной величиной, а меняется под влиянием двух потоков: пополнения и гибели. Объекты, остающиеся в составе массива, также могут меняться под влиянием изменений одного или нескольких связанных с ними признаков. Таким образом, рассматриваемый массив является динамическим образованием, структура и свойства которого есть явные функции времени.

Если, как упоминалось выше, все признаки, характеризующие состояние объектов Q , количественно измеримы, то геометрическим образом любого объекта будет точка $x^j = (x_1^j, x_2^j, \dots, x_n^j)$ для $\forall j \in N$. Тогда, используя в качестве меры близости евклидову метрику, получим

$$\|x^j - x^k\| = \sqrt{\sum_{i=1}^n (x_i^j - x_i^k)^2} \quad j, k \in N, \quad (1)$$

где n – размерность пространства классифицирующих признаков. В большинстве случаев разнообразие форм кластеров ограничивается классом выпуклых множеств. Иначе говоря, в рамках традиционных метрик понятие расстояния как меры близости между объектами утрачивает свою определенность. Однако даже при выполнении условия выпуклости с увеличением размерности пространства классифицирующих признаков резко возрастает сложность аналитического описания кластерных границ и интерпретация полученных результатов. Именно по этим причинам авторы остановили свой выбор на кластерах сферической формы, поскольку сложность математического описания сферы как геометрического объекта практически не зависит от размерности пространства, а значит, эта форма как нельзя более подходит для работы с большими данными.

Построение кластерных сфероидов непосредственно в пространстве классифицирующих признаков, конечно, невозможно. Признаки, описывающие состояние объектов массива данных, могут иметь различную природу, разные единицы измерений, а потому понятие «тело» или «фигура» в пространстве таких признаков неопределенно. По этой причине необходимо для всех классифицирующих признаков осуществить переход к новым переменным, измеряемым в единичных безразмерных единицах. Поскольку массив данных в течение всего времени существования претерпевает динамические изменения, минимальные и максимальные значения классифицирующих признаков его элементов не могут быть зафиксированы и использованы для перехода к безразмерным величинам. Вследствие этого путем анализа априорной информации и физических особенностей признаков устанавливаются точная верхняя x_{\sup}^j и точная нижняя x_{\inf}^j грани для каждого из n признаков, после чего выполняется переход к безразмерным переменным y^j по формуле

$$y^j = M \frac{x^j - x_{\inf}^j}{x_{\sup}^j - x_{\inf}^j}, \quad (2)$$

где M – масштабирующий множитель, выбираемый из соображений удобства представления данных. Далее в тексте множество векторов $\{y^j\} \quad j \in N$ будем называть модифицированными значениями классифицирующих признаков.

Рассмотрим последовательно все этапы построения алгоритма кластеризации модифицированного массива данных в режиме динамических изменений. Он позволяет осуществить классификацию массива данных с большим числом классифицирующих признаков в самоорганизующемся режиме.

На множестве модифицированного массива данных $\{y^j\}$ выделим некоторое подмножество образов (точек), образующих более или менее тесную группировку, и построим кластерный сфероид $S^1(r, e_0^1)$ с центром в точке e_0^1 и радиусом r . Решения, принимаемые на этом этапе, являются исключительно эвристическими, поскольку формализованные рекомендации по выбору количества кластеров, положениям их центров, величинам радиусов дать невозможно. Поэтому выбор определяется сообразно целям исследования и во многом зависит от квалификации и субъективных предпочтений исследователя. Как бы то ни было, в состав кластера № 1 $S^1(r, e_0^1)$ войдут все элементы модифицированного массива, удовлетворяющие условию $\|e_0^1 - y^j\| \leq r$.

Положение точки e_0^1 , которая является геометрическим центром кластера $S^1(r, e_0^1)$, может совпасть с центром группировки (центром масс) элементов, находящихся в границах кластерного сфероиды, только случайно. Однако именно близость к центру группировки есть главное классификационное условие. Обозначим начальное положение центра группировки кластера № 1 как g_0^1 . Оно легко находится по формуле

$$g_0^1 = \frac{1}{|S^1(r, e_0^1)|} \sum_{y \in S^1(r, e_0^1)} y^j, \quad (3)$$

где $|S^1(r, e_0^1)|$ – мощность подмножества элементов в составе кластера $S^1(r, e_0^1)$. Перемещая геометрический центр кластера в точку g_0^1 , получим новый сфероид $S^1(r, e_1^1)$. В результате такого движения в сфероиде $S^1(r, e_1^1)$ появятся элементы массива данных, которых не было в $S^1(r, e_0^1)$, а часть элементов, входящих в состав $S^1(r, e_0^1)$, покинут кластер. Очевидно, что и в сфероиде $S^1(r, e_1^1)$, несовпадение геометрического центра e_1^1 и центра группировки g_1^1 также может иметь место. В соответствии с описанной процедурой вновь перемещаем центр сфероида в точку g_1^1 и получаем новый кластерный сфероид $S^1(r, e_2^1)$. Согласно известной теореме функционального анализа [8], последовательность точек $\{e_n^1\}$ сходится, а значит, за конечное число шагов может быть достигнуто выполнение условия

$$\|e_h^1 - e_{h-1}^1\| < \varepsilon, \quad (4)$$

где ε – любое наперед заданное положительно число, определяющее приемлемую меру близости между геометрическим центром кластерного сфероида и центром группировки находящихся в его составе элементов.

В процессе реализации данного алгоритма может оказаться, что величина радиуса кластерной сферы r и часть элементов модифицированного массива, которые должны были попасть в состав кластера $S^1(r, e_h^1)$, в него не вошли, и это не соответствует содержательному смыслу задачи. Тогда радиус сфероида увеличивается так, чтобы самый удаленный от точки e_h^1 объект из тех, которые должны войти в кластер, оказался на его границе. Обозначим $\{y^j\}$ – множество элементов модифицированного массива данных, которое необходимо дополнительно ввести в состав расширенного кластерного сфероида. Величина радиуса увеличенного кластера с центром в точке e_h^1 $S^1(R, e_h^1)$ при этом составит

$$R = \max_{\{y^j\}} \|e_h^1 - y^j\|. \quad (5)$$

Очевидно, что после изменения размеров кластера и введения в его состав новых элементов положение центра группировки опять сместится относительно точки геометрического центра e_h^1 . Однако с помощью описанного выше итерационного алгоритма это смещение устраняется.

Построение других кластерных сфероидов происходит по аналогичному сценарию. Начальные размеры кластеров, их необходимое количество m и положение их геометрических центров e_0^p $p = 1, m$ задается согласно целям проводимого исследования и особенностям обрабатываемого массива данных. После того как формирование системы кластеров по материалам достаточно представительной обучающей выборки будет завершено, можно приступить к эксплуатации системы в динамическом режиме, то есть в условиях действия потоков пополнения и гибели. Разработанный алгоритм, функционируя в режиме динамической кластеризации, проявляет важное и весьма ценное с практической точки зрения качество: самоорганизуемость. В ходе обмена элементами с внешней средой наполнение кластеров изменяется, и алгоритм без участия пользователя в автоматическом режиме непрерывно выполняет коррекцию положения кластерных сфероидов по заданной величине допустимого отклонения геометрических центров кластеров от центров их группировки. Следует заметить, что по мере наполнения кластеров смещение центров группировок заметно уменьшается и может не требовать корректирующего воздействия.

Особое внимание при реализации алгоритма следует обратить на возможность возникновения пересечений двух и более кластеров, которые могут появиться в процессе коррекции их положения и размеров. Этого ни в коем случае нельзя допускать, ибо тогда встает проблема неоднозначной классификации объектов, что дискредитирует саму идею использования кластерного анализа в классификационных процедурах. Для предотвращения подобной ситуации после каждой коррекции положения кластеров в алгоритме предусмотрена проверка условия

$$d(e_k^j, e_h^i) \leq r_k^j + r_h^i \quad \forall j \neq i. \quad (6)$$

Его выполнение означает, что расстояние между геометрическими центрами j -го и i -го кластеров оказалось меньше суммы их радиусов. В этом случае радиус одного из кластеров уменьшается до тех пор, пока пересечение не будет устранено. Подробная блок-схема алгоритма представлена на рисунке 1.

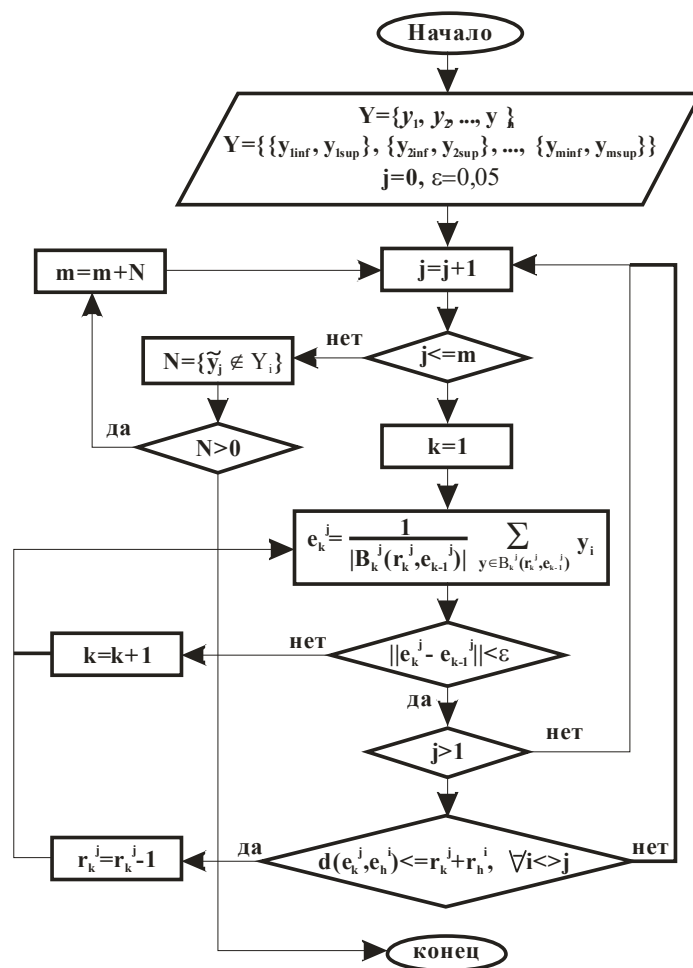


Рисунок 1 – Блок-схема алгоритма динамической кластеризации

Разработанный алгоритм показал высокие эксплуатационные характеристики на массивах больших данных, о чем будет более подробно сказано ниже при рассмотрении примеров его практического применения. Здесь же авторы считают необходимым упомянуть еще об одной особенности своей разработки. Как известно, сферы не могут заполнять пространство полностью. Более того, как указывает М. Гарднер [3], не получен ответ на вопрос о том, какую долю пространства произвольной размерности заполняют сферические тела при наиболее плотном способе упаковки. Отсюда следует, что в ходе построения классификационных категорий какая-то часть массива данных не будет классифицирована, поскольку не войдет в границы ни одного из сформированных кластерных сфероидов. В динамике доля объектов, не прошедших классификацию, может изменяться как в ту, так и в другую сторону, вместе с изменениями положения кластеров, и ликвидировать ее полностью невозможно. Этот недостаток, имманентно присущий рассмотренному алгоритму, обусловлен неизменяемыми свойствами кластеров как геометрических тел. Отмеченный факт, конечно, следует иметь в виду при проведении классификации. Однако он не может оказать сколь-нибудь существенного влияния на результаты классификации и поставить под сомнение их достоверность. С практической точки зрения, действительно, если из миллиона данных массива две или три тысячи объектов не будут классифицированы, то общую картину это никак не изменит.

Для демонстрации возможностей алгоритма и их геометрической иллюстрации был выбран сравнительно небольшой массив (несколько сотен элементов), составленный из городов России, классифицирующими признаками которых является площадь занимаемой ими территории W и численность населения X – эти сведения имеются в интернете и открыты для доступа (например, https://wiki2.org/ru/Список_городов_России_с_территорией_больше_100_квадратных_километров). В качестве точной нижней грани по признаку X , т.е. X_{inf} , была выбрана величина 10000, поскольку населенные пункты с меньшей численностью жителей не имеют в РФ статуса города. В качестве же X_{sup} было взято число 5000000, поскольку городов с населением, превышающим 5 миллионов жителей, в России только 2, и они имеют особый статус и структуру. По тем же основаниям в качестве

точной нижней границы W_{inf} принята площадь 1 км², а в качестве точной верхней границы W_{sup} – 1400 км², так как город Санкт-Петербург, население которого 5,2 миллиона человек, занимает именно такую территорию.

Применяя формулу (2), преобразуем признаки X и W к безразмерным величинам с масштабированием множителем $M = 100$. Фрагменты этого множества приведены в таблице 1. На первом этапе классификации было решено сформировать три кластерных образования, начальные положения геометрических центров которых расположены в точках $e_0^1 = (255000, 75,5)$; $e_0^2 = (750000, 425,5)$; $e_0^3 = (1350000, 850,5)$, координаты их выражены в натуральных единицах населения и площади. Радиусы же в безразмерных величинах оказались равными $r_0^1 = 7,2$; $r_0^2 = 20,2$; $r_0^3 = 12,7$. Процесс пошагового преобразования кластерных сфероидов с параметром точности $\varepsilon = 0,5$ представлен в таблице 2 и визуализирован на рисунке 2. На нем пунктирными линиями очерчены промежуточные положения кластеров, а сплошными – их финальные состояния и размеры.

Таблица 1 – Фрагмент списка городов РФ с указанием численности населения и территории

№	Город	Население, X	Y	Площадь, км ² , W	V
1	Новосибирск	1 602 915	41,92	506,67	46,14
2	Уфа	1 115 560	32,15	707	60,46
3	Орск	230 414	14,41	621,33	54,34
4	Казань	1 231 878	34,48	314,16	53,82
5	Волжский	326 055	16,33	229,12	26,30
6	Омск	1 178 391	33,41	566,9	50,45
7	Самара	1 169 719	33,24	541,382	48,62
8	Ростов-на Дону	1 125 299	32,35	348,5	34,83
...
137	Дербент	123 162	12,26	69,63	14,90
...
284	Кинешма	185 368	18,73	78,67	15,71
...
306	Энгельс	208 604	20,12	84,47	16,12

Таблица 2 – Этапы преобразования кластеров

№	Центр кластера [население, площадь]	Радиус r
1		
1	[220698, 102,5]	7,2
2	[227734, 108,1]	7,2
2		
1	[567468, 334,4]	10,2
2	[461127, 316,11]	8,2
3	[442760, 305,50]	7,2
4	[439437, 301,9]	8,2
5	[443997, 293,5]	6,2
3		
1	[10559716, 788,7]	29,7
2	[978274, 557,8]	13,7
3	[1024047, 214,5]	11,7
4	[1073611, 485,8]	10,7
5	[1117677, 480,0]	11,7
6	[1117677, 480,0]	10,7

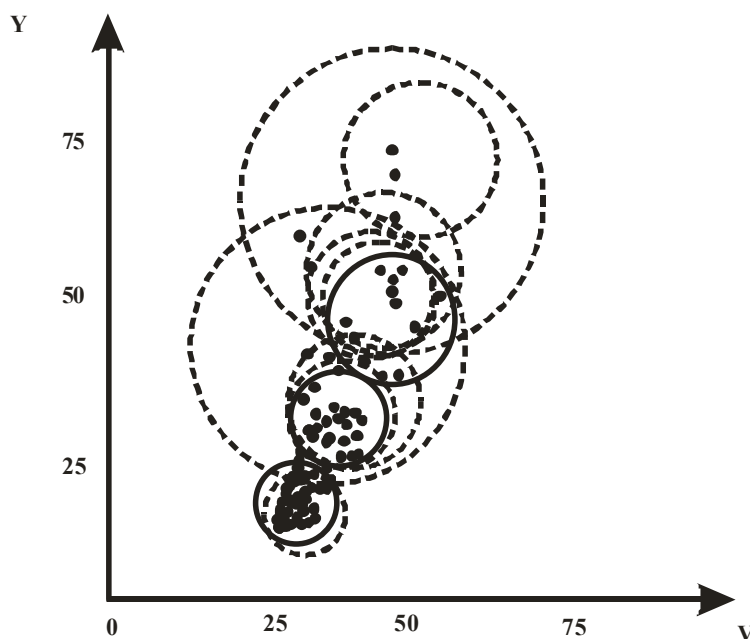


Рисунок 2 – Процесс коррекции кластерных сфероидов

Из данных таблицы видно, что финальная позиция первого кластера была достигнута после двух корректирующих воздействий, не затронувших величину радиуса. Это объясняется тем, что плотность группировки в зоне кластера S^1 наиболее высока (малые города больше похожи друг на друга, чем мегаполисы). Работа со вторым и третьим кластером потребовала пять и шесть корректирующих воздействий соответственно. Они коснулись не только положений геометрических центров, но и радиусов. Это оказалось обязательным для предотвращения пересечения кластеров. Интересным результатом, по мнению авторов, представляющим самостоятельное значение, является значительное (почти в три раза) уменьшение радиуса кластерного сфероида S^3 по сравнению с начальным и заметное смещение положения его геометрического центра в сторону меньших значений признака X -населения. Это указывает на то, что города с населением 800–900 тысяч жителей ближе по структуре к мегаполисам, чем к городам, население которых насчитывает 500–600 тысяч человек, а также подтверждает тезис о самоорганизации алгоритма.

На рисунке 2 ясно видно, что часть объектов исходного массива данных не вошла в состав ни одного из трех созданных кластеров, и так как поместить их туда не представляется возможным, то они остаются вне классификации. Отдельно для таких объектов «изгоев» алгоритм был запущен еще раз, что привело к созданию еще трех кластеров: S^4 , S^5 , S^6 , что визуализировано на рисунке 3. Их радиусы равны $r_0^4 = 11,2$; $r_0^5 = 7,2$; $r_0^6 = 1,9$. Число элементов, вошедших в состав этих кластеров, меньше чем в трех первых, что несколько не снижает значимость полученных результатов. В условиях потоков пополнения и гибели, а также естественных изменений объектов (рост или убыль населения, изменения границ городской черты и т.п.), положение кластеров и мощность множеств элементов в их составе будет меняться, что и делает актуальной идею динамической кластеризации.

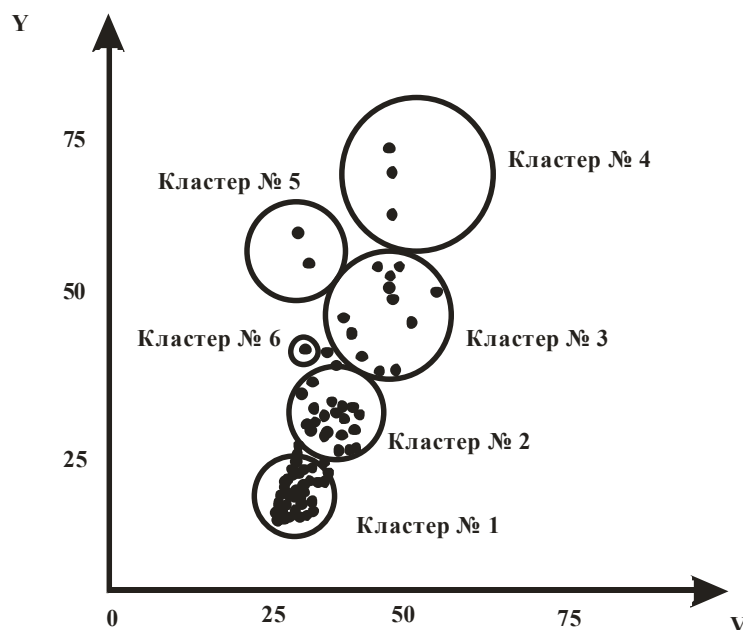


Рисунок 3 – Этапы и повторный результат расчетов, выполненные в соответствии с разработанным алгоритмом

Авторы располагают сведениями о применении описанного здесь алгоритма для обработки больших данных. По заказу одной территориальной телекоммуникационной компании была предпринята успешная попытка классификации ее клиентов. В качестве классификационных признаков, которых набралось более 100, были предложены различные услуги и сервисы, предоставляемые пользователям и некоторые данные пользователей. Общий объем клиентской базы компании составлял ≈ 1000000 единиц, имея незначительные колебания нерегулярного характера по времени. Обработка этого массива позволила выделить базы 9 категорий пользователей, различающихся по характеру их предпочтений. Результаты были использованы для создания портрета «типичного клиента» и выработки рациональной ценовой политики по комплексу предлагаемых опций. К сожалению, мы не можем опубликовать эти материалы, поскольку они являются собственностью компании, где проводилось исследование. Руководство компании не сочло возможным дать разрешение на их публикацию, считая эти материалы коммерческой тайной.

Выводы:

1. Разработан алгоритм, позволяющий осуществлять кластеризацию массивов «больших данных» в динамике их изменений.
2. Как показала практика, алгоритм эффективен для обработки данных, характеризуемых большим числом количественно измеримых признаков.
3. Алгоритм обладает свойствами самоорганизующести и самообучаемости, т.е. в автоматическом режиме осуществляет коррекцию положения и размеров кластеров в факторном пространстве с учетом пространственных ограничений.
4. В работе представлен и проиллюстрирован демонстрационный пример с подробной интерпретацией результатов.

Библиографический список

1. Аль-Хашеди А. А. Экспертиза методов, используемых в различных задачах распознавания образов / А. А. Аль-Хашеди, Е. А. Печеный, Н. К. Нуриев // Вестник технологического университета. – 2017. – Т. 20, № 1. – С. 125–127.
2. Воронцов К. В. Алгоритмы кластеризации и многомерного шкалирования / К. В. Воронцов. – Москва : МГУ, 2007. – 234 с.
3. Гарднер М. Математические головоломки и развлечения / М. Гарднер. – Москва : Мир, 1971. – 510 с.
4. Дюран Б. Кластерный анализ / Б. Дюран, П. Оделл. – Москва : Статистика, 1977. – 128 с.
5. Ершов К. С. Анализ и классификация алгоритмов кластеризации / К. С. Ершов, Т. Н. Романова // Новые информационные технологии в автоматизированных системах. – 2016. – № 19. – С. 274–279.
6. Журавлёв Ю. И. Распознавание. Классификация. Прогноз. Математические методы и их применение / Ю. И. Журавлёв. – Москва : Наука, 1989. – 302 с.
7. Игнатъев Н. А. Кластерный анализ данных и выбор объектов-эталонов в задачах распознавания с учителем / Н. А. Игнатъев // Вычислительные технологии. – 2015. – Т. 20, № 6. – С. 36–45.

8. Колмогоров А. Н. Элементы теории функций и функционального анализа / А. Н. Колмогоров, С. В. Фомин. – Москва : Физматлит, 2004. – 572 с.
9. Низаметдинов Ш. У. Анализ данных : учебное пособие / Ш. У. Низметдинов, А. П. Румянцев. – Москва : МИФИ, 2012. – 286 с.
10. Нурiev Н. К. Математическое моделирование эволюции кластерных образований / Н. К. Нурiev, А. А. Аль-Хашеди, Е. А. Печеный // *Современные наукоемкие технологии*. – 2018. – № 8. – С. 110–116.
11. Coates A. Learning Feature Representations with K-means / A. Coates, A. Y. Ng. // *Neural Networks: Tricks of the Trade*. – 2012. – P. 561–580.
12. Berkhim P. Survey of Clustering Data Mining Techniques / P. Berkhim // *Accue Software*. – 2002. – 160 p.
13. Griffin G. Learning and using taxonomies for fast visual categorization / G. Griffin, P. Perona // *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. – 2008. – P. 1–8.
14. Ilango M. A survey of grid based clustering algorithms / M. Ilango, V. Mohan // *Intern. J. of Eng. Sci. and Technology*. – 2010. – Vol. 2 (8). – P. 3441–3446.
15. Jain A. K. Data Clustering: A review / A. K. Jain, M. N. Marty, P. J. Flynn // *ACM Computing Surveys*, 1999. – Vol. 31, № 3. – P. 264–323.
16. Kagan J. Clustering Lange and High Dimensional Data / J. Kagan, C. Nicholas, M. Teloulle. – Режим доступа: <http://www.csee.umtc.edu/nicholas/clustering/tutorial.pdf>, свободный. – Заглавие с экрана. – Яз. рус.
17. Kumar J. Parallel using large data sets / J. Kumar, R. T. Mills, F. M. Hoffman, W. W. Hargrove // *Procedia Computer Science*. – 2011. – № 4. – P. 1602–1611.
18. Shaukat K. Dengue Fever in Perspective of Clustering Algorithms / K. Shaukat, N. Masood, A. B. Shafaat, K. Jabbar, H. Shabbir et al. // *Data Mining Genomics Proteomics*. – 2015. – Vol. 6, № 176. DOI:10.4172/2153-0602.1000176.
19. Ximing Lv. Research on P2P Network Loan Risk Evaluation Based on Generalized DEA Model and R-Type Clustering Analysis under the Background of Big Data / Lv. Ximing, Lan Zhou, Xiaona Guo // *Financial Risk Management*. – 2017. – Vol. 6, № 2. – P. 163–190.
20. Yin J. A. Developmental approach to structural self-organization in reservoir computing / J. Yin, Y. Meng, Y. Jin // *IEEE Transactions on Autonomous Mental Development*. – 2012. – № 4 (4). – P. 273–289.

References

1. Al-Khashedi A. A., Pecheny E. A., Nuriev N. K. Ekspertiza metodov, ispolzuyemykh v razlichnykh zadachakh raspoznavaniya obrazov [Examination of methods used in various pattern recognition tasks]. *Vestnik tekhnologicheskogo universiteta* [Bulletin of the University of Technology], 2017, vol. 20, no. 1, pp. 125–127.
2. Vorontsov K. V. *Algoritmy klasterizatsii i mnogomernogo shkalirovaniya* [Algorithms for clustering and multidimensional scaling]. Moscow, Moscow State University Publ., 2007. 234 p.
3. Gardner M. *Matematicheskiye golovolomki i razvlecheniya* [Math puzzles and fun]. Moscow, Mir Publ., 1971. 510 p.
4. Durant B., Odell P. *Klasternyy analiz* [Cluster analysis]. Moscow, Statistika Publ., 1977. 128 p.
5. Ershov K. S., Romanova T. N. Analiz i klassifikatsiya algoritmov klasterizatsii [Analysis and classification of clustering algorithms]. *Novyye informatsionnyye tekhnologii v avtomatizirovannykh sistemakh* [New information technologies in automated systems], 2016, no. 19, pp. 274–279.
6. Zhuravlev Yu. I. *Raspoznavaniye. Klassifikatsiya. Prognoz. Matematicheskiye metody i ikh primeneniye* [Recognition. Classification. Forecast. Mathematical methods and their application]. Moscow, Nauka Publ., 1989. 302 p.
7. Ignatiev N. A. Klasternyy analiz dannykh i vybor obektov-etalonov v zadachakh raspoznavaniya s uchitelem [Cluster data analysis and selection of objects-standards in tasks of recognition with a teacher]. *Vychislitel'nyye tekhnologii* [Computational Technologies], 2015, vol. 20, no. 6, pp. 36–45.
8. Kolmagorov A. N., Fomin S. V. *Elementy teorii funktsiy i funktsionalnogo analiza* [Elements of the theory of functions and functional analysis]. Moscow, Fizmatlit Publ., 2004. 572 p.
9. Nizametdinov W. W., Romyantsev A. P. Analiz dannykh : uchebnoye posobiye [Data Analysis : Tutorial]. Moscow, Moscow Engineering Physics Institute, 2012. 286 p.
10. Nuriev N. K., Al-Hashedi A. A., Pecheny E. A. Matematicheskoye modelirovaniye evolyutsii klasternykh obrazovaniy [Mathematical modeling of the evolution of cluster formations]. *Sovremennyye naukoymkiye tekhnologii* [Modern High Technologies], 2018, no. 8, pp. 110–116.
11. Coates A., Ng A. Y. Learning Feature Representations with K-means. *Neural Networks: Tricks of the Trade*, 2012, pp. 561–580.
12. Berkhim P. Survey of Clustering Data Mining Techniques. *Accue Software*, 2002. 160 p.
13. Griffin G., Perona P. Learning and using taxonomies for fast visual categorization. *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2008, pp. 1–8.
14. Ilango M., Mohan V. A survey of grid based clustering algorithms. *Intern. J. of Eng. Sci. and Technology*, 2010, vol. 2 (8), pp. 3441–3446.
15. Jain A. K., Marty M. N., Flynn P. J. Data Clustering: A review. *ACM Computing Surveys*, 1999, vol. 31, no. 3, pp. 264–323.
16. Kagan J., Nicholas C., Teloulle M. *Clustering Lange and High Dimensional Data*. Available at: <http://www.csee.umtc.edu/nicholas/clustering/tutorial.pdf>.

17. Kumar J., Mills R. T., Hoffman F. M., Hargrove W. W. Parallel using large data sets. *Procedia Computer Science*, 2011, no. 4, pp. 1602–1611.
18. Shaukat K., Masood N., Shafaat A. B., Jabbar K., Shabbir H. et al. Dengue Fever in Perspective of Clustering Algorithms. *Data Mining Genomics Proteomics*, 2015, vol. 6, no. 176. DOI:10.4172/2153-0602.1000176.
19. Ximing Lv, Zhou Lan, Guo Xiaona. Research on P2P Network Loan Risk Evaluation Based on Generalized DEA Model and R-Type Clustering Analysis under the Background of Big Data. *Financial Risk Management*, 2017, vol. 6, no. 2, pp. 163–190.
20. Yin J. A., Meng Y., Jin Y. Developmental approach to structural self-organization in reservoir computing. *IEEE Transactions on Autonomous Mental Development*, 2012, no. 4 (4), pp. 273–289.

DOI 10.21672/2074-1707.2020.49.4.020-032

УДК 004:614

НЕЙРОСЕТЕВАЯ ТЕХНОЛОГИЯ ОБНАРУЖЕНИЯ АНОМАЛЬНОГО СЕТЕВОГО ТРАФИКА

Статья поступила в редакцию 05.12.2019, в окончательном варианте – 26.02.2020.

Частикова Вера Аркадьевна, Кубанский государственный технологический университет, 350072, Российская Федерация, г. Краснодар, ул. Московская, 2, кандидат технических наук, доцент, e-mail: chastikova_va@mail.ru

Жерлицын Сергей Анатольевич, Кубанский государственный технологический университет, 350072, Российская Федерация, г. Краснодар, ул. Московская, 2, студент, e-mail: kpytooooo@gmail.com

Воля Яна Игоревна, Кубанский государственный технологический университет, 350072, Российская Федерация, г. Краснодар, ул. Московская, 2, студент, e-mail: volya_y@mail.ru

Сотников Владимир Владимирович, Кубанский государственный технологический университет, 350072, Российская Федерация, г. Краснодар, ул. Московская, 2, студент, e-mail: bubert9@mail.ru

Рассмотрены существующие методы анализа сетевого трафика, указаны их возможности и ограничения. Продемонстрирована актуальность решаемой задачи. Обоснована целесообразность использования нейросетевого подхода к обнаружению аномалий сетевого трафика. Исследована эффективность использования алгоритмов роевого интеллекта применительно к задаче обучения нейронных сетей, выявлены особенности данных алгоритмов. Реализована объектно-ориентированная библиотека для выявления сетевых атак с использованием нейросети с архитектурой многослойного перцептрона. На данном этапе исследования для обучения нейросети и оценки качества распознавания трафика был применен датасет KDD Cup 1999 Data. Описаны преимущества и недостатки реализованного решения. Представлен способ устранения распространенного недостатка датасетов, связанного с несбалансированностью обучающих данных. Описаны используемые технологии: архитектура нейронной сети, алгоритм обучения, способ уменьшения размерности обрабатываемых данных. На втором этапе был использован набор данных CSE-CIC-IDS2018. Предложена нейросетевая модель, построенная на базе архитектуры LSTM и эмбединговой сетей. Для обучения разработанной системы предложено применение алгоритма Adam, основанного на градиентном спуске. На основе использования названных алгоритмов, моделей и технологий был реализован, а затем и протестирован программный комплекс для обнаружения сетевых атак.

Ключевые слова: нейронная сеть, сетевая атака, многослойный перцептрон, роевой интеллект, LSTM-сеть, эмбединговая сеть, Focal Loss, алгоритм Adam

NEURAL NETWORK TECHNOLOGY FOR DETECTING ANOMALOUS NETWORK TRAFFIC

The article was received by the editorial board on 05.12.2019, in the final version – 26.02.2020.

Chastikova Vera A., Kuban State Technological University, 2 Moskovskaya St., Krasnodar, 350072, Russian Federation,

Cand. Sci. (Engineering), Associate Professor, e-mail: chastikova_va@mail.ru

Zherlitsyn Sergey A., Kuban State Technological University, 2 Moskovskaya St., Krasnodar, 350072, Russian Federation,

student, e-mail: kpytooooo@gmail.com

Volya Yana I., Kuban State Technological University, 2 Moskovskaya St., Krasnodar, 350072, Russian Federation,

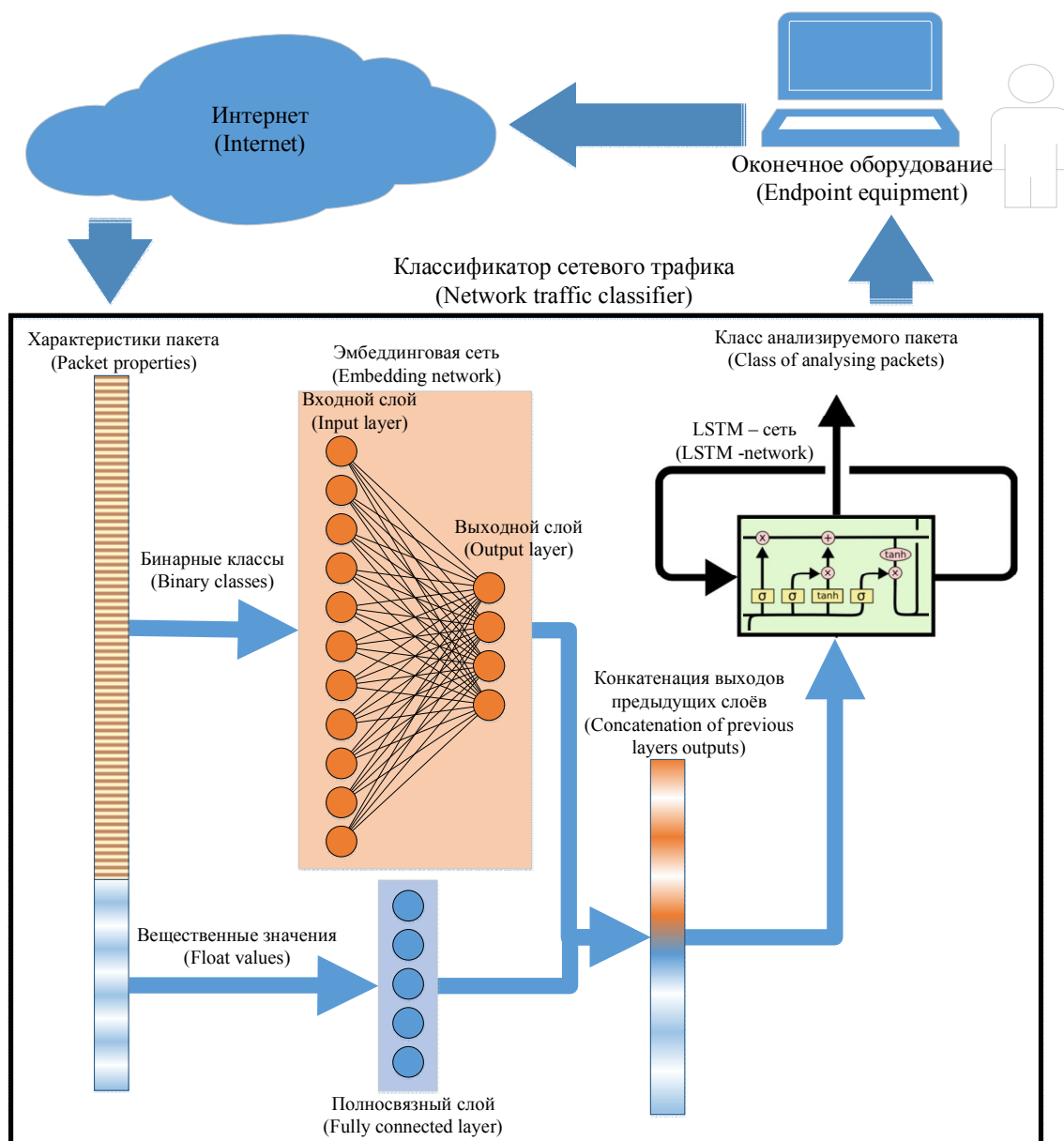
student, e-mail: volya_y@mail.ru

Sotnikov Vladimir V., Kuban State Technological University, 2 Moskovskaya St., Krasnodar, 350072, Russian Federation, student, e-mail: bubert9@mail.ru

Existing network traffic analysis methods are considered. The relevance of the problem is demonstrated. The efficiency of swarm intelligence algorithms as applied to the task of training neural networks is analyzed, the features of these algorithms are revealed. An object-oriented library for detecting network attacks using a multi-layer perceptron neural network architecture has been implemented. The advantages and disadvantages of the implemented solution are described. A method is proposed for eliminating the widespread lack of datasets related to the imbalance of training data. The technologies used are described: the architecture of the neural network, the learning algorithm, a method of reducing the dimension of the processed data. A neural network model based on the LSTM architecture and embedding networks is proposed. To train the developed system, the use of the Adam algorithm based on gradient descent is proposed. Based on the use of the above algorithms, models and technologies, a software package for detecting network attacks was implemented and then tested.

Key words: neural network, network attack, perceptron, swarm intelligence, LSTM network, embedding network, Focal Loss, Adam algorithm

Graphical annotation (Графическая аннотация)



Введение. Функционирование современного информационного общества требует обеспечения стабильного доступа физических лиц и организаций к базам данных и различным сетевым ресурсам в любое время. Однако такой доступ может прерываться, в том числе и из-за сетевых атак. Они производятся с целью нарушения работы информационной системы или затруднения доступа к ней; могут привести к значительным убыткам пострадавших лиц и компаний.

Актуальность данной проблемы определяется большим количеством происходящих в мире сетевых атак, что видно из графика на рисунке 1 [9].

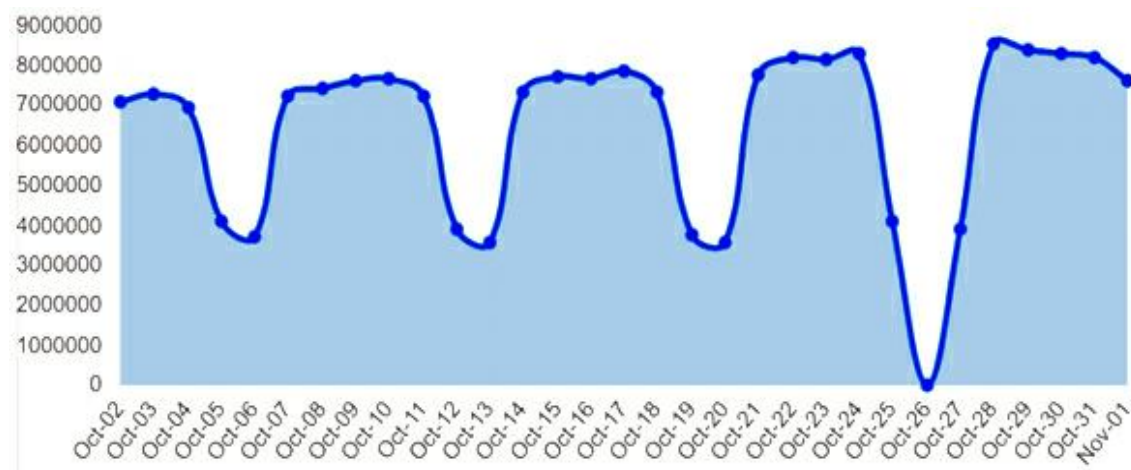


Рисунок 1 – Количество сетевых атак за период с 02.10.2019 до 01.11.2019. Сетевым атакам подвергаются устройства жителей и организаций по всему миру. Согласно статистике «Лаборатории Касперского», количество атак с каждым годом увеличивается, при этом их сложность также растет [9]. Стандартные средства детектирования сетевых атак не всегда способны обнаружить угрозы, что может привести к утрате работоспособности интернет-ресурсов и значительному ущербу для компании

Общая характеристика существующих методов обнаружения сетевых атак. Один из распространенных способов обнаружения сетевых атак – это анализ IP-адресов, с которых приходят запросы на веб-сервер, и анализ данных сетевого трафика на полное совпадение с имеющейся базой признаков сетевых атак. В связи с возрастающей сложностью угроз данный метод имеет низкую эффективность. Одна из возможностей обойти такой алгоритм обнаружения – скрыть или проксировать IP-адрес злоумышленника. При этом в случае реализации угрозы средствами защиты будут заблокированы хосты обычных пользователей, и они не смогут получить доступ к ресурсам.

Также существуют системы поиска сетевых атак, основанные на проверке соответствия сетевого трафика определенным гибким правилам. Данный способ позволяет отфильтровать пакеты данных с целенаправленно искаженным содержанием, переполнение буфера определённых сервисов, но потребует значительное время, чтобы проанализировать новые варианты сетевых атак и дополнить систему правил.

Помимо вышеперечисленного в настоящее время используются такие методы обнаружения сетевых атак, как статистические методы, экспертные системы и нейронные сети. Они применяются как комплексно, так и в качестве отдельных средств анализа сетевого трафика. Такой подход является более гибким и позволяет упростить процесс актуализации системы безопасности [1].

Данная работа посвящена исследованию и разработке средств защиты на базе нейронных сетей, способных отличить вредоносные подключения от нормальных с целью последующего принятия мер по обеспечению безопасности защищаемой системы.

Подготовительный этап построения модели нейросетевой системы. Выбор архитектуры нейронной сети определяется параметрами анализируемой информации. Для обучения модели была использована база данных сетевых атак, находящаяся в открытом доступе KDD Cup 1999 Data [13], находящаяся в открытом доступе в интернете. Из всех приведенных на каждый пункт параметров сетевого трафика были выбраны 28 [5], на основе которых можно определить наличие вредоносной активности в сети. Все данные были конвертированы из текстового формата в цифровой и нормализованы для корректной работы нейронной сети.

Для названного датасета нормализация заключалась в приведении каждого параметра к диапазону от 0 до 1 в соответствии с формулой:

$$p_{in} = \frac{p_i - p_{min}}{p_{max} - p_{min}}$$

На основе анализа различных архитектур нейронных сетей для решения поставленной задачи был выбран многослойный перцептрон, так как его архитектура способна симулировать множественные условные взаимосвязи.

Подобная нейронная сеть принимает на вход вектор входных значений – параметры сетевого трафика. Каждый нейрон связан со всеми элементами последующего слоя, нейроны, в нашем случае, имеют сигмоидальную функцию активации [5].

Способ обучения нейросетевой системы. Для обучения сети были применены: генетический алгоритм, алгоритм серых волков, алгоритм светлячков [2–5].

Алгоритм светлячков [2] и алгоритм серых волков [3] показали высокую эффективность при решении задачи глобальной оптимизации многоэкстремальных функций. Это и стало критерием выбора их в качестве алгоритмов обучения нейронных сетей.

Каждый из рассматриваемых методов опирается на вычисление фитнес-функции и направлен на поиск её глобального минимума. В данной задаче эту роль выполняет среднееквадратическое отклонение результатов, полученных на выходе нейронной сети, от ожидаемых значений, хранящихся в базе. Общая схема полного цикла обучения нейронной сети представлена на рисунке 2.

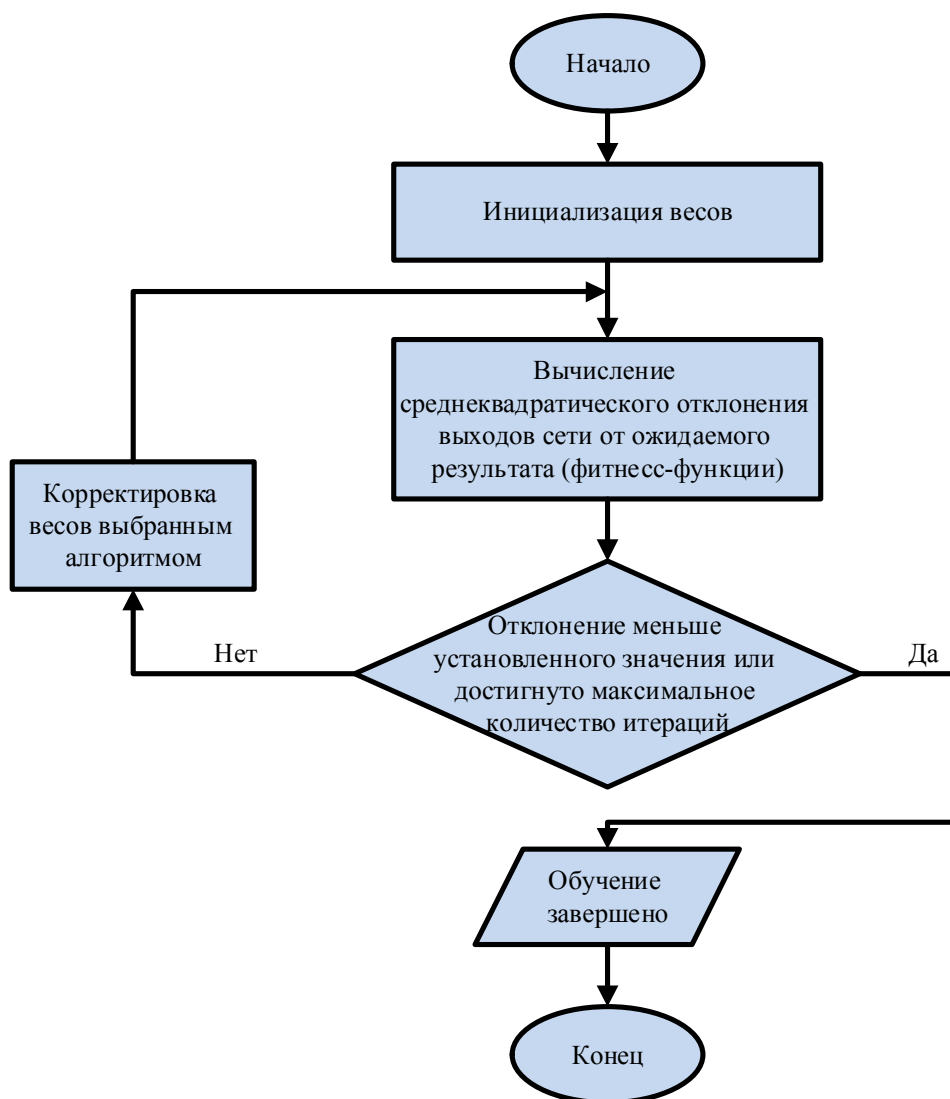


Рисунок 2 – Схема полного цикла обучения сети

Для предварительного тестирования эффективности алгоритмов при решении задачи обучения нейронной сети были разработаны и реализованы на языке программирования C# объектно-ориентированная библиотека и программный комплекс для создания, настройки, сохранения и обучения нейронных сетей различных видов, на основе которой проводились исследования [2–5]. Согласно их результатам, алгоритм светлячков показал высокое быстродействие и относительно низкую степень отклонения от точки глобального экстремума.

В рамках данного этапа исследования алгоритм светлячков использовался в качестве основного алгоритма обучения. Он продемонстрировал устойчивость при усложнении структуры нейронной сети: как при увеличении количества нейронов на скрытых слоях, так и при увеличении глубины сети. Результаты применения алгоритма при разных конфигурациях сети представлены на рисунке 3.

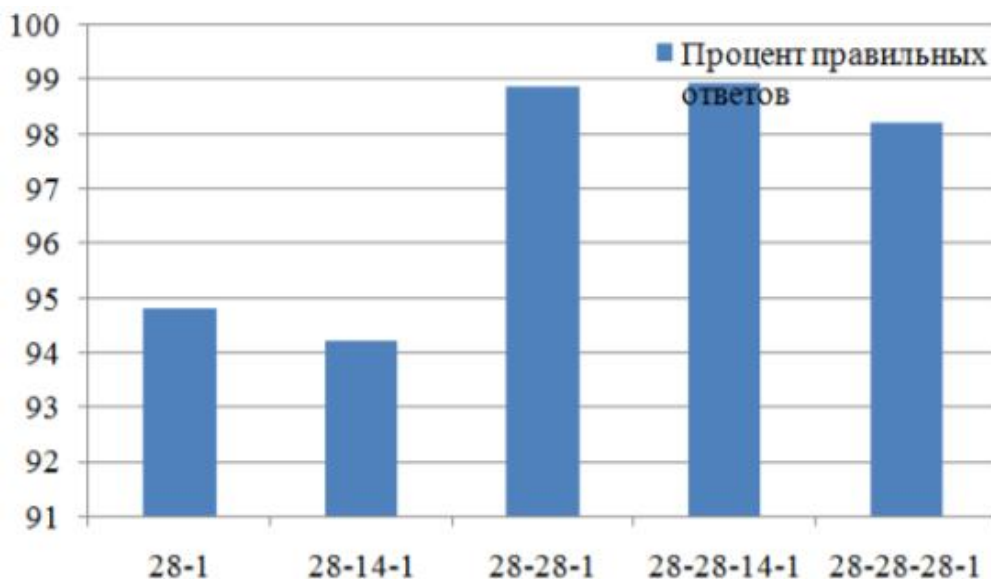


Рисунок 3 – Эффективность обучения различных конфигураций нейронной сети алгоритмом светлячков

В данной диаграмме подписи по оси ординат являются условными обозначениями архитектуры нейросети. Каждое из чисел каждой подписи столбца обозначает слой нейросети, а значение числа – количество нейронов на данном слое. Так, например, запись 28-14-1 указывает, что нейросеть обладает входным слоем с 28 нейронами, одним скрытым слоем с 14 нейронами и выходным слоем, состоящим из одного нейрона.

Результат промежуточного этапа позволил выявить недостатки разработанной модели. Однако более подробный анализ провести было невозможно в силу малого объема примеров отдельных видов атак, отсутствия средств мониторинга и неполноты (отдельные типы атак были представлены в критически малом объеме, менее 50 примеров) обучающего датасета.

Новая обучающая выборка. Продолжение исследования стало возможным благодаря появлению новой, более полной, актуальной обучающей выборки. Поэтому на втором этапе данной работы используется набор данных Канадского университета кибербезопасности CSE-CIC-IDS2018 – это один из крупнейших открытых датасетов на текущий момент [8]. Он был сгенерирован в специально развернутой сети, состоящей из 50 хостов злоумышленников, 420 клиентских хостов и 30 серверов, и включает в себя 7 различных векторов атак: брутфорс, эксплуатацию уязвимости Heartbleed, ботнет-атаку, атаку на отказ в обслуживании, распределенную атаку на отказ в обслуживании, атаки на веб-приложения, а также атаку на сеть изнутри с помощью бэkdора в одном из клиентских хостов [1].

Проблема несбалансированности обучающей выборки и предлагаемый подход к ее решению. Проблемой, характерной для всех датасетов с аномальным сетевым трафиком, является крайняя степень несбалансированности примеров для различных классов данных: доброкачественный трафик может составлять до 80 % от общего объема обучающей выборки, в то время как отдельные векторы атак могут быть представлены в менее чем одном проценте сетевых пакетов. Легко классифицируемые доброкачественные пакеты вносят наибольший вклад в градиент функции потерь, на которой основаны применяемые алгоритмы обучения, негативным образом влияя на обучение.

Для борьбы с этой проблемой было принято решение использовать функцию потерь Focal Loss (FL) [20], изначально предложенную группой Facebook AI Research для использования в задачах распознавания образов, но уже успешно апробированную в других областях классификации в условиях несбалансированности классов. Она применяет веса не к классам сетевых атак в целом, а к конкретным пакетам, классификация которых вызвала наибольшие затруднения у модели [11].

FL основана на подсчете перекрестной энтропии между предсказанным и истинным распределением вероятности появления класса в выборке. При этом стандартная формула

$$CE(P_t) = -\log \log (P_t)$$

дополняется множителем:

$$FL(P_t) = -(1 - P_t)^\gamma * \log \log (P_t),$$

где γ – гиперпараметр функции.

Это позволяет сбалансировать влияние на результат функции потерь каждого из классов данных, представленных в обучающей выборке.

Проблема разреженности данных. Некоторые из используемых при обучении характеристик пакетов из датасета [8] хранились не в числовой форме, а в виде названий и терминов, в общем виде – строк. Для предотвращения потери данных, оптимальным решением является применение унитарного кодирования. Каждый такой строковый параметр заменяется набором колонок, каждая из которых соответствует одному из возможных текстовых значений параметра. При этом в колонку, соответствующую содержащемуся в конкретном примере значению раскладываемого параметра, вносится число 1, а остальные колонки заполняются нулями.

Ввиду того, что достаточно большое количество информационных полей каждого сетевого пакета закодировано унитарным кодом, обучающая выборка представляет собой разреженную матрицу. В целях снижения размерности обрабатываемых данных часть столбцов данной матрицы отображается в компактные вектора-эмбединги (от англ. embedding – вложение) [19] средствами самой нейронной сети, которые затем конкатенируются с остальными информационными признаками.

Эмбединговый фрагмент нейросети представлен в виде 2-х полносвязных слоёв, первый из которых обладает значительно большей размерностью, чем второй.

Данная конструкция является первым функциональным блоком модели. Он располагается параллельно с одним полносвязным слоем, на который подаётся небинарная часть характеристик пакета сетевого трафика. Графическая визуализация эмбединговой нейросети приведена на рисунке 4.

Приведенная схема демонстрирует главное свойство нейросети названного типа – количество входных параметров значительно больше количества выходных, что способствует повышению их компактности и понижению размерности в целях экономии вычислительных ресурсов.

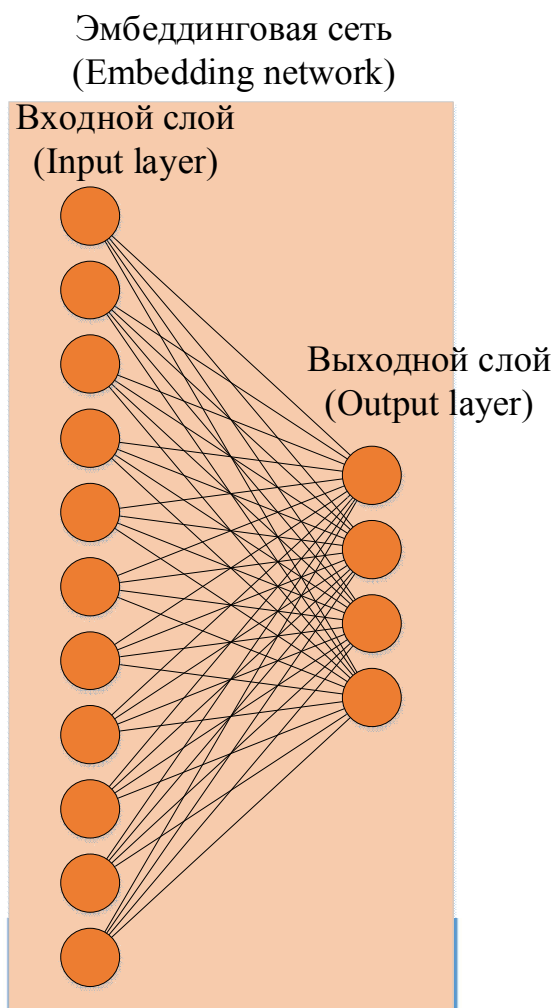


Рисунок 4 – Схема эмбединговой сети

Рекуррентные нейронные сети с долгой кратковременной памятью (LSTM). LSTM-сети (англ. long short-term memory – долгая кратковременная память) [12] являются одной из наиболее эффективных существующих нейросетевых моделей для обработки упорядоченных наборов данных (наряду с Gated Recurred Units сетями) [10, 12]. В отличие от классических рекуррентных сетей, они позволяют аккумулировать в себе информацию за продолжительное время, позволяя избежать проблему взрывающихся и затухающих градиентов [10], путем хранения своего состояния $s_i^{(t)}$ в нейронах одних скрытых слоев и управления этим состоянием с помощью нейронов других слоев, скрытых внутри LSTM-ячеек. Эти нейроны управляются рекуррентными связями внутри каждой ячейки. Схематичное изображение LSTM-сети приведено на рисунке 5 [12].

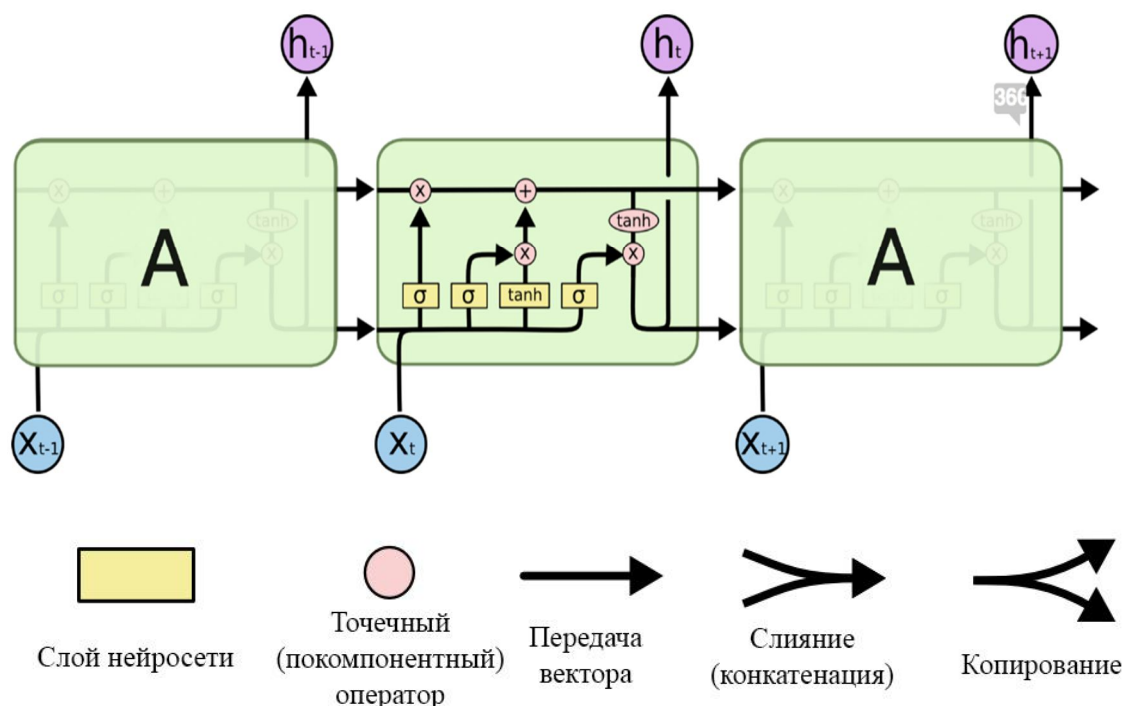


Рисунок 5 – Пример однослойной LSTM-сети: x_t – t -й вектор входных данных; h_t – t -й вектор выходных данных; σ – сигма-функция активации; \tanh – функция активации – гиперболический тангенс; + и \times – операторы покомпонентного сложения и умножения векторов соответственно

LSTM-сеть состоит из рекуррентно соединенных ячеек, заменяющих собой привычный для классических рекуррентных сетей скрытый слой. Входные данные обрабатываются обычными искусственными нейронами. Значения их активаций могут быть аккумулированы нейронами состояния (нейронами, конвертирующими входные данные и данные с предыдущего состояния в вектор текущего состояния), если это позволяет значение входного вентиля. Нейроны состояния соединены линейной (то есть не содержащей функции активации) рекуррентной связью друг с другом, которая управляется вентиляем забывания (рис. 5). Выходное значение каждой LSTM-ячейки управляется соответствующим вентиляем забывания. Каждый вентиль снабжен сигмоидальной функцией активации, в то время как входной слой может быть реализован с помощью любой нелинейной функции.

Подобная архитектура позволяет LSTM-сети аккумулировать информацию на протяжении долгого времени. Однако, после того как информация была использована, ячейка может «забыть» свое старое состояние с помощью соответствующего вентиля. К примеру, если обрабатываемый трафик состоит из нескольких сессий и для качественного анализа нейросети необходимо помнить контекст каждой из сессий в отдельности, то LSTM-сеть может обучиться автоматически забывать старый контекст при запуске обработки новой сетевой сессии.

Алгоритм обучения нейронной сети. Для оптимизации нейронной сети был выбран алгоритм Adam (“adaptive moments”), являющийся градиентным алгоритмом оптимизации первого порядка с адаптивной скоростью обучения [16]. Он совмещает преимущества таких методов оптимизации, как RMSProp и momentum [16], обладая при этом большей вычислительной устойчивостью, особенно на ранних этапах обучения. Это достигается коррекцией смещения оценки моментов первого и второго порядков [11]. Алгоритм работает следующим образом:

1. Задать входные параметры:
 - 1.1. Размер шага ϵ (оптимальное значение: 0,001).
 - 1.2. Скорости убывания оценок моментов первого и второго порядка $\rho_1, \rho_2 \in [0,1]$ (оптимальные значения: 0,9 и 0,999 соответственно).
 - 1.3. Константу δ , обеспечивающую вычислительную стабильность (оптимальное значение: 10^{-8}).
 - 1.4. Начальные параметры θ (инициализируются случайным образом).
2. Инициализировать моменты первого и второго порядков $s = 0, r = 0$ и шаг $t = 0$.
3. Пока не будет получена заданная точность, итеративно повторять выполнение следующих шагов:

3.1. Извлечь минибатч из m элементов из обучающей выборки $\{x^{(1)}, \dots, x^{(m)}\}$, а также соответствующие им метки $y^{(i)}$.

3.2. Вычислить градиент $g \leftarrow \frac{1}{m} \nabla_{\theta} \sum_i L(f(x^{(i)}; \theta), y^{(i)})$.

3.3. Увеличить шаг $t \leftarrow t + 1$.

3.4. Вычислить несмещенную оценку момента первого порядка $s \leftarrow \rho_1 s + (1 - \rho_1)g$.

3.5. Вычислить несмещенную оценку момента второго порядка $r \leftarrow \rho_2 r + (1 - \rho_2)g \odot g$.

3.6. Вычислить смещение момента первого порядка: $\hat{s} \leftarrow \frac{s}{1 - \rho_1^t}$.

3.7. Вычислить смещение момента второго порядка: $\hat{r} \leftarrow \frac{r}{1 - \rho_2^t}$.

3.8. Вычислить изменение параметров: $\Delta\theta = -\epsilon \frac{\hat{s}}{\sqrt{\hat{r} + \delta}}$.

3.9. Обновить параметры: $\theta \leftarrow \theta + \Delta\theta$.

Сравнительный анализ указанных алгоритмов оптимизации приведен на рисунке 6.

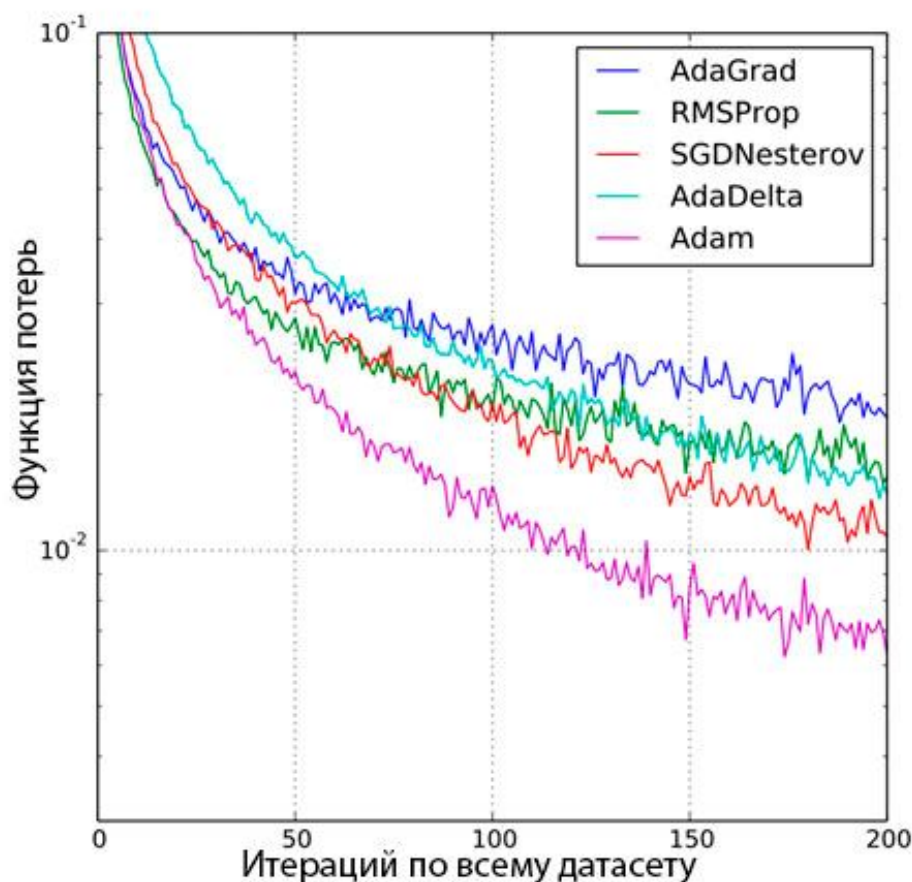


Рисунок 6 – Сравнение алгоритмов оптимизации

Из приведённого графика видно, что выбранный в качестве основного в данном исследовании алгоритм Adam демонстрирует наибольшую производительность, обеспечивая значительный отрыв в прогрессе уменьшения функции потерь уже с 30-й итерации.

Итоговая нейросетевая архитектура. С учетом особенностей, используемых для обучения данных и функций описанных элементов, была разработана нейросетевая архитектура, организованная следующим образом (табл.).

Таблица – Классы сетевого взаимодействия

Название слоя (тип слоя)	Длина выходного вектора	Число параметров	Соединен со слоем
Traffic (входной слой)	196	0	-
Lamda_1 (лямбда-выражение)	157	0	Traffic
Dense_1 (полносвязный слой)	128	20224	Lamda_1
Lamda (лямбда-выражение)	39	0	Traffic
Batch_normalizationv1 (слой батч-нормализации)	128	512	Dense_1
Dense (полносвязный слой)	16	640	Lamda
Dense_2 (полносвязный слой)	64	8256	Batch_normalizationv1
Batch_normalizationv2 (слой батч-нормализации)	16	64	Dense
Batch_normalizationv22 (слой батч-нормализации)	64	256	Dense_2
Concatenate (конкатенирующий слой)	80	0	Batch_normalizationv2, Batch_normalizationv22
Bidirectional (двунаправленная LSTM-сеть)	64	28928	Concatenate
Dense_3 (полносвязный слой)	16	1040	Bidirectional
Batch_normalizationv23 (слой батч-нормализации)	16	64	Dense_3
Dense_4 (полносвязный слой)	10	170	Batch_normalizationv23

Графическая визуализация данного набора слоёв представлена на рисунке 7.

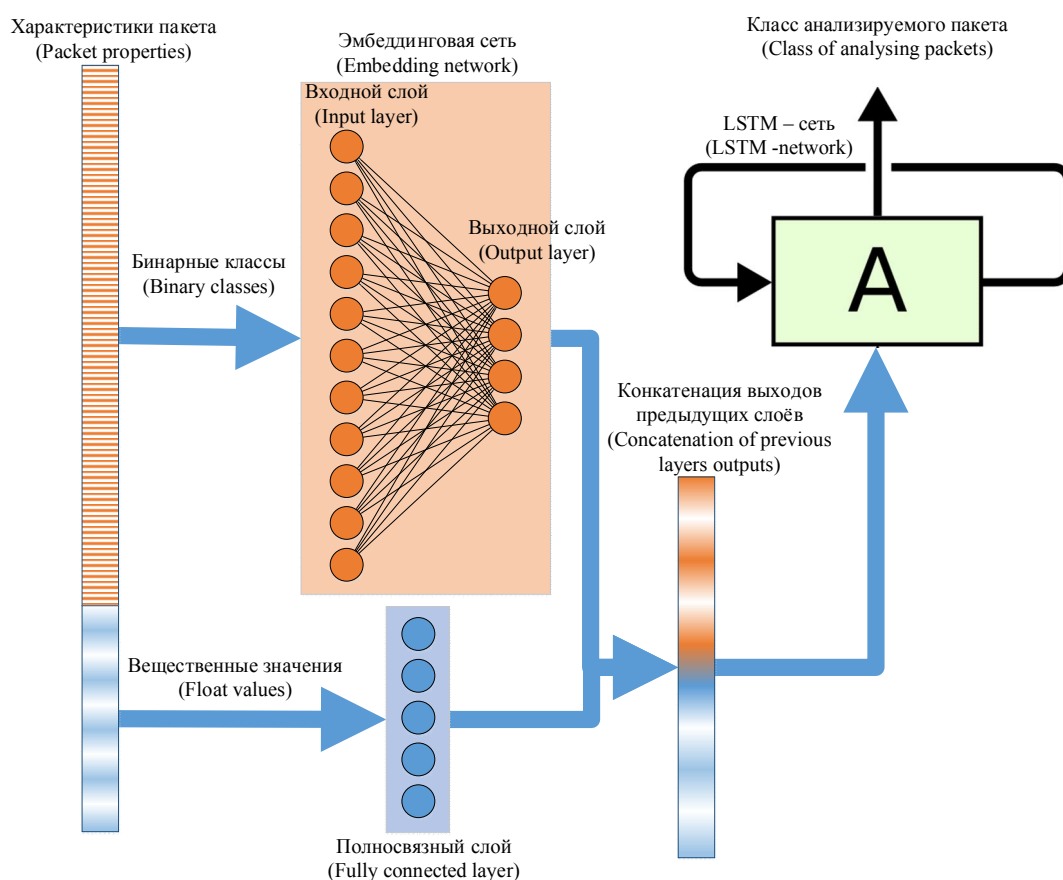


Рисунок 7 – Нейросетевая архитектура системы

Разработанная нейронная сеть содержит 14 слоев и 60154 параметров. Для программной реализации был использован язык программирования Python. В качестве основной библиотеки машинного обучения использовался Keras [14] в сочетании с бэкендом TensorFlow [17].

Практическое обучение модели. Для обеспечения эффективного мониторинга и исследования процесса обучения нейронной сети был разработан ряд функций по сбору и генерации логов, выполняющихся в фоновом режиме. Для визуализации статистических данных применялся модуль TensorBoard [18].

График уменьшения функции потерь изображен на рисунке 8.

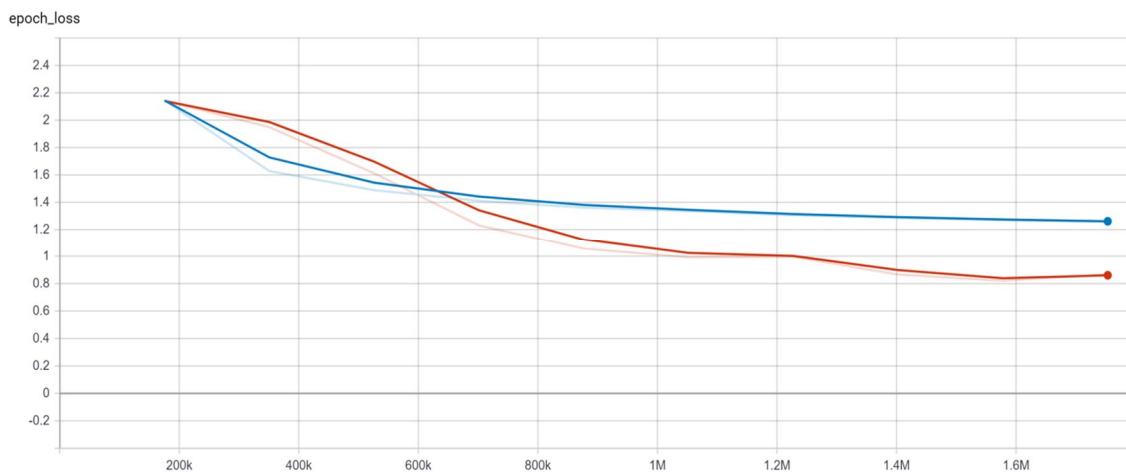


Рисунок 8 – Уменьшение функции потерь в процессе обучения нейронной сети

Полученные данные о функции потерь обладают одной нехарактерной для практики особенностью – начиная с 700000-й итерации значение функции потерь на валидационном датасете представляет меньшее значение, чем на обучающем датасете. Это обусловлено применением функции потерь FL.

Выводы. Разработана и исследована нейросетевая модель анализа трафика на основе перцептрона, алгоритмов роевого интеллекта и обучающего набора данных KDD Cup 1999 Data. Полученные результаты показали невозможность построения достаточно точной модели на основе названного датасета в силу его недостаточной репрезентативности.

Был проведен ряд тестов с различными гиперпараметрами нейронной сети глубокой архитектуры. Выявлены оптимальные значения параметров, при которых достигается максимальное значение точности определения аномального сетевого трафика и снижается количество ложных срабатываний. Проведено обучение и анализ точности работы описанной глубокой нейронной сети на новом датасете CSE-CIC-IDS2018. При этом были получены удовлетворяющие современным требованиям значения.

Библиографический список

1. Власенко А. В. Разработка алгоритмов и программ выбора оптимального набора компонент нейтрализации актуальных угроз на основе описания модели и интеграции их в WEB-приложение / А. В. Власенко, П. И. Дзьобан // Вестник Адыгейского государственного университета. Серия 4: Естественно-математические и технические науки. – 2014. – № 3 (142). – С. 189–193.
2. Частикова В. А. Анализ эффективности работы алгоритма светлячков в задачах глобальной оптимизации / В. А. Частикова, Я. И. Воля // Научные труды КубГТУ. – 2016. – № 15. – С. 105–111.
3. Частикова В. А. Исследование алгоритма серых волков / В. А. Частикова, С. А. Жерлицын // Научные труды КубГТУ. – 2016. – № 16. – С. 136–142.
4. Частикова В. А. Исследование эффективности алгоритма поиска косяком рыб в задаче глобальной оптимизации / В. А. Частикова, М. А. Дружинина, А. С. Кекало // Современные проблемы науки и образования. – 2014. – № 4.
5. Частикова В. А. Сравнительный анализ некоторых алгоритмов роевого интеллекта при обнаружении сетевых атак нейросетевыми методами / В. А. Частикова, М. П. Малыхина, С. А. Жерлицын, Я. И. Воля // Политехнический сетевой электронный научный журнал Кубанского государственного аграрного университета. – 2017. – № 129. – С. 106–115.
6. Частикова В. А. Нейросетевой подход к решению задачи построения фоторобота по словесному описанию / В. А. Частикова, С. А. Жерлицын, Я. И. Воля // Известия Волгоградского государственного технического университета. – 2018. – № 8 (218). – С. 63–67.
7. Chollet F. Deep Learning with Python / F. Chollet. – Manning Publications Co., 2017. – 384 с.
8. CSE-CIC-IDS2018 on AWS. – Режим доступа: <https://www.unb.ca/cic/datasets/ids-2018.html>, свободный. – Заглавие с экрана. – Яз. рус. (дата обращения: 10.09.2019).
9. Geron A. Hands-On Machine Learning with Scikit-Learn and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems / A. Geron. – O'Reilly Media, 2017. – 574 p.
10. Gers F. A. Learning to Forget: Continual Prediction with LSTM / F. A. Gers, J. Schmidhuber, F. Cummins // Neural Computation. – 2000. – № 12 (10). – P. 2451–2471.
11. Goodfellow I. Deep Learning / I. Goodfellow, Y. Bengio, A. Courville. – MIT Press, 2016. – 775 с.
12. Hochreiter S. Long short-term memory / S. Hochreiter, J. Schmidhuber // Neural computation. – 1997. – № 9 (8). – P. 1735–1780.
13. KDD Cup 1999 Data. – Режим доступа: <http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html>, свободный. – Заглавие с экрана. – Яз. рус. (дата обращения: 25.08.2019).
14. Keras: Deep Learning for humans. – Режим доступа: <https://github.com/keras-team/keras>, свободный. – Заглавие с экрана. – Яз. рус. (дата обращения: 21.06.2019).
15. Murphy Kevin P. Machine Learning: A Probabilistic Perspective / Kevin P. Murphy. – MIT Press, 2012. – 1096 с.
16. Ruder S. An overview of gradient descent optimization algorithms / S. Ruder // Cornell University Library. – 2016.
17. Tensorflow. An Open Source Machine Learning Framework for Everyone. – Режим доступа: <https://github.com/tensorflow/tensorflow>, свободный. – Заглавие с экрана. – Яз. рус. (дата обращения: 25.06.2019).
18. Tensorboard. TensorFlow's Visualization Toolkit. – Режим доступа: <https://github.com/tensorflow/tensorboard>, свободный. – Заглавие с экрана. – Яз. рус. (дата обращения: 01.07.2019).
19. Tomas Mikolov. Distributed Representations of Words and Phrases and their Compositionality / Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg Corrado, Jeffrey Dean // NIPS'13: Proceedings of the 26th International Conference on Neural Information Processing Systems. – 2013. – Vol. 2. – P. 3111–3119.
20. Tsung-Yi Lin. Focal Loss for Dense Object Detection / Tsung-Yi Lin, Priya Goyal, Ross Girshick, Kaiming He, Piotr Dollár // Cornell University Library. – 2017.

References

1. Vlasenko A. V., Dzoban P. I. Razrabotka algoritmov i programm vybora optimalnogo nabora komponent neytralizatsii aktualnykh ugroz na osnove opisaniya modeli i integratsii ikh v WEB-prilozhenie [Development of algorithms and programs to choose the optimal set of components of actual threat neutralization on the basis of model description and their integration in Web appendix]. *Vestnik Adygeyskogo gosudarstvennogo universiteta. Seriya 4: Estestvenno-matematicheskie i tekhnicheskie nauki* [Bulletin of Adygea State University. Series 4: Natural-Mathematical and Technical Sciences], 2014, no. 3 (142), pp. 189–193.
2. Chastikova V. A., Volya Ya. I. Analiz effektivnosti raboty algoritma svetlyachkov v zadachakh globalnoy optimizatsii [Firefly algorithm efficiency analysis in case of global optimization problem]. *Trudy Kubanskogo gosudarstvennogo tekhnicheskogo universiteta* [Proceedings of Kuban State Technical University], 2016, no. 15, pp. 105–111.
3. Chastikova V. A., Zherlitsyn S. A. Issledovanie algoritma serykh volkov [Research of the grey wolf algorithm]. *Nauchnye trudy Kubanskogo gosudarstvennogo tekhnicheskogo universiteta* [Proceedings of Kuban State Technical University], 2016, no. 16, pp. 136–142.
4. Chastikova V. A., Druzhinina M. A., Kekalo A. S. Issledovanie effektivnosti algoritma poiska kosyakov ryb v zadache globalnoy optimizatsii [Efficiency research of the fish school search algorithm in the global optimization problem]. *Sovremennyye problemy nauki i obrazovaniya* [Modern problems of Science and Education], 2014, no. 4.
5. Chastikova V. A., Malykhina M. P., Zherlitsyn S. A., Volya Yu. I. Sravnitelnyy analiz nekotorykh algoritmov roevogo intellekta pri obnaruzhenii setevykh atak neyrosetevymi metodami [Comparative analysis of some swarm intelligence algorithms with detection of network attacks using neural network methods]. *Politematicheskii setevoy elektronnyy nauchnyy zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta* [Polytematic Network Electronic Scientific Journal of Kuban State Agrarian University], 2017, no. 129, pp. 106–115.
6. Chastikova V. A., Zherlitsyn S. A., Volya Yu. I. Neyrosetevoy podkhod k resheniyu zadachi postroeniya foto-roboty po slovesnomu opisaniyu [Neural network method of identification by unformalized semantic characteristics]. *Izvestiya Volgogradskogo gosudarstvennogo tekhnicheskogo universiteta* [News of Volgograd State Technical University], 2018, no. 8 (218), pp. 63–67.
7. Chollet F. *Deep Learning with Python*. Manning Publications Co., 2017.
8. *CSE-CIC-IDS2018 on AWS*. Available at: <https://www.unb.ca/cic/datasets/ids-2018.html> (accessed: 10.09.2019).
9. Geron A. *Hands-On Machine Learning with Scikit-Learn and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*. O'Reilly Media, 2017.
10. Gers F. A., Schmidhuber J., Cummins F. Learning to Forget: Continual Prediction with LSTM. *Neural Computation*, 2000, no. 12 (10), pp. 2451–2471.
11. Goodfellow I., Bengio Y., Courville A. *Deep Learning*. MIT Press, 2016.
12. Hochreiter S., Schmidhuber J. Long short-term memory. *Neural computation*, 1997, no. 9 (8), pp. 1735–1780.
13. *KDD Cup 1999 Data*. Available at: <http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html> (accessed 25.08.2019).
14. *Keras: Deep Learning for humans*. Available at: <https://github.com/keras-team/keras> (accessed 21.06.2019).
15. Murphy K. P. *Machine Learning: A Probabilistic Perspective*. MIT Press, 2012.
16. Ruder S. *An overview of gradient descent optimization algorithms*. Cornell University Library, 2016.
17. *Tensorboard. TensorFlow's Visualization Toolkit*. Available at: <https://github.com/tensorflow/tensorboard> (accessed 01.07.2019).
18. *Tensorflow. An Open Source Machine Learning Framework for Everyone*. Available at: <https://github.com/tensorflow/tensorflow> (accessed 25.06.2019).
19. Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg Corrado, Jeffrey Dean. Distributed Representations of Words and Phrases and their Compositionality. *NIPS'13: Proceedings of the 26th International Conference on Neural Information Processing Systems*, 2013, vol. 2, pp. 3111–3119.
20. Tsung-Yi Lin, Priya Goyal, Ross Girshick, Kaiming He, Piotr Dollár. *Focal Loss for Dense Object Detection*. Cornell University Library, 2017.

DOI 10.21672/2074-1707.2020.49.4.033-048
УДК 004.89

ФОРМИРОВАНИЕ НЕЧЕТКОЙ СИСТЕМЫ ОЦЕНКИ ДЛЯ СИСТЕМЫ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ ПРИ УПРАВЛЕНИИ ПРОЦЕССОМ ПЕРЕРАБОТКИ В УСЛОВИЯХ НЕОПРЕДЕЛЕННОСТИ ИНФОРМАЦИИ О КАЧЕСТВЕ МАТЕРИАЛА

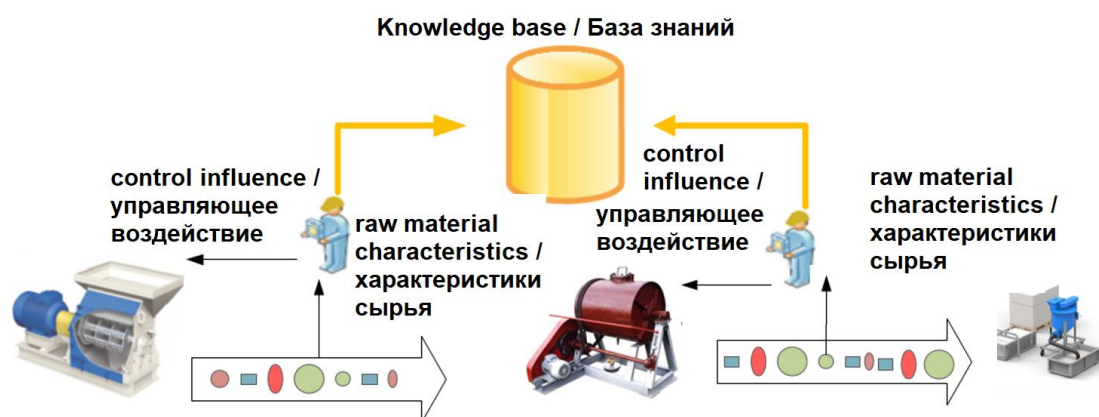
Статья поступила в редакцию 29.02.2020, в окончательном варианте – 08.03.2020.

Миловидова Анна Александровна, Государственный университет «Дубна», 141980, Российская Федерация, Московская область, г. Дубна, ул. Университетская, 19, старший преподаватель, e-mail: milanna@uni-dubna.ru, ORCID: 0000-0002-6646-4508

В работе показана необходимость учета нечеткой информации о качестве материала в процессе его переработки с целью повышения эффективности функционирования промышленных предприятий. Подробно проанализирована роль и функции эксперта-технолога при принятии решений, относящихся к управлению процессом переработки. Представлено описание концепции системы поддержки принятия решений при управлении процессом переработки в условиях неопределенности информации о свойствах материала. Именно свойства материала (на основе нечетких характеристик) используются в качестве основы для формирования эффективных управленческих решений. Детально описан процесс формирования системы нечетких оценок и предложена модель базы данных и базы знаний системы поддержки принятия решений при управлении процессом переработки в условиях неопределенности. Показано применение предложенной системы оценок и модели при формировании базы нечетких знаний для системы поддержки принятия решений при управлении режимами работы классификатора в процессе обогащения железорудного концентрата.

Ключевые слова: поддержка принятия решений, управление процессом переработки, неопределенность информации о качестве материала, нечеткая система оценки, база знаний

Графическая аннотация (Graphical abstract)



CREATION OF THE FUZZY MEASUREMENT SYSTEM FOR DECISION SUPPORT SYSTEM WHEN MANAGING THE PRODUCTION PROCESS IN AN UNCERTAIN INFORMATION ABOUT THE QUALITY OF THE MATERIAL BEING PROCESSED

The article was received by the editorial board on 29.02.2020, in the final version – 08.03.2020.

Milovidova Anna A., Dubna State University, 19 Universitetskaya St., Dubna, Moscow region, 141980, Russian Federation, Senior Teacher, e-mail: milanna@uni-dubna.ru, ORCID: 0000-0002-6646-4508

The necessity of accounting for fuzzy information about the quality of the material in the processing process in order to improve the functioning of industrial enterprises is shown. The role and functions of an expert technologist in making decisions in managing the production process is analyzed in detail. Description of the concept of the decision support system in the process of production control in conditions of uncertainty of information on material quality. Those material properties (based on fuzzy characteristics) are used as forming an effective management decision. The process of forming a fuzzy measurement system is described in detail and a model of the database and knowledge base of the decision support system for managing the production process in the face of uncertainty is proposed. The

application of the proposed measurement system and model is shown in the formation of a fuzzy knowledge base for a decision support system for managing the iron ore concentrate enrichment process.

Key words: decision support, managing the production process, uncertain information about the quality of the material, fuzzy measurement system, knowledge base

Введение. В перерабатывающей промышленности общим фактором неопределённости в производстве для процессов переработки (ПП) является неопределённость свойств материала (исходного сырья, промежуточного продукта и конечного продукта). Эта неопределённость сказывается на эффективности функционирования промышленного предприятия и свойств конечного продукта.

Качество сырья предопределяет режим работы и производительность оборудования, характер используемой технологии, влияет на себестоимость выпускаемой продукции [1].

ПП как объект управления можно отнести к трудно формализуемому классу объектов [8, 13], который характеризуется нестационарностью, инерционностью, запаздыванием, случайными возмущениями, нечёткой информацией [2, 3, 7, 11].

Трудности принятия решений для задач управления процессом переработки возникают в связи с нечёткой информацией о качестве материала; сложностью построения моделей и алгоритмов управления, базирующихся на точных математических методах; проблемами применения методологии и инструментария учёта человеческого фактора в непредвиденных нештатных ситуациях [12]. Применение методов теории нечетких множеств позволяет формализовать такие неопределённости [4, 5, 6, 9, 10, 14]. В связи с отсутствием прямых способов оценки качества материала огромную роль играют эксперты-технологи и базирующиеся на их опыте экспертные методы оценки качества.

Таким образом, целью исследования, приведенного в данной статье, является разработка подхода к формированию нечеткой системы оценки как инструмента поддержки принятия решений при управлении процессом переработки в условиях неопределённости информации о качестве материала.

Управление процессом переработки в условиях неопределённости. Как уже было сказано выше, в современной промышленности важную роль в управлении ПП (УПП) играет эксперт-технолог. Выполняемые им функции отражены на рисунке 1.

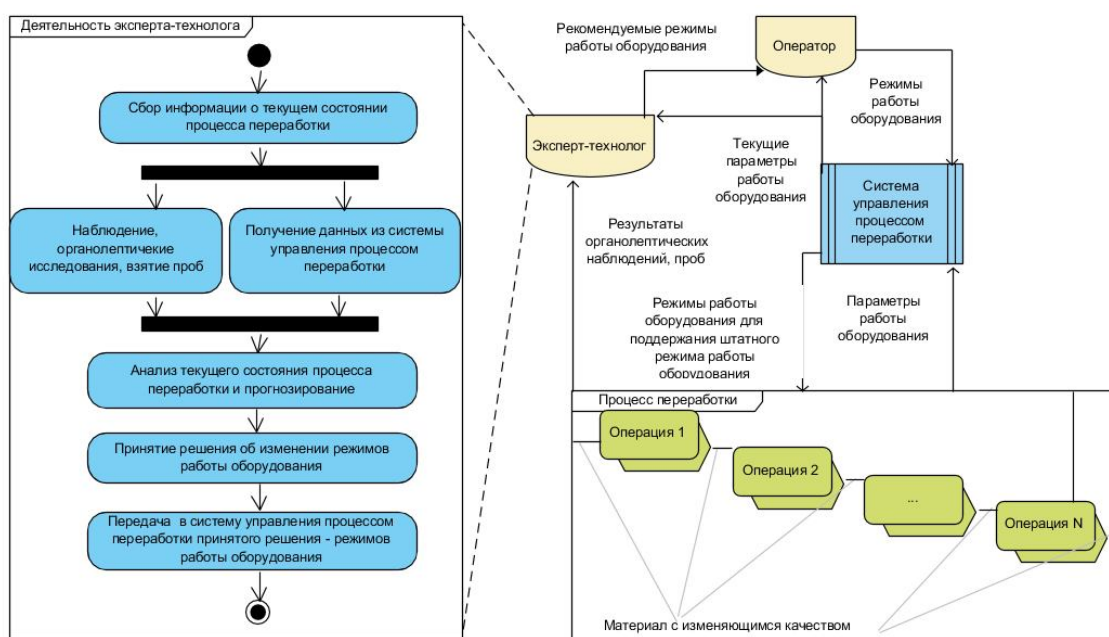


Рисунок 1 – Общая схема управления процессом переработки в условиях неопределённости информации о качестве материала

Используя свой опыт, он определяет возможное изменение качества поступающего материала или наличие форс-мажорной ситуации в производственной цепи в ходе переработки. По результатам такого определения принимается решение об изменении режимов работы оборудования. Знания, включающие основные эмпирические принципы принятия управленческих решений технологом, являются основой для системы управления ПП. Современные методы цифрового производства, обработки данных и моделирования позволяют использовать программный инструментарий извлечения, представления, обработки и формирования объективных баз знаний,

создавать имитационные модели. Исключение эксперта-технолога из непосредственного участия в УПП за счёт использования проблемно-ориентированных систем управления и предоставление технологу возможности развития и адаптации таких систем является одним из факторов перспективности данного исследования. Применение нечеткой логики в принятии решений при УПП позволит учитывать опыт и знания экспертов.

Для пространственно-распределённых процессов группы аппаратов ПП распределены на территории предприятия и связаны техническими средствами (конвейеры, спецтранспорт, трубопроводы и т.д.) переноса материала (материалов) от одной группы к другой, образуя тем самым транспортную сеть. Особенность переноса материала по сети состоит в неизменности свойств материала. В узлах ПП происходит изменение свойств материала аппаратами ПП. При этом в некоторых узлах аппараты ПП снабжены системами локального управления. Особенность принятия решений при УПП состоит в согласованном выборе локальных эффективных управляющих воздействий, обеспечивающих эффективность процесса в целом. При этом выбор управляющих переменных зависит от качества материала, движущегося от одних узлов к другим. Материал переносится указанными техническими средствами за различные времена, что приводит к необходимости прогнозировать эффективные согласованные управления для узлов (например, контактное расписание железнодорожного транспорта для металлургических предприятий). Поставим каждому узлу в соответствие агента, функции которого состоят в определении текущего состояния аппаратов ПП, свойств обрабатываемого материала, определении режимов работы аппаратов, оценке прибытия к некоторому моменту времени новой порции материала, формирования локального эффективного решения к моменту прибытия материала и согласования выработанного решения с другими агентами. Тогда создаётся возможность коллективного принятия комплексного упреждающего эффективного УПП.

Применение существующего опыта управления позволит повысить качество принятия управленческих решений за счет саморегулирования системы управления и выработки упреждающих воздействий. Таким образом, одной из важнейших компонент системы поддержки принятия решений (СППР) при УПП является система нечеткой оценки и база знаний.

Концепция системы поддержки принятия решений при управлении процессом переработки. Процесс переработки может быть представлен множеством единиц оборудования (M); множеством связей между единицами оборудования (S); множеством локальных систем управления (MU) оборудованием (M) (например, совокупность ПИД-регуляторов, исполнительных органов); системой оценок текущего состояния ПП; системой оценки свойств материала; множеством переменных состояния для оборудования ПП $Y = (Y_1, Y_2, \dots, Y_i, \dots, Y_n)$; множеством приборов, измеряющих текущее состояние оборудования $D = (D_1, D_2, \dots, D_i, \dots, D_n)$; множеством шкал оценки $H = (h_1, h_2, \dots, h_i, \dots, h_n)$; типом шкал $E = \{\text{порядковая, арифметическая, целая}\}$; матрицей «прибор – шкала» (DH)

$$dh(i, j) \in DH, \quad \text{где } dh(i, j) = \begin{cases} 1, & \text{если } i\text{-ому прибору соответствует } h_j \text{ шкала,} \\ 0, & \text{если } i\text{-ому прибору не соответствует } h_j \text{ шкала;} \end{cases}$$

матрицей «прибор – узел ответственности» (DR);

$$dr(i, j) \in DR, \quad \text{где } dr(i, j) = \begin{cases} 1, & \text{если } i\text{-ый прибор находится в } j\text{-ом узле,} \\ 0, & \text{если } i\text{-ый прибор не находится в } j\text{-ом узле;} \end{cases}$$

матрицей «система управления – оборудование» (WM)

$$wm(i, j) \in WM, \quad \text{где } wm(i, j) = \begin{cases} 1, & \text{если } i\text{-ое оборудование управляется} \\ & j\text{-ой локальной системой,} \\ 0, & \text{если } i\text{-ое оборудование не управляется} \\ & j\text{-ой локальной системой;} \end{cases}$$

множеством управляющих переменных $U = (u_1, u_2, \dots, u_i, \dots, u_m)$, $u_i \in V_i$, где V_i – допустимое значение u_i ; матрицей «управление – система управления» (UW)

$$uw(i, j) \in UW, \quad \text{где } uw(i, j) = \begin{cases} 1, & \text{если } u_i\text{-ое управление связано с} \\ & j\text{-ой локальной системой,} \\ 0, & \text{если } u_i\text{-ое управление не связано с} \\ & j\text{-ой локальной системой;} \end{cases}$$

характеристики исходного материала (X). $X = (x_1, x_2, \dots, x_i, \dots, x_l)$, $x_i \in G_i$, где G_i – допустимое множество x_i .

Характеристики входного продукта являются нечёткими параметрическими переменными. Процедура фазификации входных переменных (X) осуществляется экспертами на основе опыта и группы приборов измерения (D), характеризующих состояния оборудования (Y):

- системой показателей качества процесса переработки (Q);
- системой оценки показателей качества ПП;
- матрицей «показатель – оборудование» (QM)

$$qm(i, j) \in QM, \text{ где } qm(i, j) = \begin{cases} 1, & \text{если } i\text{-ый показатель характеризует состояние} \\ & j\text{-ого оборудования,} \\ 0, & \text{если } i\text{-ый показатель не характеризует состояние} \\ & j\text{-ой оборудования.} \end{cases}$$

Основные положения задачи принятия решений при управлении ПП:

1) Показатели качества (Q) являются функционалами от текущего состояния всех единиц оборудования (M), текущих характеристик материала ПП (X).

2) Если известны характеристики входного материала (X) на заданном интервале времени и при этом известно начальное состояние ПП, то тогда известны эффективные режимы работы всех единиц оборудования ПП, т.е. известны эффективные значения показателей (Q^*).

3) Если входное сырьё не поступает в ПП и известно распределение материала по всей цепи ПП, тогда известны эффективные режимы работы всех единиц оборудования.

4) Если входное сырьё поступает в ПП с известными нечёткими значениями и известно распределение для качества материала, то тогда агентная система способна формировать эффективное управление режимами работы оборудования.

Для принятия решений при УПП определены: вектор текущего состояния ПП $Y(t)$; система оценки текущего состояния ПП $S_1(Y(t))$; вектор текущего состояния входного материала $X(t)$; определён класс функций $X(t)$, т.е. $X(t) \in \Omega(X(t))$; система оценки текущего качества материала в ПП $S_2(X(t))$; шкала системы оценки определена как арифметическая; вектор управления аппаратами ПП $U(t)$; переменные управления определены как нечёткие переменные; система оценки управляющих переменных вектора $U(t)$ $S_3(U(t))$; определён класс функций $U(t)$, т.е. $U(t) \in \Psi(U(t))$; вектор показателя качества ПП P_1 ; показатель качества ПП определён как нечёткая переменная; система оценки качества ПП $S_4(P_1)$; качество ПП определено как нечёткая переменная; вектор показателя качества конечного продукта ПП P_2 ; качество конечного продукта определено как нечёткая переменная; система оценки качества конечного продукта $S_5(P_2)$; достигнутый показатель качества ПП (P_1^*); достигнутый показатель качества определён как нечёткая переменная; достигнутый показатель качества конечного продукта (P_2^*); множество аппаратов ПП (R); состояние аппаратов ПП ($r_i(t), i=1,2,\dots,k,\dots,N$); связи между аппаратами ПП; узлы в технологической сети, в которых изменяются переменные состояния материала; управляемые аппараты по всей схеме ПП; начальные условия ПП.

Общий критерий эффективности качества ПП (Q_1) определён как

$$Q_1 = (P_1^* - P_1)^2. \quad (1)$$

Ограничение на вектора $X(t), Y(t)$:

$$W_1(X(t), Y(t)) = 0. \quad (2)$$

Общий критерий эффективности материала

$$Q_2 = (P_2^* - P_2)^2. \quad (3)$$

Общий критерий эффективности качества конечного продукта (Q_2)

$$Q_3 = (P_3^* - P_3)^2. \quad (4)$$

Принята следующая зависимость

$$\begin{aligned} P_1 &= P_1(X(t), U(t)) \\ P_2 &= P_2(X(t), U(t)), \\ P_3 &= P_3(X(t), U(t)) \end{aligned} \quad (5)$$

где P_1, P_2 и P_3 являются функционалами.

Задача принятия решений при УПП формулируется следующим образом. Для заданного начального состояния аппаратов ПП и начального состояния материала в узлах ответственности агентов и произвольного состояния материала $X(t) \in \Omega(X(t))$ найти

$$U^*(t), U^*(t) \in \Psi(U(t)), \quad (6)$$

которое обеспечивает минимум Q_1 , Q_2 и Q_3 , т.е. требуется

$$\min_{U(t) \in \Psi(U(t))} Q_1 = (P_1^* - P_1)^2$$

$$\min_{U(t) \in \Psi(U(t))} Q_2 = (P_2^* - P_2)^2$$

$$\min_{U(t) \in \Psi(U(t))} Q_3 = (P_3^* - P_3)^2.$$

Причём оптимальное $U^*(t), U^*(t) \in \Psi(U(t))$ является упреждающим вектором управления.

Условия функционирования ПП:

1. На вход ПП поступает исходное сырьё, характеризующееся определёнными свойствами и определёнными объёмами (порциями). Свойства и порции исходного материала задаются нечёткими переменными.

2. Интервалы между поступлениями порций материала задаются нечёткими переменными.

3. Определена экспертная таблица, устанавливающая зависимость между свойствами материала и эффективными показателями качества аппаратов ПП (показатели качества заданы нечёткими переменными).

4. Агенты-соседи формируют согласованные решения о качестве функционирования аппаратов ПП и текущих характеристиках материала и передают принятые решения глобальному агенту. В свою очередь «агенты-не соседи» передают сведения о текущем состоянии аппаратов ПП и качестве материала.

5. Функционирование агентов состоит в принятии согласованных решений о свойствах материала.

Согласно вышесказанному, особенность поставленной задачи поддержки принятия решения состоит в следующем:

1) Для известных свойств материала экспертными методами определены эффективные значения управляемых переменных аппаратов ПП. Свойства материала являются нечёткими переменными и определяются системой нечёткой оценки. Экспертные эффективные управляемые переменные аппаратов ПП являются нечёткими переменными.

2) Коллективная задача агентов состоит в выработке согласованного общего управляющего упреждающего решения для всех аппаратов ПП с учётом движения материала и возможного изменения его свойств в активных узлах ТС.

3) Агентная система принятия упреждающего управленческого решения является надстройкой над системой управления производством. Она позволит повысить надёжность, ритмичность, стабильность, качество ПП по сравнению с ручным и локальным управлением. Агенты размещаются в узлах ответственности, в которых по косвенным признакам в реальном времени эксперты-технологи определяют качество материала и на основе этого принимают согласованные упреждающие решения.

В силу распределённости локальных систем управления и изменения во времени свойств порций материала, а также необходимости формирования комплексного управления (т.е. управляющих воздействий в момент t для всех локальных систем управления), целесообразна разработка процедуры упреждающего управления. Эта процедура в общем виде может быть описана следующим образом.

Если для i -й управляемой единицы оборудования выполняется условие

$$\bar{X}_{t_0-\tau_i} = \bar{X}_{t_0+\tau_i}, \quad (7)$$

тогда $\bar{U}_{t_0+\tau_i} = \bar{U}_{t_0-\tau_i}$.

Если условие (7) не выполняется, тогда, согласно экспертной таблице эффективности, определяется эффективное управление как $\bar{U}_{t_0+\tau_i} = \Lambda(\bar{X}_{t_0+\tau_i})$, где Λ – система правил, устанавливающая зависимость $\bar{U}_{t_0+\tau_i}$ от $\bar{X}_{t_0+\tau_i}$.

Определяем $\bar{Y}_{t_0+\tau_i} = (\bar{U}_{t_0+\tau_i}, \bar{X}_{t_0+\tau_i})$, затем $K_\Delta = (\bar{Y}_{t_0+\tau_i})$, если $\mu(K_\Delta, \bar{K}^*) \leq \delta$, где μ – процедура сравнения показателей и δ – близость показателей, тогда $\bar{U}_{t_0+\tau_i} = \bar{U}_{t_0-\tau_i}$.

Для ситуации, когда время переходного процесса исполнительного органа управления равно нулю в моменты $t_0 + \tau_i$ необходимо применить $\bar{U}_{t_0+\tau_i}$. В противном случае управление $\bar{U}_{t_0+\tau_i}$

необходимо применить в момент $t_0 + \tau_i + \omega_i - \lambda_i$, $\lambda_i \leq \tau_i + \omega_i$, где λ_i – время переходного процесса, ω_i – время переработки оборудованием материала с характеристиками $\bar{X}_{t_0-\tau_i}$.

В случае, когда $\lambda_i > \tau_i + \omega_i$, то $\bar{U}_{t_0+\tau_i} = \bar{U}_{t_0-\tau_i}$, т.е. происходит потеря эффективности и возможны форс-мажорные ситуации.

Таким образом, в системе управления может быть БД, в которой накапливаются форс-мажорные ситуации. На основе анализа форс-мажорных ситуаций возможна коррекция нечёткой системы оценки свойств материала. Тем самым в процессе функционирования возможна её адаптация. Проявлением адаптивности является снижение числа форс-мажорных ситуаций и улучшение фактических показателей эффективности ПП.

На рисунке 3 представлена модель системы принятия решений при УПП как надстройка над автоматизированной системой управления технологическим процессом (далее – АСУ ТП). Модель агентного принятия решений при УПП отражает динамику функционирования ПП по принципу «вход – процесс – выход» при получении информации об изменении качества перерабатываемого материала, где:

- 1) задан интервал времени $[t_0, T]$;
- 2) определен вектор нечётких входных переменных $\bar{X}(t)$, $t \in [t_0, T]$, нечёткий вектор текущего состояния оборудования $\bar{Y}(t)$, $t \in [t_0, T]$ и нечёткий вектор управления $\bar{U}(t)$, $t \in [t_0, T]$;
- 3) определён показатель качества ПП \bar{K}_Δ , $t \in [t_0, T]$;
- 4) определена параметрическая база знаний (далее – ПБЗ-1), устанавливающая связь $\bar{Y}(t) = \Omega_1(\bar{X}(t))$;
- 5) определена параметрическая база знаний (далее – ПБЗ-2), устанавливающая связь $\bar{U}(t) = \Omega_2(\bar{X}(t), \bar{Y}(t))$, где $\bar{U}(t)$ – вектор нечетких переменных, ПБЗ-1, ПБЗ-2 представлены системами продукций;
- 6) задана процедура определения фактического вектора показателя качества ПП $K_\Delta = K_\Delta(\bar{U}(t)) \Delta \in [0, T]$, где K_Δ – нечёткий вектор. Нечёткие векторы $\bar{K}_\Delta(t)$, $\bar{Y}(t)$, $\bar{X}(t)$, $\bar{U}(t)$ процедурой дефазификации могут быть переведены в четкие переменные.
- 7) определена процедура смешивания материала как нечёткая функция, зависящая от объема материала и его свойств $\Psi_{t_0+\tau_i}(\bar{X}_0) = \bar{X}_{t_0+\tau_i}$.

На рисунке 2 представлена модель системы принятия решений при управлении ПП как надстройки над АСУ ТП.



Рисунок 2 – Модель агентной системы принятия решений

Система нечеткой оценки при поддержке принятия решений для управления процессом переработки. Формирование системы нечёткой оценки включает в себя следующие стадии:

1. Формулировка проблемы (задачи), в рамках которой необходима система нечёткой оценки состояния процесса и свойств перерабатываемого материала.
2. Определение узлов, в которых происходят изменения состояния аппаратов ПП (в зависимости от свойств материала).
3. Для каждого узла определяется подсистема нечётких измерителей.
4. Для каждого измерителя определяется нечёткая переменная (на основе результатов интервью и фокус-группы с экспертами-технологами).
5. Для каждой нечёткой переменной формируется экспертная таблица нечёткой переменной (на основе результатов интервью с экспертами-технологами, являющимися фокус-группой специалистов).
6. Поиск вида и параметров функции принадлежности на основе экспертной таблицы.
7. Оценка полученной функции принадлежности.

База нечётких знаний включает следующее:

- 1) нечёткую систему оценки свойств материала, конечного продукта;
- 2) нечёткую систему оценки состояния состояний аппаратов ПП;
- 3) нечёткую систему оценки управляющих переменных;
- 4) нечёткую систему оценки показателей качества ПП;
- 5) функции, описывающие логику принятия решений.

1. Нечёткая система оценки свойств материала. Нечёткая система оценки представлена библиотекой модулей, предполагаемых для включения в СППР. Все модули унифицированы по входу – выходу. Входом (Y) является совокупность физических переменных для датчиков, измеряющих текущее состояние материала (массу, объём, цвет, температуру и т.д.). Выходом (X) является нечёткая переменная, значением которой является одно из свойств материала и т.д. Нечёткая переменная свойств является нечёткой функцией от датчиков, измеряющих состояния перерабатываемого сырья. На рисунке 3 представлен модуль нечёткой переменной свойств материала.

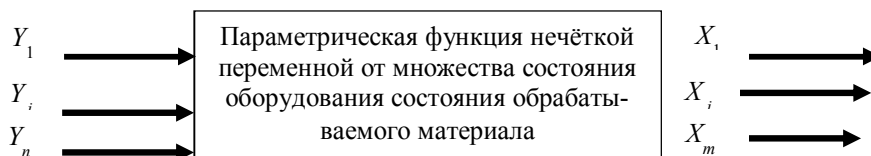


Рисунок 3 – Модуль нечёткой переменной свойств материала

2. Нечёткая система оценки состояния ПП. Текущее состояние ПП определяется на основе сведений, поступающих из системы датчиков. Переменные, характеризующие состояние ПП, являются нечёткими функциями от переменных датчиков.

3. Нечёткая система оценки показателей качества ПП. Фактические показатели качества процесса переработки определяются нечёткими переменными на момент t_i для всех узлов технологической сети. На основе этих показателей качества определяется значение критерия эффективности качества ПП.

4. Нечёткая система оценки управляющих переменных. Управляющие переменные оборудования характеризуются областями допустимых значений и являются нечёткими переменными. Экспертами для управляемого оборудования определена таблица зависимости нечётких значений управлений от нечётких свойств материала.

Эксперты-технологи определяют узлы ответственности, покрывающие всё пространство ПП. При этом узлы ответственности не пересекаются. В каждом узле ответственности определяется группа первичных датчиков, измеряющих текущее состояние аппаратов ПП и свойства материала

$$D = \{d_{ij}\}, d_{ij} - j - \text{й прибор } i - \text{й узла ответственности.}$$

Для каждого прибора определяется шкала измерений (h_{ij}), диапазон измерений (a_{ij}, b_{ij}), погрешность измерений Δ_{ij} , частота (ω_{ij}) измерения.

В каждом узле ответственности определяется группа аппаратов ПП

$$M = \{m_{ik}\}, m_{ik} - k - \text{й аппарат } i - \text{й зоны ответственности.}$$

Для i -го узла ответственности определяется матрица связи «аппарат – прибор» (R)

$$R = (r_{jk}), r_{jk} - \text{связь } j - \text{го прибора с } k - \text{м аппаратом.}$$

$$r_{jk} = \begin{cases} 0, & \text{если связь отсутствует,} \\ 1, & \text{если есть связь между } j - \text{м прибором и } k - \text{м аппаратом.} \end{cases}$$

Для каждого узла ответственности определяется группа нечётких переменных (G), характеризующих качество ПП (G_1), качество материала (G_2), конечного продукта (G_3).

Для каждого узла ответственности определяется группа регуляторов управления аппаратами ПП. Для каждого регулятора определяется управляющая переменная. Каждая управляющая переменная характеризуется диапазоном допустимых значений, шкалой измерения, погрешностью. Каждый регулятор характеризуется длительностью переходного процесса. Для каждого регулятора определена нечёткая управляющая переменная. Для каждой нечёткой переменной (x) определяется группа первичных датчиков (d), задающих универсальное множество (диапазоны значений датчиков) для значений нечёткой переменной.

Таким образом, в БЗ содержатся описания нечётких переменных, характеризующие состояние оборудования, качества материала, управляющих переменных. С БЗ связана библиотека модулей, каждый из которых реализует дефазификацию нечёткой переменной в значение переменной согласно заданному диапазону и шкале.

На рисунке 4 представлена логическая модель БД и БЗ для агентной СППР принятия решений при управлении ПП.

Данная модель включает следующие сущности, которые предназначены для хранения:

- 1) значений параметров функции принадлежности нечетких переменных;
- 2) взаимосвязей правил и нечетких переменных;
- 3) методов дефазификации нечетких переменных;
- 4) нечетких переменных;
- 5) продукционных правил преобразования входных нечетких переменных в выходные;
- 6) параметров функций принадлежностей;
- 7) хранения функций принадлежности;
- 8) величин, измеряемых датчиками;
- 9) значений управляемых параметров;
- 10) связей между агентами и оборудованием в узле их ответственности;
- 11) сведений об оборудовании ПП;
- 12) связей между датчиками и измеряемыми величинами;
- 13) информации об агентах;
- 14) сообщений, которыми обмениваются агенты;
- 15) связей между агентами-соседями;
- 16) сведений о датчиках;
- 17) истории измерений датчиками;
- 18) истории изменения управляющих параметров работы оборудования.

Нечеткая модель управления режимом работы классификатора процесса обогащения.

В качестве примера применения такой базы нечетких оценок рассмотрим нечеткую модель определения режима работы классификатора (далее – КСН) на участке обогащения железорудного концентрата, реализованную в среде *Matlab, fuzzy logic toolbox*.

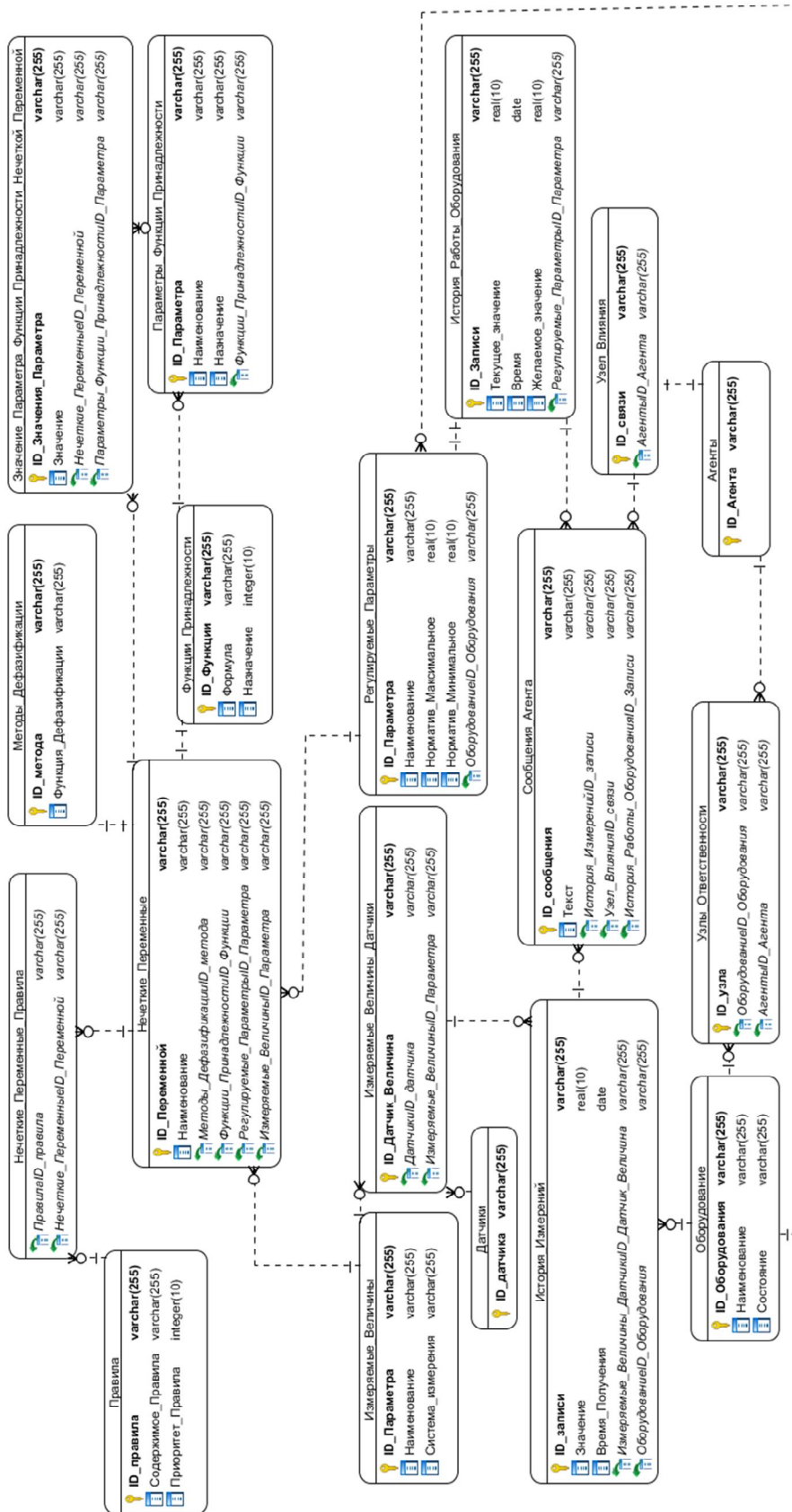


Рисунок 4 – Модель БД и БЗ агентной системы принятия решений

Классификация представляет собой процесс разделения материала на классы крупности в жидкой фазе. В таблице 1 приведены основные параметры классификатора. На рисунке 5 представлена технологическая схема классификации материала в процессе обогащения железорудного концентрата.

Таблица 1 – Основные параметры КСН

Основные параметры управления	Характеристики входного продукта	Характеристики выходного продукта
Скорость вращения винта	Плотность и гранулометрический состав слива	Производительность по пескам (плотность слива материала на следующий этап переработки – сепарацию)
Угол наклона	Интенсивность расхода – слив мельницы за единицу времени	Количество сырья, поступающего на рецикл – повторное измельчение в шаровую мельницу

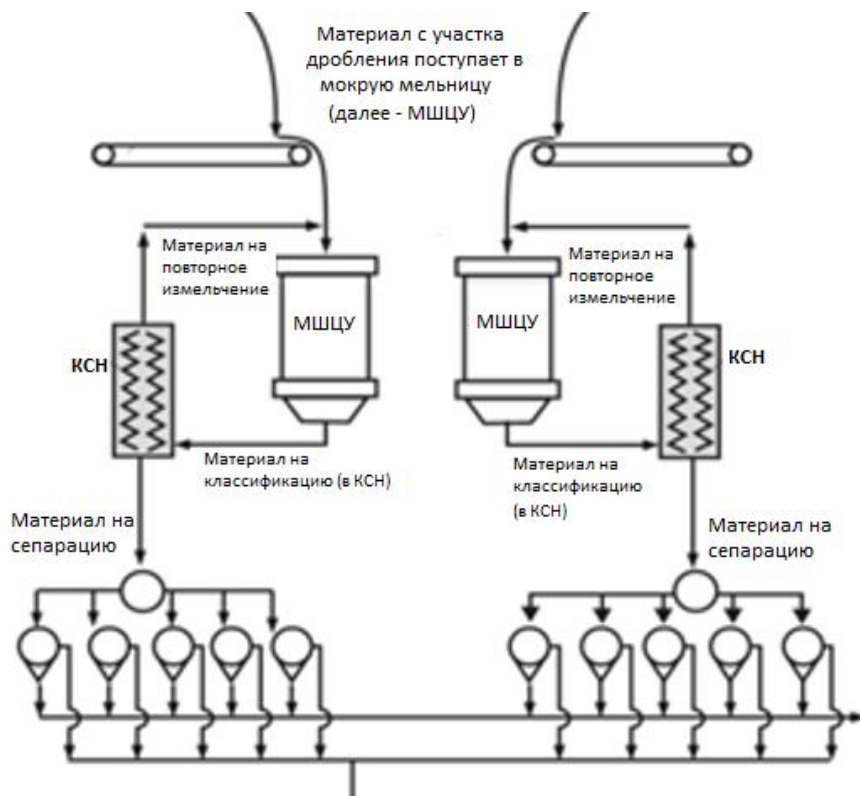


Рисунок 5 – Технологическая схема классификации материала в процессе обогащения железорудного концентрата

Нечеткая модель системы управления режимами работы КСН (рис. 6) включает в себя описание нечетких, входных и выходных параметров и систему продукций.

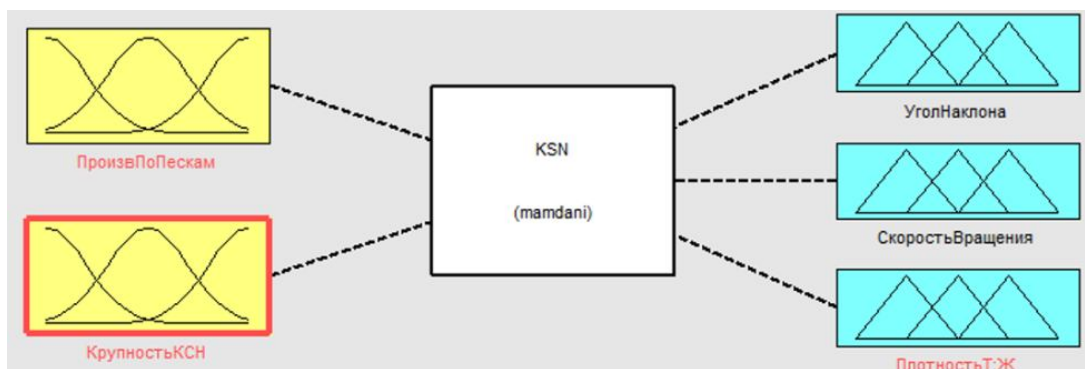


Рисунок 6 – Мнемоническое представление системы управления режимами работы КСН

Нечёткие входные переменные:

- 1) крупность материала в КСН (x_1), значения x_1 – мелкая, средняя, крупная (рис. 7а);
- 2) производительность КСН по пескам (x_2), значения x_2 – низкая, средняя, высокая (рис. 7б).

Итак, имеем две переменных, каждая из которых принимает три возможных значения, например, мелкая, средняя, крупная/высокая.

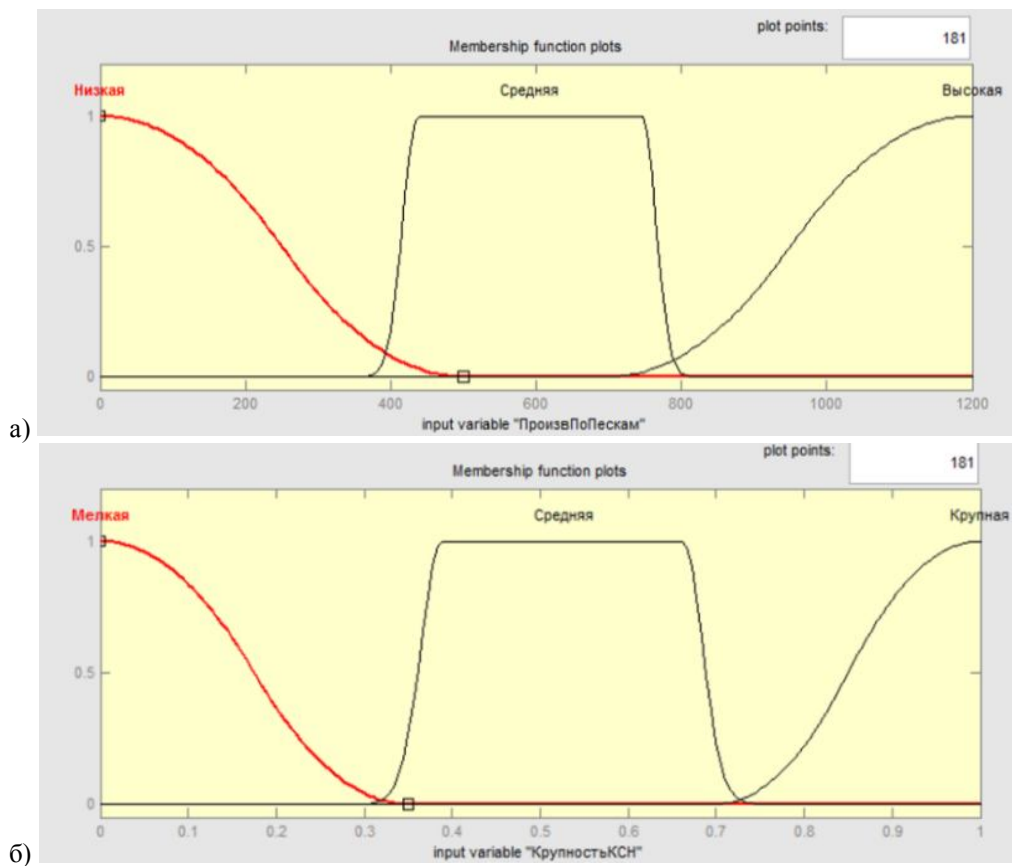


Рисунок 7 – Функции принадлежности, описывающие входные параметры: а) производительность классификатора по пескам; б) крупность материала, поступающего в классификатор

Нечеткие выходные переменные:

- 1) угол наклона классификатора – маленький, средний, большой (рис. 8а);
- 2) скорость вращения винта классификатора – низкая, средняя, высокая (рис. 8б);
- 3) плотность Т:Ж (соотношение твердого к жидкому) – низкая, средняя, высокая (рис. 8в).

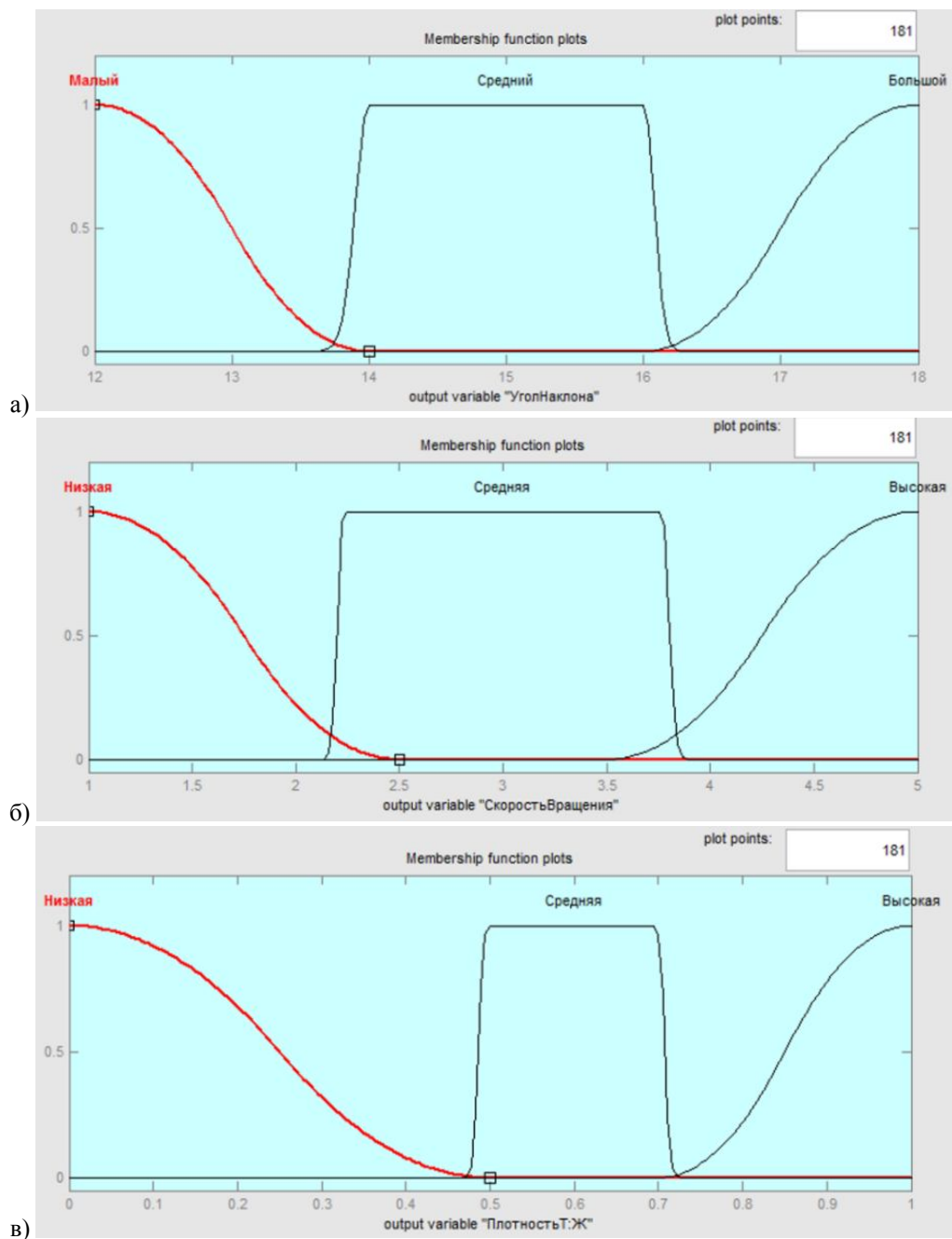


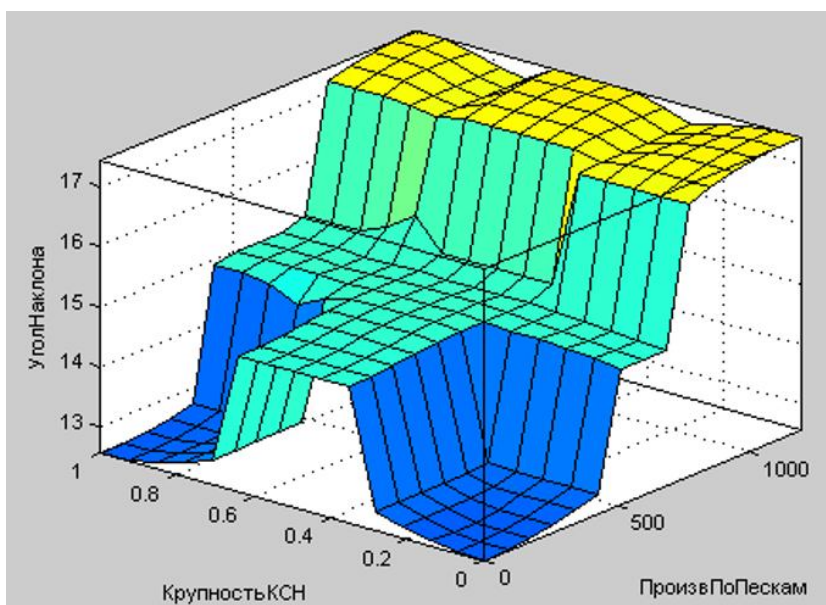
Рисунок 8 – Функции принадлежности, описывающие выходные параметры: а) угол наклона классификатора; б) скорость вращения винта классификатора; в) плотность Т:Ж

Сформирована база правил (табл. 2), которые отражают зависимости между входными и выходными нечеткими переменными.

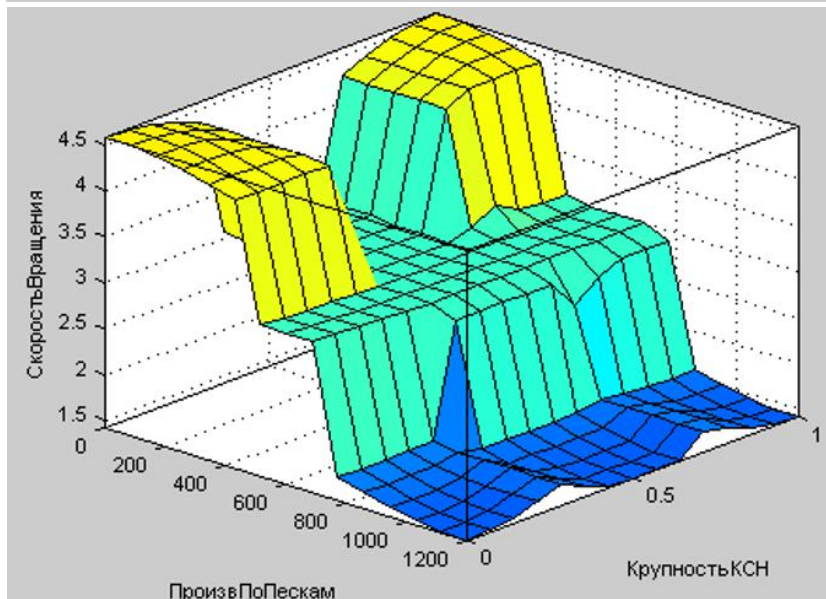
Таблица 2 – Правила дефазификации нечеткого угла наклона, скорости вращения и подачи воды в КСН (фрагмент)

1. If (ПроизвПоПескам is Низкая) and (КрупностьКСН is Мелкая) then (УголНаклона is Малый)(СкоростьВращения is Высокая)(ПлотностьТ:Ж is Низкая) (1)
2. If (ПроизвПоПескам is Высокая) and (КрупностьКСН is Крупная) then (УголНаклона is Большой)(СкоростьВращения is Низкая)(ПлотностьТ:Ж is Высокая) (1)
3. If (ПроизвПоПескам is Средняя) and (КрупностьКСН is Средняя) then (УголНаклона is Средний)(СкоростьВращения is Средняя)(ПлотностьТ:Ж is Средняя) (1)
4. If (ПроизвПоПескам is Высокая) and (КрупностьКСН is Средняя) then (УголНаклона is Большой)(СкоростьВращения is Низкая)(ПлотностьТ:Ж is Средняя) (1)
5. If (ПроизвПоПескам is Высокая) and (КрупностьКСН is Мелкая) then (УголНаклона is Большой)(СкоростьВращения is Низкая)(ПлотностьТ:Ж is Низкая) (1)
6. If (ПроизвПоПескам is Низкая) and (КрупностьКСН is Крупная) then (УголНаклона is Малый)(СкоростьВращения is Высокая)(ПлотностьТ:Ж is Высокая) (1)
7. If (ПроизвПоПескам is Средняя) and (КрупностьКСН is Крупная) then (УголНаклона is Средний)(СкоростьВращения is Средняя)(ПлотностьТ:Ж is Высокая) (1)

В результате формируются поверхности решений (рис. 9). Результаты дефазификации нечеткого угла наклона, скорости вращения и подачи воды в КСН представлены на рисунке 10.



а)



б)

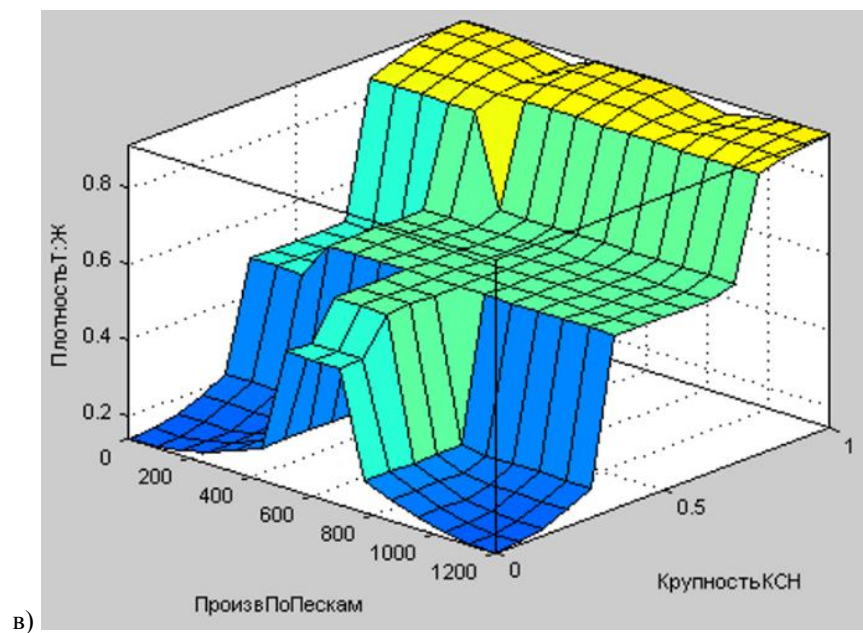


Рисунок 9 – Поверхности решений: а) по оси x – производительность КСН по пескам, по оси y – крупность материала в КСН, по оси z – угол наклона; б) по оси x – производительность КСН по пескам, по оси y – крупность материала в КСН, по оси z – скорость вращения; в) по оси x – производительность КСН по пескам, по оси y – крупность материала в КСН, по оси z – плотность Т:Ж

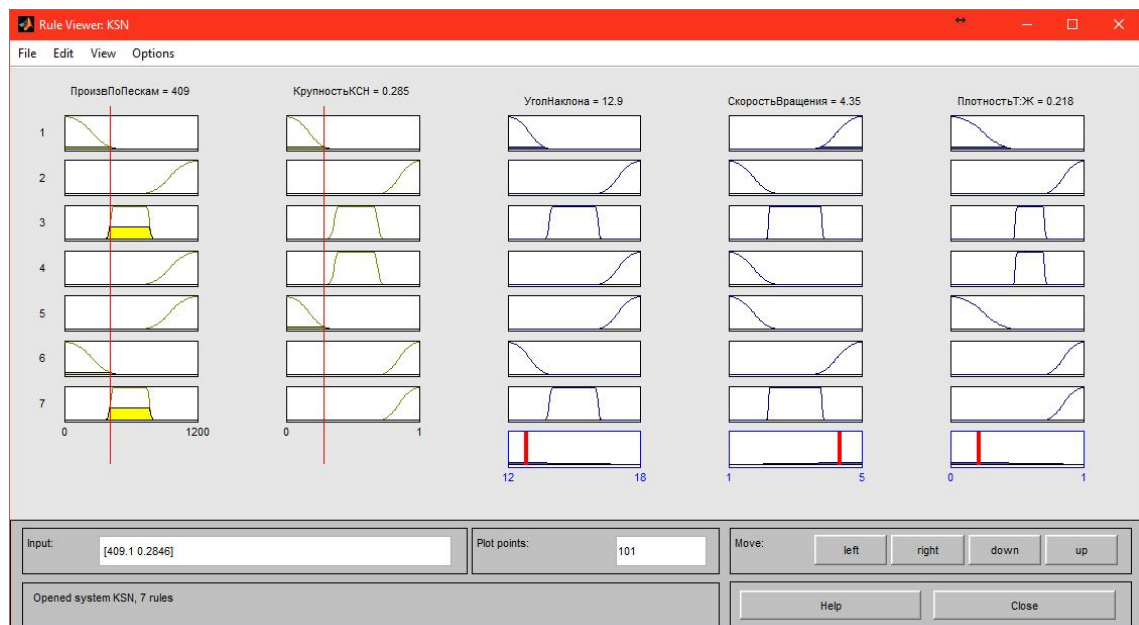


Рисунок 10 – Визуализация дефазификации нечеткого угла наклона, скорости вращения винта и подачи воды в КСН

На примере модели управления режимом работы классификатора процесса обогащения рассмотрен процесс формирования системы нечетких оценок как базовой компоненты СППР при УПП в условиях неопределенности.

Здесь необходимо отметить, что важнейшую роль в развитии СППР играет технолог. Опираясь на его знания и опыт, осуществляется адаптация и развитие системы в процессе ее функционирования (рис. 11).

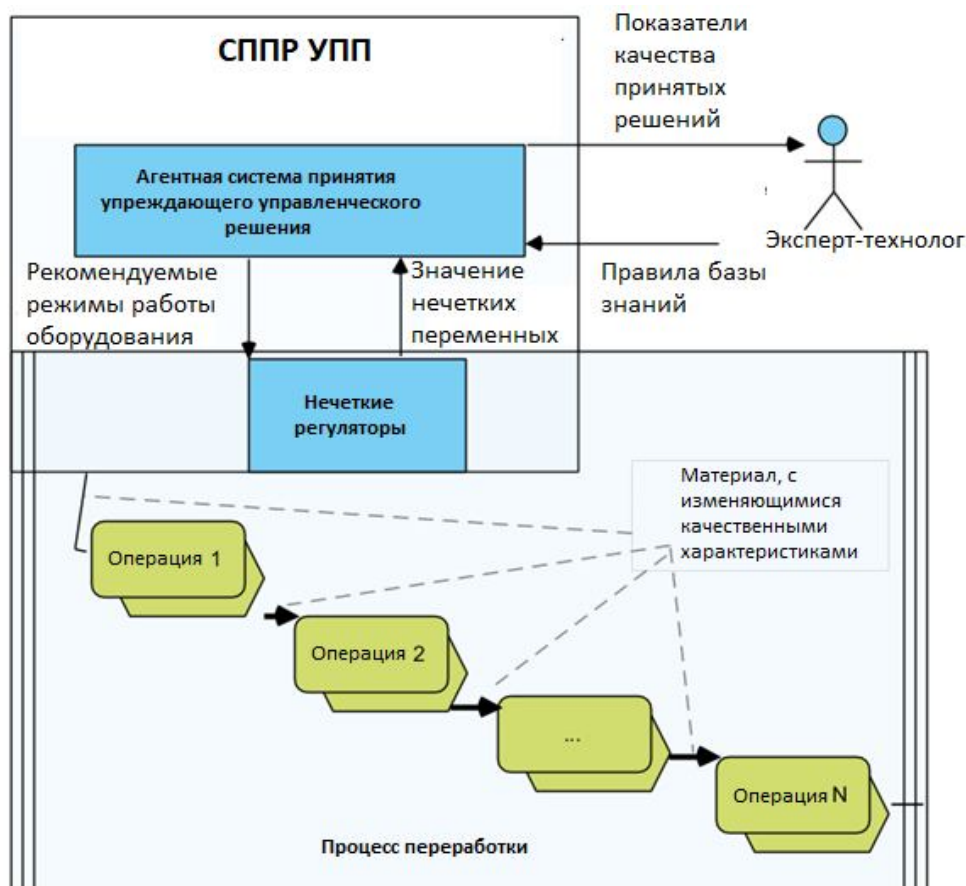


Рисунок 11 – СППР при УПП в условиях неопределенности информации о качестве материала

В рамках нечеткой системы оценки СППР при управлении процессом классификации железорудного концентрата было разработано 17 нечетких моделей (НМ) управления оборудованием:

- 1) НМ управления разгрузкой материала в бункер крупного дробления;
- 2) НМ управления производительностью крупной дробилки;
- 3) НМ управления скоростью питателя участка крупного дробления;
- 4) НМ определения результатов перемешивания руды в бункере среднего дробления;
- 5) НМ управления производительностью средней дробилки;
- 6) НМ управления скоростью питателя участка среднего дробления;
- 7) НМ определения результатов перемешивания в бункере мелкого дробления;
- 8) НМ управления производительностью мелкой дробилки;
- 9) НМ определения результатов перемешивания на конвейере передачи материала на повторное мелкое дробление;
- 10) НМ определения результатов перемешивания руды в бункере обогатительной фабрики;
- 11) НМ управления скоростью питателей на участке измельчения;
- 12) НМ управления производительностью шаровой мельницы;
- 13) НМ определения режима работы классификатора;
- 14) НМ определения режима работы магнитного сепаратора;
- 15) НМ определения результатов перемешивания материала в зумпфе;
- 16) НМ определения режима работы гидроциклона;
- 17) НМ определения режима работы дешламатора.

Сформирована первичная база знаний, включающая более 2000 правил.

Заключение. В результате проделанной работы представлена концепция СППР при управлении процессом переработки. В качестве одной из базовых компонент предложена система нечеткой оценки для поддержки принятия решения в управлении ПП в условиях неопределенности информации о качестве материала. Разработана модель базы данных и базы знаний СППР при управлении процессом переработки в условиях неопределенности. В качестве примера применения разработанного инструментария разработана модель управления режимами работы классификатора в процессе обогащения железорудного концентрата.

Библиографический список

1. Новиков С. И. Методы нечеткой логики в задачах автоматизации тепловых процессов электростанций / С. И. Новиков, В. П. Шахнович, А. В. Сафронов // Вестник ИГЭУ. – 2010. – Вып. 4. – С. 1–4.
2. Aydogan Savran. A fuzzy model based adaptive PID controller design for nonlinear and uncertain processes / Aydogan Savran, Gokalp Kahraman // ISA Transactions. – 2014. – № 53. – P. 280–288.
3. Bartolomeo Cosenza. Nonlinear fuzzy control of a fed-batch reactor for penicillin production / Bartolomeo Cosenza, Mosè Galluzzo // Computers and Chemical Engineering. – 2012. – № 36. – P. 273–281.
4. Carlo Bagnoli. The Theory of Fuzzy Logic and its Application to Real Estate Valuation / Carlo Bagnoli & Halbert C. Smith // Journal of Real Estate Research, American Real Estate Society. – 1998. – Vol. 16 (2). – P. 169–200.
5. Del Giudice V. Valuation of Real Estate Investments through Fuzzy Logic / V. Del Giudice, P. De Paola, G. B. Cantisani // Buildings. – 2017. – № 7. – P. 26.
6. Didier D. Resolution principles in possibilistic logic / D. Didier, H. Prade // International Journal of Approximate Reasoning. – 1990. – Vol. 4, issue 1. – P. 1–21.
7. Escaco J. M. Neurofuzzy model based predictive control for thermal batch processes / J. M. Escaco, C. Bordons, C. Vilas, M. R. Garcia, A. A. Alonso // Journal of Process Control. – 2009. – № 19. – P. 1566–1575.
8. Kang Jun, Wenjun Meng, Ajith Abraham, Hongbo Liu. An adaptive PID neural network for complex nonlinear system control // Neurocomputing. – 2014. – № 135. – P. 79–85.
9. McNeill M. F. Fuzzy Logic A Practical Approach / M. F. McNeill, E. Thro. – AP Professional Boston, 1994.
10. O'Hagan N. K. Decision Making with a Fuzzy Logic Inference Engine / N. K. O'Hagan, M. O'Hagan // Proc. Applications of Fuzzy Logic Technology. Sept. 1993. Soc. of Photo-Optical Instrumentation Engineers. – 1993. – P. 320.
11. Pahola T. Benavides, Urmila Diwekar. Studying various optimal control problems in biodiesel production in a batch reactor under uncertainty / Pahola T. Benavides, Urmila Diwekar // Fuel. – 2013. – № 103. – P. 585–592.
12. William L. Oberkampf. Challenge problems: uncertainty in system response given uncertain parameters / William L. Oberkampf, Jon C. Helton, Cliff A. Joslyn, Steven F. Wojtkiewicz, Scott Ferson // Reliability Engineering and System Safety. – 2004. – № 85. – P. 11–19.
13. Yao-Qing Ren. Dynamic switching based fuzzy control strategy for a class of distributed parameter system / Yao-Qing Ren, Xiao-Gang Duan, Han-Xiong Li, C. L. Philip Chen. // Journal of Process Control. – 2014. – № 24. – P. 88–97.
14. Zadeh L. A. The role of fuzzy logic in the management of uncertainty in expert systems / L. A. Zadeh // Fuzzy sets and systems. – 1983. – Vol. 11, iss. 13. – P. 197–198.

References

1. Novikov S. I., Shahnovich V. R., Safronov A. V. Metody nechetkoy logiki v zadachakh avtomatizatsii teplovykh protsessov elektrostantsiy [Fuzzy Logic Methods in the Tasks of Thermal Processes Automation of Power Plant]. *Vestnik Ivanovskogo gosudarstvennogo energeticheskogo universiteta* [Bulletin of Ivanovo State Power Engineering University], 2010, no. 4. pp. 1–4.
2. Aydogan Savran, Gokalp Kahraman. A fuzzy model based adaptive PID controller design for nonlinear and uncertain processes. *ISA Transactions*, 2014, no. 53, pp. 280–288.
3. Bartolomeo Cosenza, Mosè Galluzzo. Nonlinear fuzzy control of a fed-batch reactor for penicillin production. *Computers and Chemical Engineering*, 2012, no. 36, pp. 273–281.
4. Carlo Bagnoli & Halbert C. Smith, The Theory of Fuzzy Logic and its Application to Real Estate Valuation. *Journal of Real Estate Research, American Real Estate Society*, 1998., vol. 16 (2), p. 169–200.
5. Del Giudice, V.; De Paola, P.; Cantisani, G.B. Valuation of Real Estate Investments through Fuzzy Logic. *Buildings*, 2017, no. 7, pp. 26.
6. Didier D., Prade H. Resolution principles in possibilistic logic. *International Journal of Approximate Reasoning*, 1990, no. 4, Issue 1, pp. 1–21.
7. Escaco J.M., Bordons C., Vilas C., Garcia M.R., Alonso A.A. Neurofuzzy model based predictive control for thermal batch processes. *Journal of Process Control*, 2009, no. 19, pp. 1566–1575.
8. Kang Jun, Wenjun Meng, Ajith Abraham, Hongbo Liu. An adaptive PID neural network for complex nonlinear system control. *Neurocomputing*, 2014, no. 135, pp. 79–85.
9. McNeill M. F., Thro E. *Fuzzy Logic A Practical Approach*. AP Professional Boston, 1994.
10. O'Hagan N. K., O'Hagan M. Decision Making with a Fuzzy Logic Inference Engine. *Proc. Applications of Fuzzy Logic Technology, Sept. 1993, Soc. of Photo-Optical Instrumentation Engineers*, 1993, p. 320.
11. Pahola T. Benavides, Urmila Diwekar. Studying various optimal control problems in biodiesel production in a batch reactor under uncertainty. *Fuel*, 2013, no. 103, pp. 585–592.
12. William L. Oberkampf, Jon C. Helton, Cliff A. Joslyn, Steven F. Wojtkiewicz, Scott Ferson. Challenge problems: uncertainty in system response given uncertain parameters. *Reliability Engineering and System Safety*, 2004, no. 85, pp. 11–19.
13. Yao-Qing Ren, Xiao-Gang Duan, Han-Xiong Li, C. L. Philip Chen. Dynamic switching based fuzzy control strategy for a class of distributed parameter system. *Journal of Process Control*, 2014, no. 24, pp. 88–97.
14. Zadeh L. A. The role of fuzzy logic in the management of uncertainty in expert systems. *Fuzzy sets and systems*, 1983, vol. 11, iss. 13, pp. 197–198.

DOI 10.21672/2074-1707.2020.49.4.049-060
УДК 004.02

ПРОГНОЗИРОВАНИЕ ОТКАЗОВ СЛОЖНЫХ МНОГООБЪЕКТНЫХ СИСТЕМ НА ОСНОВЕ КОМБИНАЦИИ НЕЙРОСЕТЕЙ: ПУТИ ПОВЫШЕНИЯ ТОЧНОСТИ ПРОГНОЗИРОВАНИЯ¹

Статья поступила в редакцию 05.03.2020, в окончательном варианте – 11.03.2020.

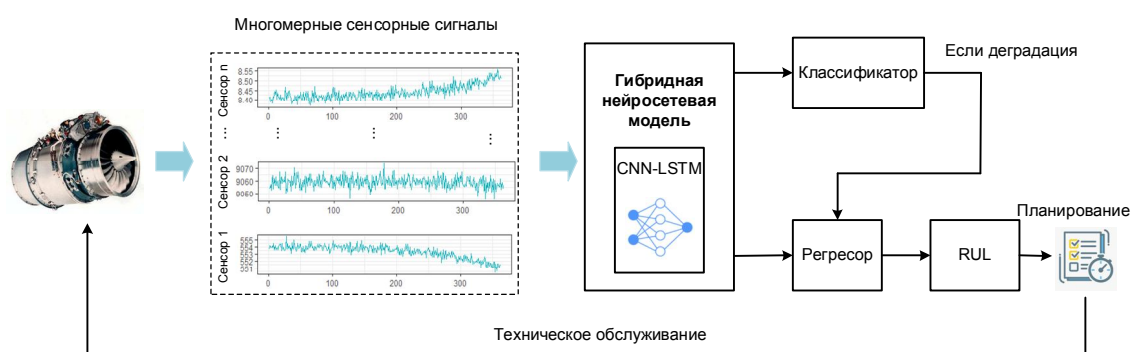
Сай Ван Квонг, Волгоградский государственный технический университет, 400005, Российская Федерация, г. Волгоград, пр. Ленина, 28,
аспирант, e-mail: svcuonghvkts@gmail.com

Щербаков Максим Владимирович, Волгоградский государственный технический университет, 400005, Российская Федерация, г. Волгоград, пр. Ленина, 28,
доктор технических наук, профессор, e-mail: maxim.shcherbakov@vstu.ru

Для прогнозирования отказов сложных многообъектных систем в работе предложена гибридная нейросетевая модель с двумя выходами на основе сверточных нейронных сетей (*convolutional neural network, CNN*) и сетей долгой краткосрочной памяти (*long short-term memory; LSTM*). Сети CNN используются для извлечения пространственных свойств из многомерных сенсорных данных, а сети LSTM – для темпорального моделирования долговременных зависимостей. Первый выход предложенной модели является классификатором, позволяющим предсказать отказ системы в следующих n -шагах. Другими словами, он является идентификатором стадии деградации оборудования. Второй выход является регрессором, позволяющим оценить величину остаточного ресурса (*remaining useful life, RUL*) оборудования. Представлены результаты вычислительных экспериментов, подтверждающие высокую эффективность предложенного решения.

Ключевые слова: предсказательное обслуживание, остаточный ресурс, глубокие нейронные сети, ансамблевые методы, гиперпараметрическая оптимизация

Графическая аннотация (Graphical annotation)



FAILURE PREDICTION OF COMPLEX MULTIPLE-COMPONENT SYSTEMS BASED ON A COMBINATION OF NEURAL NETWORKS: WAYS TO IMPROVE THE ACCURACY OF FORECASTING

The article was received by the editorial board on 05.03.2020, in the final version – 11.03.2020.

Sai Van Cuong, Volgograd State Technical University, 28 Lenin Ave., Volgograd, 400005, Russian Federation,
post-graduate student, e-mail: svcuonghvkts@gmail.com

Shcherbakov Maksim V., Volgograd State Technical University, 28 Lenin Ave., Volgograd, 400005, Russian Federation,
Doct. Sci. (Engineering), Professor, e-mail: maxim.shcherbakov@gmail.com

The paper proposes a hybrid neural network model with two outputs based on convolutional neural networks (CNN) and long short-term memory networks (LSTM) for predicting failures of complex multi-component systems. CNN networks are used to extract spatial properties from multidimensional sensor data, and LSTM networks are used for temporal modeling of long-term dependencies. The first output of the proposed model is a classifier that allows

¹ Работа выполнена при поддержке РФФИ, проект № 19-47-340010_p_a.

you to predict whether the system may fail in the next n -steps of time in the future, in other words, it is an identifier of the stage of degradation of the equipment. The second output is a regressor that allows to predict the number of the remaining useful life (RUL) of the equipment at each time step. The results of computational experiments confirming the high efficiency of the proposed solution are presented.

Key words: predictive maintenance, remaining useful life (RUL), deep neural network, ensemble method, hyperparameter optimization

Введение. В современных экономических условиях обеспечение технического обслуживания и ремонта (ТОиР) оборудования на промышленных предприятиях [2], на транспорте, в высокотехнологичных медицинских организациях [1] и пр. играет важнейшую роль. Одним из основных факторов, оказывающих влияние на рентабельность деятельности предприятий в различных отраслях промышленности, являются простои и отказы оборудования, сопровождающиеся большими производственными потерями. Это сказывается на повышении себестоимости производства продукции и снижает конкурентоспособность предприятий, работающих в различных предметных областях. Аналогично для медицинских организаций простои медоборудования ухудшают «доступность и качество» медицинской помощи и, как следствие, конкурентоспособность на рынке оказания медицинских услуг.

Обеспечение надежного функционирования оборудования на этапе эксплуатации с минимальными затратами является актуальной задачей для различных производств/организаций с учетом следующих факторов: 1) быстрого роста сложности систем; 2) возрастающих рисков, связанных с простоями оборудования; 3) ужесточения условий к выполнению требований техники безопасности на производстве и охране окружающей среды.

В последнее время в рамках внедрения концепции «Индустрия 4.0» в области организации ТОиР сложной техники значительный интерес вызывают подходы, основанные на использовании передовых методов анализа больших массивов данных на базе использования инновационных технологий искусственного интеллекта. В основном речь идет о концепции предсказательного технического обслуживания (*predictive maintenance (PdM)*), а именно о создании предсказательных моделей для предотвращения отказов оборудования. Под прогнозированием отказа здесь понимается оценка длительности временного интервала, по истечении которого произойдет отказ оборудования. Этот интервал получил название «остаточный срок полезного использования» или «остаточный ресурс» (*remaining useful life (RUL)*). Достоверная оценка RUL и вероятности наступления отказа оборудования является основанием для принятия оптимального решения при выборе стратегии воздействия на оборудование [3]. Такие решения, принятые на основе достоверного прогноза RUL, обеспечивают соблюдение оптимального баланса между затратами на проведение технологических воздействий (операций) и величиной потенциальных ущербов и рисков от отказов оборудования.

В рамках техник *PdM* исторически изучаются различные подходы к моделированию процесса деградации оборудования для оценки RUL: а) подходы, основанные на физической модели деградации (*physical model-based methodology (MBM)*); б) подходы, основанные на использовании формализованных знаний (*knowledge-based methodology (KBM)*); в) подходы, основанные на обработке и анализе данных (*data-driven methodology (DDM)*). Тенденции последних лет доказывают, что именно DDM обладает наибольшим потенциалом и практической эффективностью, которая продолжает наращиваться на основе использования как статистических методов анализа многомерных данных, так и методов машинного обучения. Статистические подходы включают винеровский процесс [13] и методы стохастической фильтрации [5]. Среди методов машинного обучения наибольшее распространение получили такие методы, как метод опорных векторов (*support vector machine (SVM)*) [8], метод случайного леса (*random forest (RF)*) и градиентный бустинг (*gradient boosting (GB)*) [9]. В последнее время методы глубокого машинного обучения показали отличную производительность при прогнозировании RUL. Здесь отметим сверточные нейронные сети (*Convolutional neural network (CNN)*) [10] и рекуррентные нейронные сети (*recurrent neural network (RNN)*) [6, 7]. Одна из причин успешного применения глубоких нейронных сетей заключается в том, что они автоматически выделяют из данных важные признаки, необходимые для решения задачи [4]. В работе [3] предложенный нами гибридный метод на основе сетей CNN и LSTM дал превосходный результат по сравнению с альтернативными методами при прогнозировании остаточного ресурса авиационных газотурбинных двигателей.

Настоящая работа посвящена совершенствованию архитектуры предложенной нами гибридной модели, оптимизации гиперпараметров и преобразованию целевой переменной RUL для получения более высокой точности при прогнозировании отказов сложных многообъектных систем.

Постановка задачи. Пусть эксплуатируется некоторый парк однотипного оборудования типа β . Условно будем называть такой парк оборудования парком оборудования и обозначать его как D^β .

Пусть каждая единица оборудования в парке $id \in D^\beta$, $id = 1, \dots, m$ оснащена сенсорной сетью, состоящей из n сенсоров $s_i^{(id)}$, $i = 1, \dots, n_s$, при этом сенсоры являются источниками информации о состоянии узлов и агрегатов оборудования и об условиях его эксплуатации. Тогда парк однотипного оборудования D^β представляется на основе сенсорных данных в следующем виде:

$$D^\beta = \langle id, \{s_i^{(id)}\}_{i=1}^{n_s}, \{r_i^{(id)}(T^{(id)})\}_{i=1}^{n_s} \rangle,$$

где id – уникальный идентификатор оборудования, $\{s_i^{(id)}\}_{i=1}^{n_s}$ – множество сенсоров для оборудования $id \in D^\beta | id = 1, \dots, m$, $r_i^{(id)}(T^{(id)})$ – данные, поступившие от сенсора $s_i^{(id)}$ для множества моментов измерения $T^{(id)} = t_1^{(id)}, t_2^{(id)}, \dots, t_{T^{(id)}}^{(id)}$, будем называть траекторией i -го сенсора для -го экземпляра оборудования $id \in D^\beta$. Тогда данные, поступившие от n_s сенсоров для единицы $id \in D^\beta$, представляют собой многомерные временные ряды $R^{(id)}$:

$$R^{(id)} = \begin{pmatrix} r_1^{(id)}(t_1^{(id)}) & \dots & r_{n_s}^{(id)}(t_1^{(id)}) \\ \vdots & \ddots & \vdots \\ r_1^{(id)}(t_{T^{(id)}}^{(id)}) & \dots & r_{n_s}^{(id)}(t_{T^{(id)}}^{(id)}) \end{pmatrix}.$$

Рассмотрим ситуацию, когда имеется конечное множество $D_q^\beta \subseteq D^\beta$, состоящее из q единиц оборудования типа β , для которых массив показаний сенсоров $R_q(T(D_q^\beta)) = \{r_i^{(id)}(T^{(id)}) | id = 1, \dots, q, \forall id : id \in D_q^\beta, i = 1, \dots, n_s\}$, состоящий из прошедших наблюдение траекторий $r_i^{(id)}$, каждая из которых заканчивается выходом из строя, соответствующего экземпляра $id \in D_q^\beta$. Другими словами, в момент времени $t_{T^{(id)}}^{(id)}$ оборудование $id \in D_q^\beta$ выходит из строя.

При изучении прошедших сенсорных данных из q единиц оборудования из множества D_q^β требуется построить эффективную модель, которая позволяет предсказать в момент времени « t », выйдет ли оборудование типа β из строя в ближайшем будущем с приемлемой точностью. Эта задача может быть сформулирована в виде одного из двух подходов. Первый подход предсказывает, существует ли вероятность отказа в следующих n -шагах по времени. Второй подход предсказывает, сколько времени осталось до следующего отказа (остаточный ресурс, RUL).

Создание целевой переменной RUL для обучения моделей. Для q единиц оборудования из множества D_q^β RUL представляет собой линейную функцию затухания и вычисляется по формуле $RUL = T^{(id)} - t^{(id)}$, где $T^{(id)}$ – время, на котором произошёл отказ для оборудования id , а $t^{(id)}$ – текущее время наблюдения за его работой (рис. 1).

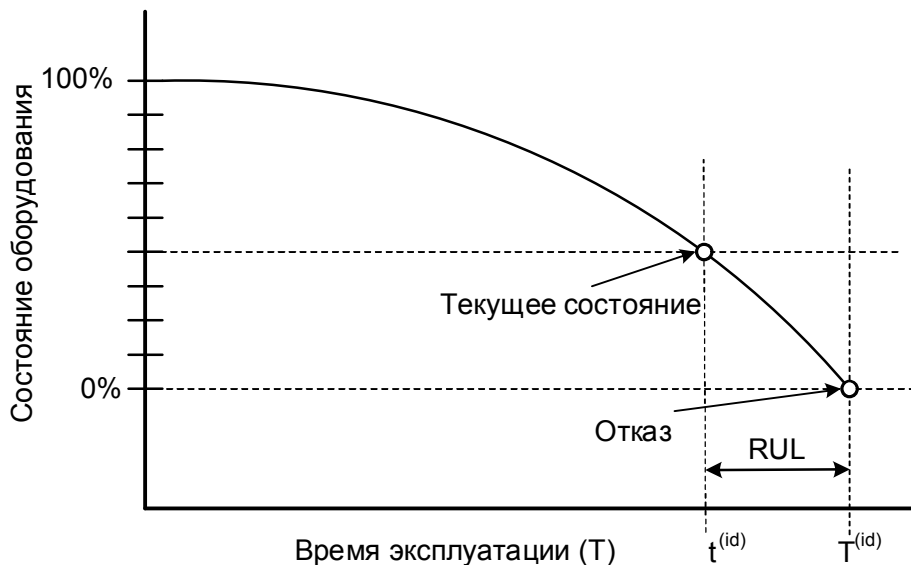


Рисунок 1 – Иллюстрация к изменению состояния оборудования и наступления отказа

Однако в начале, когда оборудование работает в совершенно «здоровом» состоянии, его состояние не может быть принято как деградирующее (рис. 2). Поэтому данные мониторинга перед тем, как оборудование начинает быстро деградировать, не следует использовать для обучения модели, потому что это может привести к тому, что модель будет ошибаться при оценке RUL. Для этого целевая переменная RUL преобразуется в виде кусочно-линейной функции следующим образом:

$$RUL(t) = \begin{cases} RUL_c & \text{для } 0 \leq t \leq T_d \\ T - t & \text{для } T_d \leq t \leq T \end{cases}$$

где T_d – момент времени начала деградация (рис. 3); RUL_c – постоянная величина RUL. Кусочно-линейная функция RUL затем используется в качестве цели для обучения алгоритмов.

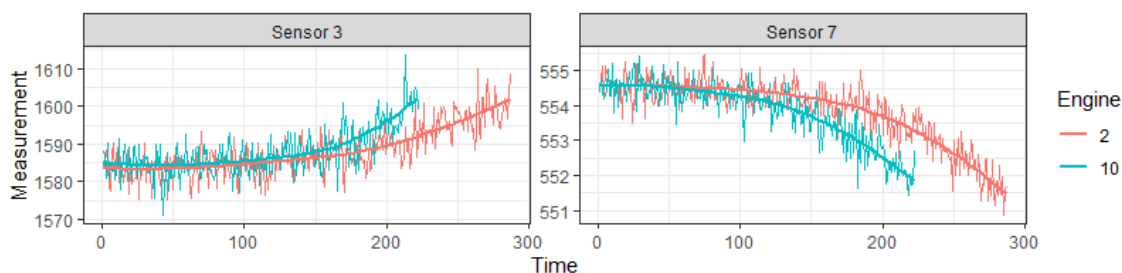


Рисунок 2 – Пример данных, полученных из двух сенсоров на двух экземплярах однотипного оборудования

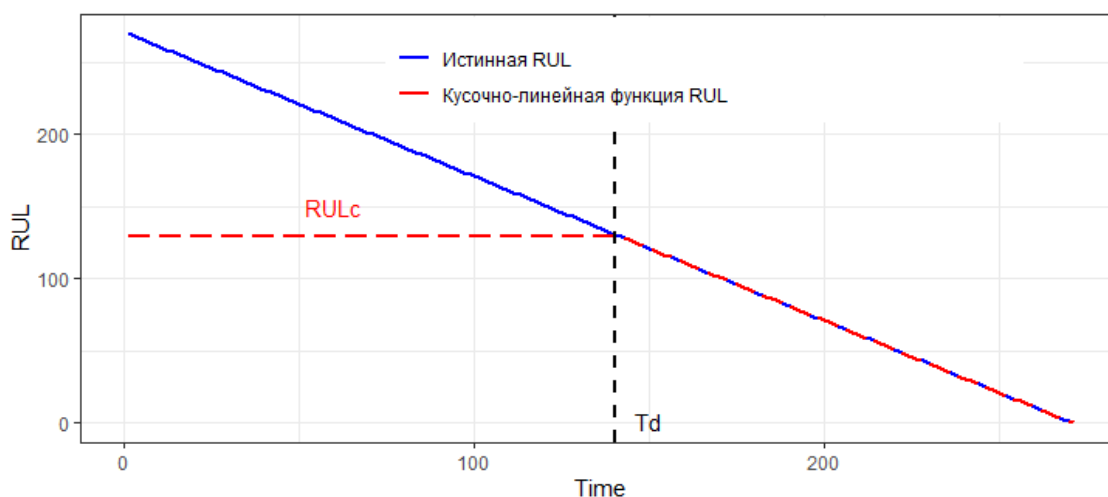


Рисунок 3 – Кусочно-линейная функция RUL

Предложенная гибридная модель CNN-LSTM. Архитектура предложенной гибридной модели CNN-LSTM представлена на рисунке 4.

Предложенная модель состоит из двух основных компонентов: а) сети 1D CNN для извлечения нужных признаков; б) сети LSTM для темпорального моделирования долговременных зависимостей признаков. Модель включает полностью связанную нейронную сеть для формирования двух выходов: 1-й предсказывает, существует ли вероятность отказа на следующих n -шагах по времени – задача классификации; 2-й предсказывает, сколько времени осталось до следующего отказа (остаточный ресурс, RUL) – задача регрессии.

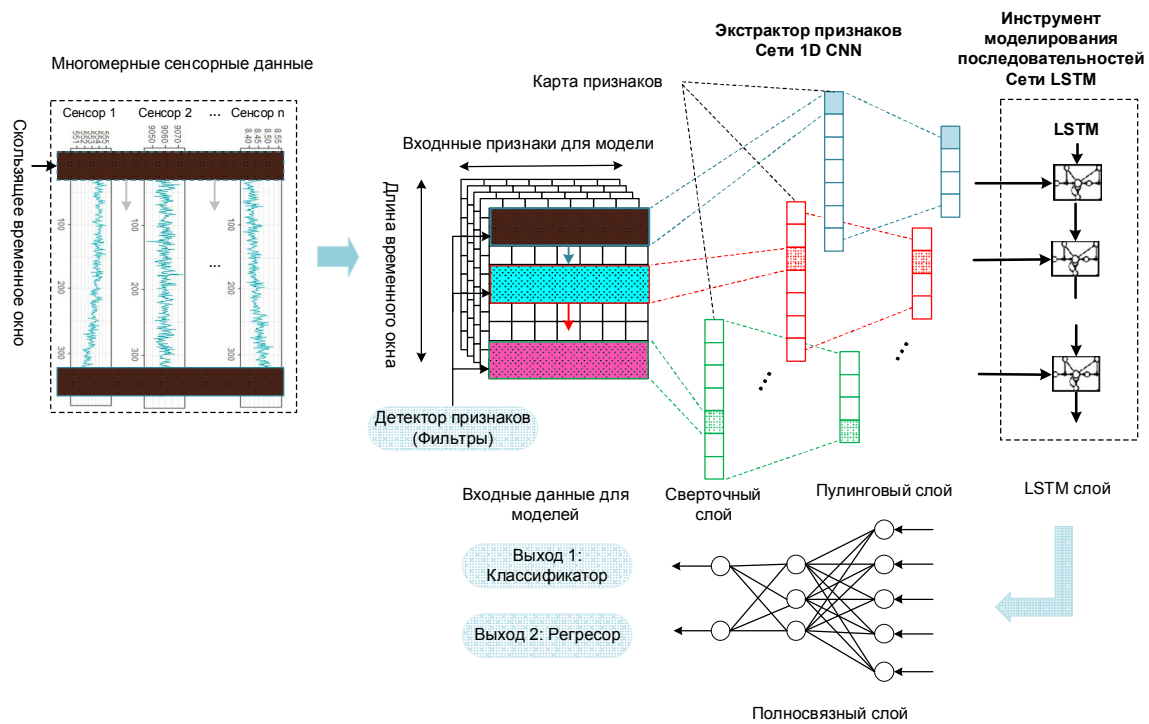


Рисунок 4 – Архитектура предложенной гибридной модели CNN-LSTM

Первый выход дает только логический ответ, но может обеспечить высокую точность при меньшем количестве данных. Второй выход предоставляет больше информации о том, когда произойдет отказ. Главным недостатком второго выхода является то, что он даёт более точный ответ тогда, когда остаточный ресурс оборудования ближе к нулю, то есть когда оборудование находится на стадии деградации. Однако на практике мы не знаем, когда оборудование находится на этой стадии. Следовательно, не можем оценить, насколько точен ответ, который дает второй выход для принятия оптимального решения. И в этом контексте важную играет первый выход как идентификатор стадии деградации оборудования.

Свёрточные нейронные сети (*convolutional neural network (CNN)*) являются лучшим инструментом машинного обучения для обработки изображений и видео. Основная причина успеха в этом случае определяется высокой способностью CNN по извлечению локальной информации из больших данных путем «скольжения» сверточных фильтров по двумерным входным данным. В таких задачах используются так называемые двумерные сверточные нейронные сети (2D CNN), так как они извлекают двумерные шаблоны из изображений и применяют идентичные преобразования к каждому такому шаблону. Аналогично можно использовать одномерные свертки для извлечения одномерных шаблонов (подпоследовательностей) из данных с сенсоров в виде последовательности. Рисунок 5 отражает разницу между 1D CNN и 2D CNN.

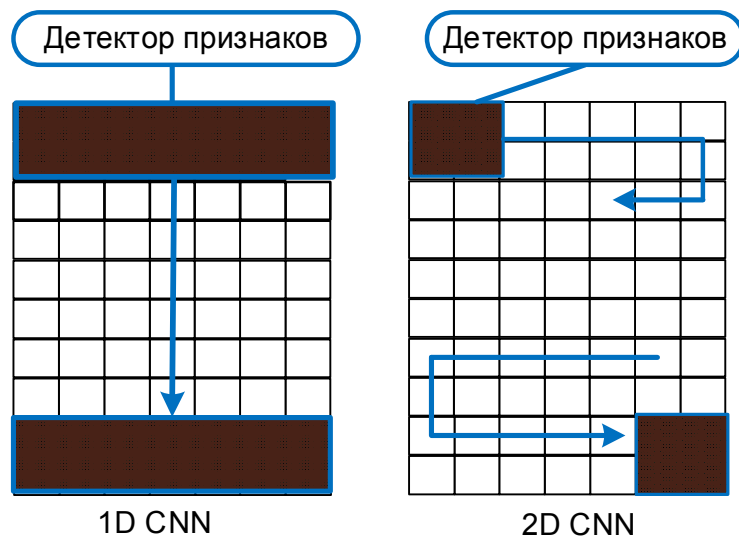


Рисунок 5 – Разница между 1D CNN и 2D CNN

Сету LSTM – особая разновидность архитектуры рекуррентных нейронных сетей (recurrent neural networks, RNN), разработанная Хохрейтером и Шмидхубером в 1997 г. LSTM имеют цепную структуру повторяющихся модулей (repeating module) нейронной сети, как и стандартные RNN. Однако повторяющиеся модули в LSTM имеют более сложную структуру (рис. 6а) по сравнению с стандартными RNN. Благодаря этой структуре LSTM способны запоминать информацию в течение длительных периодов времени. По этой причине LSTM являются мощным инструментом для моделирования данных последовательности, в частности сенсорных данных в виде многомерных временных рядов. Рисунок 6б иллюстрирует применение сетей LSTM для моделирования данных многомерных временных рядов.

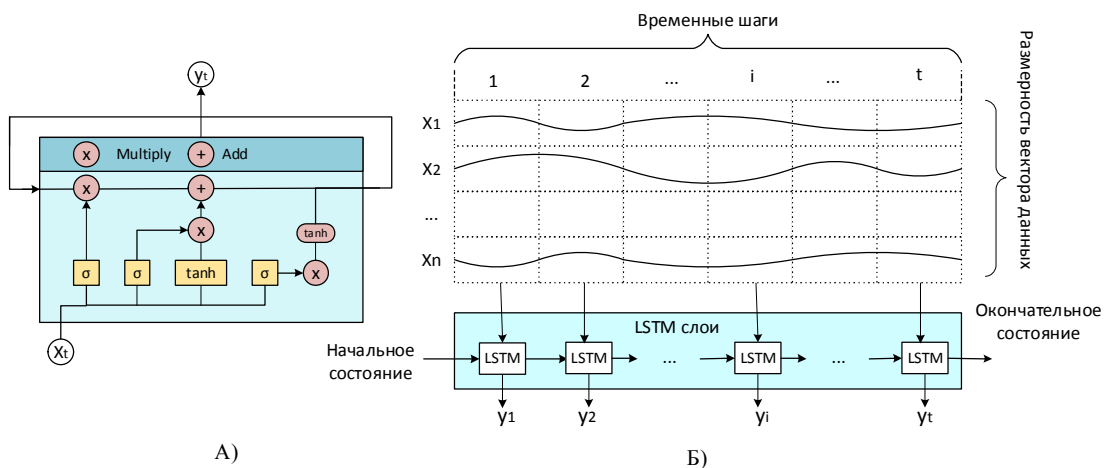


Рисунок 6 – LSTM для моделирования данных многомерных временных рядов

Алгоритм обучения предложенной гибридной модели CNN-LSTM показан на рисунке 7.

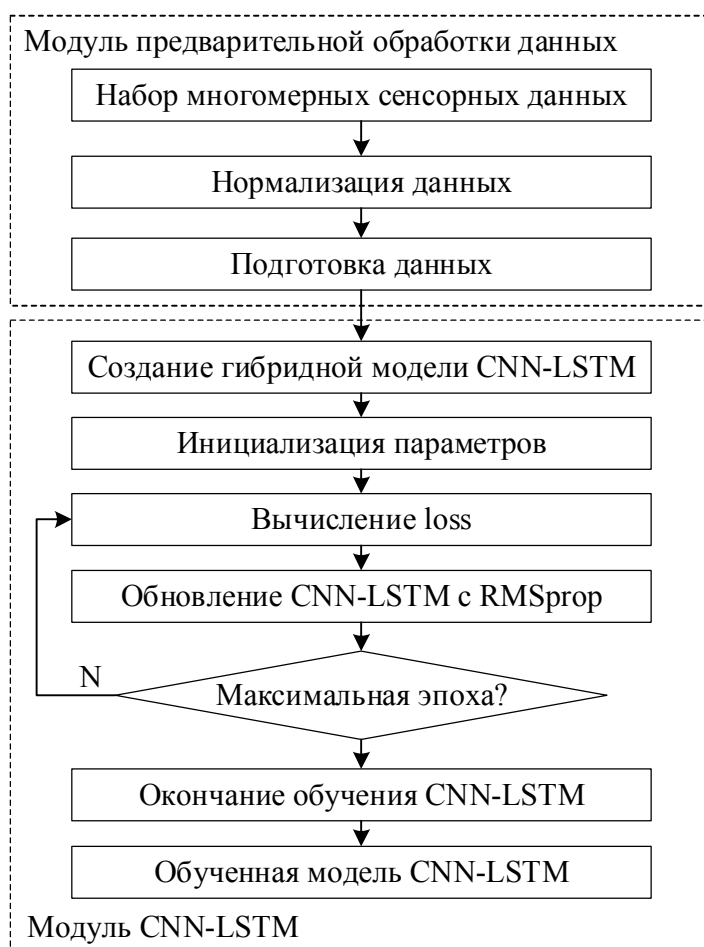


Рисунок 7 – Алгоритм обучения гибридной модели CNN-LSTM

Результаты и обсуждение. Для обоснования эффективности предложенной модели были проведены вычислительные эксперименты, реализация которых осуществлялась с помощью программы, написанной на языке Python. Нейросетевые модели были реализованы с использованием модульной библиотеки Keras [14], функционирующей поверх библиотек TensorFlow от Google [16].

Для проведения экспериментов был использован набор данных FD001 из широко известной базы данных C-MAPPS Dataset (NASA) [11], который содержит показания сенсоров в полете, имитирующих поведение однотипных авиационных газотурбинных двигателей и условия их работы. Этот набор включает в себя две выборки: обучающую выборку *train_FD001.txt* и тестовую выборку *test_FD001.txt*. Обе они содержат информацию по 100 двигателям. Каждый двигатель в обучающей выборке работает до отказа, а каждый двигатель в тестовой выборке – до определенного момента времени, предшествующего отказу. На основе данных из обучающей выборки необходимо предсказать отказ для каждого двигателя в тестовой выборке.

Подготовка данных для обучения сетей. Процесс подготовки данных для обучения моделей целесообразно представить в виде следующей последовательности шагов.

Шаг 1. Выбор признаков. Многомерные сенсорные данные в наборе FD001 содержат показания от 24 сенсоров. Однако при анализе и визуализации данных было обнаружено, что некоторые сенсоры имеют постоянные значения во времени. Поэтому следует исключить эти сенсорные показания из набора данных, так как они не дают ценной информации. В результате получается набор данных от 17 сенсоров.

Шаг 2. Нормализация данных. Следует отметить, что значения, полученные от разных сенсоров, имеют величины, относящиеся к самым разным диапазонам (рис. 2). Нейронные сети, конечно, смогут автоматически адаптироваться к таким разнородным данным, однако это усложнит их обучение и может негативно повлиять на качество результатов обучения. Поэтому для улучшения работы нейронных сетей воспользуемся min-max нормализацией данных в пределах [0, 1]:

$$X'_i = \frac{X_i - X_{min}}{X_{max} - X_{min}}$$

Шаг 3. Создание кусочно-линейной функции RUL для обучающей выборки (как описано в предыдущем разделе).

Шаг 4. Создание меток для задачи классификации. Набор данных помечается двумя классами в соответствии со значением величины RUL. Для временных шагов, где значение $RUL \geq 40$, помечается «классом 0», а там, где $RUL < 40$, помечается как «класс 1».

Шаг 5. Подготовка данных для обучения сетей. Нейронные сети будут принимать в качестве входных данных последовательность значений. В данном случае в качестве входных данных использована матрица, состоящая из 17 признаков и фиксированного скользящего окна длиной $L = 30$. В результате были получены трехмерные тензоры с формой (образцы, метки времени, признаки).

Обучение нейросетевых моделей производилось на облачном сервисе *Google Colaboratory* (также известного как *Colab*) с графическим процессором *GPU NVidia Tesla K80* с 12 Гб видеопамяти.

Оптимизация гиперпараметров нейросети с помощью фреймворка Keras-Tuner. Для выбора набора оптимальных гиперпараметров для обучающих сетей применялся приём случайного поиска с использованием фреймворка Keras-Tuner [15]. Keras-Tuner – это фреймворк для оптимизации гиперпараметров для моделей Keras, созданный Google. В данном случае выполняется поиск комбинации следующих гиперпараметров: количество нейронов в каждом слое; функции активации, которые используются в слоях; тип оптимизатора при обучении нейронной сети. Процедура оптимизации гиперпараметров для нейросетевых моделей представлена на рисунке 8.

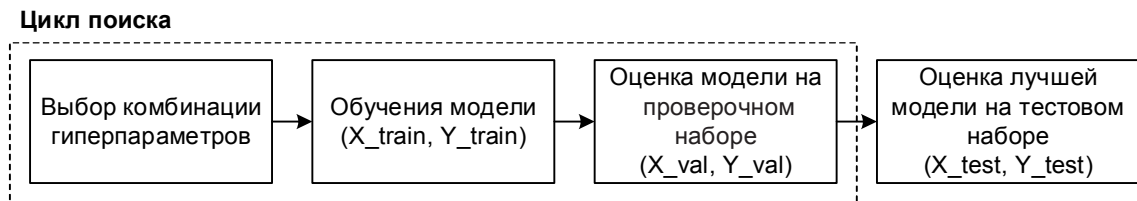


Рисунок 8 – Процедура оптимизации гиперпараметров для нейросетевых моделей

Результат в виде оптимальных гиперпараметров для предложенной модели показан в таблице 1.

Таблица 1 – Набор полученных оптимальных гиперпараметров для предложенной модели

№	Тип слоя	Фильтры/нейроны	Размер фильтра	Регион	Функция активации	Коэффициент прореживания
1	Conv1D	32	5	–	relu	–
2	Conv1D	64	3	–	relu	–
3	MaxPooling1D	–	–	3	–	–
4	LSTM	70	-	–	–	0,2
5	Dense	80	-	–	–	–
6	Dense (Классификатор)	1	-	–	sigmoid	–
7	Dense (Перцептор)	1	-	–	linear	–

На рисунке 9 представлены графики изменения значений функции ошибки по эпохе при обучении предложенной модели с использованием полученных оптимальных гиперпараметров.

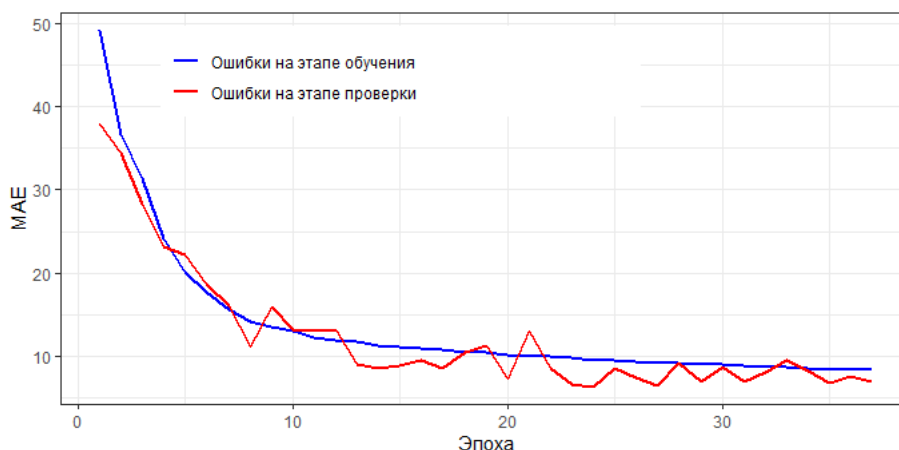


Рисунок 9 – Графики изменения значений функции ошибки по эпохе при обучении

Метрики оценки качества моделей. В данной работе для оценки точности регрессора используются средняя абсолютная ошибка (mean absolute error (MAE)), средняя квадратичная ошибка (root mean square error (RMSE)) и средняя абсолютная процентная ошибка (mean absolute percentage error (MAPE)). Они рассчитываются по формулам:

$$MAE = \frac{1}{n} \sum_{i=1}^n |e_i| = \text{mean}_{(i=1:n)}(|e_i|),$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (e_i^2)} = \sqrt{\text{mean}_{i=1:n}(e_i^2)},$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n 100 \cdot \frac{|e_i|}{y_i} = \text{mean}_{i=1:n}(100 \cdot \frac{e_i}{y_i}),$$

где $e_i = (y_i - \hat{y}_i)$, \hat{y}_i и y_i – прогнозные и истинные значения RUL для двигателя i соответственно.

Для оценки точности классификатора используется доля правильно классифицированных объектов (accuracy), точность (precision) и полнота (recall):

$accuracy = \frac{TP+TN}{TP+TN+FP+FN}$, по сути, это вероятность того, что класс будет предсказан правильно;

$precision = \frac{TP}{TP+FP}$ показывает, какую долю объектов, распознанных как объекты положительного класса, мы предсказали верно;

$recall = \frac{TP}{TP+FN}$ показывает, какую долю объектов, реально относящихся к положительному классу, мы предсказали верно,

где TP, TN, FP, FN – значения матрицы ошибок бинарной классификации (табл. 2).

Таблица 2 – Матрица ошибок бинарной классификации

	$y = 1$	$y = 0$
$\hat{y} = 1$	True Positive (TP)	False Positive (FP)
$\hat{y} = 0$	False Negative (FN)	True Negative (TN)

Примечание. \hat{y} – это ответ алгоритма на объекте; y – истинная метка класса на этом объекте. TP – истинно-положительное решение; TN – истинно-отрицательное решение; FP – ложно-положительное решение (ошибка типа 1); FN – ложно-отрицательное решение (ошибка типа 2).

На рисунке 10 представлено сравнение прогнозных значений RUL и истинных значений RUL на двигателе id = 2 в обучающей выборке для линейной целевой функции RUL и кусочно-линейной целевой функции.

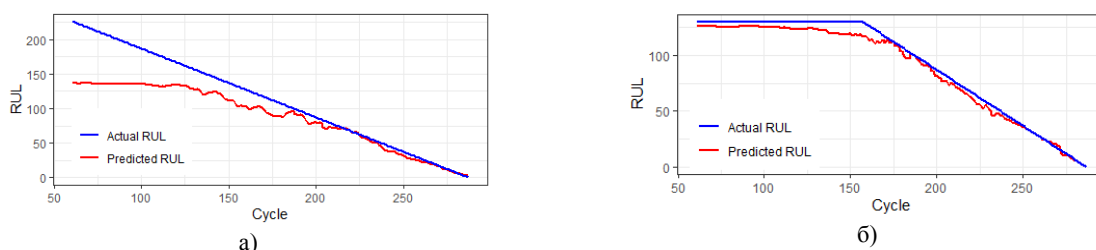


Рисунок 10 – Сравнение прогнозных значений RUL и истинных значений RUL на двигателе id = 2 в обучающей выборке: а) для линейной целевой функции RUL; б) для кусочно-линейной целевой функции RUL

RMSE, по тестовым данным полученной сети, составило 12,478. Для сравнения RMSE, по тестовым данным для предложенной нами гибридной модели в работе [3], составило 15,076. Таким образом, точность предлагаемого алгоритма повысилась на 17 %.

Таблица 3 показывает сравнение результатов, полученных для предложенной модели с результатами других исследований применительно к решению рассматриваемой задачи.

Таблица 3 – Сравнение результатов, полученных для предложенной модели с другими исследованиями

Название модели	RMSE	Год публикации
SVM [7]	29,82	2013
CNN [9]	18,45	2016
LSTM [5]	17,84	2017
Глубокие LSTM [6]	16,74	2018
DenseStacking Ensemble [11]	16,67	2019
Предложенная CNN-LSTM-модель	12,48	–

На рисунке 11 представлена визуализация реальных значений RUL и прогнозных значений RUL, полученных с помощью предложенной модели.

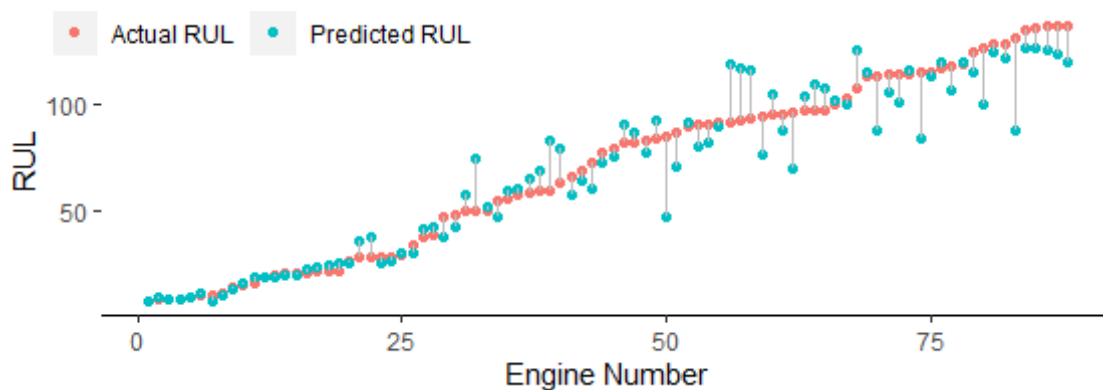


Рисунок 11 – Результаты прогнозирования RUL тестовых двигателей (по возрастанию RUL)

Следует отметить, что при возрастании величины RUL точность прогнозирования уменьшается. Другими словами, модель даёт более точные прогнозы, когда двигатель находится ближе к моменту отказа. Это подтверждается результатами, показанными в таблице 4. Она содержит результаты ошибки прогнозирования, полученные с помощью предложенной модели для прогнозирования остаточного ресурса на двигателях в тестовом наборе по различным группам.

Таблица 4 – Ошибки прогнозирования RUL предложенной модели на различных группах двигателей

Ошибка прогнозирования	Группа 1 ($RUL \leq 40$)	Группа 2 ($40 < RUL \leq 80$)	Группа 3 ($RUL \geq 80$)	Для всех двигателей
MAE	2,043977	9,047644	12,61425	8,561976
RMSE	3,024680	11,13613	16,23768	12,47810
MAPE	9,323707	15,74691	12,10987	11,92597

Точности классификации для предложенной гибридной модели на тестовых двигателях показаны в таблице 5.

Таблица 5 – Результаты для точности классификации состояний оборудования на данных по тестовым двигателям

	Accuracy	Precision	Recall
Классификатор CNN-LSTM	99 %	100 %	98,7 %

Заключение. Результаты моделирования показывают, что предложенная гибридная нейросетевая модель CNN-LSTM с двумя выходами, основанная на использовании сверточной нейронной сети (CNN) и сети долгой краткосрочной памяти (LSTM), превосходит аналоги при тестировании на базе данных C-MAPPS. Полученные результаты и алгоритмы могут быть использованы в системах превентивного обслуживания оборудования с целью высоконадежной идентификации стадии его деградации и прогнозирования отказов сложных многообъектных систем.

Библиографический список

1. Брумштейн Ю. М. Управление парком оборудования в российских медицинских учреждениях: технологии принятия и реализации решений, оценки их результативности / Ю. М. Брумштейн, Е. В. Пфандер // Прикаспийский журнал: управление и высокие технологии. – 2017. – № 4 (40). – С. 110–125. ([http://hi-tech.asu.edu.ru/files/4\(40\)/110-125.pdf](http://hi-tech.asu.edu.ru/files/4(40)/110-125.pdf))
2. Ерохин Е. А. Особенности организации служб технического обслуживания оборудования на предприятиях / Е. А. Ерохин // Экономинфо. – 2009. – № 12. – С. 11–13.
3. Сай Ван Квонг. Метод прогнозирования остаточного ресурса на основе обработки данных многообъектных сложных систем / Сай Ван Квонг, М. В. Щербаков // Прикаспийский журнал: управление и высокие технологии. – 2019. – № 1 (45). – С. 33–44. ([http://hi-tech.asu.edu.ru/files/1\(45\)/33-44.pdf](http://hi-tech.asu.edu.ru/files/1(45)/33-44.pdf))
4. Сай Ван Квонг. Глубокие нейронные сети для предсказательного технического обслуживания / Сай Ван Квонг // Моделирование, оптимизация и информационные технологии (МОИТ). – 2019. – Т. 7, № 4 (27). – 11 с. DOI: 10.26102/2310-6018/2019.27.4.011.
5. Cui L. A novel Switching Unscented Kalman Filter method for remaining useful life prediction of rolling bearing / L. Cui, X. Wang, Y. Xu, H. Jiang, J. Zhou // Measurement. – 2019. – № 135. – P. 678–684.
6. Dong D. Life prediction of jet engines based on LSTM-recurrent neural networks / D. Dong, X. Li, F. Sun // Prognostics and System Health Management Conference. – 2017. – P. 1–6.
7. Hsu C. S. Remaining useful life estimation using Long short-term memory deep learning / C. S. Hsu, J. R. Jiang // IEEE International Conference on Applied System Invention. – 2018. – P. 58–61.
8. Louen C. A new framework for remaining useful life estimation using Support Vector Machine classifier / C. Louen, S. X. Ding, C. Kandler // Conference on control and Fault-Tolerant Systems. – 2013. – P. 228–233.
9. Patil S. Remaining Useful Life (RUL) Prediction of Rolling Element Bearing Using Random Forest and Gradient Boosting Technique / S. Patil, A. Patil, V. Handikherkar, S. Desai, V. M. Phalle & F. S. Kazi // ASME 2018 International Mechanical Engineering Congress and Exposition. American Society of Mechanical Engineers Digital Collection. – 2018.
10. Sateesh Babu G. Deep Convolutional Neural Network Based Regression Approach for Estimation of Remaining Useful Life / G. Sateesh Babu, P. Zhao, X. L. Li // International Conference on Database Systems for Advanced Applications. – 2016. – P. 214–228.
11. Saxena A., Goebel K. Turbofan Engine Degradation Simulation Data Set. – Режим доступа: <https://ti.arc.nasa.gov/tech/dash/groups/pcoe/prognostic-data-repository>, свободный. – Заглавие с экрана. – Яз. англ. (дата обращения: 14.08.2019).
12. Singh S. K. A novel soft computing method for engine RUL prediction / S. K. Singh, S. Kumar, J. P. Dwivedi // Multimed. Tools Appl. – 2019 – № 78. – P. 4065–4087.
13. Zhang Z. Degradation data analysis and remaining useful life estimation: A review on Wiener-process-based methods / Z. Zhang, X. Si, C. Hu, Y. Lei // Eur. J. Oper. Res. – 2018. – № 271. – P. 775–796.
14. Keras: The Python Deep Learning library. – Режим доступа: <https://keras.io/>, свободный. – Заглавие с экрана. – Яз. англ. (дата обращения: 12.12.2019).
15. Keras Tuner. – Режим доступа: <https://keras-team.github.io/keras-tuner/>, свободный. – Заглавие с экрана. – Яз. англ. (дата обращения: 15.01.2020).
16. Tensorflow. – Режим доступа: <https://www.tensorflow.org/>, свободный. – Заглавие с экрана. – Яз. англ. (дата обращения: 08.12.2019).

References

1. Brumshteyn Yu. M., Pfander E.V. Upravlenie parkom oborudovaniya v rossiyskikh meditsinskikh uchrezhdeniyakh: tekhnologii prinyatiya i realizatsii resheniy, otsenki ikh rezul'tativnosti [Management of efficiency of operation of the park of the medical equipment in stationary medical institutions: the system analysis of the purposes, methods of acceptance and implementation of decisions, criteria for evaluation of the achieved results].

Prikaspiyskiy zhurnal: upravlenie i vysokie tekhnologii [Caspian Journal: Control and High Technologies], 2017, no. 4 (40), pp. 110–125. ([http://hi-tech.asu.edu.ru/files/4\(40\)/110-125.pdf](http://hi-tech.asu.edu.ru/files/4(40)/110-125.pdf))

2. Erokhin E. A. Osobennosti organizatsii sluzhb tekhnicheskogo osbluzhivaniya oborudovaniya na predpriyatiyakh [Characteristics of organization of equipment maintenance services at enterprises]. *Ekonominfo* [Econominfo], 2009, no. 12, pp. 11–13.

3. Say Van Kvang, Shcherbakov M. V. Metod prognozirovaniya ostatochnogo resursa na osnove obrabotki dannykh mnogoobektnykh slozhnykh system [A data-driven method for remaining useful life Prediction of multiple-component systems]. *Prikaspiyskiy zhurnal: upravlenie i vysokie tekhnologii* [Caspian Journal: Control and High Technologies], 2019, no. 1 (45), pp. 33–44. ([http://hi-tech.asu.edu.ru/files/1\(45\)/33-44.pdf](http://hi-tech.asu.edu.ru/files/1(45)/33-44.pdf))

4. Say Van Kvang. Glubokie neyronnye seti dlya predskazatel'nogo tekhnicheskogo obsluzhivaniya [Deep neural networks for predictive maintenance]. *Modelirovanie, optimizatsiya i informatsionnye tekhnologii (MOIT)* [Modeling, optimization and information technology], 2019, vol. 7, no. 4 (27). 11. p. DOI: 10.26102/2310-6018/2019.27.4.011.

5. Cui L., Wang X., Xu Y., Jiang H., Zhou J. A novel Switching Unscented Kalman Filter method for remaining useful life prediction of rolling bearing. *Measurement*, 2019, no. 135, pp. 678–684.

6. Dong D., Li X., Sun F. Life prediction of jet engines based on LSTM-recurrent neural networks. *Prognostics and System Health Management Conference*, 2017, pp. 1–6.

7. Hsu C. S., Jiang J. R., Remaining useful life estimation using Long short-term memory deep learning. *IEEE International Conference on Applied System Invention*, 2018, pp. 58–61.

8. Louen C., Ding S. X., Kandler C. A new framework for remaining useful life estimation using Support Vector Machine classifier. *Conference on control and Fault-Tolerant Systems*, 2013, pp. 228–233.

9. Patil S., Patil A., Handikherkar V., Desa S., Phalle V. M. & Kazi F. S. Remaining Useful Life (RUL) Prediction of Rolling Element Bearing Using Random Forest and Gradient Boosting Technique. *ASME 2018 International Mechanical Engineering Congress and Exposition. American Society of Mechanical Engineers Digital Collection*, 2018.

10. Sateesh Babu G., Zhao P., Li, X. L. Deep Convolutional Neural Network Based Regression Approach for Estimation of Remaining Useful Life. *International Conference on Database Systems for Advanced Applications*, 2016, pp. 214–228.

11. Saxena A., Goebel K. *Turbofan Engine Degradation Simulation Data Set*. Available at: <https://ti.arc.nasa.gov/tech/dash/groups/pcoe/prognostic-data-repository> (accessed 14.08.2019).

12. Singh S. K., Kumar S., Dwivedi J. P. A novel soft computing method for engine RUL prediction. *Multimed. Tools Appl.*, 2019, no. 78, pp. 4065–4087.

13. Zhang Z., Si X., Hu C., Lei Y., Degradation data analysis and remaining useful life estimation: A review on Wiener-process-based methods. *Eur. J. Oper. Res.*, 2018, no. 271, pp. 775–796.

14. *Keras: The Python Deep Learning library*. Available at: <https://keras.io/> (accessed 12.12.2019).

15. *Keras Tuner*. Available at: <https://keras-team.github.io/keras-tuner/> (accessed 15.01.2020).

16. *Tensorflow*. Available at: <https://www.tensorflow.org/> (accessed 08.12.2019).

УПРАВЛЕНИЕ В СОЦИАЛЬНЫХ И ЭКОНОМИЧЕСКИХ СИСТЕМАХ

УДК 004.413+004.415+004.416

РАЗРАБОТКА, ВВЕДЕНИЕ В ЭКСПЛУАТАЦИЮ И ИСПОЛЬЗОВАНИЕ «САЙТОВ-ВИЗИТОК»: АНАЛИЗ ЦЕЛЕЙ, ВОЗМОЖНЫХ ПОДХОДОВ, ТРУДОЕМКОСТИ, СТРУКТУРЫ РИСКОВ

Статья поступила в редакцию 20.11.2019, в окончательном варианте – 11.03.2020.

Бондарев Андрей Андреевич, ГБУЗ АО «Приволжская РБ», 414018, Российская Федерация, г. Астрахань, ул. Александрова, 9а,

специалист по защите информации, e-mail: andrey.a.bondarev@gmail.com

Васильев Никита Вячеславович, Национальный исследовательский технологический университет «МИСиС», 119049, Российская Федерация, г. Москва, Ленинский пр-т, 4, студент, ORCID <https://orcid.org/0000-0001-6305-1938>; e-mail: nikivas97@mail.ru

Коновалова Дарья Игоревна, Национальный исследовательский технологический университет «МИСиС», 119049, Российская Федерация, г. Москва, Ленинский пр-т, 4, магистрант, ORCID <https://orcid.org/0000-0002-8348-0171>, e-mail: haizbq@gmail.com

Брумштейн Юрий Моисеевич, Астраханский государственный университет, 414056, Российская Федерация, г. Астрахань, ул. Татищева, 20а, кандидат технических наук, доцент, ORCID <http://orcid.org/0000-0002-0016-7295>; https://elibrary.ru/author_profile.asp?authorid=280533; e-mail: brum2003@mail.ru

Выполнен комплексный анализ проблематики, связанной с созданием, внедрением, поддержкой эксплуатации, обеспечением эффективности использования и информационной безопасности сайтов-визиток. Типичные цели создания и использования сайтов-визиток проанализированы с позиций их владельцев: отдельных физических лиц, их групп, небольших организаций. Охарактеризован состав информации, размещаемой на сайтах-визитках различных типов; композиционные решения сайтов-визиток; цветовое и шрифтовое оформление; цели и принципы использования гиперссылок, графических объектов. Показано, что для физических лиц использование сайтов-визиток, в том числе на иностранных языках, может быть предпочтительным по сравнению со страничками в социальных сетях; персональными страницами на сайтах организаций. Рассмотрены возможности и ограничения создания сайтов-визиток с помощью программных модулей пакета Microsoft Office, других офисных пакетов. Показано, что имеющиеся на интернет-ресурсах средства «автоматизированного создания» сайтов позволяют обойтись без помощи профессиональных специалистов-разработчиков. Однако такие средства во многих случаях не обеспечивают достаточного качества и функциональности сайтов-визиток. В связи с этим проанализирована номенклатура и возможности инструментальных средств разработки, применение которых может быть оправданным при создании сайтов-визиток. Оценены трудозатраты на профессиональную разработку типичных «казачных» сайтов-визиток, их внедрение, при необходимости – сопровождение. Описан состав затрат на получение доменных имен для сайтов-визиток, продление сроков их действия. Исследованы вопросы хостинга сайтов-визиток (в том числе вне России для иноязычных сайтов-визиток); рациональные подходы к обеспечению «видимости» сайтов-визиток поисковыми системами интернета, поддержки необходимых уровней информационной безопасности сайтов-визиток.

Ключевые слова: сайты-визитки, иноязычные сайты, направления использования, разработка, офисные пакеты, конструкторы сайтов, инструментальные средства, хостинг, введение в эксплуатацию, информационная безопасность

DEVELOPMENT, COMMISSIONING AND USE OF “BUSINESS CARD SITES”: ANALYSIS OF TARGETS, POSSIBLE APPROACHES, LABOR CONSUMPTION, RISK STRUCTURE

The article was received by the editorial board on 20.11.2019, in the final version – 07.03.2020.

Bondarev Andrey A., State Budgetary Health Institution of the Astrakhan region «Volga regional hospital», 9a Aleksandrov St., Astrakhan, 414018, Russian Federation, information protection specialist, e-mail: andrey.a.bondarev@gmail.com

Vasilyev Nikita V., National Research Technological University “MISiS”, 4 Leninskiy Prospect, Moscow, 119049, Russian Federation, master student, ORCID <https://orcid.org/0000-0001-6305-1938>; e-mail: nikivas97@mail.ru

Konovalova Daria I., National Research Technological University “MISiS”, 4 Leninskiy Prospect, Moscow, 119049, Russian Federation, master student, ORCID <https://orcid.org/0000-0002-8348-0171>, e-mail: haizbq@gmail.com

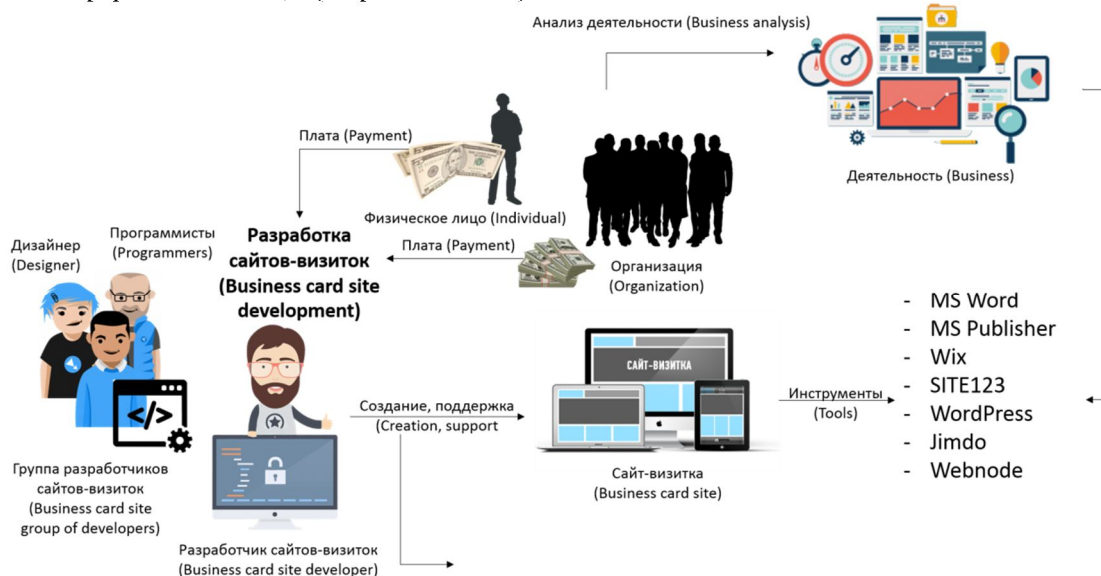
Brumshteyn Yuri M., Astrakhan State University, 20a Tatishchev St., Astrakhan, 414056, Russian Federation,

Cand. Sci. (Engineering), Associate Professor, ORCID <http://orcid.org/0000-0002-0016-7295>; https://elibrary.ru/author_profile.asp?authorid=280533; e-mail: brum2003@mail.ru

A comprehensive analysis of the problems associated with the creation, implementation, operation support, ensuring the efficient use and information security of business card sites has been carried out. Typical goals of the creation and use of business card sites were analyzed from the perspective of their owners: individuals, their groups, small organizations. The composition of the information posted on various types of business card sites; business card sites compositional solutions, color and font design, goals and principles of hyperlinks using, graphic objects are characterized. It was shown that for individuals using of business card sites, incl. in foreign languages, may be preferable compared to pages on social networks; personal pages on the websites of organizations. The possibilities and limitations of creating business card sites using the software modules of the Microsoft Office suite and other office suites are considered. It was shown that with help of available Internet tools that allow creating sites automatically, it is possible to refuse assistance of professional developers. However, such tools do not provide sufficient quality and functionality of business card sites in many cases. In this regard, the nomenclature and capabilities of such development tools, the use of which may be justified when creating business card sites, were analyzed. Labor costs for the professional development of regular custom business card sites, their implementation, if necessary – maintenance were estimated. The composition of the costs for business card sites domain names obtaining, the extension of their validity period were described. The issues of business card sites hosting (including outside of Russia for foreign-language business card sites); rational approaches of ensuring the business card sites visibility by Internet search engines, supporting the necessary levels of information security of business card sites were investigated.

Key words: business card sites, foreign-language sites, directions of use, development, office suites, site designers, tools, hosting, commissioning, information security

Графическая аннотация (Graphical annotation)



Введение. Развитие интернета привело к появлению большого количества сайтов, размещению на них различных видов информации. Одним из типов интернет-сайтов являются «сайты-визитки» (СВ). Обычно они имеют простую структуру, минимальную функциональность, небольшой объем. В свою очередь это обеспечивает следующие преимущества: возможности достаточно быстрого создания СВ, низкую трудоемкость их внедрения, небольшие затраты на хостинг. На практике создание «заказных» СВ часто осуществляют не только специализированные фирмы, действующие в сфере «сайтостроительства» (веб-разработок), но и индивидуальные разработчики, включая даже студентов вузов и техникумов. Несмотря на распространенность СВ, в существующей литературе фактически отсутствуют работы, посвященные комплексному анализу относящейся к ним проблематики. Поэтому в данной статье ставилась цель устранить данный пробел.

Общая характеристика места российских сайтов в интернет-пространстве. В соответствии с [16], «сейчас на доменной зоне .com зарегистрировано более 150 миллионов сайтов. Из них

активными являются приблизительно 140 миллионов». Для доменной зоны «.de», по данным на [16] со ссылкой на данные Verisign (www.verisign.com), «текущее количество доменов составляет порядка 100 миллионов». На третьем месте по популярности находится домен «.net» – на начало 2020 г., по данным [16], на нем «числится ... порядка 14 миллионов сайтов». Укажем также статистику иных доменных имен «первого уровня» [16]: «.uk» – 12 миллионов; «.cn» – 11 млн имен; «.org» – 10 млн; «.info» – 9 млн; «.nl» (нидерландский национальный домен) – «6 миллионов имён», причем «около 70 % из них принадлежат сайтам коммерческих организаций» [16]; «.eu» – «более 5 миллионов имён», причем эта зона «закрывается для стран, не входящих в Европейский союз»; «.ru» – «более 5 миллионов зарегистрированных имён». Эта информация необходимо учитывать при создании и размещении не русскоязычных сайтов.

На <https://cctld.ru/files/stats/Stat-Q119.pdf> (дата доступа 23.02.2020) приведена статистика по количеству имен в национальных доменных зонах: «.cn» – 20 868 593; «.de» – 16 215 736; «.uk» – «12 197 103»; «.nl» – 5 854 118; «.ru» – 5 017 021; «.br» – 4 004 821; «.it» – 3 196 818. Отдельная статистика по СВ в мировом масштабе не ведется, так как сам по себе этот термин не имеет четкого определения. Обычно считается, что СВ имеет не более нескольких страниц.

Соответственно данным, на <https://cctld.ru/> для России на 23.02.2020 г. в доменной зоне «.ru» было зарегистрировано 4 971 607 доменных имен (здесь и далее – «второго доменного уровня»); 740 770 имен – в доменной зоне «.рф». При этом количество регистраторов доменных имен в России составляет «52» и эти организации расположены в 10 городах РФ. Кроме того, согласно сведениям, на <https://cctld.ru/ru/statistics/> на апрель 2019 г. имелось еще 110 738 доменов в зоне «.su». На <https://cctld.ru/> домен «.ru» является «пятым по популярности в мире» среди «национальных доменных имен». Согласно сведениям, на <https://cctld.ru/domains/about/#ru> за 2019 г. в домене «.рф» «...насчитывается около 800 тысяч доменных имен, домен стабильно растет». Кроме того, на той же странице сайта указывается, что имеется «домен верхнего уровня. ДЕТИ». Он «предназначен для адресации русскоязычных сайтов для детей и о детях». В этом домене также могут быть размещены СВ.

Доменные имена могут регистрироваться как для использования непосредственно теми лицами, которые их регистрируют, так и для последующей продажи другим юридическим или физическим лицам. В связи с этим отметим следующее. 1. Использование СВ является простым и удобным способом «занять» зарегистрированное доменное имя для последующей перепродажи, разместить информацию об условиях такой продажи и необходимые контактные данные продавца. 2. Имеется специальная методология расчета стоимости доменных имен на «вторичном рынке» [15] для России. Однако на практике стоимости перепродажи сайтов назначаются и с учетом тех факторов, которые в этой методике не упоминаются. 3. На <https://cctld.ru/service/dellist/> имеются списки имен «освобождающихся доменов». 4. На <https://cctld.ru/service/whois/> размещен сервис, позволяющий проверить не только «занятость» имени домена, но и контактные данные «регистранта», которому он принадлежит, т.е. текущего владельца сайта.

С позиций информационной безопасности отметим, что СВ подвержены угрозам, свойственным и другим типам веб-сайтов. В частности, СВ могут быть использованы как площадки для размещения «вредоносных программных средств». В связи с этим приведем статистику, отраженную на сайте <https://netoscope.ru/> – он применяется для проверки сайтов в отношении наличия угроз информационной безопасности (ИБ). На декабрь 2019г: всего доменов (4 348 845); вредоносные – 3 792 860 (при этом увеличение количества сайтов в этих двух группах практически одинаково); спам – 211 400; фишинг – 92 831; ботнеты – 24 065. Конечно, все современные антивирусные/антиспамовые программные средства содержат «базы данных» вредоносных сайтов и на основе этих баз данных либо автоматически блокируют попытки доступа к таким сайтам, либо выдают предупреждающие сообщения. Однако существуют и угрозы «нулевого дня», т.е. появление вредоносных сайтов, информацию по которым разработчики еще не успели внести в «базы данных»; возможность появления новых типов угроз, которые не распознаются антивирусными средствами.

Интернет-пользователи, применяющие мобильные устройства для доступа в интернет, считаются менее защищенными, чем пользователи стационарных ПЭВМ. Основная причина – мобильные устройства имеют более короткий «жизненный цикл», предусмотренный производителем – в среднем 12–18 месяцев, и, как правило, без декларируемых сроков дальнейшей поддержки. Поэтому установленная на таких устройствах версия операционной системы получает обновления в «активной фазе продаж», а затем остается без поддержки после замещения производителем новой моделью на рынке. При этом производители лишают пользователей возможности самостоятельной «перепрошивки» своих устройств, используя блокировку загрузчика, привязку прошивок к региону, оператору и т.п. Отметим, что в России дорогостоящие мобильные устройства обычно используются несколько лет. При этом существует развитый рынок ремонта таких устройств, а также рынок продаж устройств, «бывших в употреблении», что обеспечивает продление их жизненного цикла.

Проблемы безопасности мобильных устройств наслаиваются на проблемы совместимости старых версий мобильных операционных систем с новыми обновленными версиями прикладного программного обеспечения. Также некоторые устройства остаются без поддержки в результате введения санкций со стороны страны-производителя операционной системы (Huawei и Google).

Антивирусные программы на мобильных устройствах пока являются редкостью, хотя они присутствуют в магазинах приложений, а всё, что оттуда устанавливается, всегда предварительно проверяется. Также свои встроенные антивирусы имеют программы онлайн-банкинга (это требование Центрального Банка РФ). Например, в «Сбербанк Онлайн» есть встроенная антивирусная программа, работающая на основе «планировщика» и осуществляющая проверку самого устройства на различные типы кастомизации. Если устройство было «рутовано» (т.е. по нему были получены права «суперпользователя»), то установка «Сбербанк Онлайн» невозможна, так как безопасная работа этого приложения не гарантируется. В частности, для «рутованных устройств» могут быть навязаны неконтролируемые перехваты данных из/в приложение.

В то же время по имеющейся статистике количество «мобильных интернет-пользователей» продолжает расти, и они могут использовать для доступа СВ. Согласно информации исследовательской компании Mediascope, размещенной на сайте <https://www.seonews.ru/events/mediascope-78-naseleniya-rossii-polzuetsya-internetom/> (на 23.02.2020) «по данным исследования WEB-Index за февраль-июль 2019 года, число пользователей интернета в России составляет 95,8 млн человек старше 12 лет или 78 % населения». При этом количество «мобильных пользователей» составляет 68,9 %, т.е. немногим меньше 78 %, а «в городах с населением от 100 тыс. человек мобильным Интернетом пользуются 47,1 млн человек или 74 % населения» – т.е. большинство жителей.

Авторам настоящей статьи не удалось найти в интернете следующей информации, важной для темы статьи: а) по количеству СВ в указанных выше «российских доменных зонах». Самостоятельное проведение такого исследования в принципе возможно (с выделением сайтов с количеством страниц не выше заданного), но крайне трудоемко; б) по количеству сайтов в этих доменных зонах, имеющих иноязычные страницы; в) о проведении «онлайн аукционов» по перепродаже доменных имен в российских доменных зонах.

Понятие «сайта-визитки», основные особенности создания таких сайтов. Целью использования СВ в типичных случаях является предоставление ограниченного объема информации о ФЛ (реже – о небольшой организации) для обеспечения минимального объема информационного присутствия в интернет-пространстве. При этом состав информации, находящейся на СВ, как правило, изменяется достаточно редко.

Существенно, что СВ могут быть представлены на иностранных языках и размещены в различных доменных зонах – для обеспечения информационного присутствия ФЛ и организаций вне русскоязычного информационного пространства (ИнП) [4]. Это особенно существенно, если у ФЛ нет других возможностей представления иноязычной информации о себе – например, из-за отсутствия на корпоративных сайтах англоязычных версий для персональных страниц ФЛ.

У лиц, занимающихся научной деятельностью, возможности «самопрезентации» в иноязычном ИнП несколько шире. Отметим, в частности, сайт www.orgcid.org. Однако на нем состав размещаемой исследователями «информации о себе» регламентирован и ограничен по объему. Для них также доступны сайты «профессиональных социальных сетей», фондов-грантодателей; репозитории научной информации и пр.

Основные особенности СВ с позиций разработчиков: небольшое количество страниц – нередко одна; обычно не применяются сколько-нибудь сложные анимации; как правило, не используется анализ cookies и др. В то же время «форма обратной связи используется» – с позиций информационной безопасности такое решение предпочтительно по сравнению с указанием e-mail (последний может быть выявлен автоматическими системами и использован для рассылки спама).

Многие СВ изначально разрабатываются в виде одностраничного приложения (SPA – Single Page Application). Взаимодействие СВ с пользователем организуется через динамически подгружаемые HTML, CSS, JavaScript – чаще всего посредством AJAX. Особым вариантом является одностраничный СВ, используемый только для загрузки Flash-приложения, а вся навигация и содержание сайта выполнены «внутри» Flash-ролика. В связи с этим отметим, что с переходом к открытой платформе HTML5 флэш-анимации начали считаться устаревшими решениями.

Как следствие, создание СВ предъявляет относительно невысокие требования к квалификации разработчиков и в ряде случаев доступно даже для рядовых пользователей, не обладающих знаниями в сфере выполнения веб-разработок, веб-дизайна и пр. При этом могут быть использованы возможности офисных пакетов; бесплатные или платные конструкторы сайтов (в том числе и тематически специализированных), которые доступны на различных интернет-ресурсах [14]. Эти конструкторы могут использовать различные шаблоны (включая многостраничные), а также типо-

вые решения для определенных предметных областей – в том числе и для обеспечения двустороннего взаимодействия с интернет-пользователями.

С позиций интернет-пользователей преимуществом применения СВ отдельными ФЛ является наличие «сгруппированной» информации о них, которую легче выявить (найти) с использованием поисковых систем интернета. По сравнению с использованием персональных страниц ФЛ и организаций в соцсетях использование СВ имеет как достоинства, так и недостатки (эти вопросы рассматриваются далее).

Состав информации и цели использования сайтов-визиток. В типичных случаях на СВ могут размещаться следующие категории информации. 1. О ФЛ (или небольших организациях), которые являются владельцами сайтов. 2. О других ФЛ или организациях – в том числе с целью их рекламы или, наоборот, компрометации. 3. Совокупности ссылок на другие интернет-сайты. Это характерно, например, для сайтов-двореев [1, 2]. Кроме того, такой прием может использоваться отдельными ФЛ для «группировки» ссылок на информацию о них, которая размещена на различных интернет-ресурсах. 4. Рекламная и информационно-рекламная информация. 5. Обзорная, компилятивная или «сводная» информация, представленная в краткой форме. 6. Справочная и/или фактографическая информация в краткой форме. 7. Статистическая информация ограниченного объема. 8. Научные статьи и иные материалы, в том числе такие, которые отказались публиковать журнальные издания, средства массовой информации и пр. 9. Ссылки для запуска программных разработок, выполненных в инициативном порядке ФЛ или их группами. Такое решение может, в частности, применяться для тестирования разработок. 10. Специальные формы для оформления подписки на новостные рассылки по электронной почте. 11. Материалы учебного или учебно-методического характера. 12. Записи собственных выступлений ФЛ, в том числе политического характера – это могут быть аудио и видеоролики, аудиозаписи и пр. 13. Фотоматериалы или ссылки на фотоматериалы, относящиеся к ФЛ или их группам. 14. Некоторые несложные «интернет-опросники», предназначенные для разового или кратковременного использования.

С позиций физических и юридических лиц можно выделить несколько типичных целей создания и применения СВ, которые последовательно рассматриваются ниже.

1. Для сколько-нибудь крупных российских организаций использование русскоязычных СВ в целом не характерно. Причина – такие сайты обычно не обеспечивают необходимую функциональность, возможности структурированного размещения значительных объемов информации и пр. Поэтому указанные организации могут применять СВ преимущественно как временный вариант – вплоть до выполнения разработки и введения в действие полнофункциональных русскоязычных сайтов. При этом доменное имя, зарегистрированное организацией для сайта, при переходе на его полнофункциональную версию обычно сохраняется.

2. В то же время для небольших организаций, а также для «индивидуальных предпринимателей», создание и использование русскоязычных СВ может быть вполне рациональным с точки зрения соотношения «функциональность/затраты» [12]. Экономически рентабельное повышение популярности (посещаемости) таких сайтов может быть обеспечено, в частности, с помощью оплаченной контекстной рекламы, видимой только региональными пользователями из тех же регионов или «узкими» целевыми группами пользователей; за счет иных средств продвижения сайта [17].

Для объективности отметим, что с целью приближения рекламы производимой продукции или оказываемых услуг к потребителям сейчас все чаще используются сайты социальных сетей (включая «ВКонтакте»), а интернет-пользователи свои запросы делают не в поисковой системе Google или Yandex, а в YouTube.

3. Для ФЛ, не являющихся частными предпринимателями, использование СВ может быть альтернативой размещению необходимой информации на персональных страничках в социальных сетях, на личных страницах сайтов организаций, на некоторых специализированных сайтах для определенных категорий лиц (например, исследователей) и пр.

Преимущества СВ по сравнению с персональными страницами на сайтах организаций: а) нет необходимости строгого соблюдения «корпоративных» правил в отношении объема и состава информации, форм ее представления, ограничений на применение графических объектов, гиперссылок и пр.; б) легче обеспечить возможности обнаружения размещенной на СВ информации с помощью поисковых систем интернета.

По сравнению со страничками в социальных сетях у СВ есть некоторые преимущества: возможность представления информации в произвольной форме и в необходимых объемах; при необходимости – обеспечение возможности дифференцированного доступа пользователей к размещенной информации за счет использования «логинов-паролей» для входа на СВ в целом или на их отдельные сегменты.

При необходимости СВ могут также применяться для инициативного размещения аудио- и видеороликов больших объемов – больших, чем допускается правилами для корпоративных сайтов, страниц в социальных сетях и пр. Это же может относиться к программным средствам, в том числе разработанным пользователями.

4. Создание СВ и их размещение на интернет-ресурсах может быть также способом быстро «заявить о себе» для некоторых вновь созданных общественных организаций, неформальных групп ФЛ и пр.

5. При создании СВ на их стартовых страницах могут применяться короткие тексты, переведенные на несколько языков, что делает информацию доступной для более широкого контингента пользователей, чем в случае только русскоязычных вариантов сайтов [2, 4].

Альтернативой может быть применение многостраничных СВ с переключателями между несколькими языками – при этом целесообразно использование кнопок-переключателей в виде национальных флагов соответствующих стран. Еще одним решением может быть автоматический выбор «языковой» страницы для входа на сайт с использованием анализа того IP-адреса, с которого пользователем произведен вход.

6. Возможно также создание российскими гражданами и организациями СВ на различных иностранных языках для размещения в «интернациональных доменных зонах» и в «национальных доменных зонах» соответствующих стран. Целью такого размещения является обеспечение минимального информационного присутствия в иноязычном информационном пространстве (ИИП). Это делает такие СВ видимыми для иноязычных поисковых запросов (на английском языке или на национальных языках соответствующих стран) – в том числе при использовании иноязычных вариантов поисковых систем интернета.

7. Создание и применение СВ достаточно характерно для целей киберсквоттинга, в том числе в доменной зоне «.ru». При этом регистрируются доменные имена, которые интуитивно ассоциируются с названиями известных организаций или их аббревиатурами (на английском или русском языках), а также с транслитерациями этих названий на латиницу. Затем «регистрантами» для этих имен создаются простейшие одностраничные сайты, и на них размещается информация о продаже имени сайта; о стоимости и иных условиях такой продажи; контактная информация для проведения переговоров о такой продаже и пр.

Такие «сайты-нэймхолдеры» (сайты-заклушки), в том числе и вне доменных зон «.ru», «.su», «.rf», могут находиться в интернет-пространстве достаточно долго. Этому способствуют следующие факторы: а) стоимость первоначальной регистрации доменных имен и их сохранения за «регистрантами» (при условии наличия сайтов) на порядки меньше доходов, которые они могут получить от продажи этих имен; б) количество доменных имен, регистрируемых на одно ФЛ или юридическое лицо обычно жестко не ограничивается; в) «регистранты» могут размещать на своих СВ «счетчики посещений» и отслеживать количества входов посетителей. Это позволяет им оценивать перспективность сохранения прав на имена сайтов на последующие периоды, определять целесообразные стоимости продаж. Отметим также, что в интернете отсутствуют сообщения о «вторичном рынке» доменных имен, по крайней мере в Рунете, т.е. об их многократной перепродаже с целью получения прибыли.

8. Еще одним направлением использования «сайтов-нэймхолдеров» может быть их применение в конкурентной борьбе между организациями – для исключения использования заинтересованными организациями (в том числе и зарубежными, которые работают в России), интуитивно ожидаемых для них доменных имен.

Аналогичное решение может также применяться политическими деятелями, а чаще лицами, действующими по их поручениям (с оговорками также политическими партиями), для заблаговременного исключения возможностей использования подходящих доменных имен их политическими конкурентами (противниками) – реальными или потенциальными.

9. Создание одностраничных сайтов вне доменных зон .rf, .ru, .su для размещения негативной информации о российских политических деятелях, политических партиях, различных организациях, руководстве регионов и пр. Такая информация обычно предназначается для «компрометации» ФЛ или юридических лиц, в том числе и в рамках применения «грязных» предвыборных технологий. При этом фактические владельцы сайтов чаще всего являются для интернет-пользователей неизвестными, а информация может иметь не только тенденциозный характер, но и просто вымышленный (фейковый). В силу размещения СВ вне доменных зон, относящихся к российской юрисдикции, для ФЛ или юридических лиц, которые компрометируются, во многих случаях просто отсутствуют или сильно затруднены возможности принудительного прекращения деятельности таких сайтов, исключения доступа к ним пользователей или принудительного удаления с сайтов недостоверной информации. Отметим, что такого рода сайты могут специально создаваться для использования в

течение только короткого времени – для решения поставленных «оперативных» задач политического характера, включая «вброс» в интернет-пространство ложной информации.

10. Одностраничные сайты, размещенные вне доменных зон «.ru», «.su», «.rf», могут также использоваться и для оперативного «вброса» в интернет-пространство недостоверной информации о чрезвычайных происшествиях (включая техногенные аварии и катастрофы); о массовых беспорядках; о межэтнических или межконфессиональных конфликтах в странах, в их регионах и пр. Отметим, что указанная информация часто имеет высокую вероятность дальнейшего «самораспространения» в интернет-пространстве, в том числе через социальные сети; реже – в виде сообщений новостных агентств и некоторых иных организаций со ссылками на интернет-источники.

При этом официальные опровержения такой информации далеко не всегда могут быть эффективны; иногда приводят к неоправданному привлечению внимания к указанным видам ложной информации. Таким образом, потенциально интернет-сайты могут быть средством проведения «информационных атак» в отношении отдельных регионов и стран. При этом высокая скорость формирования СВ не требует их заблаговременного создания, а доменные имена могут быть получены заранее.

11. Одностраничные сайты могут применяться и в целях «вбросов» в интернет-пространство фейковой информации о серьезных заболеваниях или смерти известных политических деятелей, в том числе участвующих (или предполагающих участвовать) в избирательных компаниях; о возбуждении против них уголовных дел; о чрезвычайных происшествиях с их родственниками, ближайшими помощниками и пр.

Для сравнения укажем, что распространение фейковой или тенденциозной информации о заболеваниях иных «медиа-персон» (артистов, шоуменов, спортсменов и пр.) систематически осуществляется через «новостные сообщения», публикуемые на российских интернет-сайтах массового посещения.

12. Одностраничные сайты также могут использоваться для распространения информации о кризисных проявлениях в деятельности различных партий, общественно-политических движений и пр.

13. По крайней мере ранее СВ нередко создавались для размещения совокупностей ссылок на другие сайты и выполняли функции «сайтов-дорвеев» [1, 2]. Это обеспечивало повышение «видимости» поисковыми системами интернета тех сайтов, на которые указывали ссылки на сайтах-дорвеев. Причина: поисковые системы выдавали в «начале списков поисковых выдач» те сайты, на которые ими обнаруживалось наибольшее количество ссылок с других интернет-сайтов. Однако в настоящее время в силу повышения «интеллектуальности» средств создания баз данных для поисковых систем такие искусственные методы повышения «популярности» сайтов считаются утратившими свое значение.

14. Для студентов вузов, особенно обучающихся по направлениям, связанным с «информационно-телекоммуникационными технологиями» (ИТКТ), создание СВ может преследовать также чисто учебные цели. Например, это касается изучения курсов по web-программированию, web-дизайну и некоторых иных. При этом студенты обычно стараются использовать бесплатные средства создания СВ и их хостинга, в том числе и вне вузов.

15. В целях обучения специалистов по информационной безопасности, в том числе для обучения анализу «журналов» (логов) для СВ; оценки защищенности интернет-сайтов от взлома.

16. Решение «тестовых задач» путем использования СВ обеспечивает следующие преимущества: быстроту внесения необходимых изменений на сайтах; отсутствие «побочных эффектов», связанных с наличием многочисленных интернет-страниц и пр. «Тестовые цели» создания и использования СВ могут включать в себя следующие направления: **а)** проверка работы «адаптивных режимов» работы интернет-страниц на различных типах дисплеев/устройств, которые могут применять пользователи; **б)** проверка рациональности использования выбранных цвето-шрифтовых решений и/или анимаций, в том числе путем организации «интернет-голосований» посетителей; **в)** тестирование отображения СВ при использовании различных браузеров, устройств с различными размерами дисплеев; **г)** оценка видимости СВ в интернет-пространстве, в том числе в иноязычном ИнП – с использованием иноязычных запросов. При этом может осуществляться некоторый «подбор» информационного наполнения сайтов, улучшающий их видимость для поисковых систем интернета, в том числе на иноязычных; **д)** оценки скоростей доступа к СВ, в том числе с использованием специальных тестовых программ, имитирующих открытие сайтов пользователями из различных стран мира. Отметим, что с учетом большого числа «мобильных пользователей» (см. выше) количества/объемы графических объектов на СВ должны быть ограничены; **е)** тестирование работы различных видов самостоятельно разработанных программных средств, в том числе предназначенных для взаимодействия с пользователями; для подсчетов количеств посетителей из различных регионов, стран и пр.; **ж)** экспериментальная проверка уровней ИБ сайтов путем искусственного проведения на них «хакерских атак», в том числе с учебными целями.

Таким образом, цели создания и использования СВ являются различными для разных категорий ФЛ и юридических лиц. В свою очередь, цели использования и предполагаемая длительность эксплуатации СВ значительно влияют на уровни приемлемых трудовых и/или денежных затрат на их создание; на сложность композиционных решений, на цветовое и шрифтовое оформление, на использование на таких сайтах гиперссылок, графических объектов (статических и анимированных), иных средств и пр.

Анализ достоинств и недостатков основных средств создания сайтов-визиток. Как уже было сказано ранее, создание СВ и их информационное наполнение могут осуществляться с использованием различных программных средств, предъявляющих разные требования к квалификации разработчиков.

Первое направление – это использование для создания СВ офисных пакетов.

1. В России наиболее популярным офисным пакетом является Microsoft Office (MsOffice) различных версий. В нем пользователями наиболее часто применяется текстовый редактор Microsoft Word (MsWord). При этом в большинстве версий MsWord есть встроенная возможность создания HTML-документов, которыми по существу и являются СВ. Штатными средствами MsWord удается создавать варианты СВ с элементами ввода текста на странице, графическими изображениями и гиперссылками.

Возможности бесплатной конвертации различных видов документов в HTML формат (в том числе в режиме онлайн) имеются также в многочисленных бесплатных конвертерах, размещенных на Интернет-ресурсах.

Графические объекты для СВ могут быть получены следующими способами (методами): а) фотографирование ФЛ для получения их изображений – возможно с последующей компьютерной обработкой, монтажом и пр.; б) фотографирование других объектов; в) создание графических объектов с помощью средств рисования или коррективки изображений, включенных в состав пакета MsOffice или иных офисных пакетов; г) использование для создания графических изображений специализированных программных средств для работы с графикой, в том числе средств, имеющихся в составе операционных систем; д) путем заимствования изображений с различных интернет-ресурсов с их коррективкой, компоновкой нескольких изображений в единое целое, снабжения изображений необходимыми «группирующими рамками», надписями и пр.

Однако в целом возможности СВ, созданных с использованием MsWord (а также конвертации doc или docx файлов в HTML документы), являются достаточно ограниченными. Основная причина – MsWord изначально предназначен для создания «печатных» документов, которые имеют фиксированный размер страницы, шрифт и макет. При формировании в MsWord кода веб-страницы, информация представляется в «нестандартном бумажном стиле». При этом HTML страница будет отображаться «правильно» только при использовании браузера Microsoft Internet Explorer – а он в России является малопопулярным, в том числе и у многочисленных мобильных пользователей, использующих смартфоны. В любых иных браузерах созданные HTML страницы могут отображаться совсем не так, как подразумевает разработчик.

2. Другой инструмент фирмы Microsoft, при помощи которого можно создавать СВ – MS Publisher. Это настольная издательская система, отличающаяся от MsWord тем, что акцент делается на проектировании разметки страницы, а не на оформлении и проверке текста. MS Publisher располагает выбором макетов в виде шаблонов веб-страниц. Однако у MsPublisher (аналогично MsWord) на выходе получается статичный СВ, который во многих браузерах будет отображаться неправильно, либо даже не отображаться вообще. На практике для решения таких проблем будут требоваться постоянные ручные коррективки СВ, так как обеспечить «адаптивный интерфейс» сайтов в MsPublisher нельзя.

Аналогичные возможности формирования HTML страниц есть и в соответствующих модулях других офисных пакетов. При этом недостатки отображения страниц, как правило, те же самые. Поэтому рассмотренные в пунктах 1 и 2 программные средства (а также их аналоги) по крайней мере для профессионального создания СВ не используются.

Второе направление использования программных средств для создания СВ – это применение «конструкторов сайтов» [3, 14]. Такие конструкторы обычно доступны дистанционно на интернет-ресурсах, например [21, 23]. Они могут быть как универсальными, так и специализированными. Причем ряд конструкторов позиционируются как предназначенные именно для создания СВ [5, 6, 11, 25]. Ниже достоинства и недостатки конкретных «конструкторов» сайтов рассматриваются по следующим направлениям: а) бесплатность (или платность) использования самого конструктора, хостинга/домена; б) простота/удобство интерфейса; в) требования к навыкам программирования у пользователя; г) количество предлагаемых шаблонов, тем и пр.; д) принудительное включение рекламы в созданный сайт; е) использование «дополнительных инструментов»,

включая сбор статистики посещаемости, средства SEO оптимизации. В отношении наличия средств SEO подчеркнем, что для владельцев СВ их продвижение в интернет-пространстве (в том числе и иноязычном) является важнейшей задачей. Основные цели внутренней и внешней оптимизации сайтов: «поднятие» их позиций в списках, выдаваемых поисковыми системами; увеличение количества потенциальных посетителей, особенно регулярных посетителей; увеличение средней продолжительности нахождения пользователей на сайте; увеличение сетевого трафика и т.п.

Типичный состав информации, используемой для оптимизации СВ: анализ трафика; мониторинг выдач поисковых систем и ключевых слов (для получения информации о запросах пользователей), поиск и анализ внешних ссылок на сайты, проведение конкурентного анализа сайтов-аналогов. При этом отработку ряда «технологических решений» относящихся к SEO удобно делать именно с использованием СВ.

Отметим, что в России (и за рубежом) существуют и специализированные фирмы, предлагающие услуги по «продвижению» сайтов в интернет-пространстве. Предлагаются также «*бесплатные курсы по созданию сайтов*», например, на <https://design-gym.ru/>.

Характеристика наиболее распространенных «конструкторов сайтов». В настоящее время сведения о возможностях, достоинствах и недостатках конструкторов интернет-сайтов есть в следующих типах источников: **а)** обзоры в виде статей, публикуемых в различных научных изданиях, включая «молодежные» [3, 8, 10, 24]; **б)** обзоры или перечни конструкторов на интернет-сайтах – например, [14], на <https://sitebuilders.club/besplatnyie-konstruktoriyi-saytov> – в последнем случае имеются ссылки на ресурсы в количестве «41», предназначенные для создания СВ; **в)** материалы посвященные отдельным конструкторам сайтов [9, 26]; **г)** отдельное внимание уделяется созданию и использованию сайтов для педагогов [13, 18], а также иных видов тематически специализированных сайтов [11, 20]; вопросам использования «*дизайн-интеллекта*» для создания сайтов [9], продвижения СВ в интернет-пространстве [17]. В Рунете имеется также достаточно большое количество отзывов о различных конструкторах сайтов как «хвалебного», так и «критического» характера.

Для пользователей Рунета предлагаются только русскоязычные конструкторы сайтов, без возможности переключения на англоязычный интерфейс. Однако при желании есть и англоязычные конструкторы, например, на домене «.com».

1. В настоящее время одним из самых разрекламированных (популярных) в России «*универсальных*» конструкторов сайтов является Wix (<https://ru.wix.com/>) [14, 22, 26]. Это международная облачная платформа для создания и развития интернет-проектов, которая позволяет конструировать сайты и их версии для мобильных пользователей на HTML5 с помощью инструментов drag-and-drop. Достоинства Wix: **а)** предоставление персонального домена пользователю, что облегчает задачу создания СВ. При этом использование доменного имени и хостинг на неограниченное время являются бесплатными); **б)** наличие ADI (Artificial Design Intelligence), т.е. элементов искусственного интеллекта (машинного обучения) для создания дизайна сайта. Средство ADI сканирует указанный пользователем список сайтов-аналогов и на основе собранной информации создает «кастомизированный» (персонализированный в отношении дизайна) сайт [14]. При необходимости этот дизайн затем может быть «вручную» доработан (скорректирован) разработчиком. Также на сайт может быть вставлена «индивидуальная» информация о ФЛ или организации. Использование ADI позволяет автоматически адаптировать дизайн вновь создаваемого сайта к текущим версиям сайтов-аналогов (что позволяет учесть имеющиеся тенденции web-дизайна) или к архивным версиям сайтов (например, на www.arxiv.org). Таким образом, ADI позволяет существенно снизить трудоемкость создания новых сайтов с не слишком сложной структурой.

Помимо этого, Wix предоставляет разработчику и иные инструменты для быстрого создания и запуска СВ, обладает обширной галереей шаблонов, обеспечивает возможности интеграции с другими приложениями и сторонними сервисами, позволяет проводить SEO-оптимизацию. Wix обладает следующими встроенными SEO функциями: индексация сайта в поисковых системах Яндекс и Google; рациональный выбор ключевых слов для обеспечения возможности нахождения разработанного сайта (в том числе и СВ) среди других сайтов; возможность отслеживать позицию сайта в поисковых выдачах для поисковых систем интернета; добавление разметки сайта для отображения его контента в расширенном описании сайта в поисковой выдаче; создание карты сайта для индексации и пр.

Недостатки WiX: **а)** представительство Wix в России отсутствует, а любые претензии клиентов из нашей страны компания Wix предлагает направлять в офис в Нью-Йорке. Это может быть существенным, в частности, с позиций ИБ – в том числе и при создании «подставных» сайтов для компрометации ФЛ или организаций; **б)** отсутствие в WiX «штатных» средств сбора и сохранения статистики посещаемости сайта, что может быть важным для владельцев СВ. Хотя на СВ могут быть «вставлены» необходимые средства сбора статистики посещаемости от «сторонних разра-

ботчиков», но большинство их носит платный характер; в) отсутствие возможности редактирования HTML страниц с использованием другого шаблона, по сравнению с тем, который был первоначально использован при их создании; г) автоматическая вставка на СВ рекламы сайта Wix в виде соответствующего логотипа; д) ограниченные объемы памяти для хранения данных в случае использования бесплатной версии – однако для СВ это ограничение может быть не существенным. Кроме того, на основе «платных планов обслуживания» пользователей это ограничение может быть устранено, а также получен доступ к «специальным лимитированным функциям».

2. Конструктор сайтов SITE123 (<https://app.site123.com/?aff=370034&sid=RU-508666&l=ru&landingSystem=1>) [14, 19], считается ориентированным на новичков. Его специализацией являются сайты для ведения блогов, сайты-визитки, лендинги (в том числе веб-страницы для сбора контактных данных целевой аудитории) и интернет-магазины. Достоинства: простой интерфейс; наличие инструментов SEO; функционал включает возможности интеграции с платежными системами (например, PayPal), инструментами в социальных сетях пр.; значительное количество шаблонов и интегрированных плагинов; наличие круглосуточной технической поддержки; возможности бесплатного использования (доменного имени и услуг хостинга); наличие встроенного инструмента для перевода текстов на иностранные языки. Недостатки: ограниченный функционал бесплатной версии; невозможность использования «кодинга» для сайтов, созданных с помощью SITE123; отсутствие возможностей экспорта/импорта веб-сайтов. Считается, что «система подходит для сборки технически несложных, адаптивных, приятных внешне сайтов с небольшим количеством страниц» [14]. Под такое определение СВ вполне подходят.

3. WordPress (<https://ru.wordpress.com/>) – система управления содержимым сайта с открытым исходным кодом. Может применяться для создания сайтов разной сложности: от сайтов, предназначенных для ведения блогов до достаточно сложных по структуре «новостных ресурсов». Достоинства: бесплатное использование ресурса для создания сайтов; большое количество тем, шаблонов, плагинов, различных функциональных инструментов; открытый исходный код, что обеспечивает возможности использования разработчиками «собственного кода»; возможности формирования резервных копий сайтов; наличие инструментов SEO. Недостатки: платный хостинг, если необходим свой домен 2-го уровня; необходимы знания программирования, если требуется выход за рамки стандартного функционала.

4. Jimdo (<https://www.jimdo.com/>) представляет собой «облачный сервис» для создания сайтов и управления ими. При этом знаний HTML, CSS и программирования от пользователей не требуется. Достоинства: применение конструктора с визуальным, интерфейсом; наличие техподдержки; использование для создания «платформ» Creator (drag&drop технология) и Dolphin (ADI технология), бесплатность использования. Недостатки: ограниченный набор шаблонов; относительно сложный редактор; необходимость наличия у пользователя навыков программирования; принудительное включение элементов рекламы в созданный сайт; платное использование домена.

5. Webnode (<https://www.webnode.ru/>) – это онлайн конструктор сайтов с использованием технологии drag-and-drop и возможностями элементов типа блога, форума, фотогалереи, виджетов (элементов интерфейса, имеющих стандартный внешний вид и выполняющих стандартные действия) и т.д. Достоинства: возможность создания/редактирования сайта с помощью смартфона (это важно с учетом большого количества мобильных пользователей – см. выше); наличие специальных инструментов для создания многоязычных сайтов; бесплатность использования (при соблюдении некоторых ограничений); удобный редактор; значительная коллекция шаблонов; наличие инструментов SEO; использование средств сбора «статистики по сайту» и интеграции с Google Analytics (<https://analytics.google.com/>); возможности резервного копирования данных и восстановления сайта. Недостатки: нельзя изменять «шаблон» после того, как сайт уже создан; домен в «бесплатном тарифе» имеет обязательное вложение webnode.com. Домен помимо названия, которое пользователь даст СВ, будет содержать текстовый логотип webnode.com – если пользователя это не устраивает, то он должен использовать «платный тариф» обслуживания на сайте.

Ниже для описанных средств разработки сайтов приводится сравнительная таблица с использованием следующих обозначений для критериев: к1) бесплатное пользование; к2) наличие возможности использования бесплатного хостинга/домена второго уровня; к3) простой и удобный интерфейс; к4) наличие у пользователя навыков программирования; к5) присутствие рекламы на созданном сайте; к6) наличие инструментов для анализа сайта; к7) наличие шаблонов (тем) для создания сайта; к8) наличие инструментов для SEO-оптимизации сайта; к9) наличие инструментов для поддержки создания многоязычных сайтов; к10) возможность пройти регистрацию с помощью уже имеющихся аккаунтов (Facebook, Google).

Таблица 1 – Сравнение «конструкторов сайтов» по выбранным критериям

	к1	к2	к3	к4	к5	к6	к7	к8	к9	к10
Wix	+	+	+	–	+	–	+	+	+	+
SITE123	+	+	+	–	–	неизв.	–	+	+	+
WordPress	+	–	+	+	неизв.	–	неизв.	+	+	+
Jimdo	+	–	+	+	+	–	–	неизв.	–	+
Webnode	+	–	+/-	неизв.	–	+	+	+	+	+

В выдачах поисковых систем интернета и интернет-обзорах конструкторов сайтов упоминаются и целый ряд иных разработок. При этом те сайты, которые размещены в домене «.ru», имеют только русскоязычные интерфейсы:

а) Flazio (<https://www.flazio.com>) с рекламным слоганом «Начните сейчас создавать свой сайт с Flazio». При этом для привлечения пользователей используется текст «Как создать сайт за 3 минуты в 3 простых шага». Предусматриваются следующие функциональные направления использования конструируемых сайтов: витрина (примерно соответствует СВ); бизнес; интернет-магазин; система бронирования, в том числе для гостиниц, хостелов и пр.;

б) SimpleSite (https://lnd.simplesite.com/ru_websiteplanet). Предусматривается три «предназначения» создания сайтов: «персональный/блог», «бизнес», «другое» за «три простых шага». При этом указывается, что «мы единственный конструктор сайтов, который полностью поддерживает редактирование Вашего сайта в мобильном режиме – в мобильном браузере, при этом установка специальных приложений не требуется». При этом пользователям предлагаются «примеры красивых шаблонов сайтов». Имеется также служба поддержки пользователей;

в) IM Creator!Free WebSite Builder (<https://www.imcreator.com>) – англоязычный сайт с рекламным слоганом «Built by Creators, for Creators». При этом на стартовой странице на 09.03.2020 было указано, что «17 420 822 sites built». Сайт содержит достаточно большой набор шаблонов для создания сайтов;

г) Flexbe (<https://flexbe.ru>). Рекламуется как конструктор сайтов для бизнеса с рекламным слоганом «Создайте сайт самостоятельно за несколько часов, даже если вы никогда не делали этого раньше». Помимо собственно «сайтов» (<https://flexbe.ru/site/>), поддерживается также создание «лендингов» (<https://flexbe.ru/landing>), «магазинов», «квизов». При этом в Рунете под «лендингом» понимается «страница для сбора контактов посетителей или продажи товара». На сайте предлагается также учебный «Базовый курс по созданию и продвижению сайта»;

д) NetHouse (<https://nethouse.ru/sait-vizitka>) – рекламируется как средство создания сайтов-визиток для лиц не умеющих программировать, причем за «42 минуты» и за «4 шага». На стартовой странице приводятся примеры интерфейсов созданных сайтов, некоторые расценки на работы, перечисляются сервисы, доступные для пользователей и пр.;

е) webs.com (Create a website with our professionally designed templates and easy to use free website builder) [7], URL: <https://www.webs.com/>. Этот сайт имеет только англоязычную версию; предусматриваются возможности платного и бесплатного использования;

ж) uCoz (Слоган: «Конструктор сайтов на Ваш выбор»). Сайт www.ucoz.ru. На стартовой странице предлагается три вида создания сайта: «для себя»; «для бизнеса»; «под ключ» – т.е. заказная разработка, причем цены на нее начинаются (на 09.03.2020) от 3900 руб.

з) TILDA (Сайт <https://tilda.cc/ru/>). Бесплатный конструктор сайтов TILDA. Рекламный слоган «Создайте впечатляющий сайт на Tilda для бизнеса и медиа» и, дополнительно, «450+ профессиональных блоков готовы для вашего сайта и интернет-магазина»;

и) Mobirize (сайт <https://mobirise.com/ru/>) – бесплатное программное обеспечение для создания сайтов. Имеет интерфейс типа «drag and drop» для перетаскивания готовых блоков, что облегчает создание сайтов «пользователями-не программистами». В интернете позиционируется как «... оффлайн приложение для Windows и Mac, с помощью которого можно быстро создавать небольшие вебсайты, лендинги, онлайн-резюме и пр.». При этом в интернете встречаются утверждения, что сайт можно «создать за 4 минуты».

Лендинги (Landing Page – в дословном переводе «посадочная страница») – это, как правило, одностраничные сайты. В основном они применяются для выполнения пользователями определенных действий.

Упомянем также еще несколько конструкторов сайтов: <https://www.bitrix24.ru> – для создания сайтов и лендингов (в отношении ИБ это средство выгодно отличается – оно имеет интегрированный файервол с высоким уровнем надежности); LPmotor (<https://lpmotor.ru/>) – также конструктор сайтов и лендингов (рекламный слоган «Запустите уже готовый сайт за 1 час и получите пер-

вых клиентов»; бесплатный конструктор сайтов for.ru (<https://for.ru/>) – рекламный слоган «Создайте сайт за 10 минут бесплатно».

Математическая модель оптимизации выбора средства конструирования сайтов. Рассмотрим математическую модель оптимального выбора «конструктора сайтов» с позиций разработчика, которому нужно создать единственный новый интернет-сайт.

Примем, что для конкретного разработчика/эксплуатанта, которому необходимо сконструировать и использовать определенный сайт, вектор «желательных» значений параметров «конструктора» представляет собой $\{W_j\}_{j=1\dots J}$. При этом предпоследней « $J-1$ »-й характеристикой является предполагаемая трудоемкость выполнения разработки, а J -й характеристикой – ожидаемый «положительный эффект» за весь предполагаемый период эксплуатации СВ. Таким образом, набор показателей включает как характеристики самих «конструкторов», так и некоторые дополнительные.

Для совокупности из I «конструкторов» матрица показателей (с позиций разработчика) может быть представлена как $[V_{i,j}]_{i=1\dots I, j=1\dots J}$, причем предпоследний столбец является оценками вероятной трудоемкости создания сайта для рассматриваемого «конструктора» и конкретного разработчика, а J -й столбец – оценки положительных эффектов создания/использования сайтов с позиций их эксплуатантов (эти оценки могут отличаться, так как средства разработки имеют различную функциональность). Для остальных столбцов матрицы значения коэффициентов могут быть взяты, например, как экспертные оценки показателей по различным средствам разработки. Таким образом, значки «+» и «-» из таблицы 1 необходимо будет заменить некоторыми количественными характеристиками.

Для решения задачи оптимального выбора «конструктора» необходимо, прежде всего, использовать систему некоторых «барьерных» ограничений, которые могут «отсекать» неприемлемые для пользователя варианты – например, по сложности освоения редакторов в конструкторах сайтов; по уровням требований к владению навыками программирования; по предельной трудоемкости выполнения разработки сайта и пр.

Запишем эту систему ограничений в общем виде, имея в виду, что для некоторых ограничений нижнее значение потенциально может быть нулевым, а для других ограничений, наоборот, верхнее значение «очень большое». Фактически в этих случаях происходит переход от двусторонних ограничений к односторонним.

$$\{V_j^{(\min)} \leq V_{i,j} \leq V_j^{(\max)}\}_{j=1\dots J} \quad (1)$$

Из числа «неотсеченных» вариантов выбора оптимальное решение соответствует максимуму целевой функции

$$\Psi = \left\{ \sum_{j=1}^J \lambda_j (V_{i,j} - W_j) \right\}_{i=1\dots I^*} \quad (2)$$

где $I^* \leq I$ – это количество оставшихся (не отсеженных по формуле (1)) вариантов выбора «конструктора»; $\{\lambda_j\}_{j=1\dots J}$ – система весовых коэффициентов, характеризующих субъективную «важность» для определенного разработчика и конкретного разрабатываемого сайта отклонений от желаемых значений. При этом целесообразно использовать различные величины этих коэффициентов при отклонениях в «плюс» и в «минус» по отношению к «желаемым» значениям

$$\left\{ \lambda_j = \frac{\lambda_j^{(+)}}{\lambda_j^{(-)}}; \text{если } V_{i,j} > W_j \right\}_{j=1\dots J} \quad (3)$$

Таким образом, если значения конкретных параметров для определенного сайта (например, трудоемкости разработки) «лучше», чем желаемые значения, то это будет давать положительный вклад в значение целевой функции, а если «хуже» – то отрицательный. Подчеркнем, что коэффициенты в (3) должны быть согласованы по абсолютным величинам, причем часть коэффициентов являются размерными.

На основе представленной модели может осуществляться «оптимальный» выбор решения. Его рациональность во многом определяется следующими факторами: точностями экспертных оценок параметров «конструкторов»; точностями выбора разработчиками значений коэффициентов $\{\lambda_j\}_{j=1\dots J}$.

Инструментальные средства разработки сайтов. Такие средства позволяют обеспечить большие функциональные возможности для СВ, но требуют значительно более высокой квалификации разработчиков.

При создании сайтов необходимо уделять внимание веб-серверу, при помощи которого будет реализован хостинг сайта. В настоящее время наиболее востребованными являются два: Apache и Nginx. В силу своих архитектурных особенностей они служат для различных целей: Nginx хорошо раздает статический контент, в то время как Apache справляется с этим хуже, но «не проксирует» запросы, а использует встроенные механизмы выполнения скриптов.

Наиболее распространенным языком программирования для создания интернет-сайтов сейчас является PHP. Считается, что на нем «написано» около 80 % сайтов. Для разработки сайтов часто используются веб-фреймворки различных языков программирования. Они реализуют уже готовые паттерны проектирования веб-сайтов и позволяют «переиспользовать» повторяющиеся в сайтах конструкции, такие как: «админка», модуль авторизации, средство отправки писем, протоколы взаимодействия с сервисами и др. Для выполнения разработок на PHP можно использовать любой текстовый редактор, но наиболее популярными являются IDE: VS Code, PhpStorm.

Оценки трудозатрат, связанных с разработкой, внедрением и сопровождением СВ. В случае простейшего одностраничного СВ с минимальными функциональностью и информационным наполнением только на русском языке его создание с использованием конструктора-сайтов (см. выше) может быть выполнено всего за несколько часов одним человеком – при условии, что он владеет навыками использования выбранного «конструктора».

Для более серьезных разработок СВ, насчитывающих 3–5 страниц, типичным является привлечение двух специалистов: веб-дизайнера и веб-программиста. В необходимых случаях может также привлекаться независимый тестировщик или эксперт, в том числе по информационной безопасности.

В типичных случаях цикл собственно разработки и тестирования СВ, особенно при значительном использовании средств программирования, может составлять 6–12 рабочих дней. Этот цикл может быть значительно сокращен, если у разработчиков уже есть апробированное «решение по созданию сайта», которое может быть взято за основу. Такое «решение» создателей СВ может представлять либо их собственную более раннюю разработку, либо заимствованный из интернета сайт, который предполагается скорректировать. В связи с этим отметим, что использованные на сайтах «программные коды» не защищены и в принципе могут быть скорректированы. Однако комментарии они, как правило, не содержат, что может затруднять выполнение их последующих корректировок для повторного использования.

Окончательная сдача СВ внешним заказчиком может затягиваться из-за предъявления ими дополнительных требований, которые могут заставить серьезно менять уже выполненную разработку.

Внедрение разработанного СВ может включать в себя следующие группы операций: а) получение (регистрацию) доменного имени второго уровня; б) обеспечение хостинга сайта – обычно на серверах специализированных организаций, гарантирующих круглосуточный доступ за счет использования серверов-зеркал; в) обучение владельцев сайтов простейшим операциям по корректировке информации – если такие корректировки предусматриваются.

Точность априорной оценки трудозатрат на создание сайтов с небольшим количеством страниц возрастает с накоплением разработчиками опыта. Специально отметим, что «1С-битрикс» удобен для оценки трудозатрат – он стабилен, документирован, имеет систему сертификации разработчиков.

Сопровождение внедренного СВ может включать следующие действия: а) ответы на вопросы (консультации) владельцев сайтов; б) внесение изменений на СВ по просьбам их владельцев – обычно за некоторую дополнительную плату; в) продление сроков действия доменных имен – по просьбам владельцев сайтов, которые не хотят этим заниматься самостоятельно; г) периодический аудит информационной безопасности СВ – с учетом возникновения новых видов угроз. Для этой цели может использоваться, в частности, инструмент, размещенный на www.netoscope.ru; д) независимый контроль «видимости» СВ поисковыми системами интернета; е) анализ журналов доступа к СВ для выявления контингентов посетителей, особенностей их действий на сайте, попыток взлома СВ и пр.

Прекращение эксплуатации СВ может происходить по следующим причинам: а) вместо СВ был разработан и внедрен полнофункциональный сайт или СВ с большей функциональностью; б) организация прекратила свою деятельность по той или иной причине; в) организация сочла целесообразным вместо собственного СВ использовать страницу сайта другой организации (или некоторого объединения организаций). При этом на СВ может быть размещена гиперссылка для перехода по новому адресу (для переадресации), причем такая гиперссылка может обрабатываться автоматически при входе пользователей на СВ; г) организация вместо СВ начала использовать страницы в социальных сетях.

Выводы. 1. Развитие интернета сопровождается ростом количества зарегистрированных доменных имен 2-го уровня; сайтов, использующих эти имена. Это относится и к российским доменным зонам «.ru» и «.рф», но не к «.su». 2. При решении вопросов о создании и использовании СВ важное значение имеют правильные оценки необходимой оперативности их разработки; требуемой функциональности и информационного наполнения; предполагаемой продолжительности

разработки; вероятной посещаемости интернет-пользователями, включая мобильных. 3. При выборе доменного имени 2-го уровня для регистрации сайта целесообразно отражать не только название организации, но и регион ее размещения. Поэтому в настоящие время в названиях сайтов систематически встречаются «цифровые коды» регионов. 4. Создание СВ может осуществляться как целиком на основе использования «конструкторов сайтов», так и с применением «программирования вручную». Решение обычно принимается с учетом факторов, перечисленных в пункте «2», а также приемлемых затрат на создание СВ. 5. В общем случае СВ могут использоваться не только для решения «созидательных задач», но и для распространения ложной (фейковой) информации; поддержки использования вредоносных программных средств; применения информационных средств политической борьбы и пр. 5. Расширение международных связей организаций и отдельных ФЛ заставляет обратить серьезное внимание на создание «не русскоязычных» СВ, размещение их вне доменных зон России для обеспечения хотя бы минимального ИнП в зарубежном интернет-пространстве.

Библиографический список

1. Васьковский Е. Ю. Системный анализ вопросов, связанных с востребованностью информации на web-сайтах / Е. Ю. Васьковский, Ю. М. Брумштейн // Прикаспийский журнал: управление и высокие технологии. – 2015. – № 1. – С. 59–74. ([http://hi-tech.asu.edu.ru/files/1\(29\)/59-74.pdf](http://hi-tech.asu.edu.ru/files/1(29)/59-74.pdf))
2. Васьковский Е. Ю. Системный анализ функциональных возможностей счетчиков посещаемости сайтов / Е. Ю. Васьковский, Ю. М. Брумштейн // Прикаспийский журнал: управление и высокие технологии. – 2015. – № 3. – С. 96–113. ([http://hi-tech.asu.edu.ru/files/3\(31\)/96-113.pdf](http://hi-tech.asu.edu.ru/files/3(31)/96-113.pdf))
3. Басова Т. А. Сравнительный анализ конструкторов сайтов, представленных в сети Интернет / Т. А. Басова // Студенческая наука и XXI век. – 2017. – № 15. – С. 426–427.
4. Брумштейн Ю. М. Поиск информации в Интернете: анализ влияющих факторов и моделей поведения пользователей / Ю. М. Брумштейн, Е. Ю. Васьковский, Т. Х. Куаншкалиев // Известия Волгоградского государственного технического университета. Сер. Актуальные проблемы управления, вычислительной техники и информатики в технических системах. – 2017. – № 1 (196). – С. 50–55.
5. Сафарова Э. Сайт-визитка компании как необходимый инструмент привлечения покупателей / Э. Сафарова, В. С. Исмагилова // Символ науки. – 2016. – № 8-1 (20). – С. 125–127.
6. Борсук Н. А. Сравнительный анализ средств разработки сайтов-визиток / Н. А. Борсук, Т. Д. Карпова, А. В. Егорова // Символ науки. – 2017. – Т. 2, № 3. – С. 25–28.
7. Дадаев Я. Э. Обзор конструктора сайтов webs.com / Я. Э. Дадаев // Электронный мультидисциплинарный научный журнал с порталом международных научно-практических конференций Интернет-наука. – 2017. – № 10. – С. 10–16.
8. Закутаева К. В. Современные технологии разработки сайта и управления им / К. В. Закутаева, И. В. Смагина // Экономическая среда. – 2019. – № 2 (28). – С. 15–24.
9. Искусственный дизайн-интеллект вместо конструкторов сайтов. – Режим доступа: <https://vc.ru/design/60372-iskusstvennyy-dizayn-intellekt-vmesto-konstruktorov-saytov>, свободный. – Заглавие с экрана. – Яз. рус. (дата обращения: 04.02.2020).
10. Киселева Е. А. Обзор онлайн конструкторов для создания сайтов / Е. А. Киселева // Постулат. – 2019. – № 1-1 (39). – С. 119.
11. Киселева Е. А. Разработка одностраничного сайта по продажам кондитерских изделий с помощью конструктора TILDA / Е. А. Киселева // Постулат. – 2019. – № 1-1 (39). – С. 136.
12. Копычев Ю. С. Выбор вида сайта: корпоративный сайт или сайт-визитка / Ю. С. Копычев, И. Б. Тесленко // Студенческая наука Подмосквю : материалы Международной научной конференции молодых ученых. – 2016. – С. 329–332.
13. Красовский Ю. Н. Персональный сайт как средство оценки ИКТ-компетентности педагога / Ю. Н. Красовский // Методист. – 2012. – № 5. – С. 30–32.
14. Лучшие конструкторы сайтов 2020 Россия. – Режим доступа: <https://www.luchshiekonstruktoiry.ru/>, свободный. – Заглавие с экрана. – Яз. рус. (дата обращения: 23.02.2020).
15. Методология оценки стоимости доменов .RU и .RF на вторичном рынке – шаг к эффективной работе бизнеса с цифровыми активами. – Режим доступа: https://www.cnews.ru/news/line/2020-02-13_metodologiya_otsenki_stoimosti, свободный. – Заглавие с экрана. – Яз. рус. (дата обращения: 23.02.2020).
16. Мировая статистика доменных имен. – Режим доступа: <https://timeweb.com/ru/community/articles/mirovaya-statistika-domennyh-zon>, свободный. – Заглавие с экрана. – Яз. рус. (дата обращения: 23.02.2020).
17. Новикова Н. Е. Особенности продвижения сайта-визитки с помощью инструментов CRO / Н. Е. Новикова, Д. А. Шуклин // Альманах научных работ молодых ученых университета ИТМО : в 5 т. – 2016. – С. 52–55.
18. Новиков М. Ю. Веб-сайт педагога как инструмент профессионального развития / М. Ю. Новиков // Международный научно-исследовательский журнал. – 2017. – № 4-3 (58). – С. 42–44.
19. Обзор и отзывы о конструкторе сайтов SITE123. – Режим доступа: <https://uguide.ru/konstruktor-sajtov-site123-obzor-otzyvy-primery-sajtov>, свободный. – Заглавие с экрана. – Яз. рус. (дата обращения: 04.02.2020).

20. Самолетов Р. В. Конструктор интернет-магазина – верный шаг к цифровой экономике / Р. В. Самолетов, А. А. Шурупов // *Фундаментальные и прикладные исследования кооперативного сектора экономики*. – 2019. – № 2. – С. 124–132.
21. Санникова Н. А. Рекомендации по выбору онлайн конструктора для создания интерактивных сайтов / Н. А. Санникова, М. Н. Оленькова // *Международный студенческий научный вестник*. – 2015. – № 3-2. – С. 246–247.
22. Суслова С. А. Создание интернет-ресурса педагога в конструкторе wix.com / С. А. Суслова, Л. С. Белан // *Лучшая научно-исследовательская работа 2017 : сборник статей победителей VII Международного научно-практического конкурса*. – 2017. – С. 67–70.
23. Терентьев И. А. Сравнительный анализ популярных русскоязычных конструкторов сайтов / И. А. Терентьев, А. Е. Иванова // *Вестник МГУП имени Ивана Федорова*. – 2012. – № 9. – С. 166–171.
24. Феськова М. И. Методы создания и администрирования авторских сайтов / М. И. Феськова, Д. О. Неговора // *Форум молодых ученых*. – 2019. – № 4 (32). – С. 1076–1081.
25. Хабибуллин Р. Р. Экономическая сущность сайта-визитки / Р. Р. Хабибуллин // *От синергии знаний к синергии бизнеса : сборник статей и тезисов докладов V Международной научно-практической конференции студентов, магистрантов и преподавателей*. – 2018. – С. 783–787.
26. Шайдуров А. А. WIX в качестве выбора конструктора для создания сайта / А. А. Шайдуров, А. В. Ерьс // *Постулат*. – 2019. – № 3 (41). – С. 87.

References

1. Vaskovskiy Ye. Yu., Brumshteyn Yu. M. Sistemnyy analiz voprosov, svyazannykh s vostrebovannostyu informatsii na web-saytakh [System analysis of issues related to the demand for information on websites]. *Prikaspiyskiy zhurnal: upravlenie i vysokie tekhnologii* [Caspian Journal: Control and High Technologies], 2015, no. 1, pp. 59–74. ([http://hi-tech.asu.edu.ru/files/1\(29\)/59-74.pdf](http://hi-tech.asu.edu.ru/files/1(29)/59-74.pdf))
2. Vaskovskiy Ye. Yu., Brumshteyn Yu. M. Sistemnyy analiz funktsionalnykh vozmozhnostey schetchikov poseshchaemosti saytov [System Analysis of Functional Capabilities of Site Attendance Meters]. *Prikaspiyskiy zhurnal: upravlenie i vysokie tekhnologii* [Caspian Journal: Control and High Technologies], 2015, no. 3, pp. 96–113. ([http://hi-tech.asu.edu.ru/files/3\(31\)/96-113.pdf](http://hi-tech.asu.edu.ru/files/3(31)/96-113.pdf))
3. Basova T. A. Sravnitelnyy analiz konstruktorov saytov, predstavlennykh v seti Internet [Comparative analysis of designers of sites presented on the Internet]. *Studencheskaya nauka i XXI vek* [Student science and the XXI century], 2017, no. 15, pp. 426–427.
4. Brumshteyn Yu. M., Vaskovskiy Ye. Yu., Kuanshaliev T. Kh. Poisk informatsii v Internete: analiz vliyayushchikh faktorov i modeley povedeniya polzovateley [Search for information on the Internet: analysis of influencing factors and patterns of behavior of users]. *Izvestiya Volgogradskogo gosudarstvennogo tekhnicheskogo universiteta. Seriya. Aktualnye problemy upravleniya, vychislitelnoy tekhniki i informatiki v tekhnicheskikh sistemakh* [News of the Volgograd State Technical University. Series. Current problems of management, computer engineering and informatics in technical systems], 2017, no. 1 (196), pp. 50–55.
5. Safarova E., Ismagilova V. S. Sayt-vizitka kompanii kak neobkhodimyy instrument privlecheniya pokupateley [Site-business card of the company as a necessary tool to attract buyers]. *Simvol nauki* [Symbol of Science], 2016, no. 8-1 (20), pp. 125–127.
6. Borsuk N. A., Karpova T. D., Yegorova A. V. Sravnitelnyy analiz sredstv razrabotki saytov-vizitok [Comparative analysis of means of development of business card sites]. *Simvol nauki* [Symbol of Science], 2017, vol. 2, no. 3, pp. 25–28.
7. Dadaev Ya. E. Obzor konstruktora saytov webs.com [Review of website designer webs.com]. *Elektronnyy multidistsiplinarynyy nauchnyy zhurnal s portalom mezhdunarodnykh nauchno-prakticheskikh konferentsiy Internetnauka* [Electronic multidisciplinary scientific journal with portal of international scientific and practical conferences of the Internet], 2017, no. 10, pp. 10–16.
8. Zakutaeva K. V., Smagina I. V. Sovremennyye tekhnologii razrabotki sayta i upravleniya im [Modern technologies of site development and management]. *Ekonomicheskaya sreda* [Economic Environment], 2019, no. 2 (28), pp. 15–24.
9. *Iskusstvennyy dizayn-intellekt vmesto konstruktorov saytov* [Artificial design intelligence instead of site designers]. Available at: <https://vc.ru/design/60372-iskusstvennyy-dizayn-intellekt-vmesto-konstruktorov-saytov> (accessed 04.02.2020).
10. Kiseleva Ye. A. Obzor onlayn konstruktorov dlya sozdaniya saytov [Review of online designers for site creation]. *Postulat* [Postulate], 2019, no. 1-1 (39), p. 119.
11. Kiseleva Ye. A. Razrabotka odnostranichnogo sayta po prodazham konditerskikh izdeliy s pomoshchyu konstruktora TILDA [Development of a one-page confectionery sales site with the help of designer TILDA]. *Postulat* [Postulate], 2019, no. 1-1 (39), p. 136.
12. Kopychev Yu. S., Teslenko I. B. Vybor vida sayta: korporativnyy sayt ili sayt-vizitka [Choice of type of site: corporate site or site-business card]. *Studencheskaya nauka Podmoskovyu: materialy Mezhdunarodnoy nauchnoy konferentsii molodykh uchennykh* [Student science of Moscow region: Proceedings of the International Scientific Conference of Young Scientists], 2016, pp. 329–332.
13. Krasovskiy Yu. N. Personalnyy sayt kak sredstvo otsenki ikt-kompetentnosti pedagoga [Personal site as a means of assessing the ICT competence of the teacher]. *Metodist* [Methodist], 2012, № 5, pp. 30–32.

14. *Luchshie konstruktory saytov 2020 Rossiya* [Best site designers 2020 Russia]. Available at: <https://www.luchshiekonstruktory.ru/> (accessed 23.02.2020).
15. *Metodologiya otsenki stoimosti domenov .RU i .RF na vtorichnom rynke – shag k effektivnoy rabote biznesa s tsifrovymi aktivami* [The methodology for estimating the value of .RU and .RF domains in the secondary market is a step towards effective business operation with digital assets]. Available at: https://www.cnews.ru/news/line/2020-02-13_metodologiya_otsenki_stoimosti (accessed 23.02.2020).
16. *Mirovaya statistika domennykh imen* [World domain name statistics]. Available at: <https://timeweb.com/ru/community/articles/mirovaya-statistika-domennykh-zon> (accessed 23.02.2020).
17. Novikova N. Ye., Shuklin D. A. Osobennosti prodvizheniya sayta-vizitki s pomoshchyu instrumentov CRO [Peculiarities of promotion of business card site with the help of CRO tools]. *Almanakh nauchnykh rabot molodykh uchenykh universiteta ITMO* [Almanac of scientific works of young scientists of ITMO University], in 5 vol., 2016, pp. 52–55.
18. Novikov M. Yu. Veb-sayt pedagoga kak instrument professionalnogo razvitiya [Teacher Website as a Professional Development Tool]. *Mezhdunarodnyy nauchno-issledovatel'skiy zhurnal* [International Research Journal], 2017, no. 4-3 (58), pp. 42–44.
19. *Obzor i otzvyvy o konstruktore saytov SITE123* [Reviews and responses about site designer SITE123]. Available at: <https://uguide.ru/konstruktor-sajtov-site123-obzor-otzvyvy-primery-sajtov> (accessed 04.02.2020).
20. Samoletov R. V., Shurupov A. A. Konstruktor internet-magazina – vernyy shag k tsifrovoy ekonomike [Designer of the online store – the right step to the digital economy]. *Fundamentalnye i prikladnye issledovaniya kooperativnogo sektora ekonomiki* [Fundamental and Applied Research of the Cooperative Sector of the Economy], 2019, no. 2, pp. 124–132.
21. Sannikova N. A., Olenkova M. N. Rekomendatsii po vyboru onlayn konstruktora dlya sozdaniya interaktivnykh saytov [Recommendations on selection of online designer for creation of interactive sites]. *Mezhdunarodnyy studentcheskiy nauchnyy vestnik* [International Student Scientific Journal], 2015, no. 3-2, pp. 246–247.
22. Suslova S. A., Belan L. S. Sozdanie internet-resursa pedagoga v konstruktore wix.com [Creating an Internet resource of a teacher in the designer wix.com]. *Luchshaya nauchno-issledovatel'skaya rabota 2017: Sbornik statey pobediteley VII Mezhdunarodnogo nauchno-prakticheskogo konkursa* [Best research work 2017: Proceedings of the Winners of the VII International Scientific and Practical Competition], 2017, pp. 67–70.
23. Terentev I. A., Ivanova A. Ye. Sravnitelnyy analiz populyarnykh russkoyazychnykh konstruktorov saytov [Comparative analysis of popular Russian-speaking site designers]. *Vestnik Moskovskogo gosudarstvennogo universiteta pechati imeni Ivana Fedorova* [Bulletin of Moscow State University of Printing Arts named by Ivan Fedorov], 2012, no. 9, pp. 166–171.
24. Feskova M. I., Negovora D. O. Metody sozdaniya i administrirovaniya avtorskikh saytov [Methods of creation and administration of author's sites]. *Forum molodykh uchenykh* [Forum of Young Scientists], 2019, no. 4 (32), pp. 1076–1081.
25. Khabibullin R. R. Ekonomicheskaya sushchnost sayta-vizitki [Economic Essence of the Business Card Site]. *Ot sinergii znaniy k sinergii biznesa: sbornik statey i tezisov dokladov V Mezhdunarodnoy nauchno-prakticheskoy konferentsii studentov, magistrantov i prepodavateley* [From Synergy of Knowledge to Synergy of Business: The Collection of Articles and Theses of Reports of the V International Scientific and Practical Conference of Students, Masters and Teachers], 2018, pp. 783–787.
26. Shaydurov A. A., Yeres A. V. WIX v kachestve vybora konstruktora dlya sozdaniya sayta [WIX as a constructor choice to create the site]. *Postulat* [Postulate], 2019, no. 3 (41), p. 87.

УДК: [004.5+004.4+ 51-7]:331.5

**НЕЛОЯЛЬНОСТЬ ПЕРСОНАЛА ОРГАНИЗАЦИЙ В УСЛОВИЯХ РАЗВИТИЯ
ИНФОРМАЦИОННО-ТЕЛЕКОММУНИКАЦИОННЫХ ТЕХНОЛОГИЙ:
АНАЛИЗ ВОЗМОЖНЫХ ИСТОЧНИКОВ ИНФОРМАЦИИ,
ПОДХОДОВ К ВЫЯВЛЕНИЮ МОДЕЛЕЙ И МЕТОДОВ УПРАВЛЕНИЯ**

Статья получена редакцией 15.01.2020, в окончательном варианте – 09.03.2020.

Иглина Наталия Анатольевна, Астраханский государственный технический университет, 414025, Российская Федерация, г. Астрахань, ул. Татищева, 16, кандидат экономических наук, e-mail: eaglenat@mail.ru; ORCID <https://orcid.org/0000-0003-3176-2529>, РИНЦ: AuthorID: 555026

Брумштейн Юрий Моисеевич, Астраханский государственный университет, 414056, Российская Федерация, г. Астрахань, Татищева, 20а, кандидат технических наук, доцент, e-mail: brum2003@mail.ru; ORCID <http://orcid.org/0000-0002-0016-7295>, РИНЦ: AuthorID 280533

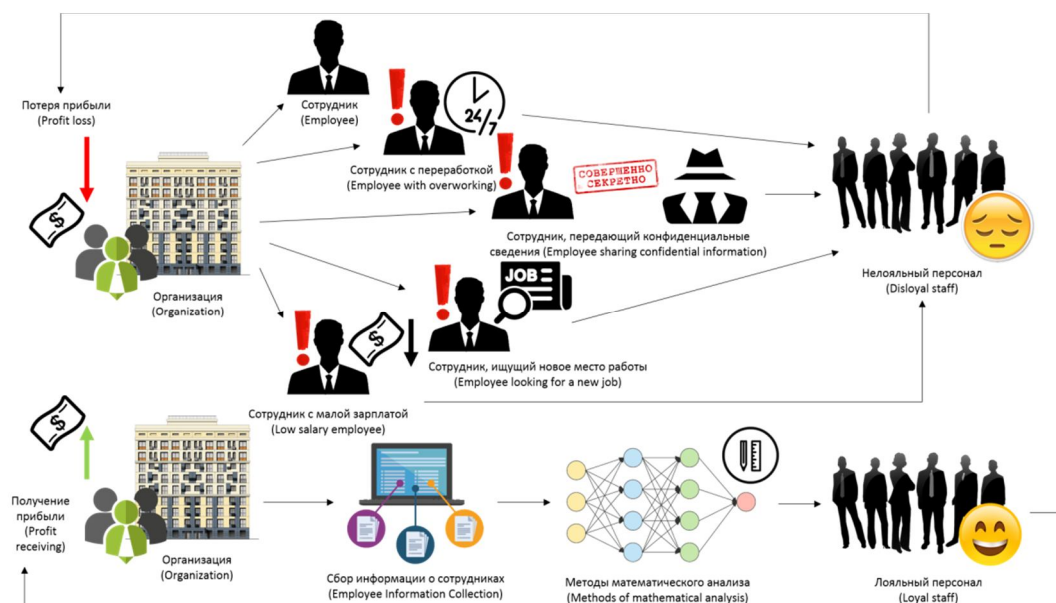
Врублевская Кира Александровна, Астраханский государственный университет, 414056, Российская Федерация, г. Астрахань, Татищева, 20а, магистрант, e-mail: vrublevskayakira@gmail.com, ORCID <https://orcid.org/0000-0002-4725-0387>, РИНЦ AuthorID: 1046895

Коновалова Дарья Игоревна, Национальный исследовательский технологический университет «МИСиС», 119049, Российская Федерация, г. Москва, Ленинский проспект, 4, магистрант, e-mail: haizbq@gmail.com, ORCID 0000-0002-8348-0171

Обоснована целесообразность использования понятия «нелояльность персонала» (НП). Показаны преимущества применения термина НП для выполнения текущих оценок и прогнозов «нелояльности» отдельных сотрудников, их групп, подразделений организаций. Указаны направления использования информационно-телекоммуникационных технологий (ИТКТ) в работе организаций, в частности, в обеспечении информационных и коммуникационных потребностей физических лиц (ФЛ) в процессе их деятельности. Показано, что развитие ИТКТ значительно повлияло на следующее: условия поиска организациями кандидатов на вакантные должности, а ФЛ – мест работы; методы оценки профессиональной пригодности кандидатов и работающего персонала; способы обеспечения его лояльности в условиях высоких информационных, коммуникационных, а иногда и эмоционально-психологических нагрузок, конкуренции между организациями; на структуру угроз информационной безопасности для организаций, в том числе связанных с использованием ИТКТ персоналом. Представлена комплексная математическая модель «затрат на персонал», позволяющая обосновать рациональность расходов на выявление фактов НП, на предотвращение и/или снижение вероятности случаев проявления НП; на уменьшение ущерба при проявлениях НП – в том числе с учетом рыночных условий деятельности организаций. Показано, что НП может быть «направлена», на организации в целом, на руководителей организаций, на отдельные структурные подразделения и на их руководителей, на группы сотрудников, на технологии организации работы. Рассмотрены различия между индивидуальной и групповой НП; «пассивной» и «активной» формами проявления НП. Выполнен системный анализ причин НП, видов НП, источников информации о проявлениях НП, следствий проявлений НП для организаций. Проанализированы возможности использования работодателями ИТКТ для следующих целей: оценки и прогноза НП при приеме сотрудников на работу; выявления случаев НП работающих сотрудников; управления уровнями НП с учетом имеющихся ограничений по ресурсам, факторов неопределенности и риска; снижения вероятности НП при его увольнении из организации. Показана целесообразность для организаций применения комплексной системы мониторинга и управления НП.

Ключевые слова: информационно-телекоммуникационные технологии, математические модели, нелояльность персонала, направленность нелояльности, формы нелояльности, причины нелояльности, виды нелояльности, следствия нелояльности, выявление нелояльности, оценки нелояльности, управление нелояльностью

Графическая аннотация (Graphical annotation)



**NON-LIABILITY OF ORGANIZATION PERSONNEL IN THE CONDITIONS
OF INFORMATION-TELECOMMUNICATION TECHNOLOGIES DEVELOPMENT:
ANALYSIS OF POSSIBLE INFORMATION SOURCES,
IDENTIFICATION APPROACHES, MODELS AND MANAGEMENT METHODS**

The article was received by the editorial board on 15.01.2020, in the final version – 09.03.2020.

Iglina Natalia A., Astrakhan State Technical University, 16 Tatishchev St., Astrakhan, 414025, Russian Federation,

Cand. Sci. (Economics), e-mail: eaglenat@mail.ru; ORCID <https://orcid.org/0000-0003-3176-2529>, RSCI: AuthorID: 555026

Brumshteyn Yuri M., Astrakhan State University, 20a Tatishchev St., Astrakhan, 414056, Russian Federation,

Cand. Sci. (Engineering), Associate Professor, ORCID <http://orcid.org/0000-0002-0016-7295>; https://elibrary.ru/author_profile.asp?authorid=280533; e-mail: brum2003@mail.ru

Vrublevskaya Kira A., Astrakhan State University, 20a Tatishchev St., Astrakhan, 414056, Russian Federation,

master student, e-mail: vrublevskayakira@gmail.com, ORCID <https://orcid.org/0000-0002-4725-0387>, RSCI AuthorID: 1046895

Konovalova Daria I., National Research Technological University «MISiS», 4 Leninskiy Prospect, Moscow, 119049, Russian Federation,

master student, ORCID <https://orcid.org/0000-0002-8348-0171>, e-mail: haizbq@gmail.com

The expediency of applying the concept of «person disloyalty» (PD) is substantiated. The advantages of using in order to perform theoretical assessments and forecasts of «disloyalty» of individual employees, their groups, organizational units are substantiated. The directions of the information-telecommunication technologies (ITCT) application in the work of organizations, in particular, in ensuring the information needs of individuals in the process of their activities, are indicated. It is shown that the development of ITCT significantly influenced the following: the conditions for organizations to search for candidates for vacant posts, and for individuals for job places; methods for assessing the professional suitability of candidates and staff; ways to ensure its loyalty in conditions of high information, communication, and sometimes emotional and psychological stress, competition between organizations; the structure of information security threats to organizations, including associated with the use of ITKT staff. A comprehensive mathematical model of «personnel costs» is presented, it allows justifying the rationality of expenses for identifying facts of PD, preventing or reducing the likelihood of occurrences of PD; reducing damage during manifestations of PD - including taking into account the market conditions for the activities of organizations. It is shown that the PD can be «targeted» to the organization as a whole, to the heads of organizations, to individual structural divisions and to their leaders, to groups of employees, to technology for organizing work. The differences between individual and group PD; «passive» and «active» forms of manifestation of PD are considered. A systematic analysis of

the PD causes, types, sources of information about the manifestations, and the consequences of manifestations for organizations has been performed. The possibilities of employers using ИТКТ for the following purposes: assessment and forecast of PD when hiring employees; identify cases of PD working employees; management of PD levels taking into account existing resource limitations, factors of uncertainty and risk; exclusion of PD when it is dismissed from the organization are analyzed. It is shown that it is expedient for organizations to use an integrated monitoring and control system for PD.

Key words: information and telecommunication technologies, mathematical models, personnel disloyalty, orientation of disloyalty, forms of disloyalty, causes of disloyalty, types of disloyalty, consequences of disloyalty, identification of disloyalty, disloyalty assessment, disloyalty management

Введение. Развитие экономики, основанной на знаниях, усилило зависимость эффективности работы организаций различных типов от компетентности персонала, его мотивации к труду, сохранения его лояльности организации-работодателю (ОР). Вопросам оценки и обеспечения «лояльности персонала» посвящено большое количество работ, например, [2, 3, 7, 9, 12, 27, 28, 29, 31]. Однако в теоретическом и практическом плане важно решение и класса «смежных» задач, связанных с выявлением фактов «нелояльности» персонала (НП) – в том числе с использованием информационно-телекоммуникационных технологий (ИТКТ); анализом объективных и субъективных причин, приводящих к НП; предотвращением или уменьшением вероятности фактов НП; снижением возможных ущербов от НП. Этот класс задач имеет определенные особенности, недостаточно отраженные в существующей литературе. Поэтому целью настоящей работы является попытка устранения указанного недостатка.

Понятие «нелояльность персонала» – анализ особенностей содержания. В настоящее время лояльность персонала изучается с различных точек зрения: психологической, социологической, экономической, юридической и др. [16, 17, 18]. При этом «лояльность» рассматривается как характеристика индивидуума (только!), определяющая его поведение по отношению к ОР в целом.

Типичная трактовка понятия «лояльность персонала» в литературе [23] – *«совокупность качеств работника, обусловленных его личными особенностями (внутренние факторы) и параметрами организации (внешние факторы)», характеризующих отношение работника к организации индикатируемых через поведенческие признаки, как в текущем периоде, так и в долгосрочной перспективе*. Далее в [23] со ссылками на различные литературные источники приводятся «10 трактовок ключевых признаков лояльного поведения сотрудников», причем шесть из них прямо соответствуют отсутствию признаков проявления НП и еще как минимум два – «косвенным» проявлениям НП. Подход к характеристике лояльности персонала, основанный на учете «проявления» признаков его нелояльности является типичным для литературы по менеджменту, управлению персоналом и пр. [2, 5, 16, 18, 23, 27, 30]. При этом детального анализа признаков НП, как правило, не проводится, так как это не является темой исследований. Поэтому в данной статье мы сосредоточимся на анализе НП и начнем рассмотрение с вопросов ее классификации.

(1) Классификация по направленности НП. (н1) НП по отношению к ОР в целом – такая трактовка является традиционной практически для всех опубликованных работ. (н2) Нелояльность по отношению к «бренду» или репутации холдинга (или иного вида объединения), в который входит ОР. (н3) Нелояльность по отношению к определенному руководителю ОР или к высшему руководству ОР в целом. (н4) Нелояльность по отношению к подразделению организации, в котором работает сотрудник. Причины: структурное подразделение ОР может обладать собственным «психологическим микроклиматом», сложными взаимоотношениями между сотрудниками внутри него, конфликтами между ними. (н5) Нелояльность по отношению к руководителю подразделения ОР, в котором сотрудник работает. Такой вид нелояльности может быть вызван экономическими причинами, психологическими и пр. Устранение нелояльности по пунктам «н4», «н5» возможно, в частности, за счет перераспределения сотрудников между подразделениями ОР [10]. (н6) Нелояльность по отношению к другим сотрудникам ОР, не соответствующим пунктам «н3» и «н5». (н7) Нелояльность по отношению к используемым в ОР методам организационного и/или экономического управления [6, 9], технологиям организации работ, принципам планирования карьерного роста сотрудников [33] и пр. (н8) Нелояльность сотрудников по отношению к «контрагентам» ОР, включая ее клиентов [3, 21], органы государственного контроля и управления. (н9) Нелояльность сотрудников по отношению к акционерам ОР (для акционерных обществ).

Предложенная классификация позволяет отдельно анализировать причины возможных фактов НП с учетом особенностей отдельных сотрудников; виды их действий, соответствующих НП; последствия этих действий и пр.

(2) Классификация по «внешним условиям» деятельности ОР, в которых имеют место проявления НП. (а) «Нормальные» условия деятельности ОР. (б) Предкризисные условия деятельности, несущие потенциальную угрозу интересам ОР и/или отдельным сотрудникам. (в) Кризисные условия деятель-

ности ОР [20], в том числе связанные с экономическими процессами в стране, регионе, населенном пункте, отрасли; возросшей конкуренцией и пр.

(3) Классификация по характеру действий сотрудников, относимых к НП. (д1) Активные (инициативные) действия отдельных работников или их групп, прямо направленные против интересов ОР; преследующие цель нанести вред ОР в целом или их отдельным подразделениям. (д2) Действия сотрудников, преследующие цель получения «личной выгоды» (например, за счет использования «служебной информации») и при этом наносящие косвенный вред ОР. (д3) Отсутствие необходимых действий сотрудников, предусмотренных соответствующими должностными инструкциями ОР, регламентами работы подразделений и пр. – это можно назвать «умышленным бездействием». (д4) Отсутствие проявления сотрудниками необходимой личной инициативы по выработке/реализации действий в условиях изменяющихся ситуаций. Обычно это не является прямым нарушением должностных обязанностей сотрудников, но часто может рассматриваться как проявление ими НП. (д5) Формальное выполнение сотрудниками служебных обязанностей, но с ненадлежащим качеством и/или с нарушением контрольных сроков. Частным случаем этого направления НП могут быть нарушения правил информационной безопасности (ИБ) организаций, в том числе и вузов [4]. Принято считать, что основными категориями «похищаемой» информации являются следующие. (к1) Персональные данные, в т.ч. сотрудников и клиентов ОР. В отношении персональных данных в России действует специальный Закон [22]. Его нарушение может приводить не только к штрафам для юридических лиц (ст. 13.11 КоАП РФ [19]), но в некоторых случаях и к отзыву у ОР лицензий на осуществление определенных видов деятельности. (к2) «Финансово-платежная» информация: сведения о перечислении ОР платежей или поступления средств на их счета; данные о выплатах сотрудникам денежных средств; сведения о предложениях ОР при участии в тендерах и пр. Утечки такой информации могут наносить прямой финансовый ущерб ОР [15]. (к3) Коммерческие тайны и служебная информация ОР: их планы развития в конкурентной среде; назначения и изменения цен на товары и услуги; технологические и конструкторские разработки, не защищенные патентами; ноу-хау и т.д. [15].

(д6) Попытки «блокирования» сотрудниками исполнения тех решений руководства ОР и/или подразделений, которые могут наносить вред их личным интересам. На практике это может делаться со ссылками на объективные причины, на некоторые инструкции, на объективные обстоятельства и пр. (д7) Выполнение сотрудниками в «не служебное» время работ в интересах других организаций, являющихся конкурентами ОР по тем или иным направлениям деятельности.

Конкретные виды действий сотрудников, которые могут рассматриваться как проявления НП, значительно зависят от типа организации; должностей, занимаемых в них сотрудниками; их личных целей деятельности и пр.

(4) Классификация НП по «массовости» действий или бездействия сотрудников. (м1) Индивидуальные действия (или бездействие) отдельных сотрудников. (м2) Действия/бездействие неформальных групп сотрудников, в том числе прямо согласованные ими друг с другом. (м3) Действия/бездействие сотрудников подразделений, направленные на обеспечение интересов этих подразделений в ущерб интересам ОР в целом и/или руководителей ОР.

(5) Классификация по «частоте» действий, которые можно отнести к НП. (ч1) Разовые (однократные) действия, в том числе и под влиянием особых ситуаций. (ч2) Неоднократные проявления сотрудниками нелояльности. (ч3) Систематические, в том числе периодические проявления НП. (ч4) Постоянное проявление сотрудниками нелояльности.

Характеристика направлений влияния ИТКТ на деятельность организаций, связанную с подбором персонала. Развитие ИТКТ значительно повлияло на возможности ОР по подбору персонала, а также на условия конкуренции организаций за персонал. Перечислим основные направления. (1) ОР широко используют собственные сайты для размещения объявлений об имеющихся вакансиях; о требованиях к кандидатам; в ряде случаев – для проведения «предварительных тестирований» кандидатов. (2) Также ОР могут использовать для поиска кандидатов специализированные сайты, содержащие объявления организаций о вакансиях и резюме тех ФЛ, которые ищут работу. В отношении ФЛ может определяться круг потенциально подходящих кандидатов для последующих контактов с ними. В отношении выбора ФЛ мест работы важна информация о предлагаемых ОР на сайтах условиях труда, включая уровни оплаты. (3) При личных собеседованиях с кандидатами ОР могут применять специальные тестовые программы для выяснения уровней профессиональной пригодности ФЛ. Это уменьшает вероятность принятия неверных (или не оптимальных) кадровых решений, в том числе в отношении распределения принимаемых лиц между подразделениями ОР. (4) Кроме того, ИТКТ могут использоваться кадровыми службами организаций для сбора информации о кандидатах из различных источников – прежде всего, с интернет-сайтов. Это особенно касается лиц, занимающихся научной деятельностью – информация об их публикациях, наукометриче-

ских показателях публикационной активности и пр. доступна в различных источниках. Информация о кандидатах может собираться и через социальные сети, в том числе специализированные. (5) Переписка с представляющими интерес для ОР кандидатами по электронной почте, в ряде случаев – использование систем мгновенного обмена сообщениями, средств видеоконференцсвязи. (6) Иногда ОР могут использовать средства электронных коммуникаций для организации взаимодействия и с так называемыми «рекрутинговыми агентствами» (РА), однако в России они не получили большого распространения. Для темы статьи существенно, что РА зачастую занимаются трудоустройством не безработных, а уже работающих ФЛ, которые хотят улучшить условия своего труда, включая уровень оплаты. Таким образом, деятельность РА может способствовать повышению уровней НП за счет переходов их в другие организации, а также ведения переговоров о таких переходах.

В свою очередь, ФЛ могут использовать ИТКТ при поиске мест работы по следующим направлениям. (1) Анализ объявлений о вакансиях на сайтах потенциально интересных для них организаций, в том числе в отношении уровней оплаты, возможностей дистанционной работы и пр. (2) Изучение объявлений о вакансиях, которые ОР дают на специализированных сайтах. (3) Коммуникации с потенциальными ОР в электронной форме: использование электронной почты (в том числе и для пересылки копий документов), видеоконференцсвязи, систем мгновенного обмена сообщениями, SMS и пр.

По умолчанию в русскоязычных публикациях по «мобильности трудовых ресурсов» предполагается переход сотрудников в другие российские организации. Возможен переезд персонала и за пределы России, однако эти процессы сдерживаются рядом факторов. (а) Различиями в «содержании подготовки» специалистов в вузах России и за рубежом; нормативно-правовой базы хозяйственной деятельности и пр. Как следствие, работу вне России значительно легче найти, например, специалистам по ИТКТ-направлениям, чем юристам и экономистам. (б) «Языковыми» барьерами. Они наиболее существенны в случае «восточных» стран, чем «западных», в которых широко используется английский язык, активно изучаемый в российских вузах. При этом автоматические переводчики «с голоса в голос» (автономные устройства или использующие интернет-ресурсы) обычно полезны только для «бытового общения», но не профессионального.

Влияние развития ИТКТ на характер труда сотрудников, возможные варианты проявления нелояльности персонала. В условиях развития ИТКТ существенно изменился характер и содержание труда сотрудников многих организаций; способы индивидуального и группового планирования их деятельности; обеспечения эффективности взаимодействия персонала в процессе работы; цели, принципы и направления контроля деятельности сотрудников со стороны ОР, включая предотвращение и выявление фактов НП. Основные причины этого следующие. (1) Рост объемов работы с различными видами информации, особенно при необходимости ее «глубокой переработки», увеличивает «умственную» нагрузку на сотрудников, чаще всего повышает требования к объемам их необходимых знаний и умений. (2) Интенсивное использование работниками ИТКТ и, как следствие, увеличение возможностей утечек конфиденциальной служебной информации в результате неумышленных или преднамеренных действий персонала организаций [14]. Такие действия, как уже было сказано, могут осуществляться с целью нанесения вреда ОР, получения личной выгоды и пр. Отметим, что в настоящее время реализовать такие утечки можно с использованием персональных ЭВМ – служебных и личных; за счет применения смартфонов с фотокамерами и пр. При этом «торговля» коммерчески значимой служебной информацией может осуществляться анонимно – в том числе на основе использования «анонимных сетей», в которых покупатели и продавцы остаются неизвестными друг другу [1]. (3) Широкое использование некоторыми организациями «мобильных сотрудников», в отношении которых может быть достаточно сложно контролировать их оперативную деятельность, «контакты» и информационный обмен с различными юридическими и ФЛ, факты проявления (или попыток проявления) нелояльности. (4) Расширение применения ОР труда «дистанционных сотрудников» (ДС), на основе использования возможностей ИТКТ. Такие сотрудники физически могут находиться за сотни и тысячи километров от мест расположения центральных офисов ОР. Возможности использования ДС прямо отражены в статье 312.2 Трудового кодекса РФ [26]. При этом, в соответствии с последним абзацем этой статьи, в трудовом договоре с ДС *«может предусматриваться дополнительное условие об обязанности дистанционного работника использовать при исполнении им своих обязанностей по трудовому договору о дистанционной работе оборудование, программно-технические средства, средства защиты информации и иные средства, предоставленные или рекомендованные работодателем»*. Потенциально это обеспечивает возможности контроля проявлений НП таких сотрудников с использованием инструментальных средств, программного обеспечения, установленного на ПЭВМ этих ФЛ.

Контингент ОР, которые используют проверку собственных сотрудников на «нелояльность» с использованием специальных тестов и/или т.н. «полиграфов» достаточно ограничен – в основном это силовые структуры и правоохранительные органы.

Развитие ИТКТ расширило возможности поиска «более подходящих мест работы» уже работающими в ОР сотрудниками. Такое поведение ФЛ в литературе традиционно рассматривается как проявление НП [27], хотя оно может быть следствием отсутствия возможностей профессионального или карьерного роста внутри ОР [23, 32]. При этом ФЛ могут использовать как пассивные средства поиска мест работы (изучение объявлений организаций о вакансиях), так и активные – размещение в интернете собственных резюме без указания фамилии и с применением специально созданных адресов электронной почты.

Таким образом, потенциально развитие ИТКТ увеличивает для ОР угрозы, связанные с ростом текучести кадров. При этом, в силу традиционно низкой трудовой мобильности персонала в России, переход ФЛ (особенно среднего и пожилого возраста) в другие ОР осуществляется чаще всего в пределах тех населенных пунктов (или регионов), в которых они постоянно проживают.

Математическая модель «затрат на персонал», учитывающая расходы на управление нелояльностью персонала. Для оценки рентабельности затрат ОР на «управление персоналом», включая расходы на предотвращение и выявление фактов НП, примем следующую математическую модель (для простоты – без учета коэффициентов дисконтирования затрат и доходов)

$$R = 100\% \times (P - Z - N_1 - N_2) / Z, \quad (1)$$

где R – рентабельность затрат; P – сумма «положительных эффектов» от деятельности персонала; Z – сумма затрат ОР по направлениям, связанным с организацией и оплатой труда персонала; N_1 – сумма негативных эффектов, связанных с деятельностью персонала в ОР, включая проявления НП, реализации угроз ИБ, связанных с работой сотрудников и пр.; N_2 – сумма негативных эффектов, связанных с уходом сотрудников из ОР, включая их переходы в конкурирующие фирмы.

Для Z примем следующую структуру, которая включает в себя больше компонент, чем приведено, например, в [9]

$$Z = Z_1 + Z_2 + Z_3 + Z_4 + Z_5 + Z_6, \quad (2)$$

где Z_1 – затраты, связанные с выплатами «базовой» зарплаты персоналу; Z_2 – премиальные и иные дополнительные выплаты персоналу ОР; Z_3 – затраты, на повышение квалификации и сертификацию персонала ОР (отметим, что инвестиции ОР в повышение квалификации персонала в настоящее время юридически защищены обычно слабо); Z_4 – затраты, связанные с обеспечением нормальных условий труда персонала: соблюдение требований охраны труда и техники безопасности, эргономики рабочих мест, санитарно-гигиенических условий и т.д.; Z_5 – затраты, на сбор (получение) информации о фактах проявления НП, в том числе с использованием ИТКТ; Z_6 – прямые затраты на предотвращение и/или снижение уровней НП ОР, включая профилактические мероприятия.

Для N_1 примем формулу, соответствующую аддитивной модели

$$N_1 = \sum_{i=1}^{I_1} (W_i^{(1)} U_i^{(1)}), \quad (3)$$

где I_1 – общее количество видов негативных эффектов, относящихся к N_1 ; $W_i^{(1)}$ – вероятность проявления i -го негативного эффекта; $U_i^{(1)}$ – ущерб в случае «реализации» i -го негативного эффекта.

Для N_2 примем аналогичную формулу

$$N_2 = \sum_{i=1}^{I_2} (W_i^{(2)} U_i^{(2)}), \quad (4)$$

где I_2 – общее количество видов негативных эффектов, относящихся к N_2 ; остальные обозначения – аналогичны (3). Все перечисленные выше виды затрат в той или иной степени влияют на величины негативных эффектов от деятельности персонала в ОР или его ухода из этой организации. При этом влияние может оказываться как на $\{W_i^{(1)}\}$ и $\{W_i^{(2)}\}$, так и на $\{U_i^{(1)}\}$ и $\{U_i^{(2)}\}$.

Системный анализ причин нелояльности персонала, видов его нелояльного поведения, источников информации о НП, следствий проявления НП для организаций различных типов. Рассмотрим, прежде всего, причины НП. При этом будем увязывать «причины» с теми «факторами», которые их определяют (вызывают).

(П1) – организационно-экономическая неудовлетворенность: системой материальной и нематериальной мотивации, распределением полномочий и ответственности, особенностями «обратной связи» между результатами труда и условиями работы, корпоративной культурой, уровнем свободы (допустимой инициативой) при принятии и реализации решений, перспективами карьерного роста, возможностями «профессионального развития» сотрудников.

(П2) – неудовлетворенность персонала особенностями труда: его содержанием и характером; престижностью; профессией, должностью.

(П3) – социальная неудовлетворенность: межличностными взаимоотношениями или социально-психологическим климатом в коллективе; взаимоотношениями работника с руководством, подчиненными, коллегами и т.д.

(П4) – психофизиологическая неудовлетворенность: условиями и режимом труда, технической оснащенностью и удобством (эргономикой) рабочего места, уровнем физической и/или психоэмоциональной нагрузки и т.д.

(П5) – неудовлетворенность, обусловленная личностными характеристиками работника ОР: его компетенциями, потребностями, целями и ценностями, типом темперамента, характером, индивидуальным отношением к трудовой деятельности и т.д.

Очевидно, что НП по пунктам «П2», «П5» может быть связана с недочетами при подборе персонала. Частично эти два типа причин можно «компенсировать» перераспределением сотрудников между подразделениями ОР [10]. Причины «П3» и «П4» могут быть устранены (или, частично компенсированы) за счет принятия организационно-экономических [6] или технических мер. При этом приходится учитывать ограниченность ресурсов ОР, которые могут быть использованы для этих целей, особенно – финансово-экономических ресурсов.

В качестве «источников информации» о возможной НП, работающего в ОР, укажем следующие. **(И1)** Сведения от других сотрудников ОР. **(И2)** Внутренняя информация о деятельности персонала, собираемая с использованием систем видеонаблюдения, программно-технических средств, установленных на ПЭВМ сотрудников; на серверах организаций. Это может касаться, в частности, следующего: поведения сотрудников при общении с клиентами; переписки сотрудников по электронной почте, в том числе с использованием корпоративных систем такой почты; доступа (или попыток доступа) сотрудников на различные сайты, длительности работы их на этих сайтах; сведения о голосовых переговорах сотрудников с клиентами ОР [13], а также с организациями-контрагентами ОР и пр. **(И3)** Сведения, которые могут быть получены на общедоступных сайтах интернета в отношении лиц, ищущих работу: www.headhunter.ru, www.careerist.ru, ru.jobappmatch.org, <https://vk.com/headhunter> и др. **(И4)** Фактические результаты деятельности сотрудников ОР (оперативные и долговременные) – они косвенно могут свидетельствовать о проявлениях нелояльности. **(И5)** Сведения об информационной активности сотрудников организаций (или кандидатов для приема на работу), доступные на сайтах различных социальных сетей. **(И6)** Сведения о НП, которые могут быть извлечены из сообщений средств массовой информации, в том числе на основе анализа характера материалов об ОР, публикуемой в этих средствах.

Отметим, что в общем случае «информационную нелояльность» к конкретной ОР могут проявлять не только ее «действующие» сотрудники, но и некоторые иные категории лиц, включая следующие: а) бывшие сотрудники ОР, уволенные по инициативе ОР, в том числе из-за нарушений ими действующих правил ОР, по сокращению штатов, из-за несоответствия образования занимаемой должности, конфликтов с другими ФЛ и пр.; б) лица, которым было отказано в приеме на работу в ОР, в том числе и по результатам работы в пределах «испытательного срока»; в) сотрудники других организаций или иные лица, выполняющие специальные поручения по информационной дискредитации (компрометации) конкретных ОР. Частным случаем этого могут быть т.н. «заказные статьи» в средствах массовой информации, а также «заказные публикации» на интернет-сайтах; негативные отзывы об ОР на интернет-сайтах.

Переходим к представлению видов НП.

(В1) – Поиск сотрудниками других мест работы – традиционно указывается как вид НП в большинстве источников (например, [5]). Он может быть вызван различными причинами, в том числе и отсутствием в ОР продуманной политики «карьерного роста персонала» [33]; потерей ОР конкурентных позиций в рыночных условиях, сокращением реальных доходов сотрудников в связи с процессами инфляции и пр.

(В2) – Поиск сотрудниками дополнительных источников доходов внутри ОР, в том числе без согласия руководителей подразделений, в которых работают эти сотрудники. Отметим, что такое согласие может требоваться не во всех случаях, по крайней мере, формальное. Например, в вузах России широко распространено внутреннее совместительство, в том числе ведение преподавате-

лями учебных курсов, «закрепленных» за другими кафедрами. И это обычно не считается проявлением НП – по крайней мере, если между кафедрами отсутствуют «конфликтные ситуации».

(B3) – Поиск сотрудниками дополнительных источников доходов за пределами ОР, в том числе выполнение работ для конкурирующих организаций. Однако даже если «другая организация» не конкурирующая, то работа сотрудника на нее обычно ведет к неполному использованию его «интеллектуальных ресурсов» по основному месту работы; часто – к использованию «на стороне» интеллектуальных ресурсов, формируемых по месту основной работы.

(B4) – Незаконное обогащение за счет ресурсов организации, в том числе путем передачи (продажи) ее коммерческих и иных секретов; использования служебной информации в личных целях. К данному виду НП можно также отнести и мелкие хищения материальных ценностей из ОР.

(B5) – Использование вне ОР учебно-методических, программных и иных видов разработок, выполненных в ОР различными ФЛ; информационных и информационно-аналитических материалов, накопленных в ОР. Такие действия могут не нарушать действующего законодательства или должностных инструкций, но в конкурентных условиях наносить вред ОР. Подчеркнем, что к «B5» не относится межвузовский обмен преподавателями в рамках академической мобильности – для чтения лекций, проведения мастер-классов и пр.

(B6) – Халатное отношение персонала к собственным должностным обязанностям, в том числе нарушение принятых в организации правил работы;

(B7) – Саботаж распоряжений руководства организации или отдельных подразделений – прежде всего таких распоряжений, выполнение которых может наносить реальный или потенциальный ущерб самому работнику.

(B8) – Несанкционированное распространение сотрудниками служебной информации, прежде всего имеющей коммерческую значимость. Это может делаться как умышленно, так и в силу непонимания соответствующих правил «информационной безопасности» [4].

(B9) – Умышленное распространение «правдивой» негативной информации об ОР, ее контрагентах – в том числе сведений, носящих служебный характер.

(B10) – Умышленное распространение сотрудниками дезинформации в отношении ОР в сочетании с «инсайдерской» правдивой информацией – для повышения доверия к распространяемым сообщениям.

(B11) – Нарушения трудовой дисциплины сотрудниками, в том числе опоздания или преждевременный уход с работы, использование рабочего времени для действий «не служебного характера»: личная переписка по электронной почте, голосовые переговоры личного характера через смартфоны; компьютерные игры; поиск в интернете информации «неслужебного характера».

(B12) – Уничтожение при увольнении сотрудниками служебной информации, размещенной на их ПЭВМ, включая некоторые перспективные наработки.

(B13) – Для вузов специфическими видами НП можно считать следующие. (в13.1) Недостаточная активность преподавателей и сотрудников в отношении научной деятельности, публикационной активности – в том числе в зарубежных журналах. (в13.2) Невысокий учебно-методический уровень проведения аудиторных занятий со студентами, низкая активность преподавателей на таких занятиях. (в13.3) Небольшое количество или несвоевременность выполнения преподавателями учебно-методических разработок, включая учебные пособия. (в13.4) Нежелание сотрудников тратить время и силы на написание и защиту диссертаций. (в13.5) Формальный характер работы со студентами, особенно выпускных курсов, что снижает возможности последних в отношении участия в конкурсах и олимпиадах, подаче заявок на гранты и пр.

Как виды НП, так и их следствия могут значительно различаться в зависимости от особенностей ОР: ее целей деятельности; организационно-правовой формы, направлений экономической активности и пр. Поэтому далее приводятся следствия НП, характерные для большинства типов организаций. Дополнительно для вузов и торговых компаний указываются специфические следствия.

(C1) – Утечка кадров. Наиболее серьезный ущерб ОР может наносить уход в конкурирующие компании высококвалифицированных сотрудников, включая управленческий персонал; программистов, разрабатывающих и сопровождающих сложные программные комплексы и пр. Варианты «негативных эффектов» для ОР от утечки кадров. (а) Снижение конкурентоспособности ОР в результате потери «интеллектуальных потенциалов», связанных с персоналом; возможностей ОР выполнять сложные, наукоемкие работы. (б) Срыв уже начатых работ, в том числе связанных с разработкой программного обеспечения, реализацией различных проектов и пр. (в) Утечка из ОР «служебных секретов» (конфиденциальной информации), «ноу-хау» и пр. – в том числе в результате необходимости «обеспечения» служебной лояльности сотрудников, перешедших в новые ОР, к другому работодателю. (г) Срыв процессов регистрации интеллектуальной собственности и, как следствие, потенциальных возможностей получения ОР доходов от этой собственности. (д) Поте-

ри времени, связанные с вхождением в курс дела сотрудников, принятых ОР на места уволившихся ФЛ. (е) Финансовые и временные затраты, связанные с подготовкой и повышением квалификации вновь принятых на работу сотрудников. Это особенно важно в отношении дорогостоящей подготовки специалистов в области ИТКТ. (ж) Нарушение сложившихся деловых коммуникаций с организациями-контрагентами, которые часто во многом зависят от «личного взаимодействия» между их сотрудниками.

(С2) – Потери, связанные с экономическими преступлениями сотрудников по отношению к ОР. Считается, что они получают наибольшее распространение в условиях экономического спада и высокой «закредитованности» населения, когда сотрудники вынуждены «решать» свои финансовые проблемы. Возможные формы таких преступлений: «продажа» служебной информации конкурентам [11, 13], в том числе «бесконтактная» продажа с использованием «анонимных сетей» в интернете [1]; сговор с клиентами или поставщиками с целью получения т.н. «откатов».

Специфические следствия (варианты) НП для вузов по пункту «С2».

(С2.1) Коррупция среди преподавательского состава в форме получения денежных средств, материальных ценностей (подарков) и пр. Возможные причины. (а) Неполнота или нечеткость нормативной базы, регулирующей правила взаимодействия между студентами и преподавателями. (б) Неполнота применяемых средств контроля за деятельностью сотрудников со стороны вуза в целом, руководителей отдельных подразделений и пр. (в) Недостаточный размер оплаты труда преподавателей. (г) Неэффективность (недостаточность) проведения профилактических мероприятий по борьбе с коррупционными проявлениями среди преподавателей и сотрудников вуза.

(С2.2) Участие преподавателей и административного персонала в различных «незаконных схемах» получения дипломов о высшем образовании лицами, которые фактически обучение в вузах не проходят. Это касается, в частности, т.н. «проведенных дипломов», при которых в вузе (а чаще в его филиале) учится только «зачетка», а сам студент на занятия не появляется. Отметим, что использование таких схем в России считается характерным для коммерческих вузов, но не бюджетных. Формально такие «схемы» могут также рассматриваться как разновидность «коррупционных проявлений».

(С3) – Утечка служебных данных, не относящаяся к «продаже» служебной информации сотрудниками. (а) Разглашение сотрудниками промышленной или коммерческой тайны ОР без целей извлечения доходов. (б) Утечка из ОР персональных данных сотрудников и/или клиентов. Отметим, что накопление и использование значительных объемов таких данных характерно для медицинских учреждений, финансово-кредитных и образовательных организаций.

Специфические следствия (варианты) НП для вузов по пункту «С3».

(С3.1) Несанкционированная передача сотрудниками результатов служебной интеллектуальной деятельности конкурирующим вузам и иным организациям. Это может приводить к исключению или затруднению возможностей научных публикаций, подачи заявок на гранты и пр.

(С3.2) Несанкционированная передача сотрудниками вузов производственным организациям (включая фирмы по разработке программного обеспечения) выполненных разработок или их фрагментов для дальнейшего использования. Если такая передача осуществляется без заключения договоров с ОР, то это может лишать их определенной части доходов.

(С3.3) Использование сотрудниками плагиата при публикации научных и иных видов работ. Это может наносить ущерб не только самим авторам, но и репутации вуза; научным изданиям вузов, в которых публикуется плагиат. В результате могут снижаться возможности вузов по получению грантов, проведению научных мероприятий, дальнейшему осуществлению научных публикаций, привлечению для работы высококвалифицированных сотрудников и пр.

(С4) – Ухудшение показателей эффективности деятельности ОР, обусловленное недостаточно корректным обращением сотрудников с ее клиентами, контрагентами: снижение прибыли, рентабельности, производительности труда, потеря клиентуры [21].

Возможные специфические последствия для вузов по пункту «С4».

(С4.1) Потери «клиентской базы обучения», в том числе абитуриентов при приеме в вузы, а также отчисления студентов в процессе обучения с целью перехода в другие вузы. (С4.2) Потери возможностей привлечения в вузы для работы по совместительству сотрудников «организаций – потенциальных работодателей» для выпускников вузов, а также для еще обучающихся студентов.

Для торговых организаций последствиями НП могут быть следующие.

(С4.3) Снижение объемов продаж товаров или оказания сопутствующих услуг. (С4.4) Падение лояльности клиентов не только к конкретной торговой организации, но и к владельцу «бренда», используемого организацией.

(С5) – Нестабильная работа организации, или ее отдельных подразделений, например, вследствие аварий, простоев по вине работников, высокого уровня дефектной продукции, попадающей в продажу по вине работников.

(С6) – Репутационные потери для организации, что в конечном счете может влиять на ее рыночные позиции, возможности привлечения и удержания квалифицированного персонала.

Ниже авторы приводят типичные сочетания «причин-видов-следствий» НП для двух типов организаций: вузов (табл. 1) и предприятий торговли (табл. 2). При этом часть объектов в этих таблицах повторяются.

Таблица 1 – Характеристика типичных причин, видов нелояльности сотрудников и их следствий для вузов

Причины НП	Направленность НП	Виды НП	Следствия НП для организации
П1	Н1 – Н3, Н5, Н7, Н8	В1	С1, С4, С6
		В2	С4, С5
		В3	С2, С4, С6
		В4	С2, С2.1, С2.2, С4, С5, С6
		В6	С3, С3.1, С3.2, С3.3, С4, С5, С6
		В7	С3, С4, С5, С6
		В8	С2, С3, С4, С6
		В11	С4, С5
П2	Н1 – Н3, Н5, Н7	В1	С1, С4, С6
		В6	С3, С3.1, С3.2, С3.3, С4, С5, С6
		В11	С4, С5
П3	Н3 – Н6, Н8	В1	С1, С4, С6
		В6	С3, С3.1, С3.2, С3.3, С4, С5, С6
		В7	С3, С3.1, С3.2, С4, С5, С6
		В11	С4, С5
П4	Н1 – Н3, Н5, Н7, Н8	В1	С1, С4, С6
		В11	С4, С5
		В13	С4, С6
П5	Н1 – Н9	В1	С1, С4, С6
		В4	С2, С2.1, С2.2, С4, С5, С6
		В5	С3, С3.1, С3.2, С4
		В6	С3, С3.1, С3.2, С3.3, С4, С5, С6
		В7	С3, С3.1, С3.2, С4, С5, С6
		В8	С2, С3, С4, С6
		В9	С4, С6
		В10	С4, С6
		В11	С4, С5
		В12	С4, С5
В13	С4, С6		

В математическом отношении эти две таблицы могут рассматриваться, например, как четырехмерные массивы, в которых «заполненными» являются только некоторые из элементов – соответствующие сочетаниям, приведенным в таблицах. При этом значениями элементов массивов для конкретных организаций могут быть некоторые экспертные оценки в количественной или качественной форме; результаты вычислений статистических показателей и пр.

Использование таких массивов потенциально может быть полезным в системах поддержки принятия решений по борьбе с НП – в том числе для обеспечения «функциональной полноты» выполняемых оценок в отношении НП; при разработке комплексов мер по предотвращению НП, контроле таких мер и пр.

Таблица 2 – Характеристика типичных причин, видов нелояльности персонала и их следствий для предприятий торговли

Причины НП	Направленность НП	Виды НП	Следствия НП для организации
П1	Н1 – Н3, Н5, Н7, Н8	В1	С1, С4, С6
		В2	С4, С4.3, С4.4, С5
		В3	С2, С4, С4.3, С4.4, С6
		В4	С2, С4, С4.3, С4.4, С5, С6
		В6	С3, С4, С4.3, С4.4, С5, С6
		В7	С3, С4, С4.3, С4.4, С5, С6
		В8	С2, С3, С4, С4.3, С4.4, С6
П2	Н1 – Н3, Н5, Н7	В11	С4, С4.3, С4.4, С5
		В1	С1, С4, С4.3, С4.4, С6
		В6	С2, С3, С4, С4.3, С4.4, С6
П3	Н3 – Н6, Н8	В11	С4, С4.3, С4.4, С5
		В1	С1, С4, С4.3, С4.4, С6
		В6	С3, С4, С4.3, С4.4, С5, С6
П4	Н1 – Н3, Н5, Н7, Н8	В7	С4, С4.3, С4.4, С5, С6
		В11	С4, С4.3, С4.4, С5
		В1	С1, С4, С4.3, С4.4, С6
П5	Н1 – Н9	В11	С4, С4.3, С4.4, С5
		В1	С1, С4, С4.3, С4.4, С6
		В4	С2, С4, С4.3, С4.4, С5, С6
		В5	С3, С4, С4.3, С4.4
		В6	С3, С4, С4.3, С4.4, С5, С6
		В7	С3, С4, С4.3, С4.4, С5, С6
		В8	С2, С3, С4, С4.3, С4.4, С6
		В9	С4, С4.3, С4.4, С6
П5	Н1 – Н9	В10	С4, С4.3, С4.4, С6
		В11	С4, С4.3, С4.4, С5
П5	Н1 – Н9	В12	С4, С4.3, С4.4, С5

Комплексный мониторинг, предотвращение и уменьшение вероятности фактов проявления нелояльности персонала в организациях. В общем случае мониторинг и управление (включая профилактику) НП должны осуществляться для следующих категорий ФЛ. (Г1) Кандидаты для приема на работу в ОР. (Г2) Работающие в организации сотрудники. (Г3) Временно не работающие в ОР граждане, включая лиц, находящихся в декретных отпусках, отпусках по уходу за ребенком и пр. (Г4) Уволившиеся из организации ФЛ, которые перешли на работу в иные ОР. (Г5) Уволившиеся из организации граждане, которые перешли на пенсию и не работают. (Г6) Особая категория ФЛ, которые могут проявлять нелояльность к организации – это сотрудники организаций-контрагентов ОР. Они могут обладать определенной информацией о деятельности ОР и/или быть заинтересованными в распространении некоторых сведений об ОР. Такая заинтересованность может носить как «личный», так и «служебный» характер. В первом случае определенные меры воздействия на таких ФЛ могут быть реализованы через руководство их организаций.

Для проведения мониторинга (учета) фактов НП целесообразно использование специализированной подсистемы в общей компьютерной информационной системе (КИС) организации. При этом права доступа к такой подсистеме должны быть строго ограничены – как в отношении ввода/корректировки информации, так и возможностей ее просмотра. Целесообразность использования именно подсистемы (а не отдельной системы) определяется возможностью использования «готового» списка сотрудников ОР, ее бывших работников и, может быть, лиц, которые рассматривались как «кандидатуры» для приема на работу. При этом по сотрудникам ОР из других подсистем КИС может быть взята следующая информация: сведения о материальных вознаграждениях, взысканиях и поощрениях, об изменениях должностей, возможно – оперативная информация о «служебной активности» [12, 13]. Для каждой из групп лиц по пунктам «Г1...Г5» имеется определенная специфика в отношении направлений и технологий мониторинга информации, методов ее группировки, обобщения и оценки. Рассмотрим эту специфику подробнее.

(1) Прием ФЛ на работу. Возможные «кандидаты для приема на работу» могут выявляться, в основном, с использованием следующих источников информации. (а) Инициативная подача кандидатами заявлений, резюме и пр., в том числе через сайт ОР и по электронной почте на адрес ОР.

(б) Самостоятельный поиск ОР кандидатов на основе информации, доступной в интернете – включая сайты типа www.HeadHunter.ru и т.п. (в) Предложения со стороны рекрутинговых агентств. (г) Использование рекомендаций ФЛ-кандидатов от постоянных сотрудников ОР. (д) Рекомендации «кандидатов для приема» со стороны вузов, в которых эти лица проходят обучение, некоторых иных организаций. (е) Использование для целей отбора результатов прохождения «производственных практик» студентами вузов на базе ОР.

Для предотвращения в будущем потенциально неэтичного поведения ФЛ, принимаемых на работу в ОР, может использоваться следующее:

- поиск информации о кандидатах в интернете, в том числе и в социальных сетях. При этом целесообразно выделять две категории информации: сведения, исходящие от самих кандидатов; отзывы о них со стороны других ФЛ и юридических лиц, в том числе предыдущих или «текущих» работодателей;

- запрос мнения о кандидатах в отделах кадров их прежних ОР;
- получение отзывов о кандидатах от сотрудников с их прежних мест трудовой деятельности. Однако такие оценки могут быть достаточно субъективными, особенно при наличии каких-то конфликтов;

- анализ отзывов (рекомендаций) от сотрудников той ОР, куда предполагается прием известных им кандидатов. Такие отзывы также могут быть субъективными – если эти лица лично заинтересованы в приеме своих «знакомых» или, наоборот, не заинтересованы в появлении «конкурентов по работе»;

- изучение документов, представленных лицами, желающими поступить на работу в ОР, включая резюме, историю смены мест работы и пр. При этом для проверки таких документов сотрудниками ОР могут осуществляться «контрольные звонки», отправляться запросы – в бумажной и электронной формах, использоваться содержательный и психологический анализ резюме;

- для некоторых категорий должностей – изучение списков публикаций и наукометрических показателей для ФЛ-кандидатов, представленных на www.elibrary.ru, www.scholar.google.ru, в иных базах наукометрических показателей и пр.;

- тестирование кандидатов, в том числе для оценки уровней их знаний и умений, некоторых психофизиологических характеристик, включая концентрацию внимания, объем кратковременной зрительной и слуховой памяти. При этом могут быть применены различные виды тестовых компьютерных программ, в том числе специально разработанных самими ОР и/или доступных на различных интернет-ресурсах;

- собеседования с кандидатами, в том числе для выявления их психологических характеристик; возможной/потенциальной совместимости (или несовместимости) с другими сотрудниками ОР; способности адекватно воспринимать информацию; корректно выполнять получаемые инструкции.

По результатам анализа собранной информации могут приниматься следующие виды решений. (а) Принципиальное решение о целесообразности приема ФЛ на работу и о сроках такого приема – в том числе в связи с ожидаемым появлением вакансий в будущем. (б) О подразделении и должности для ФЛ, принимаемых на работу. (в) О заработной плате лиц, принимаемых на работу. (г) О рекомендациях кандидатам на занятие должностей в отношении профессиональной переподготовки или повышения квалификации.

Специфическим вариантом организации приема на работу можно считать конкурсный отбор. В России он прямо предусматривается в нормативных документах только для некоторых категорий должностей – например, для преподавательских и некоторых иных должностей в вузах. Однако конкурсный отбор возможен и в иных случаях – но при обязательном соблюдении правил, предусмотренных Трудовым кодексом.

(2) В отношении контроля и управления НП целесообразно различать следующие группы лиц. (а) ФЛ, трудовой договор с которыми заключен на «неопределенный срок». (б) Лица, работающие в ОР в рамках «испытательных сроков». (в) ФЛ, работающие в ОР по «срочным трудовым договорам» (СТД) или по договорам гражданско-правового характера, в том числе договорам подряда. При принятии решений о перезаключении СТД или о переводе ФЛ на «постоянную» работу учитывается и информация о фактах проявления НП, их конфликтах с другими лицами и пр.

В наглядной форме деятельность по мониторингу фактов проявления НП и управлению лояльностью ФЛ, работающих в ОР; принятию и реализации соответствующих решений показана на рисунке. По сравнению с «компонентами», перечисленными в литературе для моделей управления персоналом (например, [24, 25]), схема на рисунке 1 содержит существенно больше «объектов».

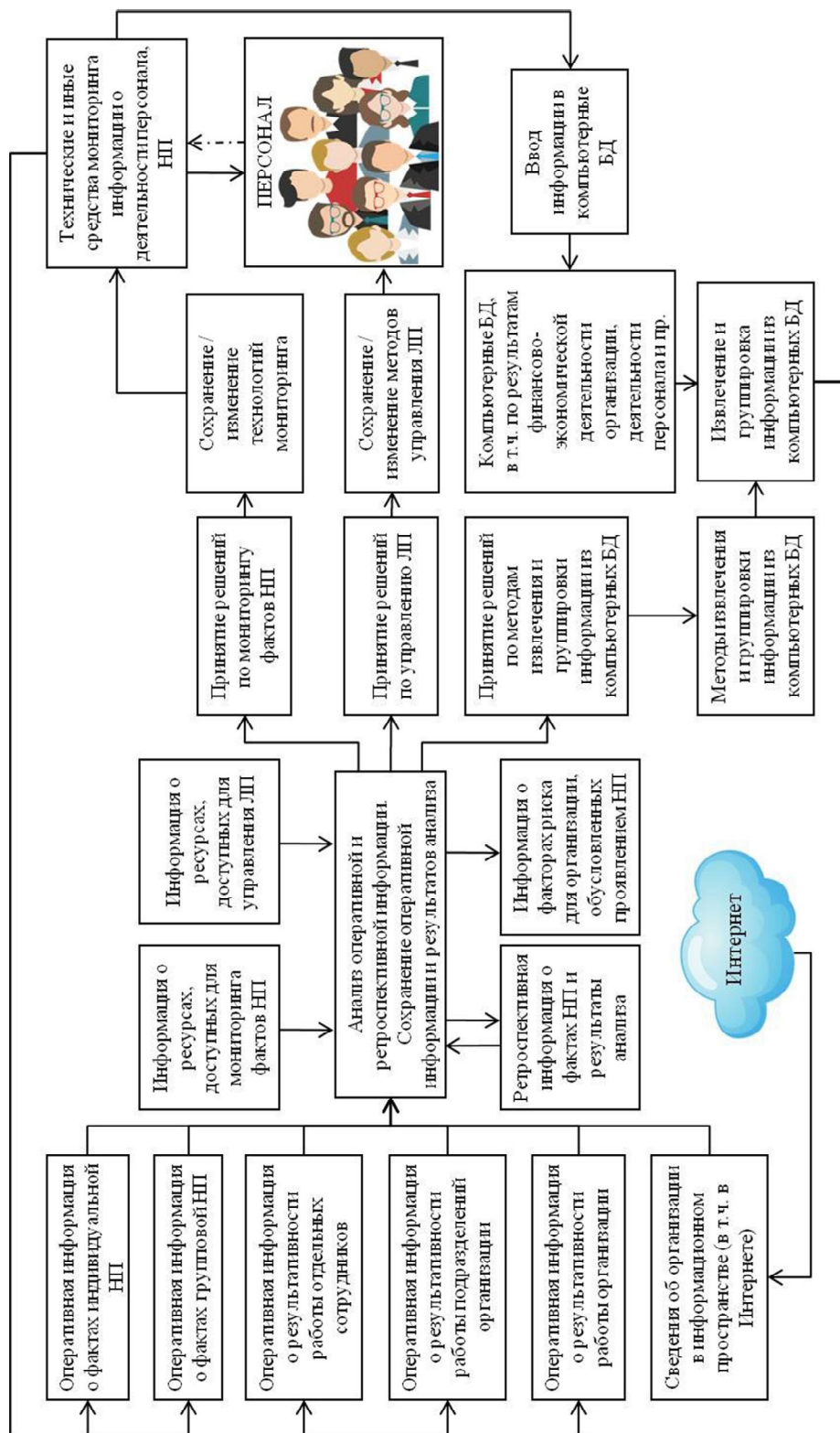


Рисунок – Схема деятельности по управлению лояльностью персонала, работающего в ОР. Обозначения: НП – лояльность персонала; ЛП – лояльность персонала, БД – базы данных

(3) Для лиц, которые временно не работают в ОР, мониторинг фактов их нелояльности может производиться потому, что они могут обладать некоторыми оперативными «служебными секретами», распространение информации о которых потенциально может нанести ущерб ОР. Наиболее вероятны факты такой НП для лиц, которые не предполагают возвращаться на работу в ОР.

(4) Уволившись из ОР ФЛ, которые перешли на работу в иные организации. Выделим такие подкатегории. (4а) Лица, «принудительно» уволенные из ОР по ее инициативе, в том числе из-за выявленных фактов НП, низкой результативности работы, при «сокращениях персонала». Для этой подкатегории ФЛ вероятны как «инициация» негативной информации о прежней ОР, так и участие в распространении (или подтверждении) негативной информации, созданной другими ФЛ или организациями. В условиях развития ИТКТ такие действия обычно осуществляются не в форме «устных контактов», а путем использования средств интернета. (4б) ФЛ, перешедшие на работу в другие организации по собственной инициативе: ради повышения зарплаты, изменения профиля деятельности, обеспечения профессионального роста и пр. В отношении НП наибольшую потенциальную опасность представляют лица, работающие в организациях-конкурентах по отношению к прежней ОР. Причина – их служебные обязанности могут прямо предусматривать действия, направленные на подрыв конкурентных позиций прежней ОР. (4в) Лица, добровольно перешедшие на работу в другие организации при посредничестве ОР, в том числе и в порядке т.н. «аутплейсмента». Для этой подкатегории ФЛ проявления фактов НП маловероятны.

(5) Для лиц, полностью прекративших трудовую деятельность (ушедших на пенсию) необходимость мониторинга информации о возможных фактах их нелояльности по отношению к прежней ОР может определяться такими причинами. (5.1) Наличием у них ретроспективной «инсайдерской» информации о деятельности ОР, распространение которой может нанести ей вред. (5.2) Организации-конкурентами этот контингент лиц может быть использован для распространения «внушенной» им информации или для распространения негативных сведений от их имени. (5.3) Исходящие от этой категории ФЛ негативные сведения об ОР (в том числе искаженные и даже вымышленные), будут восприниматься с большим доверием, чем от лиц, которые в ОР никогда не работали.

Выводы. 1. Целесообразно рассмотрение вопросов нелояльности ФЛ не только по отношению к организации-работодателю в целом, но и к некоторым другим видам объектов, указанных в статье. 2. Развитие ИТКТ оказывает значительное влияние на условия и направления деятельности персонала многих видов организаций. При этом увеличилось количество причин, которые могут повышать вероятность нелояльного поведения, в том числе нарушений правил информационной безопасности. 3. Одновременно в условиях развития ИТКТ увеличился и набор «возможностей» персонала в отношении форм проявления нелояльности по отношению к организации-работодателю и ее отдельным сотрудникам. 4. Развитие ИТКТ повлияло и на возможные «следствия» проявления нелояльности, в том числе путем «вбросов» фальсифицированной информации об организациях, утечек служебной информации и пр. 5. Развитие ИТКТ существенно расширяет возможности работодателей по выявлению, накоплению и анализу фактов проявления нелояльности. Для решения этих задач представляется целесообразным создание и использование специализированных функциональных подсистем информационных систем организаций. 6. Кроме того, такие подсистемы могут использоваться для выдачи советов по уменьшению вероятностей проявления фактов нелояльности, снижения ущербов от таких фактов.

Библиографический список

1. Авдошин С. М. Технология анонимных сетей / С. М. Авдошин, А. В. Лазаренко // Информационные технологии. – 2016. – Т. 22, № 4. – С. 284–291.
2. Алавердов А. Р. Проектный подход в обеспечении высокого уровня лояльности персонала / А. Р. Алавердов // Вестник Академии. – 2017. – № 2. – С. 19–28.
3. Алавердов А. Р. Нелояльные сотрудники: типология и их влияние на конкурентные позиции современной организации / А. Р. Алавердов // Вестник Академии. – 2017. – № 3-2. – С. 5–12.
4. Ажмухамедов И. М. Методика оценки уровня информационной безопасности вуза и ее апробация на примере Астраханского государственного технического университета / И. М. Ажмухамедов, Д. Ю. Учаев, А. И. Ажмухамедов // Прикаспийский журнал: управление и высокие технологии. – 2014. – № 1 (25). – С. 033–046. ([http://hi-tech.asu.edu.ru/files/1\(25\)/33-46.pdf](http://hi-tech.asu.edu.ru/files/1(25)/33-46.pdf))
5. Бадренкова Е. С. Формирование лояльности персонала и ее значение для успешной работы организации / Е. С. Бадренкова // Вестник РГГУ. Серия «Экономика. Управление. Право». – 2015. – № 3 (146). – С. 72–78.
6. Баранская С. С. Методы стимулирования и организационная лояльность / С. С. Баранская // Методологические основы работы психолога : материалы XII Международной конференции студентов, аспирантов и молодых ученых «Психология XXI века». – Санкт-Петербург : Изд-во С.-Петербург. ун-та, 2009. – С. 133–135.

7. Берлизова Я. В. Влияние организационной лояльности на индивидуальную производительность персонала / Я. В. Берлизова // Научное обозрение. – 2015. – № 19. – С. 329–336.
8. Блэк Джон. Экономика: Толковый словарь: Англо-русский / Джон Блэк. – Москва : ИНФРА-М: Весь Мир, 2000. – 829 с.
9. Боровских Н. В. Анализ уровня и структуры затрат на персонал / Н. В. Боровских, Е. А. Кипервар, С. Г. Полянская // Вестник Сибирского института бизнеса и информационных технологий. – 2018. – № 1 (25). – С. 11–17.
10. Брумштейн Ю. М. Оптимизация распределения персонала между подразделениями организаций на основе компетентного подхода / Ю. М. Брумштейн, И. А. Дюдиков // Прикаспийский журнал: управление и высокие технологии. – 2015. – № 2 – С. 45–58. ([http://hi-tech.asu.edu.ru/files/2\(30\)/45-58.pdf](http://hi-tech.asu.edu.ru/files/2(30)/45-58.pdf))
11. Бушуева М. С. Коммерческая тайна и ответственность за ее разглашение / М. С. Бушуева, Д. И. Шарипова, М. Л. Макаревич // Инновационная экономика: перспективы развития и совершенствования. – 2017. – № 8 (26). – С. 10–14.
12. Волкова А. В. Обзор основных концепций и подходов к проблеме формирования и оценки лояльности персонала в компании / А. В. Волкова // Актуальные тенденции и инновации в развитии российской науки : сборник научных статей. – Москва, 2019. – С. 95–97.
13. Врублевская К. А. Методы и модели управления рисками деятельности коллекторских агентств в условиях развития информационно-телекоммуникационных технологий / К. А. Врублевская, Ю. М. Брумштейн, Н. И. Мавлюдова // Прикаспийский журнал: управление и высокие технологии. – 2019. – № 4. – С. 81–98.
14. Выборнова О. Н. Управление рисками информационной безопасности в условиях неопределенности / О. Н. Выборнова, И. М. Ажмухамедов, Ю. М. Брумштейн // Проблемы информационной безопасности. Компьютерные системы. – 2016. – № 1. – С. 7–14.
15. Глобальное исследование утечек конфиденциальной информации в 2018 году. – Режим доступа: https://www.infowatch.ru/sites/default/files/report/analytics/russ/InfoWatch_Global_Report_2018_year.pdf?rel=1, свободный. – Заглавие с экрана. – Яз. рус. (дата обращения: 06.01.2020).
16. Головин Д. В. Гендерные и возрастные особенности влияния стиля управления на организационную лояльность / Д. В. Головин, Н. Т. Тришина, А. Ю. Чернов // Инновации в науке. – 2016. – № 11 (60). – С. 127–138.
17. Душкин А. С. Особенности организационной лояльности сотрудников органов внутренних дел / А. С. Душкин // Педагогика и психология в деятельности сотрудников органов внутренних дел: теория, методика, практика : материалы всероссийской научно-практической конференции. – Санкт-Петербург : Санкт-Петербургский университет Министерства внутренних дел Российской Федерации, 2018. – С. 106–110.
18. Иванова В. С. Корпоративная лояльность как фактор повышения эффективности работы государственного служащего / В. С. Иванова, С. А. Прозорова // Государственное и муниципальное управление в XXI веке: теория, методология, практика. – 2016. – № 26. – С. 295–300.
19. Кодекс Российской Федерации об Административных правонарушениях. От 30.12.2001 № 195-ФЗ (ред. от 18.03.2019). – Режим доступа: http://www.consultant.ru/document/cons_doc_LAW_34661/, свободный. – Заглавие с экрана. – Яз. рус.
20. Кошарная Г. Б. Влияние лояльности персонала на уровень текучести кадров в организации / Г. Б. Кошарная, Ю. Г. Жаткина // Качество жизни населения в условиях модернизации российского общества: правовые, социокультурные и социально-экономические аспекты : сборник статей VII Всероссийской научно-практической конференции. – Пенза : Приволжский Дом знаний, 2014. – С. 82–84.
21. Лустина Т. Н. Формирование лояльности потребителей как необходимое условие для эффективного менеджмента отношений с клиентами / Т. Н. Лустина, Т. В. Богачева, А. Г. Панова // Актуальные вопросы экономики, управления, права, социологии и педагогики в условиях информационного общества: сб. научных статей по итогам Международной конференции. – Волгоград : ООО «Волгоградское научное издательство», 2014. – С. 55–58.
22. О персональных данных. Федеральный закон РФ от 27 июля 2006 года № 152-ФЗ, ред. от 31.12.2017 по 498-ФЗ. – Режим доступа: http://www.consultant.ru/document/cons_doc_LAW_61801/, свободный. – Заглавие с экрана. – Яз. рус.
23. Позднякова Т. В. Лояльность персонала организаций: анализ теоретико-методологических и методических аспектов / Т. В. Позднякова // ПНиО. – 2013. – № 2. – Режим доступа: <https://cyberleninka.ru/article/n/loyalnost-personala-organizatsiy-analiz-teoretiko-metodologicheskikh-i-metodicheskikh-aspektov>, свободный. – Заглавие с экрана. – Яз. рус. (дата обращения: 22.02.2020).
24. Полосухина М. В. Формирование модели управления лояльностью персонала современной организации / М. В. Полосухина // Управление экономическими системами. – 2011. – № 36. – Режим доступа: <https://cyberleninka.ru/article/n/formirovanie-modeli-upravleniya-loyalnostyu-personala-sovremennoy-organizatsii>, свободный. – Заглавие с экрана. – Яз. рус. (дата обращения: 02.11.2019).
25. Сидоренкова К. С. Анализ формирования модели управления лояльностью персонала компании / К. С. Сидоренкова // Инфраструктурные отрасли экономики: проблемы и перспективы развития. – 2016. – № 13. – С. 76–79.
26. Трудовой кодекс Российской Федерации от 30.12.2001 № 197-ФЗ (ред. от 16.12.2019).
27. Харский К. В. Благонадежность и лояльность персонала / К. В. Харский. – Москва : Питер, 2003. – 495 с.
28. Allen N. J. Affective, continuance, and normative commitment to the organization: an examination of construct validity / N. J. Allen, J. P. Meyer // Journal of Vocational Behavior. – 1996. – № 49. – P. 252–276.

29. Mayer R. C. Predicting participation and production outcomes through a two-dimensional model of organizational commitment / R. C. Mayer, F. D. Schoorman // *Academy of Management Journal*. – 1992. – № 35 (3). – P. 671–684.
30. McElroy J. C. Railroad Employee Commitment and Work-Related Attitudes and Perceptions / J. C. McElroy, P. C. Morrow // *Transportation Journal*. – 1995. – № 34 (3). – P. 13–24.
31. Mowday R. T., Porter L. W., Steers R. M. Employee-organization linkages / R. T. Mowday, L. W. Porter, R. M. Steers. – New York : Academic Press, 1982. – 253 p.
32. Grusky O. Career Mobility and Organizational Commitment / O. Grusky // *Administrative Science Quarterly*. – 1966. – № 10. – P. 488–503.

References

1. Avdoshin S. M., Lazapenko A. V. Tekhnologiya anonimnykh setey [Technology of anonymous networks]. *Informatsionnye tekhnologii* [Information technologies], 2016, vol. 22, no. 4, pp. 284–291.
2. Alaverdov A. R. Proektnyy podkhod v obespechenii vysokogo urovnya loylnosti personala [Design approach in ensuring high level of staff loyalty]. *Vestnik Akademii* [Academy Bulletin], 2017, no. 2, pp. 19–28.
3. Alaverdov A. R. Neloyalnye sotrudniki: tipologiya i ikh vliyanie na konkurentnye pozitsii sovremennoy organizatsii [Disloyal employees: typology and their influence on competitive positions of modern organization]. *Vestnik Akademii* [Academy Bulletin], 2017, № 3-2, pp. 5–12.
4. Azhmukhamedov I. M., Uchaev D. Yu., Azhmukhamedov A. I. Metodika otsenki urovnya informatsionnoy bezopasnosti vuza i ee aprobatsiya na primere Astrakhanskogo gosudarstvennogo tekhnicheskogo universiteta [A technique of assessment of level of information security of higher education institution and its approbation on the example of the Astrakhan state technical university]. *Prikaspiyskiy zhurnal: upravlenie i vysokie tekhnologii* [Caspian Journal: Control and High Technologies], 2014, no. 1 (25), pp. 033–046. ([http://hi-tech.asu.edu.ru/files/1\(25\)/33-46.pdf](http://hi-tech.asu.edu.ru/files/1(25)/33-46.pdf))
5. Badrenkova Ye. S. Formirovanie loylnosti personala i ee znachenie dlya uspekhov raboty organizatsii [Formation of personnel loyalty and its importance for successful work of the organization]. *Vestnik Rossiyskogo gosudarstvennogo gumanitarnogo universiteta. Seriya «Ekonomika. Upravlenie. Pravo»* [Bulletin of the Russian State University for the Humanities. Series "Economy. Management. Right"], 2015, № 3 (146), pp. 72–78.
6. Baranskaya S. S. Metody stimulirovaniya i organizatsionnaya loylnost [Methods of Stimulation and Organizational Loyalty]. *Metodologicheskie osnovy raboty psikhologa: materialy XII Mezhdunarodnoy konferentsii studentov, aspirantov i molodykh uchennykh «Psikhologiya XXI veka»* [Methodological Foundations of Psychologist's Work: Materials of the XII International Conference of Students, Postgraduate Students and Young Scientists "Psychology of the XXI Century]. Saint Petersburg, Saint Petersburg University Publ., 2009, pp. 133–135.
7. Berlizeva Ya. V. Vliyanie organizatsionnoy loylnosti na individualnuyu proizvoditelnost personala [Impact of organizational loyalty on individual staff productivity]. *Nauchnoe obozrenie* [Scientific Review], 2015, no. 19, pp. 329–336.
8. Blek Dzhon. *Ekonomika: Tolkovyy slovar: Anglo-russkiy* [Economics: Interpreted dictionary: English-Russian]. Moscow, INFRA-M Publ., Ves Mir Publ., 2000. 829 p.
9. Borovskikh N. V., Kipervar Ye. A., Polyanskaya S. G. Analiz urovnya i struk-tury zatrat na personal [Analysis of the level and structure of staff costs]. *Vestnik Sibirskogo instituta biznesa i informatsionnykh tekhnologiy* [Journal of the Siberian Institute of Business and Information Technologies], 2018, no. 1 (25), pp. 11–17.
10. Brumshteyn Yu. M., Dyudikov I. A. Optimizatsiya raspredeleniya personala mezhdru podrazdeleniyami organizatsiy na osnove kompetentnostnogo podkhoda [Optimization of personnel distribution among organizational divisions on the basis of competence approach]. *Prikaspiyskiy zhurnal: upravlenie i vysokie tekhnologii* [Caspian Journal: Control and High Technologies], 2015, no. 2, pp. 45–58. ([http://hi-tech.asu.edu.ru/files/2\(30\)/45-58.pdf](http://hi-tech.asu.edu.ru/files/2(30)/45-58.pdf))
11. Bushueva M. S., Sharipova D. I., Makarevich M. L. Kommercheskaya tayna i otvetstvennost za ee razglashenie [Commercial secret and responsibility for its disclosure]. *Innovatsionnaya ekonomika: perspektivy razvitiya i sovershenstvovaniya* [Innovative Economy: Prospects for Development and Improvement], 2017, no. 8 (26), pp. 10–14.
12. Volkova A. V. Obzor osnovnykh kontseptsii i podkhodov k probleme formirovaniya i otsenki loylnosti personala v kompanii [Overview of the main concepts and approaches to the problem of formation and evaluation of personnel loyalty in the company]. *Aktualnye tendentsii i innovatsii v razvitiy rossiysskoy nauki: sbornik nauchnykh statey* [Current trends and innovations in the development of Russian science: the collection of scientific articles]. Moscow, 2019, pp. 95–97.
13. Vrublevskaya K. A., Brumshteyn Yu. M., Mavlyudova N. I. Metody i modeli upravleniya riskami deyatelnosti kollektorskiykh agentstv v usloviyakh razvitiya informatsionno-telekommunikatsionnykh tekhnologiy [Methods and models of risk management of collection agencies in conditions of information and telecommunication technologies development]. *Prikaspiyskiy zhurnal: upravlenie i vysokie tekhnologii* [Caspian Journal: Control and High Technologies], 2019, no. 4, pp. 81–98.
14. Vybornova O. N., Azhmukhamedov I. M., Brumshteyn Yu. M. Upravlenie riskami informatsionnoy bezopasnosti v usloviyakh neopredelennosti [Management of information security risks in conditions of uncertainty]. *Problemy informatsionnoy bezopasnosti. Kompyuternye sistemy* [Information Security Problems. Computer Systems], 2016, no. 1, pp. 7–14.
15. *Globalnoe issledovanie utechek konfidentsialnoy informatsii v 2018 godu* [Global research of leaks of confidential information in 2018]. Available at: https://www.infowatch.ru/sites/default/files/report/analytics/russ/InfoWatch_Global_Report_2018_year.pdf?rel=1 (accessed 03.01.2020).

16. Golovin D. V., Trishina N. T., Chernov A. Yu. Gendernye i vozrastnye osobennosti vliyaniya stilya upravleniya na organizatsionnyuyu loylnost [Gender and age features of influence of management style on organizational loyalty]. *Innovatsii v nauke* [Innovations in Science], 2016, no. 11 (60), pp. 127–138.
17. Dushkin A. S. Osobennosti organizatsionnoy loylnosti sotrudnikov organov vnutrennikh del [Peculiarities of organizational loyalty of employees of internal affairs agencies]. *Pedagogika i psikhologiya v deyatelnosti sotrudni-kov organov vnutrennikh del: teoriya, metodika, praktika : materialy vserossiyskoy nauchno-prakticheskoy konferentsii* [Pedagogy and psychology in the activities of employees of internal affairs agencies: theory, methodology, practice: materials of the All-Russian scientific and practical conference]. Saint Petersburg, Saint Petersburg University of the Ministry of the Interior of the Russian Federation, 2018, pp. 106–110.
18. Ivanova V. S., Prozorova S. A. Korporativnaya loylnost kak faktor povysheniya effektivnosti raboty gosudarstvennogo sluzhashchego [Corporate loyalty as a factor of increasing the efficiency of the public servant]. *Gosudarstvennoe i munitsipalnoe upravlenie v XXI veke: teoriya, metodologiya, praktika* [State and Municipal Administration in the XXI Century: Theory, Methodology, Practice], 2016, no. 26, pp. 295–300.
19. *Kodeks Rossiyskoy Federatsii ob Administrativnykh pravonarusheniyakh* [Code of Administrative Offences of the Russian Federation], of 30.12.2001 no. 195-FZ (ed. of 18.03.2019). Available at: http://www.consultant.ru/document/cons_doc_LAW_34661/ (accessed 05.03.2020).
20. Kosharnaya G. B., Zhatkina Yu. G. Vliyaniye loylnosti personala na uroven tekuchesti kadrov v organizatsii [The influence of staff loyalty on the level of staff turnover in the organization]. *Kachestvo zhizni naseleniya v usloviyakh modernizatsii rossiyskogo obshchestva: pravovye, sotsiokulturnye i sotsialno-ekonomicheskie aspekty : sbornik statey VII Vserossiyskoy nauchno-prakticheskoy konferentsii* [Quality of life of the population in the conditions of modernization of the Russian society: legal, sociocultural and socio-economic aspects: Proceedings of the VII All-Russian Scientific and Practical Conference]. Penza, Privolzhskiy Dom znaniy Publ., 2014, pp. 82–84.
21. Lustina T.N., Bogacheva T.V., Panova A.G. Formirovaniye loylnosti potrebiteley kak neobkhodimoye usloviye dlya effektivnogo menedzhmenta otosheniy s klientami [Formation of consumer loyalty as a necessary condition for effective management of relations with clients]. *Aktualnye voprosy ekonomiki, upravleniya, prava, sotsiologii i pedagogiki v usloviyakh informatsionnogo obshchestva: sbornik nauchnykh statey po itogam Mezhdunarodnoy konferentsii* [Topical issues of economy, management, law, sociology and pedagogy in the conditions of the information society: scientific articles on the results of the international conference]. Volgograd, 2014, pp. 55–58.
22. *O personalnykh dannykh. Federalnyy zakon RF* [On personal data. The federal law of the Russian Federation] of July 27, 2006, no. 152-FZ, ed. of 31.12.2017 by 498-FZ. Available at: http://www.consultant.ru/document/cons_doc_LAW_61801/ (accessed 03.03.2020).
23. Pozdnyakova T. V. Loylnost personala organizatsiy: analiz teoretiko-metodologicheskikh i metodicheskikh aspektov [Loyalty of organization personnel: analysis of theoretical-methodological and methodological aspects]. *PNiO* [PNiO], 2013, no. 2. Available at: <https://cyberleninka.ru/article/n/loylnost-personala-organizatsiy-analiz-teoretiko-metodologicheskikh-i-metodicheskikh-aspektov> (accessed 22.02.2020).
24. Polosukhina M. V. Formirovaniye modeli upravleniya loylnostyu personala sovremennoy organizatsii [Formation of a model of personnel loyalty management of a modern organization]. *Upravlenie ekonomicheskimi sistemami* [Economic Systems Management], 2011, no. 36. Available at: <https://cyberleninka.ru/article/n/formirovaniye-modeli-upravleniya-loylnostyu-personala-sovremennoy-organizatsii> (accessed 02.11.2019).
25. Sidorenkova K. S. Analiz formirovaniya modeli upravleniya loylnostyu personala kompanii [Analysis of formation of the model of management of personnel loyalty of the company]. *Infrastrukturnye otrasli ekonomiki: problemy i perspektivy razvitiya* [Infrastructure sectors of the economy: problems and prospects of development], 2016, no. 13, pp. 76–79.
26. *Trudovoy kodeks Rossiyskoy Federatsii* [Labour Code], of 30.12.2001 no. 197-FZ (ed. of 16.12.2019).
27. Kharskiy K. V. *Blagonadezhnost i loylnost personala* [Integrity and Staff Loyalty]. Moscow, Piter Publ., 2003. 495 p.
28. Allen N. J., Meyer J. P. Affective, continuance, and normative commitment to the organization: an examination of construct validity. *Journal of Vocational Behavior*, 1996, no. 49, pp. 252–276.
29. Mayer R. C., Schoorman F. D. Predicting participation and production outcomes through a two-dimensional model of organizational commitment. *Academy of Management Journal*, 1992, no. 35 (3), pp. 671–684.
30. McElroy J. C., McElroy J. C., Morrow P. C. Railroad Employee Commitment and Work-Related Attitudes and Perceptions. *Transportation Journal*, 1995, no. 34 (3), pp. 13–24.
31. Mowday R. T., Porter L. W., Steers R. M. *Employee-organization linkages*. New York, Academic Press, 1982. 253 p.
32. Grusky O. Career Mobility and Organizational Commitment. *Administrative Science Quarterly*, 1966, no. 10, pp. 488–503.

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ, ЧИСЛЕННЫЕ МЕТОДЫ И КОМПЛЕКСЫ ПРОГРАММ

DOI 10.21672/2074-1707.2020.49.4.094-111
УДК 519.178/.245:004.94

НАХОЖДЕНИЕ ПРОВОДЯЩЕГО ОСТОВА В ДВУМЕРНОЙ РЕШЕТКЕ МЕТОДОМ ЗАЛИВКИ

Статья поступила в редакцию 21.12.2019, в окончательном варианте – 02.2020.

Гордеев Иван Иванович, Астраханский государственный университет, 414056, Российская Федерация, г. Астрахань, ул. Татищева, 20а,
кандидат физико-математических наук, e-mail: g2i@mail.ru, ORCID <https://orcid.org/0000-0001-5036-4791>; elibrary: https://elibrary.ru/author_profile.asp?authorid=581330

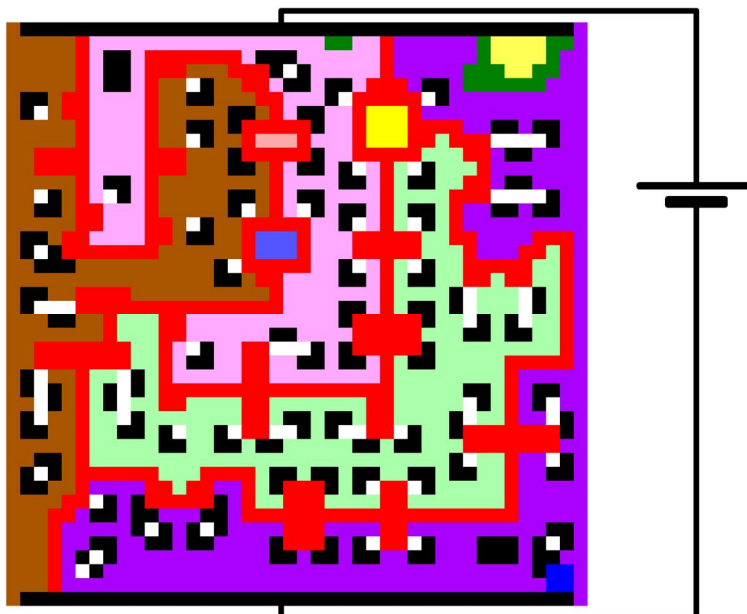
Обчаренко Сергей Сергеевич, Астраханский государственный университет, 414056, Российская Федерация, г. Астрахань, ул. Татищева, 20а,
студент, e-mail: obcharehko@mail.ru

Сизова Анастасия Александровна, Астраханский государственный университет, 414056, Российская Федерация, г. Астрахань, ул. Татищева, 20а,
студент, e-mail: lucky_girl_zz@mail.ru

Показано место проблематики, рассматриваемой в настоящей статье, в структуре исследований по физике и вычислительной математике, практическая значимость этого направления исследовательских работ. Рассмотрена модель перколяции узлов в двумерной решетке с открытыми граничными условиями. Для этой модели проведен анализ основанного на заливке алгоритма, позволяющего находить остов соединяющего кластера. Предлагаемый алгоритм позволяет отделить от соединяющего кластера так называемые висячие части, которые не проводят ток. Приводится подробная классификация висячих частей (висячие концы, висячие циклы и висячие дуги). Охарактеризовано, какие именно висячие части обрабатываются на разных этапах алгоритма. Предлагается исправление некоторых неточностей, допущенных предыдущими авторами, описывавшими алгоритм ранее. Обосновывается, что при реализации алгоритма можно обойтись рассмотрением меньшей окрестности каждой из ячеек, соответствующих узлам. Обсуждаются особые случаи расположения узлов, алгоритмическая обработка которых не была описана ранее в литературе. Представлены особенности реализации исправленного (модифицированного) алгоритма на языке программирования C++. Указаны некоторые усовершенствования в реализации алгоритма, позволяющие увеличить скорость работы программы, построенной на его основе.

Ключевые слова: идентификация остова, соединяющий кластер, перколяция узлов, двумерная решетка, открытые граничные условия, заливка, окрестность фон Неймана, окрестность Мура, информационные технологии, алгоритмы вычислений

Графическая аннотация (Graphical annotation)



THE DETERMINATION OF THE CONDUCTING BACKBONE
IN A TWO-DIMENSIONAL LATTICE BY THE FLOODING METHOD

The article was received by the editorial board on 21.12.2019, in the final version – 20.02.2020.

Gordeev Ivan I., Astrakhan State University, 20a Tatishchev St., Astrakhan, 414056, Russian Federation, Cand. Sci. (Physics and Mathematics), e-mail: g2i@mail.ru, ORCID <https://orcid.org/0000-0001-5036-4791>; https://elibrary.ru/author_profile.asp?authorid=581330

Ovcharenro Sergey S., Astrakhan State University, 20a Tatishchev St., Astrakhan, 414056, Russian Federation, student, e-mail: obchapehko@mail.ru

Sizova Anastasia A., Astrakhan State University, 20a Tatishchev St., Astrakhan, 414056, Russian Federation, student, e-mail: lucky_girl_zz@mail.ru

In this paper, we give an analysis of a flood-based algorithm that makes possible to find the backbone of spanning cluster for the site percolation model in a two-dimensional lattice with open boundary conditions. The algorithm allows separating from the spanning cluster the so-called dangling parts, which do not conduct current. A detailed classification of the dangling parts (dangling ends, dangling cycles and dangling arcs) is given and it is specified which dangling parts are processed at different stages of the algorithm. Corrections are suggested for some inaccuracies of previous algorithm description. We found possible to consider a smaller neighborhood with the implementation of the algorithm. Cases in the implementation of the algorithm, not discussed earlier, are discussed. The implementation of the corrected algorithm in the C++ programming language is described. Some improvements allowing speed up the program are made to the algorithm.

Key words: backbone identification, spanning cluster, site percolation, 2D lattice, open boundary condition, flooding, von Neumann neighborhood, Moore neighborhood, information technology, computing algorithms

Введение. В настоящее время актуальной задачей является моделирование проводимости в неупорядоченных средах, состоящих из случайно расположенных проводящих и непроводящих элементов. Эта задача важна, в частности, для оценки проводимости при прохождении электрического тока в случайной среде [22]; при прохождении газа или жидкости в пористой среде [25]. В отношении последнего направления отметим задачи геофильтрации в неоднородных породах, а также исследование характеристик геологических пород на основе методов электропроводности [3, 8].

Несмотря на значительное количество работ, посвященных проблематике перколяции [5, 6, 7, 26, 27], некоторые направления исследований и разработок остаются исследованными недостаточно полно. Поэтому основной целью данной работы было устранение указанного недочета применительно к одному из направлений исследований по перколяции.

Общая характеристика проблематики исследований. Для моделирования случайных неупорядоченных сред часто используются периодические решетки, в которых рассматривается случайное заполнение узлов (sites) проводящими и непроводящими элементами. С точки зрения теории графов периодическая решетка является *периодическим графом* (periodic graph) без петель (loops). Вершины графа в подобных моделях называют узлами, а рёбра графа называют *связями* (bonds). При этом вершины графа являются случайным образом *занятыми* (occupied) или *свободными* (vacant) [20]. Обычно занятые узлы считаются проводящими, свободные – непроводящими. Следует отметить, что в англоязычной математической литературе по теории графов, петель называют *ребро* (edge) графа, соединенное обоими концами с одной и той же *вершиной* (vertex) [23]. В то же время в англоязычной физической литературе, посвященной перколяции [25, 31, 35], термин *loop* часто используется в значении, которое в теории графов называется *циклом* (cycle) [23]. Модель, в которой рассматриваются случайным образом занятые узлы, называется моделью *перколяции узлов* (site percolation). Возможна также модель *перколяции связей* (bond percolation), в которой рассматриваются случайным образом занятые связи [20, 25]. В теории перколяции рассматриваются *связные компоненты* (connected components) графа из однотипных (занятых или незанятых) узлов или связей, называемые *кластерами* (clusters) [20, 25].

В данной статье дается анализ основанного на *заливке* (flood fill [30] или flooding [31]) алгоритма, позволяющего находить остов соединяющего кластера для модели перколяции узлов в двумерной решетке. *Соединяющий кластер* (spanning cluster) [10, 15, 25, 31] – это множество узлов, для которых существует непрерывная цепочка проводящих узлов от каждого из узлов до двух противоположных краев решётки. Иногда словосочетание *spanning cluster* переводится на русский язык как *стягивающий кластер* [18]. Однако, на наш взгляд, такой перевод не очень удачен, поскольку соединяющий кластер не обязательно «натянут», он может «висеть» достаточно

свободно. Для соединяющего кластера используются также названия *перколяционный кластер* (percolation cluster) [25] и *бесконечный кластер* (infinite cluster) [31]. Последнее название предполагает, что при увеличении размера решетки до бесконечности количество узлов, входящих в этот кластер, также стремится к бесконечности.

При моделировании конечных квадратных решеток обычно уточняют, с какими краями решётки соединяется перколяционный кластер, например, с верхним и нижним краями.

Перколяционный кластер может проводить электрический ток, жидкость или газ, если к противоположным краям прикладывается разность потенциалов или разность давлений соответственно. Для рассматриваемого класса задач в обеспечении проводимости электрического тока, жидкости или газа участвуют не все проводящие элементы перколяционного кластера, а только те, которые входят в так называемый *остов* (backbone) [4, 28] или *скелет* (в английском переводе также *backbone*) [9]. Остов можно определить как подграф перколяционного кластера, содержащий все вершины, от которых имеется как минимум два непересекающихся пути, ведущих к противоположным краям решётки. Здесь подразумеваются *вершинно-непересекающиеся* (vertex-disjoint) пути, в литературе по теории графов такие пути часто называют просто *непересекающимися* (disjoint) или *независимыми путями* (independent paths) [11]. Вершинно-непересекающиеся пути следует отличать от *рёберно-непересекающихся* (edge-disjoint) путей [11], поскольку вершинно-непересекающиеся пути всегда являются рёберно-непересекающимися, а обратное утверждение неверно. Проводящие элементы соединяющего кластера, которые не участвуют в проводимости, относятся к так называемым *висячим частям* (dangling parts) [35] перколяционного кластера.

Для плоских графов используют два варианта наглядного изображения. Первый вариант, который используется и для неплоских графов: граф изображают в виде жирных точек (кругов) и соединяющих точки линий, где жирные точки соответствуют вершинам графа, а линии соответствуют рёбрам графа. Второй вариант: граф изображают в виде многоугольников, разделяющих плоскость на части, где многоугольники соответствуют вершинам графа, а общая сторона каждой пары соседних многоугольников соответствует ребру между соответствующими вершинами. При втором варианте изображения многоугольники часто называют также *ячейками* (cells) [31, 34]. В одних публикациях по перколяции на плоских графах придерживаются первого способа изображения, например, в [16, 36]. В других – придерживаются второго способа, например, в [31, 34]. В третьих – комбинируют первый и второй способы, например, в [25].

На рисунке 1а показан первый способ изображения графа, где чёрные круги соответствуют проводящим узлам, входящим в остов, серые круги соответствуют висячим частям перколяционного кластера (проводящим узлам, входящим в перколяционный кластер, но не входящим в остов), а белые круги соответствуют непроводящим узлам. В данной статье мы будем далее придерживаться второго способа изображения графов. Поскольку рассматривается квадратная решётка, где у каждой вершины может быть четыре смежных, то вершинам графа на рисунках будут соответствовать квадраты. На рисунке 1б приведён второй способ изображения графа с чёрными, серыми и белыми квадратами соответственно.

Описание алгоритма для идентификации остова было дано в статье Тробека и Стаматович [31]. При этом идея алгоритма в свою очередь основывается на статье Иня и Тао [34]. Следует отметить, что рассматриваемый алгоритм предполагает так называемую *«геометрию шины»* (bus-bar geometry) [28]. При использовании геометрии шины для конечной решетки рассматриваются перколяционные кластеры, соединенные с двумя любыми точками на противоположных краях (шинах) решетки. С физической точки зрения *шина* (bus-bar) [28] обычно представляет собой проводящую *полосу* (bar) [25], присоединенную к краю решетки. Возможно также рассмотрение непроводящих шин, которые расположены вдоль двух других противоположных краёв. На рисунке 2а ячейки, относящиеся к проводящим шинам, помечены белыми буквами «В», а ячейки, относящиеся к непроводящим шинам, помечены чёрными буквами «В».

Непроводящие шины по краям соответствуют так называемым *«открытым граничным условиям»* (open boundary conditions) [31, 34]. В качестве синонима открытых граничных условий иногда говорят об *отсутствии периодических граничных условий* (without periodic boundary conditions) [25, 35]. С точки зрения теории графов, шина представляет собой дополнительный ряд узлов вдоль края решетки, которые не являются случайно занятыми, а всегда заняты (или всегда свободны).

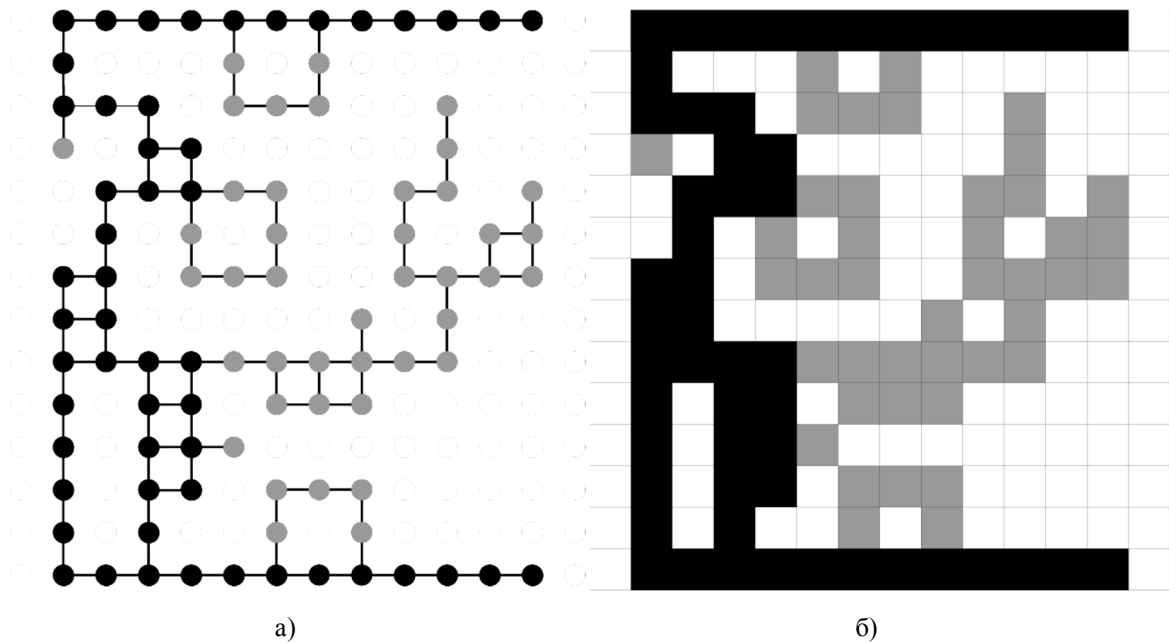


Рисунок 1 – Изображение плоского графа в виде: а) жирных точек (кругов) и соединяющих линий; б) квадратов (ячеек)

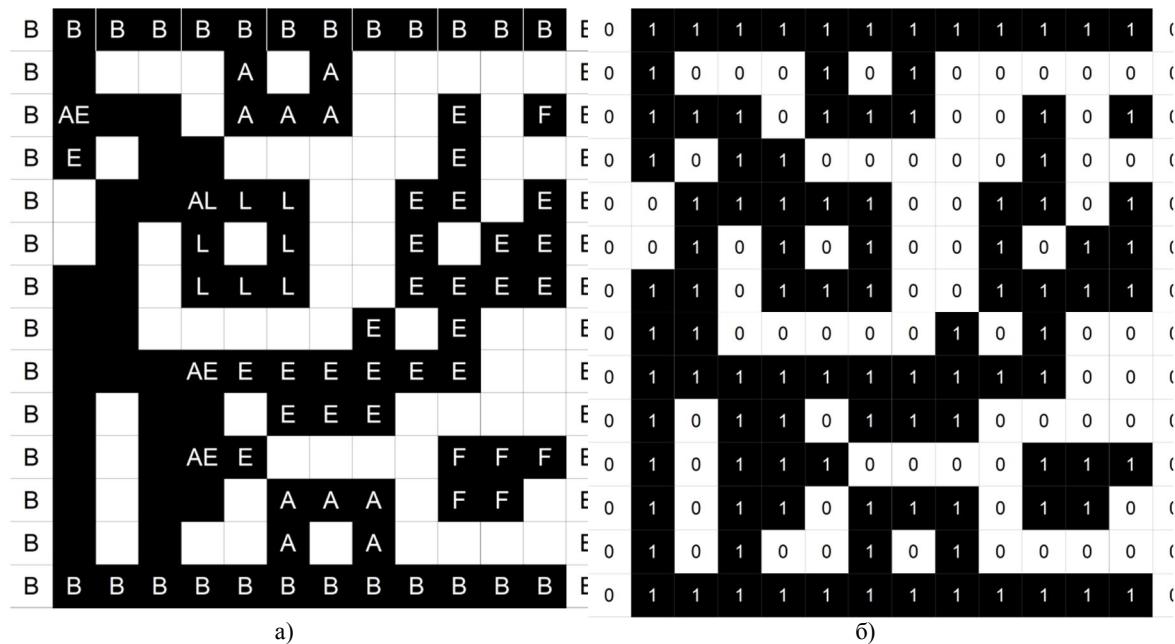


Рисунок 2 – Исходная решетка: а) ячейки, соответствующие шинам (В), висячим концам (Е), висячим циклам (L), висячим дугам (А), сочленяющие ячейки в остове для висячих концов (АЕ) и висячих циклов (AL), ячейки, не входящие в перколяционный кластер (F); б) с пометкой числами проводящих (1) и непроводящих (0) ячеек

Описываемый далее алгоритм позволяет отделить от соединяющего кластера висячие части, которые не проводят ток. В статье Иня и Тао [35] дается подробная классификация висячих частей: *висячие концы* (dangling ends), *висячие циклы* (dangling loops) и *висячие дуги* (dangling arcs). На рисунке 2а буквами помечены ячейки, соответствующие висячим концам («Е»), висячим циклам («L») и висячим дугам («А»). Висячими концами и висячими циклами считаются части перколяционного кластера, которые соединены с ним только через один *сочленяющий узел* (articulation site [34]). Сочленяющий узел в теории графов называют *точкой сочленения* (articulation point) [1, 11], (cutpoint) [17], либо *шарниром* [1]; используются также термины *разделяющая вершина* (separating vertex) или *разрезающая вершина* (cut vertex) [23], в некоторых источниках *cutvertex* пишется слитно [11]. В статье Тробека и Стаматович [31] соответствующие узлы называются

сочленяющими ячейками (articulation cells). На рисунке 2а на остовах помечены сочленяющие ячейки, к которым присоединены висячие концы («AE») и висячие циклы («AL»). При этом висячие концы не содержат циклов, а висячие циклы содержат. С точки зрения теории графов, висячие циклы содержат цикл, проходящий через точку сочленения, но не принадлежащий остову, остову принадлежит только точка сочленения. Висячие концы соединены с остовом при помощи *мостов* (bridges) [11, 17, 32], при этом конец моста, принадлежащий остову, всегда является точкой сочленения. В некоторых источниках мост называется *разделяющим ребром* (separating edge) [23] или *разрезающим ребром* (cut edge) [23], в русскоязычных публикациях и переводах используется также название *перешеек* [1, 32].

Висячими дугами называются части кластера, присоединенные через два или более узлов к одной из шин, но не имеющие соединения с другой шиной, не проходящего через первую шину. Поскольку в некоторых модификациях алгоритмов нахождения остова шины не совсем аккуратно рассматриваются тоже как часть остова, то висячие дуги могут ошибочно присоединяться к остову. В частности, в предыдущей версии того же алгоритма, предложенного в статье Стаматович и Тробека [24], висячие дуги присоединялись к остову.

Занятые ячейки, не соединённые с перколяционным кластером, а образующие отдельные кластеры, на рисунке 2б помечены белыми буквами «F».

При моделировании проводящих и непроводящих узлов их обычно кодируют различными числами. Например, проводящие узлы кодируют числом «1», а непроводящие – числом «0». При анализе различных частей решетки некоторые узлы могут помечаться и другими кодами, отличающимися от «0» и «1». Чтобы наглядно показать части решетки на рисунке, различным кодам ставят в соответствие разные цвета. На рисунке 2б показана заполненная случайным образом исходная решетка, в которой висячие части ещё не отделены, с пометкой проводящих и непроводящих узлов как цветами, так и числами.

Согласно описанию Тробека и Стаматович [31], алгоритм нахождения проводящего остова делится на 7 шагов. Нами проанализировано описание каждого шага в алгоритме Тробека, выявлены и исправлены некоторые неточности в описании этого алгоритма; выполнена реализация алгоритма на языке программирования C++ [2]. Также сделаны некоторые усовершенствования алгоритма, позволяющие ускорить работу программы.

Алгоритм основывается на анализе локальных свойств ячейки и ее соседей. На рисунке 3а показаны используемые обозначения восьми соседей ячейки (X) по сторонам света: северный (N), западный (W), южный (S), восточный (E), северо-западный (NW), юго-западный (SW), юго-восточный (SE), северо-восточный (NE), использованы стандартные сокращения по английским названиям сторон света (North, West, South, East).

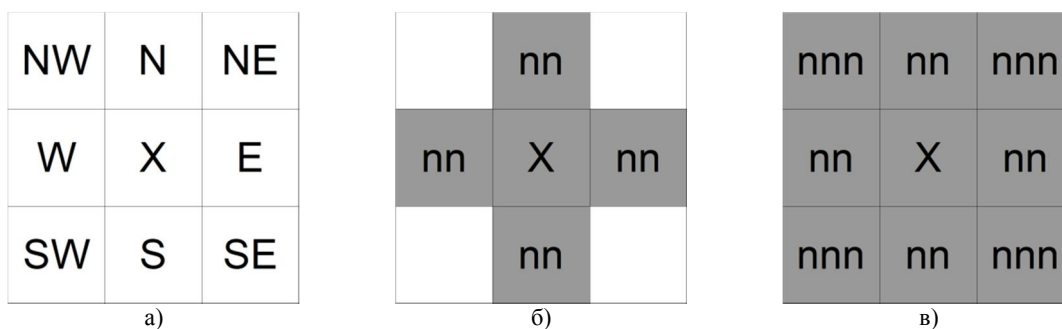


Рисунок 3 – Обозначения соседних ячеек и окрестности: а) обозначение соседних ячеек по сторонам света; б) окрестность Неймана (nn-окрестность); в) окрестность Мура (nnn-окрестность)

В теории клеточных автоматов используются понятия *окрестности фон Неймана* (von Neumann neighborhood) и *окрестности Мура* (Moore neighborhood) [29]. В окрестность фон Неймана (рис. 3б) входит ячейка X, а также её N, W, S и E соседи, называемые nn-соседями (nn – сокращение от nearest neighbors – ближайшие соседи) [31]. В окрестность Мура (рис. 3в), кроме ячеек, входящих в окрестность фон Неймана, входят также NW, SW, SE и NE соседи, называемые nnn-соседями (nnn – сокращение от next nearest neighbors – следующие ближайшие соседи) [31]. Для окрестности фон Неймана и окрестности Мура используются также названия nn-окрестность и nnn-окрестность, соответственно. В nn-окрестность (nnn-окрестность) входит сама ячейка и 4 ячейки (8 ячеек) вокруг, соответственно. Множества ячеек одного типа (занятых или незанятых), в которых любые две ячейки, либо входят в общую nn-окрестность (nnn-окрестность), либо имеют связывающую их цепочку

из соседей того же цвета, имеющих общую np -окрестность ($npnp$ -окрестность), образуют np -кластеры ($npnp$ -кластеры), соответственно. В рассматриваемой модели в качестве перколяционного рассматривается np -кластер из проводящих узлов, соединяющий верхнюю и нижнюю шины. В качестве вспомогательных рассматриваются непроводящие $npnp$ -кластеры.

В статье Тробека и Стаматович [31] предлагается рассматривать для нахождения остова также 16 еще более дальних $npnp$ -соседей (not next-nearest neighbors), которые вместе с 9 ячейками $npnp$ -окрестности образуют окрестность Мура второго порядка (или $npnp$ -окрестность) из 25 ячеек. Тробек и Стаматович, указывая необходимость использования $npnp$ -окрестности, ссылаются на статью Иня и Тао [34]. Однако в [34] не рассматриваются $npnp$ -окрестности, анализируются только $npnp$ -окрестности. Наш анализ также показал, что для нахождения остова достаточно рассмотрения $npnp$ -окрестности.

В предлагаемом алгоритме анализируется случайно заполненная квадратная решетка размера « $L * L$ », L предполагается без учёта шин, на рисунке 2 $L = 12$. Первоначально все ячейки решетки случайным образом закрашены в два цвета: белый (0 – непроводящие узлы) и черный (1 – проводящие узлы). С учётом шин размер решётки оказывается « $(L+2) * (L+2)$ », на рисунке 2 изображена решётка с шинами, где $L+2 = 14$. Чтобы у каждой анализируемой ячейки было восемь np -соседей, к решетке вокруг шин добавляется ещё один вспомогательный слой по всему периметру (постоянно заполнен значением 0, на рисунке 2 не показан), с этим слоем решетка имеет размер « $(L+4) * (L+4)$ ».

Характеристика алгоритма. При описании алгоритма Тробек и Стаматович [31] используют понятие *такта времени* (time-step), соответствующее *дискретному шагу* (discrete step) в публикациях по клеточным автоматам [29]. Клеточный автомат рассматривается Тробеком и Стаматович как вычислительная машина, которая может реализовываться на реальном компьютере, в частности на суперкомпьютере. Тробек и Стаматович предполагают, что для обработки каждой ячейки может использоваться отдельный *вычислительный элемент* (computing element) [31] высокопроизводительного компьютера, например, отдельное ядро центрального или графического процессора. Понятие вычислительного элемента применяется, чтобы пояснить синхронизацию обработки ячеек разными вычислительными элементами. Каждый вычислительный элемент имеет собственную память и арифметический блок. При достаточно большом размере решётки даже на самых высокопроизводительных суперкомпьютерах нерационально выделить отдельный вычислительный элемент для каждой ячейки, поэтому Тробек и Стаматович отмечают, что возможна обработка одним вычислительным элементом более чем одной ячейки. Однако вопрос об обработке нескольких ячеек одним вычислительным элементом не получил сколько-нибудь подробного освещения у Тробека и Стаматович.

В данной публикации мы подробно рассматриваем обработку большого числа ячеек на одном вычислительном элементе и возможность оптимизации соответствующей программной реализации. В качестве предельного случая нами рассматривается реализация обработки всех ячеек на одном вычислительном элементе.

Изначально вся решетка помечена двумя цветами – белым (0) и черным (1). За один такт каждая ячейка решетки может либо однократно менять свое состояние (цвет), либо не менять.

Также используется большое количество уникальных цветов для пометки непроводящих ячеек и нахождения непроводящих кластеров. Кроме того, применяется еще три особых цвета для раскраски остова. Эти три цвета условно называются синим, зеленым и красным. Вопрос о кодировке данных цветов в статье Тробека и Стаматович не обсуждается. Поэтому, чтобы избежать совпадений с другими цветами, нами была выбрана кодировка отрицательными числами. Далее приводится соответствующее кодировке цветов перечисляемый тип, используемый в программе.

```
enum{blue=-1, green=-2, red=-3};
```

В описываемом алгоритме выделяются 7 этапов, называемых «Шагами» (Steps). На рисунке 4 показана блок-схема алгоритма.

На «этапе 1» за один такт делается закрашка так называемых затравочных ячеек внутри непроводящих $npnp$ -кластеров уникальными цветами, зависящими от номера ячейки. В качестве затравочных выбираются все белые ячейки, у которых 3 соседа (S, SE и E) черные, а также нижние ячейки на левой и правой белых шинах. Такой способ выбора затравочных ячеек обеспечивает, что в каждом непроводящем $npnp$ -кластере будет хотя бы одна затравочная ячейка (возможно несколько затравочных ячеек в одном кластере). Номер цвета, в который закрашивается затравочная ячейка, вычисляется по формуле: $m = i \times n + j$, где i – номер строки, j – номер столбца ячейки.

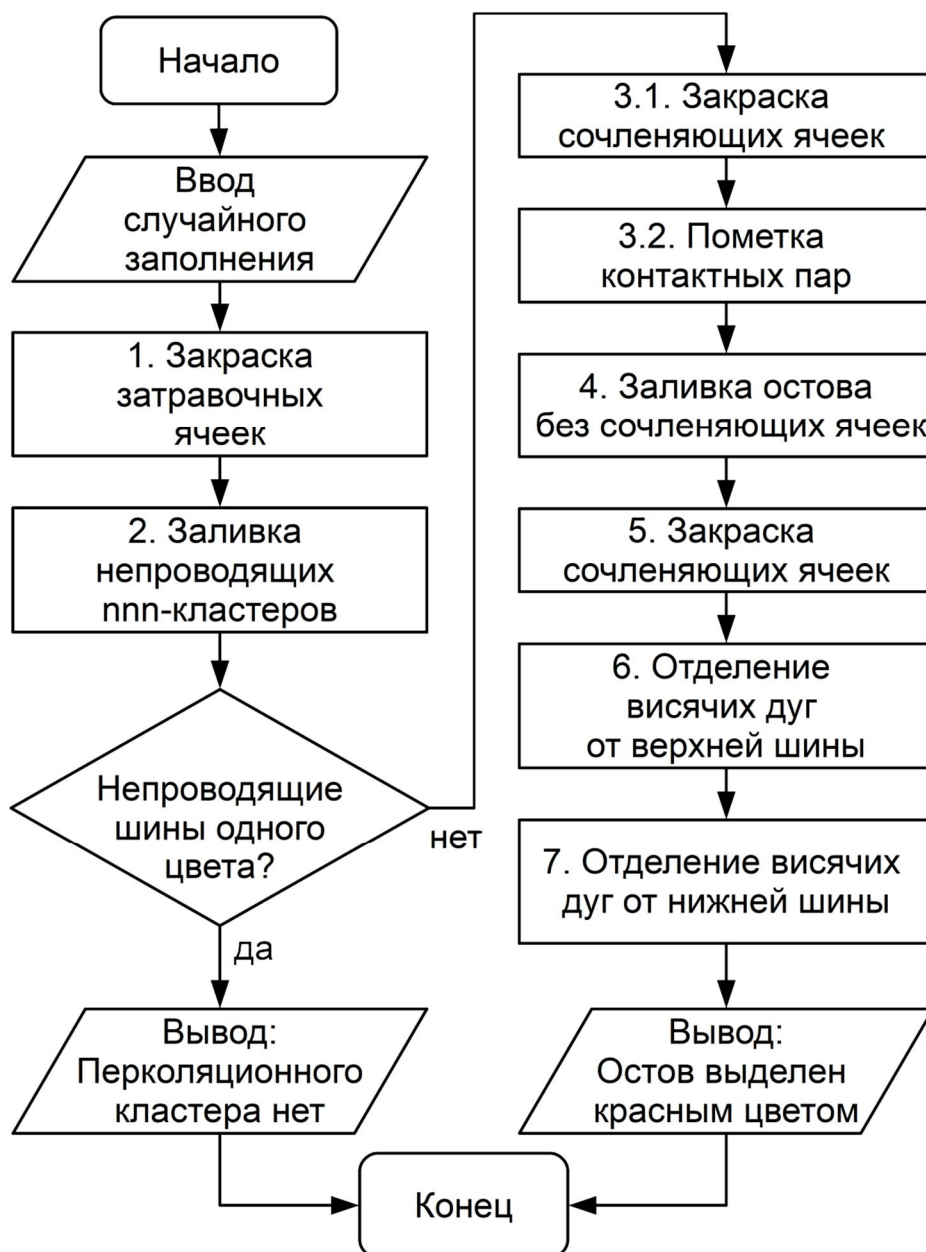


Рисунок 4 – Блок-схема алгоритма нахождения остова

На рисунке 5а показана решетка после обработки по «этапу 1» исходной решётки (рис. 2). В данном примере закрашено четырнадцать затравочных ячеек, в соответствии с описанным правилом получивших номера цветов от 39 до 238.

На «этапе 2» происходит заливка каждого непроводящего ппп-кластера максимальным номером цвета из затравочных ячеек, имеющихся в данном кластере.

Заливка непроводящих ппп-кластеров означает, что на каждом такте перекрашиваются все нечерные ячейки, которые имеют в ппп-окрестности соседей с большими номерами цветов. Этап 2 сводится к многократной обработке всей решетки, которая повторяется до тех пор, пока меняется состояние хотя бы одной ячейки в решетке. На каждом такте заливки для всех нечерных ячеек (номер цвета не равен «1») из их ппп-окрестностей выбирается наибольший номер цвета. Если этот номер больше того номера цвета, который сейчас в ячейке, то ячейка перекрашивается в цвет с наибольшим номером.

На рисунке 5б показано состояние решетки после «этапа 2». Различными номерами цветов (39, 102, 211, 216, 225, 238) помечены шесть непроводящих ппп-кластеров, соответствующих незанятым узлам. Можно заметить, что в ппп-кластере, закрашенном номером цвета 225, перекрасились затравочные ячейки с номерами цветов 67 и 98, поскольку наибольшим в данном кластере

является номер 225. То же самое имеет место для кластера, закрашенного номером цвета 238. Все занятые узлы пока по-прежнему помечены черным. Левая и правая непроводящие шины оказались закрашены разными цветами – это говорит о том, что существует перколяционный кластер.

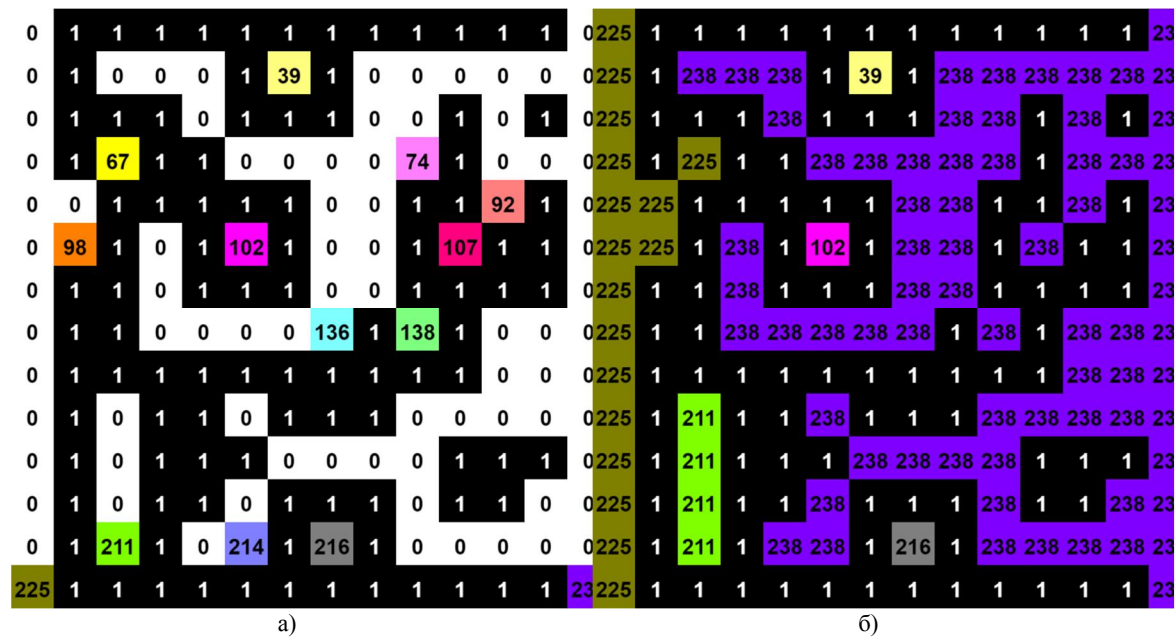


Рисунок 5 – Решетка: а) после выполнения «этапа 1»; б) после выполнения «этапа 2»

При отсутствии распараллеливания вычислений все ячейки обрабатываются одним вычислительным элементом. Если один вычислительный элемент обрабатывает больше чем одну ячейку, то работа алгоритма на «этапе 2» может быть заметно ускорена за счет выбора порядка обработки ячеек одним вычислительным элементом на одном такте. В частности, поскольку затравочные ячейки с большим номером строки и большим номером столбца имеют большие номера цветов, то более эффективно на «этапе 2» делать обработку ячеек в порядке уменьшения номеров строк и столбцов. В этом случае происходит более быстрое распространение заливки, поскольку за один такт максимальный номер цвета в непроводящем nnn -кластере может распространиться не только на соседние, но и на гораздо более дальние ячейки.

Основным результатом «этапа 2» является выяснение того, существует ли перколяционный кластер. Если в результате заливки левая и правая непроводящие шины перекрасились в один и тот же цвет, то перколяционного кластера не существует, поскольку между проводящими шинами находится непрерывный изолирующий кластер. В случае, когда перекрывающего кластера не существует, то работа алгоритма может быть закончена уже на «этапе 2». Иначе следует переход к этапу 3. Дополнительным результатом «этапа 2» является то, что висячие части перекрывающего кластера оказываются окруженными оболочками из одноцветных непроводящих узлов. Этот результат используется на следующих этапах.

Этап 3. Тробек и Стаматович предлагают разбить этот этап на два *подэтапа* (sub-steps) [31]. На первом подэтапе закрашиваются сочленяющие ячейки. Алгоритм выполняет закрашивание сочленяющих ячеек, как принадлежащих остову, так и не принадлежащих остову.

На втором подэтапе помечаются так называемые контактные пары. Контактной парой называются два проводящих nn -соседа сочленяющей ячейки, через которых сочленяющая ячейка может быть присоединена к остову.

На «этапе 3» Тробек и Стаматович предлагают использовать для анализа $nnnn$ -окрестность [24, 31]. Однако при этом они не приводят сколько-нибудь подробного описания алгоритма анализа $nnnn$ -окрестности.

В общем случае при программной реализации клеточного автомата необходимо на каждом временном шаге формировать новую копию решетки на каждом временном шаге. В нашей реализации при использовании одного вычислительного элемента для нескольких ячеек возникает необходимость делать новую копию всей решетки только на этапах 3 и 5. На «этапе 3» в копии помечаются сочленяющие ячейки, и используется анализ только nnn -окрестности. Необходимость создания копии решетки связана с тем, что при закраске сочленяющих ячеек на первом подэтапе меняется состояние решетки. Это может приводить к некорректному определению контактных пар на втором подэтапе.

Во всех проводящих черных кластерах выделяются два вида сочленяющих ячеек. Сочленяющие ячейки *первого вида* входят в висячие концы. Такие ячейки не могут принадлежать остову и закрашиваются на «этапе 3» цветом, которым залит окружающий непроводящий pnp-кластер. Более точно, закрашиваются цветом окружающего непроводящего pnp-кластера те черные ячейки, которые имеют не более трех черных np-соседей – при условии, что расположение черных ячеек и ячеек окружающего непроводящего pnp-кластера соответствует одной из конфигураций, показанных на рисунке 6, либо поворотам этих конфигураций на 90, 180 и 270 градусов.

В конфигурациях, изображенных на рисунке 6 (и далее на рис. 7, 8, 9), используются следующие метки: «1» – для проводящих ячеек черного цвета; «*» – для ячеек окружающего непроводящего кластера; «X» – для ячеек любого цвета (проводящих или непроводящих); «O» – для ячеек любого цвета, отличающегося от «*».

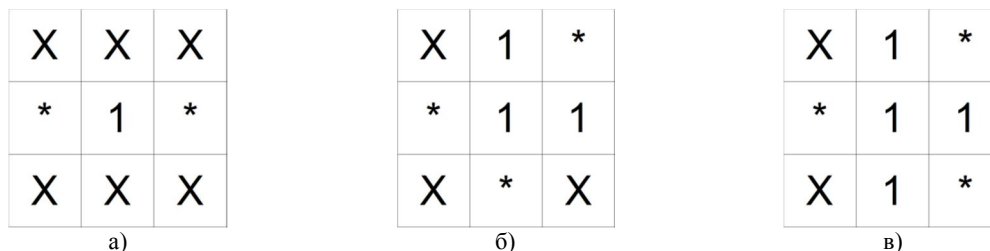


Рисунок 6 – Сочленяющие ячейки первого вида

Три изображенные на рисунке 6 конфигурации соответствуют сочленяющим ячейкам первого вида. Во всех конфигурациях центральная ячейка закрашивается цветом окружающего непроводящего кластера, поскольку эта ячейка не может принадлежать остову. После закраски цветом непроводящего кластера такие ячейки не отличаются от непроводящих.

Следует отметить, что первая из конфигураций (рис. 6а) может соответствовать не только сочленяющим ячейкам, но также висячим ячейкам (если у центральной ячейки ровно один проводящий np-сосед) или изолированным ячейкам (если у центральной ячейки нет проводящих np-соседей). Для упрощения алгоритма все висячие и изолированные ячейки также закрашиваются цветом окружающего непроводящего кластера (хотя в закраске изолированных ячеек необходимости нет).

Сочленяющие ячейки *второго вида* присоединяют висячие циклы. Для таких ячеек ответить на вопрос о принадлежности их к остову сразу нельзя. Поэтому сочленяющие ячейки второго вида помечаются цветом «0» (не используемым после второго этапа), и рядом с каждой сочленяющей ячейкой второго вида помечаются так называемые контактные пары. Цветом «0» закрашиваются те черные ячейки, которые имеют трех или четырех черных np-соседей и не менее двух pnp-соседей, относящихся к одному окружающему непроводящему кластеру при условии, что расположение черных ячеек и ячеек окружающего непроводящего pnp-кластера соответствует одной из конфигураций, показанных на рисунке 7, либо поворотам этих конфигураций на 90, 180 и 270 градусов.

Три изображенные на рисунке 7 конфигурации соответствуют сочленяющим ячейкам второго вида. Во всех конфигурациях центральная ячейка закрашивается значением «0», которое после заливки белых кластеров на «этапе 2» оказывается неиспользованным.

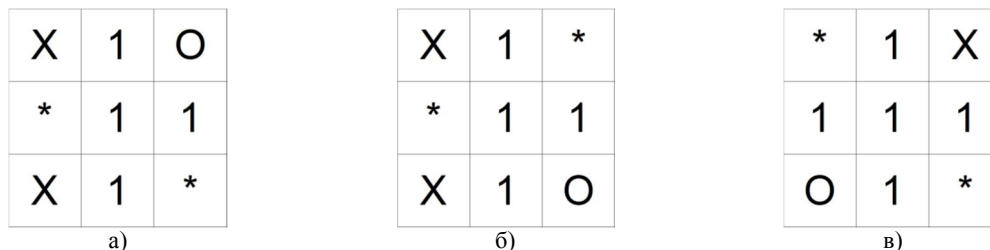


Рисунок 7 – Сочленяющие ячейки второго вида

Все ячейки, не являющиеся сочленяющими, висячими или изолированными, закрашиваются на «этапе 3» тем же цветом, что и во входной копии решетки.

Одновременно с пометкой сочленяющих ячеек на шаге 3 в нашей программе помечаются контактные пары для сочленяющих ячеек второго вида. К контактной паре принадлежат две ячейки, являющиеся одновременно np-соседями для ячейки, помеченной на рисунке 7 «O» и сочленяющей ячейки. В контактную пару может входить ячейка на шине.

Заметим, что одна ячейка может входить в более чем одну контактную пару, но этот важный

вопрос не рассмотрен Тробеком и Стаматович. Например, на рисунке 8а центральная ячейка входит в две контактные пары одновременно. Также Тробеком и Стаматович не рассмотрен важный вопрос, касающийся того, что сочленяющая ячейка может входить в контактную пару для другой сочленяющей ячейки. Например, на рисунке 8б две центральные ячейки одновременно являются сочленяющими и входят в контактные пары.

Возможны и еще более сложные сочетания вхождения в контактные пары с сочленяющими ячейками. Например, на рисунке 9 центральная ячейка входит одновременно в две контактные пары и является сочленяющей.

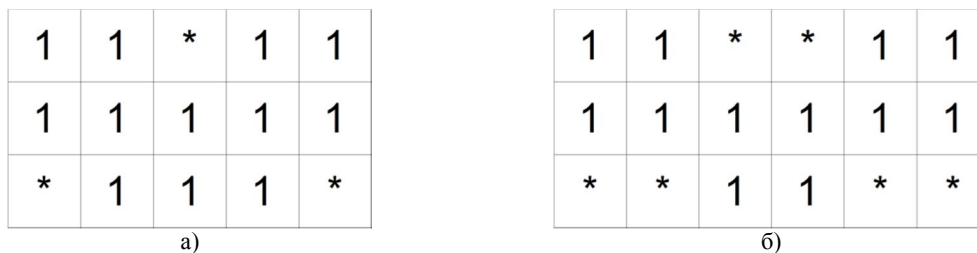


Рисунок 8 – Специфические расположения контактных пар: а) центральная ячейка входит в две контактные пары; б) две центральные ячейки одновременно являются сочленяющими и входят в контактные пары

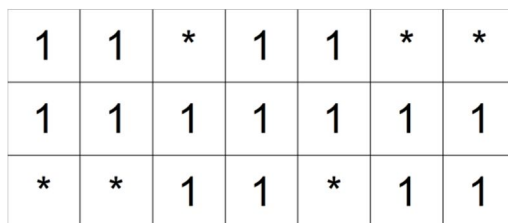


Рисунок 9 – Центральная ячейка входит в две контактные пары и одновременно является сочленяющей

В статье [24] Стаматович и Тробека упоминают о 12 случаях (вариантах) пометки контактных ячеек, но, к сожалению, не поясняют, какие именно это случаи. Исходя из приведенных выше примеров, количество возможных сочетаний (с учетом ориентации для контактных пар) может оказаться существенно больше 12. Поэтому для упрощения алгоритма в нашей программе для каждой ячейки хранится отдельная вспомогательная структура данных, которая учитывает, что одна ячейка может входить в несколько контактных пар. Далее нами приводится определение соответствующей структуры данных в виде фрагмента программного кода на языке C++ (рис. 10)

```
typedef struct
{
    char NW, SW, NE, SE;
} Pairs;
```

Рисунок 10 – Фрагмент программного кода, описывающий структуру данных

Поля каждой ячейки NW, SW, NE и SE содержат «1», если соответствующий nnn-сосед ячейки входит с ней в контактную пару и содержит «0» в противном случае. Общий размер подобной структуры в C++ составляет 4 байта, хотя, используя битовые поля, можно уложиться и в 4 бита.

На рисунке 11а показаны помеченные белым (0) сочленяющие ячейки второго вида после первого подэтапа для этапа 3. На рисунке 11б синим цветом (-1) помечены ячейки, входящие в контактные пары после второго подэтапа. Можно заметить, что на рисунке 11б имеется нечетное количество ячеек, входящих в контактные пары (15), поскольку одна из ячеек входит в две контактные пары.

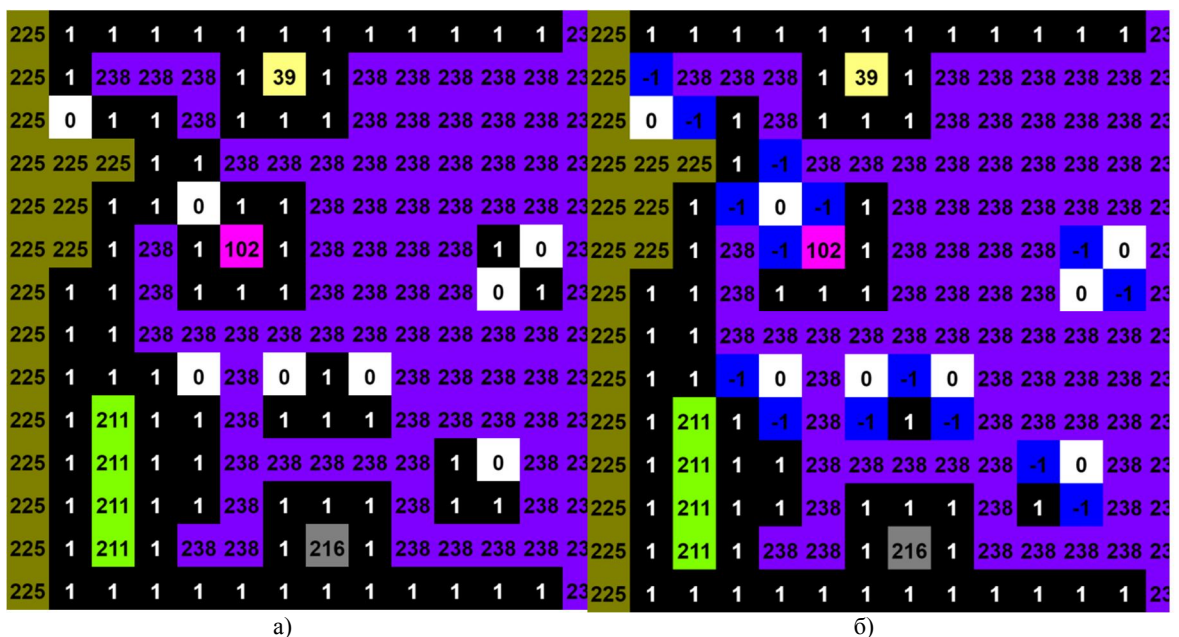


Рисунок 11 – Решетка на этапе 3: а) сочленяющие ячейки второго вида после выполнения первого подэтапа этапа 3; б) ячейки контактных пар

Наша реализация алгоритма на «этапе 3» существенно отличается от описания алгоритма, предлагаемого Тробеком и Стаматович. Поэтому приведем фрагмент кода, отвечающий за пометку контактных пар – для определенности, когда у сочленяющей ячейки второго вида есть три проводящих pnn-соседа (рис. 12).

```

if(cell[m-1]==max1&&cell[m+n+1]==max1&&cell[m-n+1]!=max1| |//пара NW иSE
cell[m+n]==max1&&cell[m-n-1]==max1&&cell[m-n+1]!=max1)
    q[m+1].NW=q[m-n].SE=1;
else if(cell[m+n]==max1&&cell[m-n+1]==max1&&cell[m-n-1]!=max1| |// -90 градусов
cell[m+1]==max1&&cell[m+n-1]==max1&&cell[m-n-1]!=max1)
    q[m-1].NE=q[m-n].SW=1;
else if(cell[m+1]==max1&&cell[m-n-1]==max1&&cell[m+n-1]!=max1| |// -180 градусов
cell[m-n]==max1&&cell[m+n+1]==max1&&cell[m+n-1]!=max1)
    q[m-1].SE=q[m+n].NW=1;
else if(cell[m-n]==max1&&cell[m+n-1]==max1&&cell[m+n+1]!=max1| |// -270 градусов
cell[m-1]==max1&&cell[m-n+1]==max1&&cell[m+n+1]!=max1)
    q[m+1].SW=q[m+n].NE=1;

```

Рисунок 12 – Фрагмент программного кода в случае, когда у сочленяющей ячейки второго вида есть три проводящих pnn-соседа

В приведенном фрагменте кода max1 содержит цвет четвертого (непроводящего) pnn-соседа. Решетка для повышения эффективности хранится в одномерном массиве cell. Одномерный массив q типа Pairs содержит для каждой ячейки информацию о том, с какими ячейками она входит в контактную пару. Алгоритмом анализируются соседи-ячейки с индексом m. При этом N, W, S, E, NW, SW, SE, NE соседи имеют соответственно индексы m-n, m-l, m+n, m+l, m-n-l, m+n-l, m+n+l, m-n+l, где n = L+4 (размер решетки с учетом дополнительных слоев).

Следующий фрагмент кода (рис. 13) обрабатывает случаи, когда у сочленяющей ячейки второго вида есть четыре проводящих pnn-соседа.

```

if(cell[m-n-1]!=1 && cell[m-n-1]==cell[m+n+1])
{
    if(cell[m+n-1]!=cell[m-n-1])
        q[m-1].SE=q[m+n].NW=1;
    if(cell[m-n+1]!=cell[m-n-1])
        q[m+1].NW=q[m-n].SE=1;
}
else if(cell[m-n+1]!=1&&cell[m-n+1]==cell[m+n-1])
{
    if(cell[m-n-1]!=cell[m-n+1])
        q[m-1].NE=q[m-n].SW=1;
    if(cell[m+n+1]!=cell[m-n+1])
        q[m+1].SW=q[m+n].NE=1;
}

```

Рисунок 13 – Фрагмент программного кода для случаев, когда у сочленяющей ячейки второго вида есть четыре проводящих pnp-соседа

На этапе 4 делается комбинированная заливка, с одновременным применением np- и pnp-правил, следующим образом. От верхней и нижней шин зеленым цветом (код цвета «-2») заливаются черные ячейки, для которых выполняется одно из следующих условий: они являются np-соседом либо одной из шин; либо они уже являются закрашенной зеленым ячейки (np-правило); либо они входят в контактную пару с уже закрашенной зеленым цветом ячейкой (pnp-правило, поскольку входящие в контактную пару ячейки являются pnp-соседями). Сами шины при этом не закрашиваются.

Так же как и на «этапе 2», «этап 4» реализуется многократным повторением описанных выше действий до тех пор, пока меняется состояние (характеристика) хотя бы одной ячейки в решетке. В статье [31] говорится о необходимости анализа pnpn-окрестностей на этапе 4. Однако предлагаемая нами реализация требует только анализа pnp-окрестностей. Поэтому приведем наш вариант реализации «этапа 4» в виде фрагмента программного кода (рис. 14).

```

int flag=1;
while(flag)
{
    flag=0;
    for(int i=2;i<n-2;i++)
        for(int j=2;j<n-2;j++)
        {
            int m=i*n+j;
            if(cell[m]==1)
            {
                if(i==(n-3) || i==2 || //отшин
                    cell[m-n]==green || cell[m-1]==green || //nn-соседи
                    cell[m+n]==green || cell[m+1]==green ||
                    cell[m-n-1]==green&&q[m].NW || //nnp-соседи
                    cell[m+n-1]==green&&q[m].SW ||
                    cell[m+n+1]==green&&q[m].SE ||
                    cell[m-n+1]==green&&q[m].NE)
                {
                    cell[m]=green;
                    flag++;
                }
            }
        }
}

```

Рисунок 14 – Фрагмент реализации программного кода для «этапа 4»

В приведенной реализации один проход для внешнего цикла while соответствует одному такту клеточного автомата. Переменная flag позволяет определить, было ли изменение характеристик (состояния) ячеек на такте. На рисунке 15а показан вид решетки после выполнения «этапа 4».

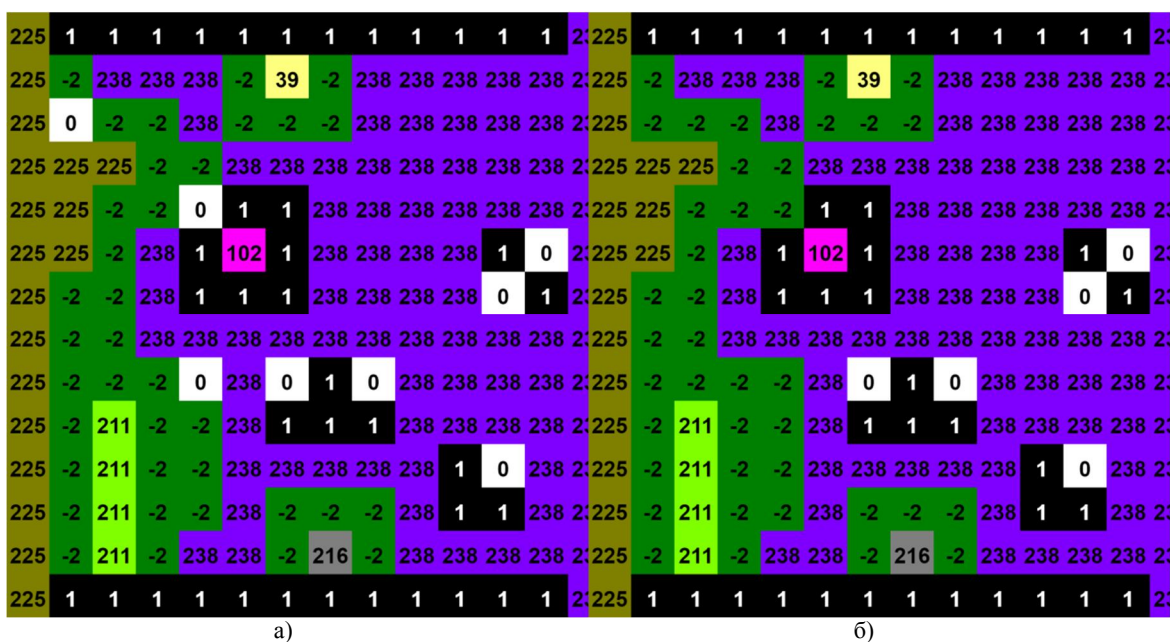


Рисунок 15 – Вид решетки после выполнения: а) «этапа 4»; б) «этапа 5»

На этапе 5 закрашиваются зеленым те ячейки, помеченные символом «0», у которых есть пп-сосед, закрашенный зеленым. Иными словами, закрашиваются сочленяющие ячейки, принадлежащие остову. Действие выполняется однократно. На рисунке 15б показан вид решетки после выполнения «этапа 5». На нем остов уже в основном определен, хотя к верхней и нижней шине могут быть прикреплены закрашенные зеленым цветом висячие дуги, не относящиеся к остову. Чтобы их отбросить, выполняются этапы 6 и 7.

На «этапе 6» от нижней шины делается пп-заливка (заливка пп-кластеров) тех ячеек, которые закрашены зеленым, синим цветом (-1). Выражаясь более точно, на «этапе 6» заливаются синим цветом те из ячеек, помеченных зеленым, которые либо прилегают к нижней шине, либо имеют пп-соседа, уже закрашенного синим цветом. В результате после закраски остова от нижней шины синим цветом этим цветом уже не будут закрашены висячие дуги, которые прикреплены к верхней шине (рис. 16а).

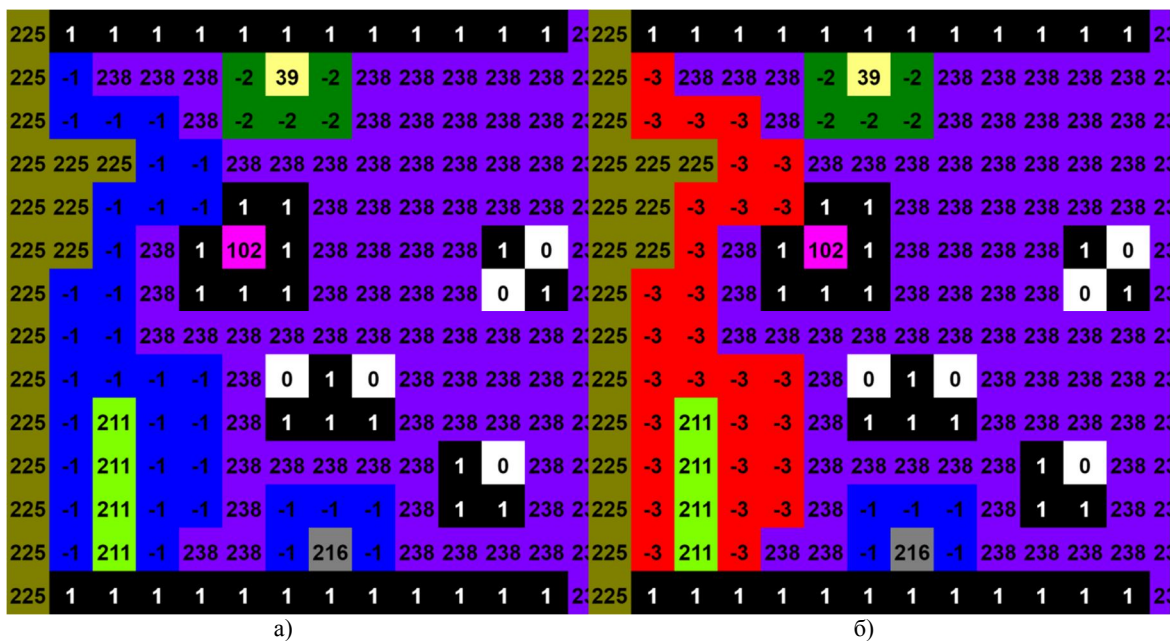


Рисунок 16 – Вид решетки после выполнения: а) «этапа 6»; б) «этапа 7»

На этапе 7 делается заливка от верхней шины красным цветом (-3) тех ячеек, которые покрашены синим цветом. В результате отбрасываются висячие дуги, которые могли быть прикреплены к нижней шине. В результате получается остов в чистом виде, покрашенный красным цветом (рис. 16б).

Результаты вычислительных экспериментов. Объём выполняемого файла написанной нами программы составляет 40 килобайт (в машинных кодах). Использовался компилятор C++ Microsoft Visual Studio 2012.

В примере заполнения решётки, который иллюстрирует этапы выполнения алгоритма на рисунках 2, 5, 11, 15, 16, из-за малого размера решётки встречаются не все варианты расположения сочленяющихся ячеек, показанные на рисунках 6–9. Для полноценного тестирования алгоритма нами специально был подготовлен файл [14] с заполнением решётки проводящими и непроводящими узлами, в котором встречаются все показанные на рисунках 6–9 случаи расположения сочленяющихся ячеек и их повороты на 90, 180 и 270 градусов. Для случаев, показанных на рисунках 8, 9, в данном файле имеются варианты крепления конфигураций ячеек к остову за угол и за середину. Для рисунка 9 в файле имеются также зеркальные конфигурации. Общий размер «случайно» заполненной части в тестовом примере $L = 40$. Однако поскольку в файле показаны шины, то общее количество строк и столбцов $L+2 = 42$. Результат обработки тестового файла на каждом этапе можно посмотреть в файле, приведенном в [13]. В графической аннотации к статье показан окончательный вариант раскраски.

Нами было протестировано среднее время выполнения алгоритма на заполненных при помощи генератора случайных чисел решетках размером $L = 200, 280, 400, 560, 800, 1120, 1600$. Шаг размера решетки был выбран так, что на каждой следующей решетке число узлов было примерно в два раза больше, чем на предыдущей. На всех размерах решеток тестировались различные доли заполнения решетки p с шагом 0,01 от 0 до 1. Для каждой доли заполнения p выполнялось по 1000 (для размеров решеток от 200 до 800) или 500 (для размеров решеток от 1120 до 1600) моделирований и определялось среднее время работы алгоритма. Тестирование проводилось на компьютере с процессором Intel Core i3-3210 с тактовой частотой 3,2 ГГц, программа задействовала одно ядро процессора. На рисунке 17 показаны пять графиков зависимости среднего времени t (в секундах) работы программы от доли заполнения решетки « p ».

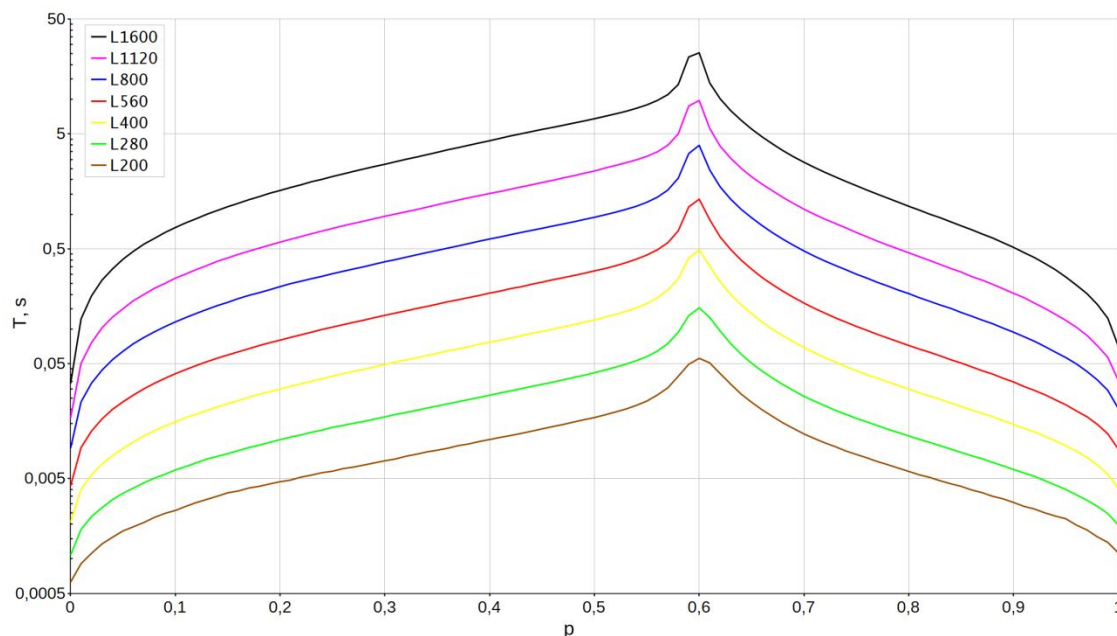


Рисунок 17 – Графики зависимости среднего времени T (в секундах) работы программы от доли заполнения решетки p для разных размеров решёток. Снизу вверх: $L = 200$ (коричневый цвет), $L = 280$ (зеленый), $L = 400$ (желтый), $L = 560$ (красный), $L = 800$ (красный), $L = 1120$ (фиолетовый), $L = 1600$ (чёрный)

Тробеком и Стаматович также делалась оценка трудоемкости алгоритма, но только для значений $0,62 \leq p \leq 0,65$ [31]. К сожалению, Тробек и Стаматович не указывают, почему был выбран именно данный диапазон вероятностей. Можно предположить, что в [31] были взяты доли заполнения решеток, для которых еще не возникало отсутствия перколяции. Нами было обнаружено при моделировании решеток с $L = 200$, что для $p = 0,61$ в 3,2 % случаев отсутствовал перколяционный

кластер, тогда как при $p = 0,62$ случаев отсутствия перколяционного кластера не встречалось.

Несмотря на то, что при $p = 0,60$ для всех размеров решеток встречались случаи отсутствия перколяционного кластера, когда алгоритм выполняет только два из семи этапов, наиболее трудоемкими оказались случаи с долей заполнения $p = 0,60$. Это значение чуть выше порога перколяции узлов на квадратной решетке, равного, согласно публикациям последних лет [12, 19, 21, 33,], с точностью до шести знаков после запятой $p = 0,592746$.

На рисунке 18 на графике с логарифмическим масштабом по обоим осям показана зависимость среднего времени расчетов «Т» от размера решетки «L» на основе экспериментальных данных для наиболее трудоемкого случая при $p = 0,60$. Нами была проделана (с использованием метода наименьших квадратов) аппроксимация трудоемкости алгоритма в зависимости от L функцией вида $T(L) \sim L^k$. В результате мы получили коэффициент $k = 3,0$, то есть $T(L) \sim L^{3,0}$. Тробек и Стаматович указывают асимптотическую сложность алгоритма вблизи порога перколяции $\Theta(L)$ при реализации на параллельных платформах [31]. Однако, поскольку Тробек и Стаматович предполагают использование отдельного вычислительного элемента для каждой ячейки, а количество ячеек растет как L^2 , то реальная трудоемкость алгоритма в расчете на один вычислительный элемент растет как $\Theta(L^3)$, хотя Тробек и Стаматович в явной форме этого не указывают. Грубая оценка трудоемкости по точкам графика, которую приводят Тробек и Стаматович для $p = 0,62$ даёт примерно $\Theta(L^{1,0})$ или, в расчете на один вычислительный элемент, $\Theta(L^{3,0})$, что совпадает с нашей оценкой.

Коэффициент детерминации для регрессии, показанной на рисунке 18, составил $R^2 = 0,9988$.

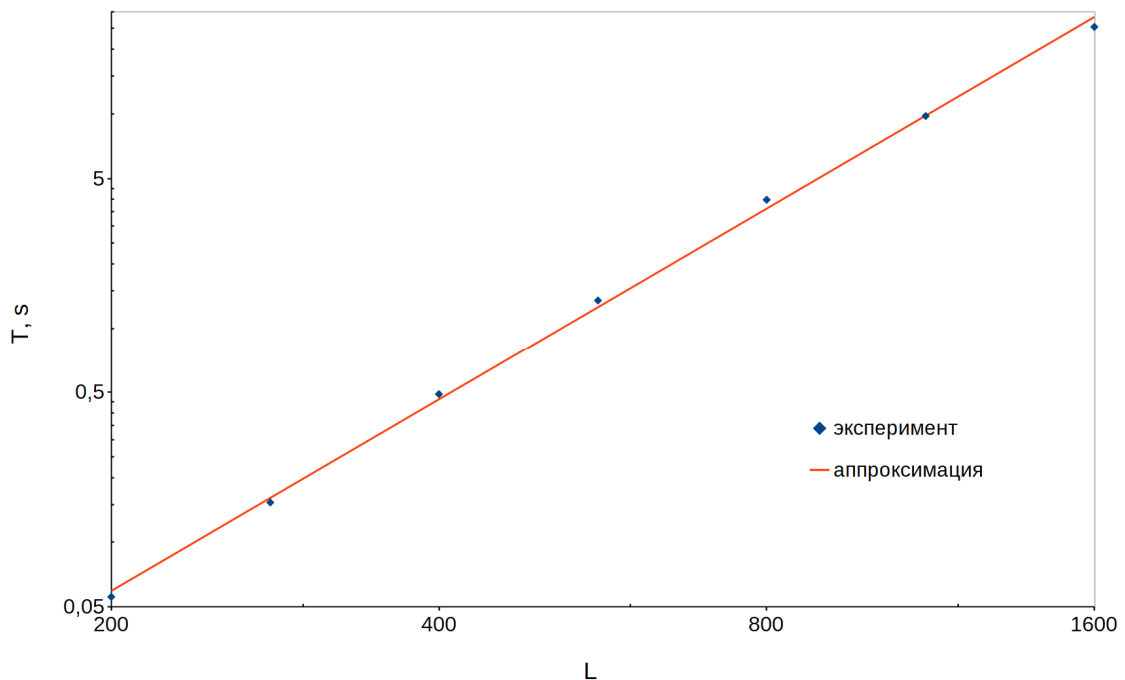


Рисунок 18 – Зависимость среднего времени T (в секундах) работы программы от размера решетки L при $p = 0,60$ (эксперимент) и аппроксимация среднего времени функцией $T(L) \sim L^k$

Заключение. Авторами был проведен анализ содержания предложенного недавно алгоритма для идентификации остова [31], основанного на заливке ячеек. Были выявлены неточности в описании алгоритма, имеющиеся в опубликованных работах. Нами предложен вариант алгоритма с исправлением этих неточностей. Кроме того, выполнена реализация алгоритма на языке программирования C++, и предложены тестовые примеры со специфическими конфигурациями ячеек.

Описанная нами реализация алгоритма на языке C++ предполагает выполнение всех вычислений на одном вычислительном элементе. Однако, как отмечают Тробек и Стаматович [31], скорость работы алгоритма должна быть существенно выше при параллельной реализации алгоритма на многих вычислительных элементах. Поэтому нами была также подготовлена параллельная реализация алгоритма с использованием интерфейса программирования MPI (Message Passing Interface). Обсуждение деталей такой реализации и её вычислительной эффективности предполагается выполнить в следующей работе.

Библиографический список

1. Алексеев В. Е. Графы и алгоритмы. Структуры данных. Модели вычислений / В. Е. Алексеев. – Москва : Интернет-университет информационных технологий ; БИНОМ. Лаборатория знаний, 2012. – 320 с.
2. Гордеев И. И. Программа для нахождения геометрического остова перколяционного кластера. Свидетельство о государственной регистрации программы для ЭВМ № 2018610093. Зарегистрировано в Реестре программ для ЭВМ 09 января 2018 года / И. И. Гордеев, А. А. Сизова.
3. Кадет В. В. Перколяционный анализ гидродинамических и электрокинетических процессов в пористых средах / В. В. Кадет. – Москва : ИНФРА-М, 2017. – 256 с.
4. Москалев П. В. Анализ структуры перколяционного кластера / П. В. Москалев // Журнал технической физики. – 2009. – Т. 79, № 6. – С. 1–7.
5. Москалев П. В. Оценки порога и мощности перколяционных кластеров на квадратных решетках с $(1, \pi)$ -окрестностью / П. В. Москалев // Компьютерные исследования и моделирование. – 2014. – Т. 6, № 3. – С. 405–414. DOI: 10.20537/2076-7633-2014-6-3-405-414.
6. Москалев П. В. Перколяционное моделирование пористых структур / П. В. Москалев. – Москва : ЛЕНАНД, 2018. – 240 с.
7. Москалев П. В. Структура моделей перколяции узлов на трехмерных квадратных решетках / П. В. Москалев // Компьютерные исследования и моделирование. – 2013. – Т. 5, № 4. – С. 607–622. DOI: 10.20537/2076-7633-2013-5-4-607-622.
8. Селяков В. И. Перколяционные модели переноса в микронеоднородных средах / В. И. Селяков, В. В. Кадет. – Москва : Недра, 1995. – 224 с.
9. Эфрос А. Л. Физика и геометрия беспорядка / А. Л. Эфрос. – Москва : Наука. Главная редакция физико-математической литературы, 1982. – 176 с.
10. Binder K. Monte Carlo simulation in statistical physics: an introduction / K. Binder, D. W. Heermann. – 4th ed. – Berlin ; New York : Springer, 2002.
11. Diestel R. Graph Theory / R. Diestel. – New York : Springer, 2000.
12. Feng X. Percolation transitions in two dimensions / X. Feng, Y. Deng, H. W. J. Blöte // Physical Review E. – 2008. – Vol. 78, 031136. DOI: 10.1103/PhysRevE.78.031136.
13. G2ii2g. Backbone-flooding/output40.txt. – Режим доступа: <https://github.com/G2ii2g/backbone-flooding/blob/master/output40.txt>, свободный. – Заглавие с экрана. – Яз. рус. (дата обращения: 27.08.2019).
14. G2ii2g. Backbone-flooding/test40.txt. – Режим доступа: <https://github.com/G2ii2g/backbone-flooding/blob/master/test40.txt>, свободный. – Заглавие с экрана. – Яз. рус. (дата обращения: 27.08.2019).
15. Gould H., Tobochnik J. An Introduction to Computer Simulation Methods: Applications to Physical Systems / H. Gould, J. Tobochnik. – Part 2. Reading. – MA: Addison-Wesley, 1987.
16. Grimmett G. Percolation / G. Grimmett // Geoffrey Grimmett. – 2nd ed. – Berlin ; Heidelberg ; New York ; Barcelona ; Hong Kong ; London ; Milan ; Paris ; Singapore ; Tokyo : Springer, 1999. (Grundlehren der mathematischen Wissenschaften; 321).
17. Harary F. Graph Theory / F. Harary. – Boca Raton ; London ; New York : CRC Press, 2018.
18. Heermann D. W. Computer Simulation Methods in Theoretical Physics / D. W. Heermann. – Berlin ; Heidelberg ; New York ; London ; Paris ; Tokyo : Springer, 1986. – 154 p.
19. Jacobsen J. L. Critical points of Potts and $O(N)$ models from eigenvalue identities in periodic Temperley-Lieb algebras / J. L. Jacobsen // Journal of Physics A: Mathematical and Theoretical. – 2015. – Vol. 48, 454003. DOI: 10.1088/1751-8113/48/45/454003.
20. Kesten H. Percolation theory for mathematicians / H. Kesten. – Boston ; Basel ; Stuttgart : Birkhauser, 1982. – 423 p.
21. Lee M. J. Pseudo-random-number generators and the square site percolation threshold / M. J. Lee // Physical Review E. – 2008. – Vol. 78, 031131. DOI: 10.1103/PhysRevE.78.031131.
22. Mutiso R. M. Electrical properties of polymer nanocomposites containing rod-like nanofillers / R. M. Mutiso, K. I. Winey // Progress in Polymer Science. – 2015. – Vol. 40. – P. 63–84. DOI: 10.1016/j.progpolymsci.2014.06.002.
23. Ore O. Theory of graphs. Providence, Rhode Island: American mathematical society / O. Ore. – 1974. – 270 p.
24. Stamatovic B. Data parallel algorithm in finding 2-D site percolation backbones / B. Stamatovic, R. Trobec // Proceeding of the First International Workshop on Sustainable Ultrascale Computing Systems (NESUS 2014). – Porto, Portugal ; Madrid : Computer Architecture Communications, and System Group (ARCOS) University Carlos III, 2014. – P. 65–70.
25. Stauffer D. Introduction to Percolation Theory / D. Stauffer, A. Aharony. – 2nd ed., rev. – London : Taylor & Francis, 2003. – 180 p.
26. Tarasevich Yu. Yu. Percolation of linear k-mers on a square lattice: From isotropic through partially ordered to completely aligned states / Yu. Yu. Tarasevich, N. I. Lebovka, V. V. Laptev // Physical Review E. – 2012. – Vol. 86, 061116. DOI: 10.1103/PhysRevE.86.061116.
27. Tarasevich Yu. Yu., Lebovka N. I., Vodolazskaya I. V. et al. Anisotropy in electrical conductivity of two-dimensional films containing aligned nonintersecting rodlike particles: Continuous and lattice models // Physical Review E. 2018. Vol. 98, 012105. DOI: 10.1103/PhysRevE.98.012105.
28. Tarasevich Yu. Yu. Identification of current-carrying part of a random resistor network: electrical approaches vs. graph theory algorithms / Yu. Yu. Tarasevich, A. S. Burmistrov, V. A. Goltseva, I. I. Gordeev, V. I. Serbin,

A. A. Sizova, I. V. Vodolazskaya, D. A. Zholobov // Journal of Physics: Conference Series. – 2018. – Vol. 955, 012021. DOI:10.1088/1742-6596/955/1/012021.

29. Toffoli T., Margolus N. Cellular Automata Machines: A new environment for modeling / T. Toffoli, N. Margolus. – Cambridge, Massachusetts ; London, England : The MIT Press, 1987.

30. Torbert S. Applied Computer Science / S. Torbert. – 2nd ed. – Springer, 2016. – 279 p.

31. Trobec R. Analysis and classification of flow-carrying backbones in two-dimensional lattices / R. Trobec, B. Stamatovic // Advances in Engineering Software. – 2017. – Vol. 103. – P. 38–45. DOI: 10.1016/j.advengsoft.2015.11.002.

32. Wilson R. J. Introduction to Graph Theory / R. J. Wilson. – 4th ed. – Edinburgh Gate, Harlow : Addison Wesley Longman Limited, 1998.

33. Yang Y. Square++: Making a connection game win-lose complementary and playing-fair / Y. Yang, S. Zhou, Y. Li // Entertainment Computing. – 2013. – Vol. 4. – P. 105–113. DOI: 10.1016/j.entcom.2012.10.004.

34. Yin W.-G. Algorithm for finding two-dimensional site percolation backbones / W.-G. Yin, R. Tao // Physica B: Condensed Matter. – 2000. – Vol. 279. – P. 84–86. DOI: 10.1016/S0921-4526(99)00675-4.

35. Yin W.-G. Rapid algorithm for identifying backbones in the two-dimensional percolation model / W.-G. Yin, R. Tao // International Journal of Modern Physics C. – 2003. – Vol. 14. – P. 1427–1437. DOI: 10.1142/S0129183103005509.

References

1. Alekseev V. E. *Grafy i algoritmy. Struktury dannykh. Modeli vychisleniy* [Graphs and algorithms. Data structures. Computation Models]. Moscow, 2012. 320 p. ISBN 978-5-94774-543-6.

2. Gordeev I. I., Sizova A. A. *Programma dlya nakhozhdeniya geometricheskogo ostova perkolyacionnogo klastera Svidetelstvo o gosudarstvennoy registratsii programmy dlya EVM № 2018610093. Zaregistrirvano v Reestre programm dlya EVM 09 yanvarya 2018 goda* [The program for finding the geometric backbone of a percolation cluster. Certificate of state registration of computer programs No. 2018610093. Registered in the Register of computer programs on January 9, 2018].

3. Kadet V. V. *Perkolyatsionnykh analiz gidrodinamicheskikh i elektrokineticheskikh protsessov v poristyykh sredakh* [Percolation analysis of hydrodynamic and electrokinetic processes in porous media]. Moscow, 2017. 256 p.

4. Moskalev P. V. Analiz struktury perkolyacionnogo klastera [Analysis of the percolation cluster structure]. *Zhurnal tekhnicheskoy fiziki* [Technical Physics], 2009, vol. 79, no. 6, pp. 1–7.

5. Moskalev P. V. Otsenki poroga i moshchnosti perkolyatsionnykh klasterov na kvadratnykh reshetkakh s $(1, \pi)$ -okrestnostyu [Estimates of the threshold and power of percolation clusters on square lattices with a $(1, \pi)$ neighborhood]. *Kompyuternye issledovaniya i modelirovaniya* [Computer Research and Modeling], 2014, vol. 6, no. 3, pp. 405–414. DOI: 10.20537/2076-7633-2014-6-3-405-414.

6. Moskalev P. V. *Perkolyatsionnoe modelirovanie poristyykh struktur* [Percolation modeling of porous structures]. Moscow, 2018. 240 p.

7. Moskalev P. V. Struktura modeley perkolyatsii uzlov na trekhmernyykh kvadratnykh reshetkakh [The structure of site percolation models on three-dimensional square lattices]. *Kompyuternye issledovaniya i modelirovanie* [Computer Research and Modeling], 2013, vol. 5, no. 4, pp. 607–622. DOI: 10.20537/2076-7633-2013-5-4-607-622.

8. Selyakov V. I., Kadet V. V. *Perkolyatsionnyye modeli perenosa v mikroneodnorodnykh sredakh* [Percolation Models for Transport in Porous Media]. Moscow, Nedra Publ., 1995. 224 p.

9. Efros A. L. *Fizika i geometriya besporyadka* [Physics and Geometry of Disorder]. Moscow, Nauka. Glavnaya redaktsiya fiziko-matematicheskoy literatury Publ., 1982. 176 p.

10. Binder K., Heermann D. W. *Monte Carlo simulation in statistical physics: an introduction*. 4th ed. Berlin; New York, Springer, 2002.

11. Diestel R. *Graph Theory*. New York, Springer, 2000.

12. Feng X., Deng Y., Blöte H. W. J. Percolation transitions in two dimensions. *Physical Review E*, 2008, vol. 78, 031136. DOI: 10.1103/PhysRevE.78.031136.

13. G2ii2g. *Backbone-flooding/output40.txt*. Available at: <https://github.com/G2ii2g/backbone-flooding/blob/master/output40.txt>.

14. G2ii2g. *Backbone-flooding/test40.txt*. Available at: <https://github.com/G2ii2g/backbone-flooding/blob/master/test40.txt>.

15. Gould H., Tobochnik J. *An Introduction to Computer Simulation Methods: Applications to Physical Systems*. Part 2. Reading, MA, Addison-Wesley, 1987.

16. Grimmet G. *Percolation*. 2nd ed. Berlin; Heidelberg; New York; Barcelona; Hong Kong; London; Milan; Paris; Singapore; Tokyo, Springer, 1999. (Grundlehren der mathematischen Wissenschaften; 321).

17. Harary F. *Graph Theory*. Boca Raton, London, New York, CRC Press, 2018.

18. Heermann D. W. *Computer Simulation Methods in Theoretical Physics*. Berlin; Heidelberg; New York; London; Paris; Tokyo, Springer, 1986. 154 p.

19. Jacobsen J. L. Critical points of Potts and $O(N)$ models from eigenvalue identities in periodic Temperley–Lieb algebras. *Journal of Physics A: Mathematical and Theoretical*, 2015, vol. 48, 454003. DOI: 10.1088/1751-8113/48/45/454003.

20. Kesten H. *Percolation theory for mathematicians*. Boston; Basel; Stuttgart, Birkhauser, 1982. 423 p.

21. Lee M. J. Pseudo-random-number generators and the square site percolation threshold. *Physical Review E*, 2008, vol. 78, 031131. DOI: 10.1103/PhysRevE.78.031131.

22. Mutiso R. M., Winey K. I. Electrical properties of polymer nanocomposites containing rod-like nanofillers. *Progress in Polymer Science*, 2015, vol. 40, pp. 63–84. DOI: 10.1016/j.progpolymsci.2014.06.002.
23. Ore O. *Theory of graphs*. Providence, Rhode Island: American mathematical society, 1974. 270 p.
24. Stamatovic B., Trobec R. Data parallel algorithm in finding 2-D site percolation backbones. *Proceeding of the First International Workshop on Sustainable Ultrascale Computing Systems (NESUS 2014)*. Porto, Portugal; Madrid, Computer Architecture Communications, and System Group (ARCOS) University Carlos III, 2014, pp. 65–70.
25. Stauffer D., Aharony A. *Introduction to Percolation Theory*. 2nd ed., rev. London, Taylor & Francis, 2003. 180 p.
26. Tarasevich Yu. Yu., Lebovka N. I., Laptev V. V. Percolation of linear k-mers on a square lattice: From isotropic through partially ordered to completely aligned states. *Physical Review E*, 2012, vol. 86, 061116. DOI: 10.1103/PhysRevE.86.061116.
27. Tarasevich Yu. Yu., Lebovka N. I., Vodolazskaya I. V. et al. Anisotropy in electrical conductivity of two-dimensional films containing aligned nonintersecting rodlike particles: Continuous and lattice models. *Physical Review E*. 2018. Vol. 98, 012105. DOI: 10.1103/PhysRevE.98.012105.
28. Tarasevich Yu. Yu., Burmistrov A.S., Goltseva V.A., Gordeev I.I., Serbin V.I., Sizova A.A., Vodolazskaya I.V. and Zholobov D.A. Identification of current-carrying part of a random resistor network: electrical approaches vs. graph theory algorithms. *Journal of Physics: Conference Series*, 2018, vol. 955, 012021. DOI:10.1088/1742-6596/955/1/012021.
29. Toffoli T., Margolus N. *Cellular Automata Machines: A new environment for modeling*. Cambridge, Massachusetts; London, England: The MIT Press, 1987.
30. Torbert S. *Applied Computer Science*. 2nd ed. Springer, 2016. 279 p.
31. Trobec R., Stamatovic B. Analysis and classification of flow-carrying backbones in two-dimensional lattices. *Advances in Engineering Software*, 2017, vol. 103, pp. 38–45. DOI: 10.1016/j.advengsoft.2015.11.002.
32. Wilson R. J. *Introduction to Graph Theory*. 4th ed. Edinburgh Gate, Harlow, Addison Wesley Longman Limited, 1998.
33. Yang Y., Zhou S., Li Y. Square++: Making a connection game win-lose complementary and playing-fair. *Entertainment Computing*, 2013, vol. 4, p. 105–113. DOI: 10.1016/j.entcom.2012.10.004.
34. Yin W.-G., Tao R. Algorithm for finding two-dimensional site percolation backbones. *Physica B: Condensed Matter*, 2000, vol. 279, pp. 84–86. DOI: 10.1016/S0921-4526(99)00675-4.
35. Yin W.-G., Tao R. Rapid algorithm for identifying backbones in the two-dimensional percolation model. *International Journal of Modern Physics C*, 2003, vol. 14, pp. 1427–1437. DOI: 10.1142/S0129183103005509.

УДК 621.548.4

РОТОР ДАРЬЕ И ОПТИМИЗИРОВАННЫЕ ВЕТРОТУРБИНЫ: СРАВНИТЕЛЬНЫЙ АНАЛИЗ

Статья поступила в редакцию 23.02.2020, в окончательном варианте – 04.03.2020.

Роткин Владимир Михайлович, НТА «Экологический императив», 3321214, Израиль, г. Хайфа, ул. Герцль, 63 алэф,
кандидат технических наук, Ph.D, доцент, e-mail: ricens@mail.ru

Соколовский Юлий Борисович, НТА «Экологический императив», 3321214, Израиль, г. Хайфа, ул. Герцль, 63 алэф,
кандидат технических наук, Ph.D, старший научный сотрудник, e-mail: sokol1937y@gmail.com

Ажмухамедов Искандар Маратович, Астраханский государственный университет, 414056, Российская Федерация, г. Астрахань, Татищева, 20а,
доктор технических наук, профессор, заведующий кафедрой информационной безопасности, e-mail: iskander_agm@mail.ru

Рассматриваются возможности повышения энергопоказателей эффективности вертикально-осевых ветроэнергетических установок как альтернативных источников энергии. При значительном потенциале масштабирования вертикально-осевых ветротурбин научно-техническая информация о них в открытых источниках существенно ограничена. Предлагается методология оценки энергоэффективности ортогональных роторов на основе сравнительного анализа путем сопоставления расчетных моделей различных типов роторов с эталонной оптимизированной ветротурбиной. Методология представлена на примере анализа энергетического потенциала ротора Дарье. Адекватность методологии подтверждается практикой применения базовых аэродинамических зависимостей в расчетных моделях как в ветроэнергетике, так и в смежных отраслях. Модель взаимодействия воздушного потока с крыльчатой лопастью основана на допущении о суперпозиции лобовых сил и подъемной силы. Энергия, извлекаемая из взаимодействующего с лопастью воздушного потока, равна работе приложенных к лопасти сил на перемещении этой лопасти за полный оборот турбины. Исследование этой зависимости на экстремум даёт оптимальную конфигурацию лопастей. Расчетная модель

взаимодействия воздушного потока с ротором Дарье рассматривается на примере симметричной лопасти с круглой лобовой поверхностью и хвостовой поверхностью треугольного сечения. Динамическая ось, проходящая через лобовую точку и вершину хвостовой части, разделяет лопасть Дарье на две несимметричные части, образующие динамическое крыло с переменной подъемной силой. Коэффициенты отбора энергии и мощности воздушного потока определяются аналогично оптимизированной крыльчатой лопасти. Сравнительный анализ выполнен на основе аэродинамического взаимодействия рабочих органов установки с воздушным потоком с учетом эффекта рассеяния потока под воздействием рабочих органов как преграды. Сопоставительный расчет показывает, что при кратном превосходстве оптимизированной турбины по вращательному моменту мощность равномасштабной установки Дарье сопоставима с оптимизированной турбиной.

Ключевые слова: ветровая турбина, лопасть, крыло, профиль, подъемная сила, лобовое сопротивление, ротор Дарье, оптимизация, энергоэффективность, сравнительный анализ

Графическая аннотация (Graphical annotation)



DARIER ROTOR AND OPTIMIZED WIND TURBINES: COMPARATIVE ANALYSIS

The article was received by the editorial board on 23.02.2020, in the final version – 04.03.2020.

Rotkin Vladimir M., NTA "Ecology Imperative", 63 Alef Herzl St., Haifa, 3321214, Israel, Ph.D., Associate Professor, e-mail ricentr@mail.ru

Sokolovsky Yuly B., NTA "Ecology Imperative", 63 Alef Herzl St., Haifa, 3321214, Israel, Ph.D., Senior Researcher, e-mail sokol1937y@gmail.com

Azhmukhamedov Iskandar M., Astrakhan State University, 20a Tatishchev St., Astrakhan, 414056, Russian Federation,

Doct. Sci. (Engineering), Professor, e-mail iskander_agm@mail.ru

The possibilities of increasing the energy indicators of the efficiency of vertical-axis wind power plants as alternative energy sources are considered. With significant potential for scaling vertical-axis wind turbines, scientific and technical information about them in open sources is significantly limited. A methodology is proposed for assessing the energy efficiency of orthogonal rotors based on a comparative analysis by comparing design models of various types of rotors with a reference optimized wind turbine. The methodology is presented as an example of the analysis of the energy potential of the Darier rotor. The adequacy of the methodology is confirmed by the practice of using basic aerodynamic dependencies in design models, both in wind energy and in related industries. The interaction model of the airflow with the winged blade is based on the assumption of a superposition of frontal forces and lift. The energy extracted from the airflow interacting with the blade is equal to the work of the forces applied to the blade to move this blade for a full revolution of the turbine. Examination of this dependence by the extremum gives the optimal configuration of the blades. The calculation model of the interaction of the airflow with the Darier rotor is considered on the example of symmetrical blade with a circular frontal surface and a tail surface of a triangular section. The dynamic axis passing through the frontal point and the apex of the tail part divides the Darier blade into two asymmetric parts, forming a dynamic wing with variable lifting force. Energy and airflow power take-offs are determined similarly to an optimized winged blade. A comparative analysis is considered on the basis of the aerodynamic interaction of the working bodies of the installation with the airflow, taking into account the effect of scattering of the flow under the influence of the working bodies as an obstacle. A comparative calculation shows that with a multiple superiority of the optimized turbine in terms of torque, the power of the Darier equal-scale installation is comparable to the optimized turbine.

Key words: wind turbine, blade, wing, profile, lifting force, drag, Darier rotor, optimization, energy efficiency, comparative analysis

Введение. Для использования энергии ветра разработано множество типов ветроэнергетических установок (ВЭУ) [1–3, 5, 8]. Как правило, ВЭУ классифицируются по расположению оси ротора ветротурбины (горизонтально- или вертикально-осевые) а также ее ориентации по отношению к направлению воздушного потока – коллинеарные (ось параллельна воздушному потоку) или ортогональные (ось перпендикулярна потоку). Наиболее распространены ВЭУ традиционной компоновки – горизонтально-осевые коллинеарные (рис. 1а). В настоящее время они составляют более 90 % парка ВЭУ. По ним накоплен значительный опыт конструирования, изготовления и эксплуатации. Такие ВЭУ характеризуются применением передовых технологий, высокой энергоэффективностью. Однако, учитывая, что экономическая эффективность ветроэнергостанций зависит от мощности единичной ВЭУ, следует отметить, что потенциал масштабирования таких установок близок к исчерпанию. Кроме того, при больших скоростях оборотов они издают значительный шум, приводят к сокращению сроков использования подшипников роторов и пр.



Рисунок 1 – Основные типы ветроэнергетических установок: а – горизонтально-осевая; б – вертикально-осевая, ротор Дарье

Масштабы и опыт применения вертикально-осевых ВЭУ (рис. 1б) значительно ниже. Информация о показателях их энергоэффективности крайне противоречива, мало верифицированных данных о результатах исследований и испытаний подобных устройств [4, 9–12]. Традиционно считалось [8, 9], что энергоэффективность таких ВЭУ существенно уступает горизонтально-осевым. В то же время, производители этих установок в своих проспектах часто приводят показатели, не уступающие и даже превосходящие характеристики для горизонтально-осевых ВЭУ.

Независимо от указанных противоречий, многие специалисты и эксперты сходятся в том, что вертикально-осевые ВЭУ, в силу своей компоновки, обладают лучшими динамическими качествами (и, как следствие, показателями надежности), а также перспективами масштабирования (рис. 2).

В подавляющем большинстве действующих вертикально-осевых ВЭУ применяются нерегулируемые лопасти (рис. 1б, 2а), чаще всего роторы Дарье [1–3]. Применение регулируемых поворотных лопастей (рис. 2б, в) способно существенно повысить энергоэффективность ВЭУ [10, 12]. Однако это приводит к усложнению конструкции, которое становится экономически оправданным только с ростом мощности единичной ВЭУ.

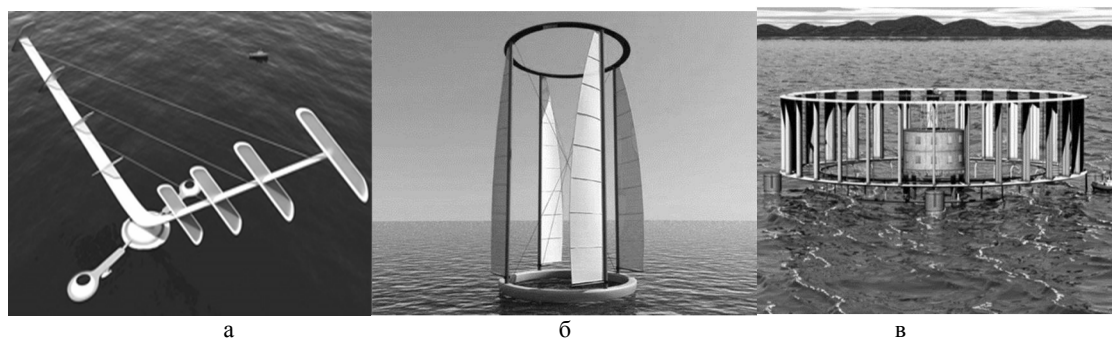


Рисунок 2 – Масштабные вертикально-осевые ветроустановки: а – со стационарно расположенными лопастями; б, в – с регулируемым положением лопастей

В настоящей работе возможности совершенствования вертикально-осевых ВЭУ рассматриваются путем сопоставления показателей энергоэффективности оптимизированной ортогональной турбины – вертикально-осевой ВЭУ, работающей в режиме оптимальной конфигурации лопастей, и неоптимизированной установки на основе ротора Дарье. Оптимизированная турбина выступает, по существу, в качестве эталонного ортогонального ветродвигателя. Адекватность данной методологии подтверждается многолетней практикой применения базовых аэродинамических зависимостей в расчетных моделях как в ветроэнергетике, так и в смежных отраслях.

Оптимизированная ортогональная турбина. Модель взаимодействия воздушного потока с крыльчатой лопастью основана на допущении о суперпозиции лобовых сил и подъемной силы [12]. Профиль линейно-выпуклой крыльчатой лопасти симметричен относительно поперечной оси и ассиметричен относительно продольной.

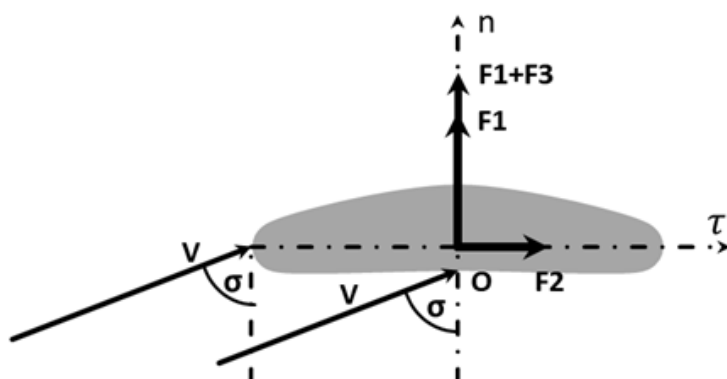


Рисунок 3 – Схема взаимодействия линейно-выпуклой лопасти с воздушным потоком

Сила поперечного лобового воздействия воздушного потока на плоское основание лопасти направлена по нормали On к поперечному сечению лопасти (рис. 3) и равна [6, 12]

$$F_1 = c_{fn} S_n \cos \sigma \frac{\rho V^2}{2}. \quad (1)$$

Соответствующая поперечная подъемная сила отсутствует в силу симметрии лопасти.

При продольном взаимодействии сила лобового воздействия направлена вдоль $O\tau$ и равна

$$F_2 = c_{f\tau} S_\tau \sin \sigma \frac{\rho V^2}{2}, \quad (2)$$

а подъемная сила – вдоль On и равна

$$F_3 = c_{l\tau} S_\tau \sin \sigma \frac{\rho V^2}{2}, \quad (3)$$

где $c_{fn}, c_{f\tau}$ – коэффициенты лобового сопротивления лопасти в поперечном (On) и продольном ($O\tau$) направлениях; $c_{l\tau}$ – коэффициент подъемной силы; S_n, S_τ – площади поперечного (ортогонального On) и продольного (ортогонального $O\tau$) сечений лопасти; σ – угол, образованный вектором относительной скорости потока с осью симметрии On ; ρ – плотность воздуха; V – относительная скорость воздушного потока.

Если соотношение скорости воздушного потока и линейной окружной скорости лопасти (рис. 4) задать параметром окружной скорости $k = V_e/V_0$, то угол $\sigma = (\vec{n}; \vec{V})$ определяется из соотношения

$$\cos(\vec{n}; \vec{V}) = \frac{\cos\varphi - k\sin(\varphi - \alpha)}{\sqrt{1 + k^2 + 2k\sin\alpha}} \quad (4)$$

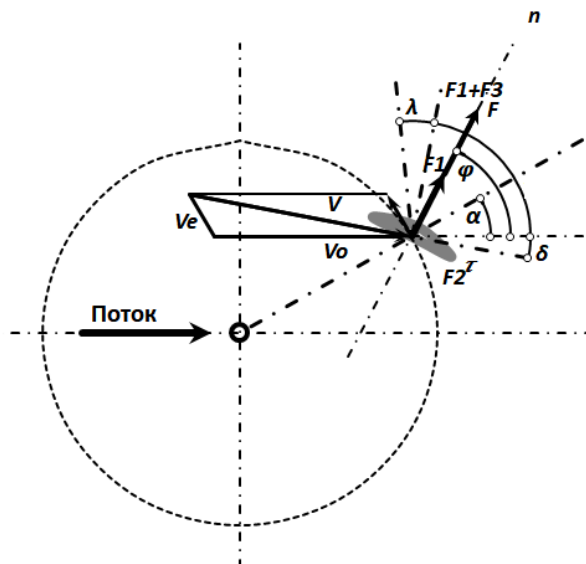


Рисунок 4 – Схема взаимодействия воздушного потока с регулируемой лопастью ортогональной ветротурбины. Углы ориентации: турбины – α , лопасти – φ

Приложенные силы воздушного потока преобразуются к виду

$$F_1 = F_0 \sqrt{1 + k^2 + 2k\sin\alpha} [\cos\varphi - k\sin(\varphi - \alpha)], \quad (5)$$

$$F_2 = s_2 F_0 \sqrt{1 + k^2 + 2k\sin\alpha} [\sin\varphi + k\cos(\varphi - \alpha)], \quad (6)$$

$$F_3 = s_3 F_0 \sqrt{1 + k^2 + 2k\sin\alpha} [\sin\varphi + k\cos(\varphi - \alpha)], \quad (7)$$

где

$$F_0 = c_{fn} S_n \rho V_0^2 / 2 \quad (8)$$

– приведенная сила действия потока (сила поперечного лобового действия, приложенная к неподвижной пластине со стороны нормально направленного потока),

$$s_2 = \frac{c_{f\tau} S_\tau}{c_{fn} S_n} \quad (9)$$

– приведенный коэффициент продольного лобового сопротивления,

$$s_3 = \frac{c_{l\tau} S_\tau}{c_{fn} S_n} \quad (10)$$

– приведенный коэффициент подъемной силы.

Соответственно, главный момент приложенных сил

$$M_\alpha = F_1 L \sin(\varphi - \alpha) - F_2 L \cos(\varphi - \alpha) + F_3 L \sin(\varphi - \alpha). \quad (11)$$

Величина энергии, извлеченной из взаимодействующего с лопастью воздушного потока за полный оборот турбины, равна работе приложенных к лопасти сил на перемещении этой лопасти, или работе момента этих сил на соответствующем угловом перемещении

$$A_0 = \int_0^{2\pi} M_\alpha d\alpha. \quad (12)$$

Функционирование ортогональной ветротурбины возможно только при определенном расположении лопастей по отношению к воздушному потоку. Анализ оптимизационной модели взаимодействия лопасти с воздушным потоком позволяет определить кинематику лопасти, обеспечивающую максимальный съем энергии воздушного потока.

Исследование зависимости $a_0(\varphi)$ на экстремум дает возможность выявить оптимальную ориентацию лопастей ортогональной ветротурбины, в зависимости от расположения их осей, то есть определить оптимальное соотношение $\varphi = \varphi(\alpha)$. Условие экстремума $da_\alpha/d\varphi = 0$ после дифференцирования и последующих преобразований дает соотношение

$$\operatorname{tg}(2\varphi - \alpha) = \frac{(1 + k\sin\alpha)}{k\cos\alpha} \pm \frac{s_3}{1 - s_2} / \left(1 \mp \frac{1 + k\sin\alpha}{k\cos\alpha} \frac{s_3}{1 - s_2}\right). \quad (13)$$

Подстановка соответствующих значений φ дает оптимальные значения коэффициентов отбора энергии и мощности воздушного потока в зависимости от параметра окружной скорости k :

$$a(k) = A_0/2\pi F_0 L, \quad (14)$$

$$n(k) = ka(k). \quad (15)$$

Соответственно, усредненная за один оборот (циклическая) абсолютная величина отбора мощности воздушного потока

$$N = nF_0 V_0. \quad (16)$$

Ротор Дарье. Главное преимущество ротора Дарье – существенная простота конструкции, связанная с жестким креплением лопастей на радиальных штангах или непосредственно на оси ротора. Такое крепление полностью исключало бы возникновение вращательного момента ротора, если бы не особая форма применяемых лопастей (рис. 5).

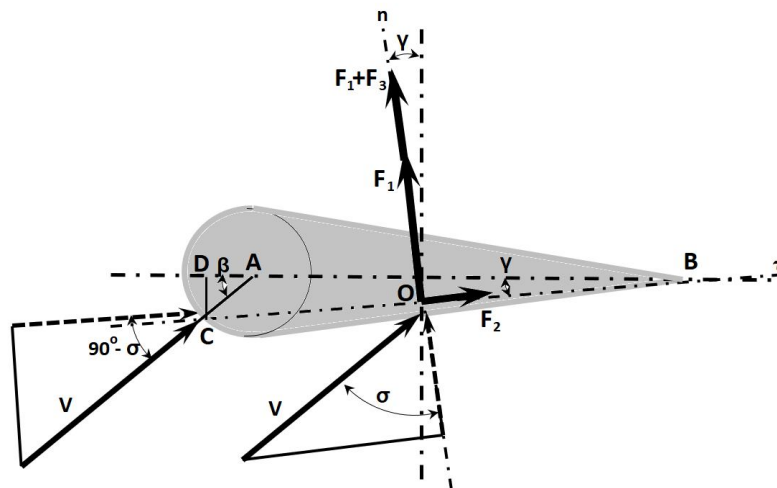


Рисунок 5 – Схема взаимодействия симметричной лопасти Дарье с воздушным потоком

Механизм и расчетная модель взаимодействия воздушного потока с ротором Дарье рассматривается на примере симметричной лопасти с лобовой поверхностью в виде кругового цилиндра радиуса $AC = r$ и хвостовой поверхностью треугольного сечения высотой $AB = b$ с основанием $2r$. Ось симметрии лопасти проходит через центр лобового сечения A и вершину B .

Центр лобового взаимодействия лопасти с потоком определяется как точка C , в которой вектор скорости потока направлен по нормали к лобовой поверхности лопасти. При этом прямая CB , проходящая через лобовую точку и вершину хвостовой части, образует динамическую ось, разделяющую лопасть на две несимметричные части, образующие динамическое крыло с переменной подъемной силой

$$c_{l\tau} = c_{l\tau 0} \sin^2 \beta. \quad (17)$$

Эксцентриситет динамической оси определяется углом γ , который зависит от направления воздушного потока по отношению к лопасти. Соответственно, сила нормального лобового сопротивления F_1 и подъемная сила F_3 направлены по нормали, перпендикулярно динамической оси лопасти, и образуют с вектором скорости потока угол

$$\sigma = 90^\circ - \beta + \gamma. \quad (18)$$

С учетом особенностей конфигурации ротора Дарье (рис. 6) силы взаимодействия потока с лопастями рассчитываются по формулам

$$F_1 = F_0 \sqrt{1 + k^2 + 2k \sin \alpha} [\cos(\alpha + \gamma) - k \sin \gamma], \quad (19)$$

$$F_2 = s_2 F_0 \sqrt{1 + k^2 + 2k \sin \alpha} [\sin(\alpha + \gamma) + k \cos \gamma], \quad (20)$$

$$F_3 = s_3 F_0 \sqrt{1 + k^2 + 2k \sin \alpha} [\sin(\alpha + \gamma) + k \cos \gamma], \quad (21)$$

где угол γ определяется соотношением (рис. 5)

$$\operatorname{tg} \gamma = (r/b) \sin \beta / ((r/b) \cos \beta + 1), \quad (22)$$

а угол β , образуемый вектором скорости с осью симметрии лопасти, связан с углом поворота ротора соотношением

$$\sin \beta = \frac{\cos \alpha}{\sqrt{1 + k^2 + 2k \sin \alpha}}. \quad (23)$$

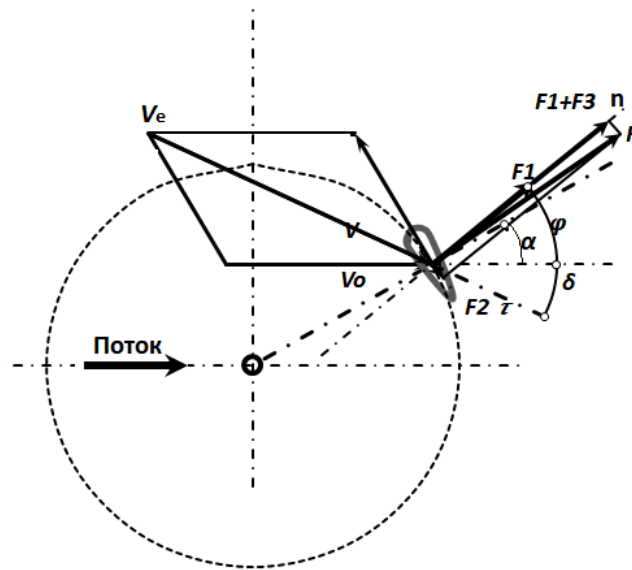


Рисунок 6 – Схема взаимодействия воздушного потока со стационарной лопастью Дарье. Углы ориентации: турбины – α , лопасти – φ

Главный момент, по аналогии с регулируемым ротором, определяется выражением

$$M_\alpha = F_1 L \sin \gamma - F_2 L \cos \gamma + F_3 L \sin \gamma. \quad (24)$$

Циклический момент, коэффициенты отбора энергии и мощности воздушного потока и циклическая величина отбора мощности также определяются аналогично, с использованием расчетных выражений (14)–(16).

Вопросы внешнего регулирования для ротора Дарье не актуальны, в связи со стационарным размещением и конфигурацией рабочих органов. Однако в существенно узком диапазоне происходит саморегулирование силового воздействия потока на лопасть за счет варьирования динамического эксцентриситета.

Сравнительный анализ энергоэффективности ротора Дарье и оптимизированной ветротурбины. Результаты расчетов показателей энергоэффективности – коэффициентов отбора энергии и мощности потока (рис. 7) показывают, что по максимальной доле извлекаемой мощности ротор Дарье кратно уступает оптимизированной ветротурбине.

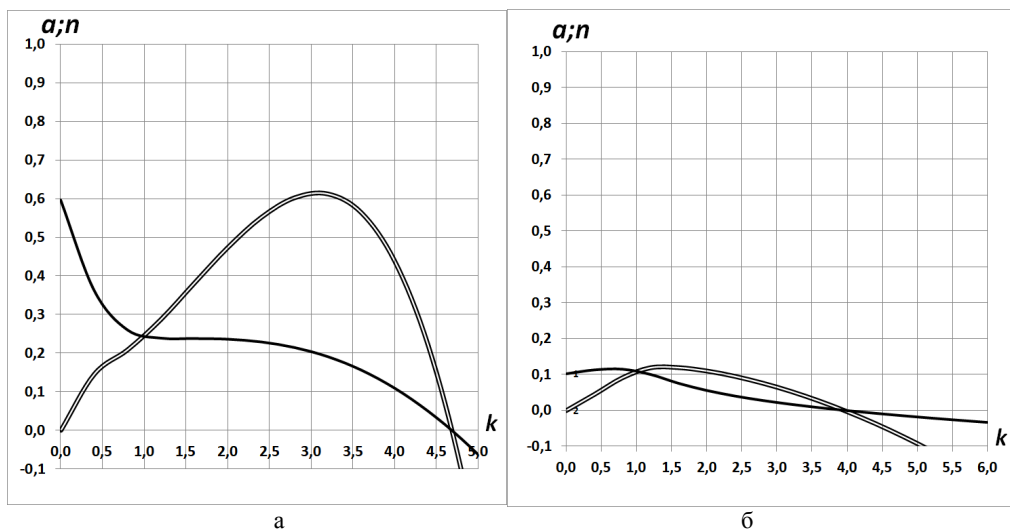


Рисунок 7 – Зависимости удельной мощности ветроустановки (двойные линии) и вращательного момента (обычные линии) от оборотной скорости ротора: а – регулируемые лопасти; б – ротор Дарье

Однако при оценке и сравнительном анализе энергоэффективности ВЭУ недостаточно рассматривать аэродинамическое взаимодействие рабочих органов установки с воздушным потоком, поскольку эта модель не учитывает эффект рассеяния воздушного потока под воздействием преграды в виде ВЭУ. Необходима расчетная методика, позволяющая осуществлять взаимное приведение параметров: с одной стороны, из энергетической модели идеальной ВЭУ, с другой – из оптимизационной модели взаимодействия рабочих органов ВЭУ с воздушным потоком, и получение совместимых (равновесных) показателей.

Учитывая сохранение массы воздуха, значения его массовых расходов (M) и скоростей (V) в начале (индекс «0») и в конце (индекс «1») потока взаимно пропорциональны:

$$M1/M0 = V1/V0 = d, \quad (25)$$

где $d \leq 1$ – коэффициент пропорциональности, определяющий степень замедления (рассеяния) воздушного потока. Выражение для коэффициента использования энергии ветра (КИЭВ), в соответствии с формулой Н.Е. Жуковского [7, 12] для идеальной ВЭУ,

$$\eta = 1/2(1 + d)(1 - d^2). \quad (26)$$

Максимальный КИЭВ достигается при $d = 1/3$ и составляет $\eta_{\max} = 16/27 = 0,593$.

В теории идеальной ВЭУ его мощность рассчитывается как

$$N_{\eta} = (1 + d)(1 - d^2)S_0\rho V_0^3/4, \quad (27)$$

где S_0 – площадь ортогонального сечения активной зоны, рассчитываемая как площадь сечения ветроколеса (ветротурбины), перпендикулярного вектору воздушного потока. С другой стороны, мощность, извлекаемая K лопастями ветроколеса (ротора) из потока воздуха [12], составляет

$$N_n = nc_{fn}KS_n\rho V_0^3(1 + d)^3/16. \quad (28)$$

Сопоставление (приравнивание) значений мощности N_n и N_{η} дает равновесное значение коэффициента рассеяния потока

$$d^* = (4S_0 - nc_{fn}KS_n)/(4S_0 + nc_{fn}KS_n), \quad (29)$$

а последующая подстановка в (22) – приведенное (равновесное) значение КИЭВ.

Результат действия эффекта рассеяния потока представлен на рисунке 8. С одной стороны, углы, образуемые векторами приложенных сил с радиальными штангами, существенно больше в оптимизированной турбине, что способствует увеличению крутящего момента. С другой стороны, распределение скоростей набегающего потока показывает значительное снижение скорости потока из-за рассеяния воздуха, что снижает величину приложенных к лопасти сил. Совместное действие этих факторов частично компенсирует преимущество оптимизированной турбины по отношению к ротору Дарье.

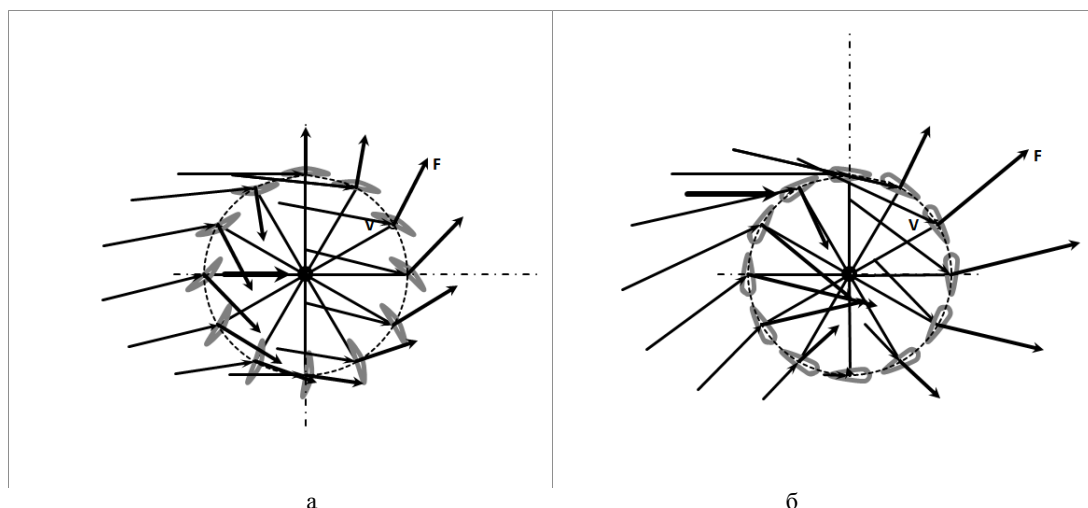


Рисунок 8 – Распределение относительных скоростей потока и сил его воздействия на регулируемые (а) и стационарные (б) лопасти

Результаты сопоставительных расчетов оптимизированной установки с регулируемыми лопастями и установки с ротором Дарье, при одинаковых габаритах ветротурбин и одинаковых скоростных параметрах ветра, представлены в таблицах 1 и 2.

Таблица 1 – Примерный расчет параметров для оптимизированной ВЭУ

Параметр	Обозначение	Формула	Значение
Плотность воздуха, кг/м ³	ρ	Относительная константа	1,20
Коэффициент лобового сопротивления.	C_{fn}	Относительная константа	1,33
КПД генератора, 0...1	η_g	Относительная константа	0,870
КПД электронного инвертора 0...1	η_i	Относительная константа	0,850
Предельный КИЭВ ветроустановки, 0...1	η_{max}	Относительная константа	0,593
Коэффициент окружной скорости/	k	Заданное значение	2,00
Скорость ветра, м/сек	V_0	Заданное значение	11
Внешний диаметр ротора, м.	D	Заданное значение	3,00
Высота ротора (лопасти), м.	H	Заданное значение	3,00
Ширина лопасти.	B	Заданное значение	1,50
Число лопастей.	K	Заданное значение	3
Площадь лопасти.	S_n	$S_n=HB$	4,50
Ометаемая площадь ветроустановки, м ²	S_0	$S_0=HD$	9,00
Массовый расход воздуха через ометаемую	M_0	$M_0=\rho S_0 V_0$	119
Мощность воздушного потока, Вт	E	$E=M_0 V_0^2/2$	7187
Коэффициент отбора мощности потока.	n	Расчетная модель	0,25
Равновесный коэффициент замедления потока	d^*	$d^*=(4S_0-nC_{fn}KS_n)/(4S_0+nC_{fn}KS_n)$	0,778
КИЭВ ветроустановки	η_t	$\eta_t=(1+d^*)(1-d^{*2})/2$	0,351
КПД ветроустановки	η	$\eta=\eta_g\eta_i\eta_t$	0,259
Мощность установки, кВт.	N	$N=E\eta/1000$	1,86
Удельная мощность установки, кВт/м ² .	n_s	$n_s=N/S_0$	0,21
Потери давления потока в ветродвигателе, Па	f	$f=(1-d^{*2})\rho V_0^2/2$	28,63

Примечание. КИЭВ – главный сравнительный показатель энергоэффективности

Таблица 2 – Примерный расчет параметров для ВЭУ с ротором Дарье

Параметр	Обозначение	Формула	Значение
Плотность воздуха, кг/м ³	ρ	Относительная константа	1,20
Коэффициент лобового сопротивления.	C_{fn}	Относительная константа	1,33
КПД генератора, 0...1	η_g	Относительная константа	0,870
КПД электронного инвертора 0...1	η_i	Относительная константа	0,850
Предельный КИЭВ ветроустановки, 0...1	η_{max}	Относительная константа	0,593
Коэффициент окружной скорости/	k	Заданное значение	2,00
Скорость ветра, м/сек	V_0	Заданное значение	11
Внешний диаметр ротора, м.	D	Заданное значение	3,00
Высота ротора (лопасти), м.	H	Заданное значение	3,00
Ширина лопасти.	B	Заданное значение	1,50
Число лопастей.	K	Заданное значение	3
Площадь лопасти.	S_n	$S_n=HB$	4,50
Ометаемая площадь ветроустановки, м ²	S_0	$S_0=HD$	9,00
Массовый расход воздуха через ометаемую	M_0	$M_0=\rho S_0 V_0$	119
Мощность воздушного потока, Вт	E	$E=M_0 V_0^2/2$	7187
Коэффициент отбора мощности потока.	n	Расчетная модель	0,12
Равновесный коэффициент замедления потока	d^*	$d^*=(4S_0-nC_{fn}KS_n)/(4S_0+nC_{fn}KS_n)$	0,887
КИЭВ ветроустановки	η_t	$\eta_t=(1+d^*)(1-d^{*2})/2$	0,201
КПД ветроустановки	η	$\eta=\eta_g\eta_i\eta_t$	0,149
Мощность установки, кВт.	N	$N=E\eta/1000$	1,07
Удельная мощность установки, кВт/м ² .	n_s	$n_s=N/S_0$	0,12
Потери давления потока в ветродвигателе, Па	f	$f=(1-d^{*2})\rho V_0^2/2$	15,47

Примечание. КИЭВ – главный сравнительный показатель энергоэффективности

Из результатов расчетов видно, что, несмотря на 5–6-тикратное преимущество оптимизированной турбины по аэродинамическим показателям (рис. 7), мощность равномасштабной установки Дарье достигает 60 % от оптимизированной турбины. При этом потери напора воздушного потока в роторе Дарье почти вдвое меньше.

Заключение. Учитывая ограниченную доступность и противоречивость данных о результатах исследований, испытаний и эксплуатации вертикально-осевых ортогональных ВЭУ, применение метода сравнительного анализа на основе верифицированных расчетных моделей ветротурбин дает дополнительные возможности для оценки эффективности и потенциала использования подобных устройств. Предварительная оценка простого в изготовлении и эксплуатации ротора Дарье, в сравнении с эталонной оптимизированной турбиной, показывает, что его энергопоказатели вполне сопоставимы с эталоном, а низкое аэродинамическое сопротивление свидетельствует о потенциале повышения энергоэффективности. Для более точной оценки потенциала такой ветротурбины нужно проведение многофакторного численного эксперимента на основе полученной расчетной методики.

Библиографический список

1. Ветроэнергетика / под ред. Д. де Рензо. – Москва : Энергоатомиздат, 1982. – 271 с.
2. Возобновляемые источники энергии. Теоретические основы, технологии, технические характеристики, экономика / З. А. Стычинский, Н. И. Воропай. – Otto-von-Guericke-Universitat, Magdeburg, 2010. – 203 с.
3. Гордон В. В. Нетрадиционные возобновляемые источники энергии / В. В. Гордон, В. Е. Губин, А. С. Матвеев. – Томск : Издательство Томского политехнического университета, 2009. – 294 с.
4. Зырянов В. М. Математическая модель взаимодействия управляемой лопастной ветротурбины с воздушным потоком / В. М. Зырянов, Л. Г. Лимонов, Ю. Б. Соколовский, В. М. Роткин // Электротехнические и компьютерные системы. – 2016. – № 23 (99). – С. 33–38.
5. Кривцов В. С. Неисчерпаемая энергия. Книга 1. Ветроэнергогенераторы / В. С. Кривцов, А. М. Олейников, А. И. Яковлев. – Харьков : ХАИ, 2003. – 382 с.
6. Мхитарян А. М. Аэродинамика : учебник / А. М. Мхитарян. – 2-е изд., перераб. и доп. – Москва : ЭКОЛИТ, 2012. – 448 с.
7. Окулов В. Л. Идеальный ветряк с конечным числом лопастей / В. Л. Окулов, Ж. Н. Соренсен // Доклады академии наук / Институт им. С.С. Кутателадзе Сибирского отделения Российской Академии наук. – 2008. – Т. 420, № 4. – С. 428–483.
8. Соколовский Ю. Б. Современные ветроэнергетические установки (обзор) / Ю. Б. Соколовский, А. Ю. Соколовский // Электротехника. – 2015. – Т. 2, № 4. – С. 27–38.
9. Соколовский Ю. Б. Ветровые энергетические установки / Ю. Б. Соколовский // Вестник дома ученых. – Хайфа (Израиль), 2014. – Vol. XXXII. – С. 80–87.
10. Rotkin V. Energy efficiency of bladed wind turbine. Optimization model / V. Rotkin, Y. Sokolovsky // Journal Scientific Israel – Technological Advantages. – 2016. – Vol. 18, № 2. – С. 45–61.
11. Rotkin V. New variants of wind energy plants / V. Rotkin, Y. Sokolovsky, A. Sokolovsky, E. Frolov // Scientific Israel – Technological Advantages. – 2016. – Vol. 18, № 4. – P. 39–50.
12. Sokolovsky Y. Theoretical and technical basis for the optimization of wind energy plants / Y. Sokolovsky, V. Rotkin. – Lulu Press, Inc., USA, 2017. – 160 p.

References

1. Renzo de D. (ed.) *Vetroenergetika* [Wind power]. Moscow, Energoatomizdat Publ., 1982. 271 p.
2. Stychinskiy Z. A., Voropay N. I. *Vozobnovlyayemye istochniki energii. Teoreticheskie osnovy, tekhnologii, tekhnicheskie kharakteristiki, ehkonomika* [Renewable energy sources. Theoretical Foundations, Technologies, Technical Specifications, Economics]. Otto-von-Guericke-Universitat, Magdeburg, 2010. 203 p.
3. Gordon V. V., Gubin V. E., Matveev A. S. *Netraditsionnye vozobnovlyayemye istochniki energii* [Unconventional renewable energy sources]. Tomsk, Publishing House of the Tomsk Polytechnic University, 2009. 294 p.
4. Zyryanov V. M., Limonov L. G., Sokolovskiy Yu. B., Rotkin V. M. *Matematicheskaya model vzaimodeystviya upravlyaemoy lopastnoy vetroturbiny s vozдушным potokom* [A mathematical model of the interaction of a driven blade wind turbine with an airstream]. *Elektrotekhnicheskie i kompyuternye sistemy* [Electrical and Computer Systems], 2016, no. 23 (99), pp. 33–38.
5. Krivtsov V. S., Oleynikov A. M., Yakovlev A. I. *Neischerpaemaya energiya: Kniga 1. Vetroenergogeneratory* [Inexhaustible energy: Book 1. Wind power generators]. Kharkiv, Kharkiv Aviation Institute, 2003. 382 p.
6. Mkhitaryan A. M. *Aerodinamika: uchebnik* [Aerodynamics: textbook], 2nd ed., rev. and add. Moscow, EKOLIT Publ., 2012. 448 p.
7. Okulov V. L., Sorensen Zh. N. *Idealnyy vetryak s konechnym chislom lopastey* [An ideal windmill with a finite number of blades]. *Doklady akademii nauk, Institut im. S.S. Kutateladze Sibirskogo otdeleniya Rossiyskoy Akademii nauk* [Reports of the Academy of Sciences Institute named after S.S. Kutateladze of the Siberian Branch of the Russian Academy of Sciences], 2008, vol. 420, no. 4, pp. 428–483.
8. Sokolovskiy Yu. B. *Sovremennyye vetroenergeticheskie ustanovki (obzor)* [Modern wind power plants (review)]. *Elektrotekhnika* [Electrical Engineering], 2015, vol. 2, no. 4, p. 27–38.

9. Sokolovskiy Yu. B. Vetrovye energeticheskie ustanovki [Wind power plants]. *Vestnik doma uchenykh* [Bulletin of the House of Scientists]. Haifa (Israel), 2014, vol. XXXII, p. 80–87.

10. Rotkin V., Sokolovsky Y. Energy efficiency of bladed wind turbine. Optimization model. *Journal Scientific Israel- Technological Advantages*, 2016, vol. 18, no. 2, pp. 45–61.

11. Rotkin V., Sokolovsky Y., Sokolovsky A., Frolov E. New variants of wind energy plants. *Journal Scientific Israel – Technological Advantages*, 2016, vol. 18, no. 4, p. 39–50.

12. Sokolovsky Y., Rotkin V. *Theoretical and technical basis for the optimization of wind energy plants*. Lulu Press, Inc., USA, 2017. 160 p.

DOI 10.21672/2074-1707.2020.49.4.121-131

УДК 004:9

МАТЕМАТИЧЕСКАЯ МОДЕЛЬ, ПРЕДНАЗНАЧЕННАЯ ДЛЯ ОЦЕНКИ КАЧЕСТВА ЭЛЕКТРОННОЙ БИБЛИОТЕКИ: СИНТЕЗ ЧИСЛА ГРАДАЦИЙ ПРОСТРАНСТВА СОСТОЯНИЙ

Статья поступила в редакцию 28.01.2020, в окончательном варианте –16.02.2020.

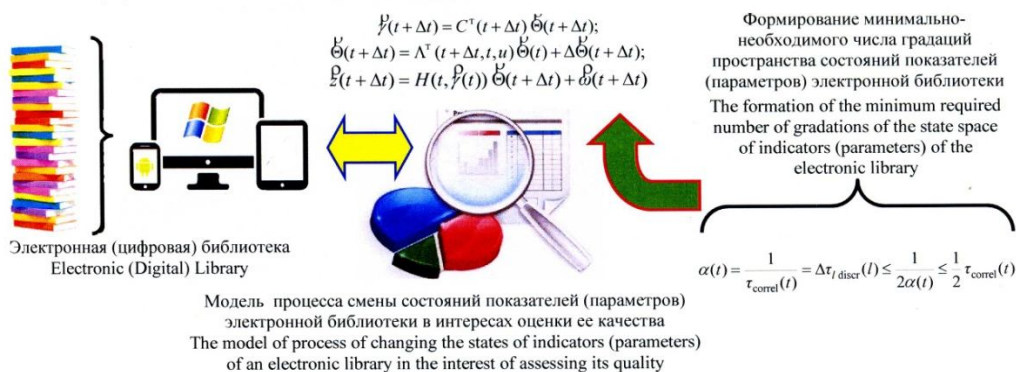
Крюкова Елена Сергеевна, Военная академия связи им. Маршала Советского Союза С.М. Буденного, 194064, Российская Федерация, г. Санкт-Петербург, пр. Тихорецкий, 3, адъюнкт, e-mail: e.krukovaab9@yandex.ru

Паращук Игорь Борисович, Военная академия связи им. Маршала Советского Союза С.М. Буденного, 194064, Российская Федерация, г. Санкт-Петербург, пр. Тихорецкий, 3, доктор технических наук, профессор, e-mail: shchuk@ Rambler.ru

Показано место электронных библиотек в информационно-образовательном пространстве и системе образования России. Охарактеризованы особенности математической модели, предназначенной для оценки качества электронной библиотеки. Рассмотрены этапы алгоритма отыскания минимально-допустимых значений приращений (числа градаций) переменных состояния, описывающих динамику изменения значений показателей качества электронной библиотеки, а также поиска оптимальных временных интервалов их дискретизации. Предложен алгоритм определения оптимального числа градаций пространств состояния для построения модели электронной библиотеки. Алгоритм учитывает требования к точности (величине отклонения) значения критерия оценки качества, к устойчивости основных оперативно-технических показателей электронной библиотеки относительно возмущающих воздействий, а также текущие требования пользователя к математической сложности оценки качества объекта такого класса. Сформулированы аналитические выражения для нахождения допустимой погрешности оценивания значения l -го показателя качества, коэффициента чувствительности и временного интервала дискретизации компонент векторного показателя качества электронной библиотеки. Предложенные подходы позволяют сформулировать достоверную (адекватную задачам управления) математическую модель в интересах оценки качества электронной библиотеки, открывают новые возможности для моделирования состояний библиотек и прогнозирования их качества, углубляют наши представления о математической природе процесса смены состояний сложных управляемых информационно-технических систем и процессов.

Ключевые слова: электронная библиотека, пространство состояний, показатель качества, оценивание, математическая модель, процесс смены состояний, марковская цепь, уравнение состояния, уравнение наблюдения, минимальное число градаций, переменные состояния, погрешность, временные интервалы дискретизации, требования к точности моделирования, управление качеством

Графическая аннотация (Graphical annotation)



MATHEMATICAL MODEL FOR ESTIMATING THE QUALITY OF ELECTRONIC LIBRARY: SYNTHESIS OF THE NUMBER OF GRADATIONS OF SPACE STATES

The article was received by the editorial board on 30.01.2020, in the final version – 16.02.2020.

Kryukova Elena S., Military Academy of Communications named after Marshal of the Soviet Union S.M. Budyonny, 3 Tikhoretskiy Ave., St. Petersburg, 194064, Russian Federation, postgraduate student, e-mail: e.kkrukovaa69@yandex.ru

Parashchuk Igor B., Military Academy of Communications named after Marshal of the Soviet Union S.M. Budyonny, 3 Tikhoretskiy Ave., St. Petersburg, 194064, Russian Federation, Doct. Sci. (Engineering), Professor, e-mail: shchuk@rambler.ru

The place of electronic libraries in the information and educational space and the education system of Russia is shown. The features of a mathematical model for evaluating the quality of an electronic library are characterized. The stages of the algorithm for finding the minimum allowable values of the increments (number of gradations) of state variables describing the dynamics of changing the values of the quality indicators of the electronic library, as well as the search for optimal time intervals for their discretization, are considered. An algorithm is proposed for determining the optimal number of gradations of state spaces for constructing a model of an electronic library. The algorithm takes into account the requirements for accuracy (deviation value) of the quality assessment criterion, for the stability of the main operational and technical indicators of the electronic library with respect to disturbing influences, as well as current user requirements for the mathematical complexity of assessing the quality of an object of this class. Analytical expressions are formulated to find the permissible error in estimating the value of the l -th quality indicator, sensitivity coefficient and the sampling time interval of the components of the vector quality indicator of the electronic library. The proposed approaches make it possible to formulate a reliable (adequate to control problems) mathematical model in the interests of assessing the quality of the electronic library, open up new possibilities for modeling the state of libraries and predict their quality, deepen our understanding of the mathematical nature of the process of changing states of complex controlled information technology systems and processes.

Key words: electronic library, state space, quality index, estimation, mathematical model, state transition process, Markov chain, equation of state, observation equation, minimum number of gradations, state variables, error, sampling time intervals, requirements for modeling accuracy, quality management

Введение. В свете тенденций развития современной науки и образования, внедрения федеральных государственных образовательных стандартов высшего образования нового поколения, сегодня в российской научно-образовательной практике очень актуальна задача создания единого информационного пространства [13].

Одним из ключевых информационно-технологических элементов единого информационного пространства являются национальные библиотеки, причем ориентированные на цифровое будущее [8]. Эти объекты получили название электронных (цифровых) библиотек и активно развиваются во всем мире [15, 19, 22–24, 29]. Электронные библиотеки (ЭБ) имеют все крупнейшие страны мира, в частности большое развитие они получили в США и России [25, 26].

Современная ЭБ предназначена для хранения информации различных видов, прежде всего текстовой. Такие ЭБ реализуются в виде сложных распределенных информационных систем, обладающих унифицированным интерфейсом, а также возможностью доступа к ним пользователей через интернет [14, 17, 22, 23, 27, 28].

Электронные библиотеки как класс информационных систем характеризуются следующими особенностями:

- глобальным доступом к ним в среде Web;
- поддержкой метаданных для системы и пользователей;
- включением неоднородных информационных ресурсов: каталогов, электронных книг, справочников, аудиовизуальных, электронных документов и т.д.;
- распределением информационных ресурсов по предназначению: государственные (национальные), корпоративные (предприятия, организации, ведомства), индивидуальные;
- наличием материализованных (в виде систем хранения данных, баз данных, дата-центров и дисков) и виртуальных коллекций информационных ресурсов (электронных версий печатных изданий, аудио и видеоданных и т.д.);
- интеграцией информационных ресурсов на различных уровнях;
- использованием пользовательских интерфейсов с возможностью их адаптации к контексту (семантике, смысловой сущности) поисковых запросов пользователей на получение конкретного информационного ресурса и др. [1].

В настоящее время ЭБ предоставляют качественно новые возможности работы с большими объемами информации. К таким возможностям можно отнести следующее:

- последовательный, выборочный или параллельный просмотр множества документов;
- многоаспектный поиск во всем объеме информации, хранящейся в ЭБ;
- копирование необходимых документов или их фрагментов как на бумагу, так и на другие носители;
- в ряде случаев (на определенных условиях) и в некоторых ЭБ – создание пользователями собственных документов, а также подборок документов и т.д. [5].

По сути ЭБ – это совокупность электронных ресурсов, организуемых по библиотечному принципу (т.е. на основе известных правил и технологий традиционного библиотековедения, включая комплектование, обработку, систематизацию, хранение и другие процессы и технологии, в том числе создание каталогов и справочно-поискового аппарата в электронной форме) [20].

При этом ЭБ является информационной системой, обеспечивающей сохранность и эффективное использование коллекции различных документов в электронной форме, размещенных в самой системе и доступных через телекоммуникационные сети в удобном для конечного пользователя виде. Иными словами, это упорядоченная коллекция разнородных электронных документов (в том числе книг журналов), снабженных средствами навигации и поиска [1, 20].

Исследуемая проблема заключается в том, что оптимальное управление ЭБ, управление ее качеством (качеством предоставляемого ею информационного ресурса) должно опираться на оценочные значения показателей качества объектов и процессов ЭБ на основе анализа результатов наблюдения за данной системой. Результаты наблюдения могут быть получены либо явно, путем измерения и диагностики реальных значений показателей ЭБ, либо, когда это невозможно или затруднено – путем моделирования динамики изменения показателей качества ЭБ в пространстве состояний. Способу нахождения важной характеристики для моделей такого типа – определению числа градаций пространства состояний посвящена данная статья.

Постановка задачи исследования. Важной задачей является оценивание качества ЭБ в интересах оптимального управления ее ресурсами [7]. При этом одним из ключевых этапов методологии оценивания качества процессов функционирования сложных информационно-технических систем, например, таких, как ЭБ, является моделирование динамики линейного или нелинейного изменения состояния (пошаговых значений) совокупности показателей качества (ПК) объектов и процессов ЭБ в пространстве состояний. Зачастую для построения моделей такого типа используют марковские или полумарковские последовательности с дискретным или непрерывным временем [2, 6, 16, 18]. Рассмотрим вариант аналитического описания процесса смены состояний ПК ЭБ.

В работах [2, 6] приведены примеры моделей на основе дискретных марковских цепей. Однако эти модели опираются на дискретные состояния ПК в дискретные моменты времени. Они относительно просты и позволяют моделировать состояния ПК сложных систем в динамике, но при этом оценки состояний будут получены точечные, пошаговые, что не всегда рационально с точки зрения принятия решений по управлению элементами и ресурсами ЭБ в целом. Зачастую достаточно интервальных оценок, усредненных за период наблюдения.

Поэтому нужна модифицированная модель, описывающая на основе марковских цепей с непрерывным временем (непрерывных цепей Маркова) динамику смены дискретных (конечных или счетных) состояний ПК ЭБ в непрерывном времени и с учетом преобразования шума возбуждения. Марковский случайный процесс называется цепью Маркова с непрерывным временем, если

переходы системы из состояния в состояние происходят не в фиксированные, а в случайные моменты времени [4].

Непрерывные марковские цепи описывают функционирование систем, принимающих в процессе работы конечное число состояний S_1, S_2, \dots, S_n . При этом переходы из одного состояния в другое осуществляются случайным образом в произвольные моменты времени t [3]. Иными словами, время пребывания системы в любом состоянии представляет непрерывную случайную величину. Как и в случае дискретной марковской цепи, необходимо выполнение марковского свойства, т.е. чтобы поведение системы в момент времени $t > t_0$ зависело только от состояния системы в момент времени t_0 и не зависело от поведения системы в моменты времени, предшествующие t_0 [30].

В результате модифицированная модель, на наш взгляд, должна представлять собой формализованное описание, в терминах непрерывной марковской цепи в форме разностных стохастических уравнений (НМЦ-РСУ), динамики изменений значений ПК ЭБ. Эти ПК, в свою очередь, характеризуют существенные свойства процесса функционирования ЭБ в пространстве состояний на непрерывном отрезке времени $t + \Delta t$.

Важным шагом при решении задач оценки качества ЭБ является формулировка совокупности ПК ЭБ – исходя из их необходимости и достаточности. При этом, поскольку эти ПК описывают разные характеристики ЭБ, они в идеале не должны быть коррелированы. Система ПК ЭБ, на наш взгляд, не должна быть громоздкой, избыточной, должна включать не более десяти локальных показателей и представлять собой иерархически взаимосвязанную совокупность этих ПК, характеризующих качество отдельных элементов ЭБ, процессов их функционирования и процесса обслуживания пользователей. Например, в качестве ПК, характеризующего своевременность предоставления услуг ЭБ, может выступать среднее время предоставления услуги пользователю. Под этим термином понимается свойство ЭБ, характеризующее ее способность обеспечивать доступ пользователей к запрашиваемым ресурсам ЭБ и предоставление им требуемого перечня услуг в установленные сроки или в реальном масштабе времени.

Достоверность предоставления услуг ЭБ – это свойство ЭБ, характеризующее ее способность обеспечивать с заданной точностью воспроизведение хранимой, обрабатываемой и передаваемой информации до пользователя. Показателем достоверности может выступать коэффициент достоверности информации, предоставляемой пользователю ЭБ: соотношение количества достоверно обработанных и предоставленных данных к общему количеству хранимых данных.

Безопасность предоставления услуг ЭБ – это свойство ЭБ, характеризующее способность ЭБ противостоять несанкционированному получению, уничтожению и (или) изменению хранимых в ней данных. Показателями, характеризующими безопасность ЭБ, могут выступать среднее время вскрытия нарушителем процессов, протекающих в ЭБ, либо вероятность компрометации предоставляемого контента.

Надежность ЭБ – ее способность сохранять в установленных пределах времени значения всех ключевых параметров, определяющих выполнение ею требуемых функций. К показателям надежности ЭБ можно отнести среднюю наработку на отказ ЭБ, среднее время восстановления ЭБ, коэффициент готовности ЭБ и коэффициент технического использования ЭБ.

Перечень этих показателей качества ЭБ приведен в качестве примера, он может быть дополнен показателями, характеризующими отказоустойчивость ЭБ, доступность контента ЭБ, объем хранения данных в ЭБ, ее масштабируемость и иными.

Как и в любой классической задаче теории оценивания, оценки могут быть представлены в двух видах: интервальные оценки и точечные оценки. Точечные оценки представляют собой оценочные значения ПК (число или числовой вектор) в конкретный момент (на конкретном временном срезе) времени, при этом моменты времени обозначаются t .

В данной статье рассматриваются интервальные оценки. Для них указывается доверительный интервал, в котором оцениваемая величина находится с определенной вероятностью, называемой коэффициентом доверия. Эти интервалы времени обозначаются через тире, например, интервал времени $t_0 - t_k$, либо как $t + \Delta t$.

Модифицированная модель смены состояний ПК ЭБ в терминах непрерывной марковской цепи может иметь вид:

$$\vec{\gamma}(t + \Delta t) = C^T(t + \Delta t) \vec{\Theta}(t + \Delta t); \quad (1)$$

$$\vec{\Theta}(t + \Delta t) = \Lambda^T(t + \Delta t, t, u) \vec{\Theta}(t) + \Delta \vec{\Theta}(t + \Delta t); \quad (2)$$

$$\vec{z}(t + \Delta t) = H(t, \vec{\gamma}(t)) \vec{\Theta}(t + \Delta t) + \vec{w}(t + \Delta t). \quad (3)$$

Выражение (1) – уравнение состояния ПК ЭБ на $(t + \Delta t)$ -м непрерывном временном интервале ее функционирования, в котором: $\vec{\gamma}(t + \Delta t)$ – вектор значений ПК ЭБ на $(t + \Delta t)$ -м временном интервале ее функционирования, в котором все элементы кроме одного (реального значения ПК ЭБ на данном интервале времени) равны 0;

$C^T(k + 1)$ – транспонированная диагональная квадратная матрица (порядка m) возможных значений ПК ЭБ на $(t + \Delta t)$ -м временном интервале ее функционирования, где число m (строк и столбцов) зависит от выбранного числа состояний (глубины моделирования);

$\vec{\Theta}(t + \Delta t)$ – вектор вспомогательных индикаторов состояния ПК ЭБ на $(t + \Delta t)$ -м временном интервале ее функционирования, который описывает динамику «тяготения» значений ПК ЭБ на $(t + \Delta t)$ -м временном интервале при переходе этих ПК из состояния в состояние.

Выражение (2) – уравнение состояния вспомогательных индикаторов состояния ПК ЭБ на $(t + \Delta t)$ -м временном интервале ее функционирования, в котором: $\Lambda^T(t + \Delta t, t, u)$ – транспонированная квадратная матрица (порядка m) плотностей вероятностей перехода λ_{ij} , представляющих собой предел отношения вероятности перехода системы за время $t + \Delta t$ из состояния i в состояние j к величине малого промежутка времени Δt (к длине этого промежутка) при стремлении Δt к нулю и обуславливающих смену состояний ПК ЭБ на $(t + \Delta t)$ -м временном интервале ее функционирования;

$\vec{\Theta}(t)$ – вектор значений индикаторов состояния ПК ЭБ на предыдущем (t) временном интервале ее функционирования; $\Delta\vec{\Theta}(t + \Delta t)$ – вектор приращений индикаторов состояния, элементы которого служат для компенсации нецелочисленной части выражения (2). Эти приращения (компенсационные добавки) могут быть получены исходя из «тяготения» значений ПК ЭБ на $(t + \Delta t)$ -м временном интервале при переходе этих ПК из состояния в состояние – в результате коррекции исходного шума возбуждения с математическим ожиданием и дисперсией, соответствующими начальному состоянию моделируемого процесса смены состояний конкретного ПК ЭБ.

Выражение (3) в системе уравнений (1)–(3), описывающих процесс смены состояний ПК ЭБ на $(t + \Delta t)$ -м непрерывном временном интервале ее функционирования, является уравнением наблюдения за процессом (процессом смены состояний ПК ЭБ), где $H(t, \vec{\gamma}(t))$ – квадратная диагональная матрица (порядка m) наблюдаемых значений процесса, а $\vec{\omega}(t + \Delta t)$ – вектор белых шумов наблюдения с нулевым средним и матрицей дисперсии $\delta_{\omega}(t + \Delta t)$.

При этом вычисление текущих значений индикаторов состояния ПК ЭБ на $(t + \Delta t)$ -м интервале времени в выражении (2) осуществляется на основе линейной процедуры:

$$\|\Delta\Theta(t + \Delta t)\| = \mathfrak{R} - \|\Theta^{\text{nai}}(t + \Delta t)\|; \quad (4)$$

$$\Delta\vec{\Theta}(t + \Delta t) = \|\Delta\Theta(t + \Delta t)\| \cdot \vec{\Theta}^{\text{prel}}(t + \Delta t); \quad (5)$$

$$\vec{\Theta}(t + \Delta t) = \vec{\Theta}^{\text{nai}}(t + \Delta t) + \Delta\vec{\Theta}(t + \Delta t). \quad (6)$$

Выражение (4) описывает правило получения элементов m -мерной матрицы приращений $\|\Delta\Theta(t + \Delta t)\|$ и содержит элементы: \mathfrak{R} – m -мерная диагональная единичная матрица; $\|\Theta^{\text{nai}}(t + \Delta t)\|$ – m -мерная матрица, столбцы которой – векторы нецелочисленных (not an integer) значений индикаторов $\vec{\Theta}^{\text{nai}}(t + \Delta t)$, повторенные m раз.

Выражение (5) описывает процедуру выбора вектора приращений $\Delta\vec{\Theta}(t + \Delta t)$ из матрицы приращений (компенсационных добавок) $\|\Delta\Theta(t + \Delta t)\|$ с помощью вектора $\vec{\Theta}^{\text{prel}}(t + \Delta t)$ предварительных значений (preliminary meaning) индикаторов моделируемого процесса смены состояний ПК ЭБ.

Выражение (6) завершает алгоритм вычислений вектора индикаторов состояний $\vec{\Theta}(t + \Delta t)$ и является модифицированной записью уравнения состояния моделируемого процесса смены состояний ПК ЭБ. При этом элементы вектора нецелочисленных значений индикаторов состояния находятся в соответствии с выражением

$$\vec{\Theta}^{\text{nai}}(t + \Delta t) = \Lambda^T(t + \Delta t, t, u)\vec{\Theta}(t). \quad (7)$$

Выражение (7) является модификацией известного уравнения – Колмогорова Чепмена [2] и получено в результате преобразования выражения (2).

Здесь $\bar{\Theta}^{\text{nai}}(t + \Delta t)$ – вектор нецелочисленных значений индикаторов состояния моделируемого процесса (процесса смены состояний ПК ЭБ), имеющий математический смысл интервальных и, в конечном итоге, финальных вероятностей нахождения случайного процесса в том или ином состоянии;

$\Lambda^T(t + \Delta t, t, u)$ – матрица (порядка m) плотностей λ_{ij} вероятностей перехода процесса из одного состояния в другое за время $t + \Delta t$;

$\bar{\Theta}(t)$ – вектор целочисленных значений индикаторов состояния моделируемого процесса на предыдущем временном интервале.

Уравнение (7) для отдельного k -го (из m возможных) нецелочисленного индикатора состояния $\Theta_k^{\text{nai}}(t + \Delta t)$ в общем случае имеет вид:

$$\Theta_k^{\text{nai}}(t + \Delta t) = \Theta_1(t)(\lambda_{k1} - \lambda_{km}) + \Theta_2(t)(\lambda_{k2} - \lambda_{km}) + \dots + \Theta_{m-1}(t)(\lambda_{km-1} - \lambda_{km}) + \lambda_{km}, \quad (8)$$

где $k = 1, 2, 3, \dots, m$, а $\lambda_{k1}, \dots, \lambda_{km}$ – элементы матрицы $\Lambda^T(t + \Delta t, t, u)$ плотностей вероятностей перехода процесса из одного состояния в другое за время $t + \Delta t$.

Таким образом, выражениями (1)–(8) представлен вариант математической модели (в терминах непрерывной марковской цепи в форме разностных стохастических уравнений), описывающей динамику изменения состояний ПК ЭБ $\bar{\gamma}(t + \Delta t)$ на $(t + \Delta t)$ -м временном интервале ее функционирования.

Важной ключевой задачей при формулировке модели ЭБ в интересах оценки качества ее функционирования является определение числа m , т.е. обоснованный выбор количества состояний ПК, характеризующих свойства ЭБ и определяющих «глубину моделирования». Под этим термином будем понимать уровень детализации анализа и описания свойств ЭБ, степень, «глубину» декомпозиции описания (моделирования) свойств объекта с помощью определенного количества параметров (показателей) их численно характеризующих. Чем больше количество ПК и состояний ПК, характеризующих (описывающих) свойства ЭБ, тем «глубже» это описание.

Алгоритм синтеза числа градаций пространства состояний. Количество состояний определяется исходя из требуемой «глубины моделирования» и точности оценивания ПК ЭБ. Оптимизация размерности пространств состояния и наблюдения процесса функционирования ЭБ с учетом требований к допустимой погрешности оценивания, представляет собой отдельную теоретическую и практическую задачу.

Действительно, одной из ключевых задач, требующих решения в рамках разработки процедур оценивания качества ЭБ, является задача отыскания адекватных условиям функционирования и применения ЭБ, достаточного, с точки зрения изменяющихся во времени задач управления ЭБ, числа градаций переменных состояния процесса функционирования ЭБ и временных интервалов их дискретизации. Решение этой задачи направлено на снижение затрат на реализацию процесса функционирования ЭБ, повышение оперативности, достоверности (точности) оценивания качества процесса функционирования ЭБ и обуславливает ценность информации, добываемой в рамках анализа, адекватную возможностям и задачам управления ЭБ в условиях конкретной ситуации.

Очевидно, что формирование достаточных градаций переменных состояния (т.е. ПК) базируется на результатах формирования векторных критериев оценивания состояния и качества процесса функционирования ЭБ – ПК ЭБ. Иными словами, ставится задача адекватного согласования числа градаций уровней (состояний) оцениваемых параметров и ПК ЭБ с текущими требованиями к основным оперативно-техническим показателям ЭБ в условиях конкретной ситуации.

При отыскании достаточного числа градаций переменных состояния и временных интервалов их дискретизации необходимо учитывать следующее:

- требования к «глубине моделирования» и точности (величине отклонения) значения критерия оценивания состояния и качества процесса функционирования ЭБ;
- требования к устойчивости основных оперативно-технических показателей ЭБ относительно возмущающих воздействий;
- требования к форме представления пространств состояния со стороны пользователей, проводящих оценивание качества ЭБ;
- требования к снижению математической сложности моделирования и оценивания качества электронной библиотеки.

Одна из традиционных методик оптимизации представления пространств состояния и наблюдения критериев оптимальности в задачах синтеза сложных управляемых систем, приведена в работе [2]. Опираясь на эту методику, учитывая перечисленные факторы, особенности формулировки систем параметров и ПК ЭБ и критериев их оценивания, сформируем вариант алгоритма отыскания оптимального и адекватного текущим задачам управления ЭБ числа градаций пространств состояния для ПК библиотеки, по которым оценивается ее качество.

Блок-схема алгоритма, описывающего последовательность решения задачи определения числа градаций переменных состояния и временных интервалов их дискретизации, приведена на рисунке.

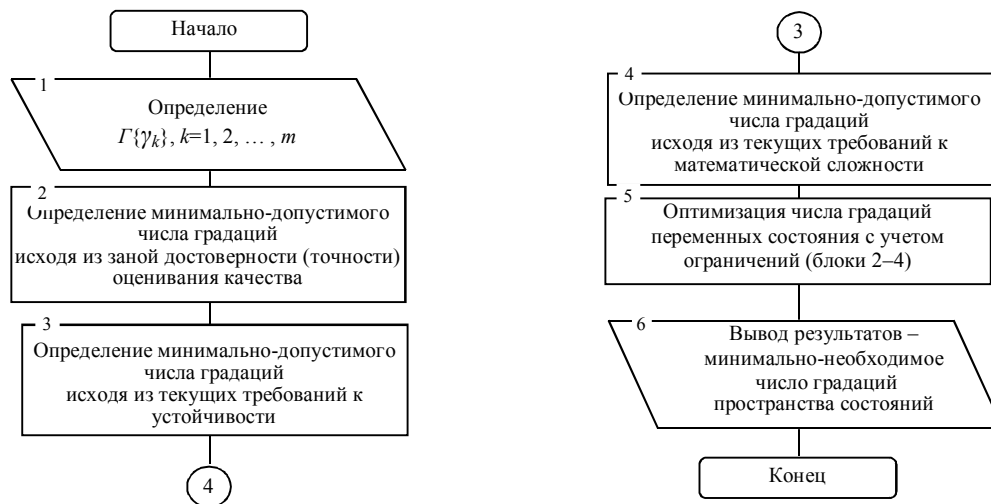


Рисунок – Блок-схема алгоритма поиска числа градаций переменных состояния и временных интервалов их дискретизации

Данный алгоритм (рис.) включает в себя следующие этапы:

1. Определение минимально-необходимой размерности пространств состояния $G = \{\gamma_k\}, k = 1, m$ процесса функционирования ЭБ.
2. Определение минимально-допустимых значений приращений (числа градаций) переменных состояния и временных интервалов их дискретизации исходя из заданной достоверности (точности) оценивания качества процесса функционирования ЭБ в конкретных условиях, определяемых в рамках текущих требований к точности управления электронной библиотекой. При этом целями управления является своевременное (в установленные сроки или в реальном масштабе времени) и безопасное обеспечение пользователей ЭБ достоверными данными (контентом, предоставляемым электронной библиотекой).
3. Определение минимально-допустимых значений приращений (числа градаций) переменных состояния и временных интервалов их дискретизации исходя из текущих требований к устойчивости процесса функционирования ЭБ и процесса управления ЭБ к возмущающим воздействиям. Под устойчивостью понимается свойство процесса возвращаться в заданный или близкий к нему установившийся режим после какого-либо внешнего или внутреннего возмущения.
4. Определение минимально-допустимых значений приращений (числа градаций) переменных состояния (ПК) и временных интервалов их дискретизации с учетом текущих требований пользователя к математической сложности оценивания качества процесса функционирования ЭБ, определяемых в рамках требований к временным и вычислительным затратам на реализацию процедуры моделирования и оценивания качества ЭБ.
5. Оптимизация значений приращений (числа градаций) переменных состояния и временных интервалов их дискретизации с учетом ограничений, введенных на 2-4 шагах алгоритма.

Для оптимизации числа градаций переменных состояния (т.е. ПК) процесса функционирования ЭБ и временных интервалов их дискретизации представляется целесообразным использовать методы теории чувствительности. Методами, позволяющими найти функцию чувствительности в задачах подобного класса, могут служить: метод построения уравнений чувствительности, прямой и регрессионный методы, метод параметрической идентификации, структурный метод, метод графов, метод конечных разностей, метод приращений, а также применение квадратичных показателей дополнительного движения [9, 21].

Трудность решения этой задачи определяется прежде всего сложностью пересчета ограничений на оперативно-технические ПК ЭБ и действующих на нее возмущений в требованиях к точности моделирования и оценивания состояния (качества) по отдельным компонентам векторного критерия оценивания. Принято считать, что критериальное оценивание это оценивание по критериям, т.е. оценка складывается из составляющих (критериев), которые отражают различные стороны (характеристики, свойства, ПК) объектов и процессов. Критерий оценки качества ЭБ может быть скалярным, т.е. характеризоваться одним единственным ПК, или векторным, характеризующимся совокупностью ПК. Это получение интегральной оценки качества при наличии различных частных, несводимых друг к другу критериев оценивания. Методы получения обобщенного показателя качества и методы многокритериального оценивания состояния, качества и эффективности функционирования сложных информационно-технических систем рассмотрены в работах [2, 6, 9, 10–12].

Процесс функционирования ЭБ, а значит, и оценивание качества ЭБ, является составной частью процесса функционирования информационного пространства. Значит, несет в себе те же требования к степени точности поддержания значений ПК ЭБ, а, следовательно, и к точности моделирования и оценивания.

Иными словами, точность моделирования и оценивания ПК ЭБ задает пользователь. В этом случае, если известна допустимая погрешность моделирования (и оценивания) значения l -го ПК ЭБ, входящего в L -ый вектор состояния (в L -ый частный показатель качества), которую можно обозначить $\Delta I_{l \text{ allow}}(\vec{\gamma}(t + \Delta t); \tau(t + \Delta t))$, соответствующая градация значений ПК $\Delta \gamma_l$ может быть определена на основе коэффициентов чувствительности из выражения

$$\Delta I_{l \text{ allow}}(\vec{\gamma}(t + \Delta t); \tau(t + \Delta t)) \geq \sum_{l=1}^n v_l(t + \Delta t) \Delta \gamma_l(t + \Delta t) + v_\tau(t + \Delta t) \Delta \tau_{\text{discr}}(t + \Delta t), \quad (9)$$

где $\Delta I_{l \text{ allow}}(\vec{\gamma}(t + \Delta t); \tau(t + \Delta t))$ – допустимое приращение (allowable increment) l -й компоненты L -го вектора состояния (векторного ПК ЭБ) на $(t + \Delta t)$ -ом интервале моделирования и оценивания качества ЭБ.

При этом коэффициент чувствительности l -й компоненты по L -й переменной (вектору) состояния (по $\Delta \gamma$ – состояниям и по $\Delta \tau$ – времени) равен

$$v_l(t + \Delta t) = \frac{\partial I_l(\vec{\gamma}(t + \Delta t); \tau(t + \Delta t))}{\partial \gamma_L}. \quad (10)$$

В качестве примера, матрица $v_L(t + \Delta t)$ коэффициентов чувствительности в явном виде для трех уровней чувствительности трех компонент L -го вектора состояния (векторного ПК ЭБ) может быть описана, как показано в выражении (11).

$$v_L(t + \Delta t) = \begin{vmatrix} 0,1 & 0,1 & 0,2 \\ 0,1 & 0,2 & 0,3 \\ 0,2 & 0,3 & 0,3 \end{vmatrix}. \quad (11)$$

Временные интервалы дискретизации (discretization) $\Delta \tau_{l \text{ discr}}(t + \Delta t)$ учтены в коэффициентах чувствительности и определяются, исходя из скорости изменения (флюктуаций) ПК ЭБ – $\alpha(t + \Delta t)$ [9]:

$$\alpha(t + \Delta t) = \frac{1}{\tau_{\text{correl}}(t + \Delta t)} = \Delta \tau_{l \text{ discr}}(t + \Delta t) \leq \frac{1}{2\alpha(t + \Delta t)} \leq \frac{1}{2} \tau_{\text{correl}}(t + \Delta t), \quad (12)$$

где $\tau_{\text{correl}}(t + \Delta t)$ – интервал корреляции (correlations) компонент векторного ПК электронной библиотеки на $(t + \Delta t)$ -м временном интервале.

Таким образом, из всех полученных в результате решения уравнений (9) и (10) и с учетом ограничения (12), допустимых временных интервалов $\Delta \tau_{l \text{ discr}}(t + \Delta t)$, необходимо выбрать минимальный с целью обеспечения заданной погрешности моделирования и оценивания по состоянию самого быстроменяющегося показателя качества процесса функционирования ЭБ.

Полученные результаты могут быть использованы в интересах оптимального управления ЭБ, управления ее качеством и качеством предоставляемого ею информационного ресурса. Найденные значения числа градаций переменных состояния и временных интервалов их дискретизации определяют требуемые границы чувствительности и устойчивости оценок, тем самым, позволяя

получать достоверные оценочные значения показателей качества ЭБ на основе моделирования динамики изменения этих ПК в пространстве состояний.

Заключение. Итак, предложенный алгоритм позволяет определить минимально-необходимое число градаций пространства состояний показателей качества электронной библиотеки. Он дополняет существующие подходы к моделированию сложных управляемых систем с учетом их представления в виде смены значений существенных параметров (показателей) объектов такого класса в пространстве состояний. Это позволяет сформулировать достоверную (адекватную задачам управления) математическую модель в интересах оценки качества электронной библиотеки, открывает новые возможности для моделирования состояний ЭБ и прогнозирования их качества, углубляет наши представления о математической природе процесса смены состояний сложных управляемых информационно-технических систем и процессов. Вместе с тем, при формировании минимально-необходимого числа градаций пространства состояний ПК для процесса функционирования ЭБ необходимо согласовывать полученные разбиения с формой представления вектора параметров (векторного ПК) ЭБ – дискретного по состояниям, но непрерывного по времени для всех входящих в него переменных, характеризующих качество электронной библиотеки.

Библиографический список

1. Богданова И. Ф. Электронные библиотеки: история и современность / И. Ф. Богданова, Н. Ф. Богданова // Информационное общество: образование, наука, культура и технологии будущего. – 2017. – Вып. 1. – С. 133–154.
2. Буренин Н. И. Методологические аспекты применения методов многокритериальной динамической оптимизации в задачах синтеза управляемых сетей многоканальной радиосвязи / Н. И. Буренин, А. П. Родинов, В. М. Терентьев. – Ленинград : ВАС, 1984. – 92 с.
3. Зейфман А. И. Марковские цепи и модели с непрерывным временем / А. И. Зейфман, В. Е. Бенинг, И. А. Соколов. – Москва : ЭЛЕКС-КМ, 2008. – 167 с.
4. Зорин А. В. Введение в общие цепи Маркова / А. В. Зорин, В. А. Зорин, Е. В. Пройдакова, М. А. Федоткин. – Нижний Новгород : Нижегородский госуниверситет, 2013. – 51 с.
5. Зуйкина К. Л., Соколова Д. В., Скалабан А. В. Электронные библиотеки в России. Текущий статус и перспективы развития / К. Л. Зуйкина, Д. В. Соколова, А. В. Скалабан. – Москва : Ваш формат, 2017. – 120 с.
6. Калинин В. Н. Теоретические основы системных исследований / В. Н. Калинин. – Санкт-Петербург : ВКА им. А.Ф. Можайского, 2013. – 278 с.
7. Крюкова Е. С. Особенности развития современных электронных библиотек и анализ подходов к оцениванию их качества / Е. С. Крюкова, И. Б. Паращук // Современные технологии: актуальные вопросы, достижения и инновации : сборник статей XXXI Международной научно-практической конференции. – Пенза : МЦНС «Наука и Просвещение», 2019. – С. 34–36.
8. Матвеев М. Ю. Национальные библиотеки: на пути к цифровому будущему / М. Ю. Матвеев // Вестник СПбГУКИ. – 2017. – № 2 (31). – С. 88–95.
9. Ненадович Д. М. Анализ чувствительности процессов фильтрации состояний управляемой радиотехнической системы / Д. М. Ненадович, И. Б. Паращук // Радиотехника. – 1997. – № 4. – С. 23–25.
10. Петухов Г. Б. Основы теории эффективности целенаправленных процессов / Г. Б. Петухов. – Москва : МО СССР, 1989. – 660 с.
11. Статистические модели и многокритериальность задачи принятия решений / под ред. И. Ф. Шахнова. – Москва : Статистика, 1984. – 563 с.
12. Терентьев В. М. Теоретические основы управления сетями многоканальной радиосвязи / В. М. Терентьев, И. Б. Паращук. – Санкт-Петербург : ВАС, 1995. – 195 с.
13. Шаповалова Г. М. Информационное общество: от цифровых архивов к цифровому культурному наследию / Г. М. Шаповалова // Международный научно-исследовательский журнал. – 2016. – № 5–6. – С. 177–181.
14. Beel J. Research-paper recommender systems: a literature survey / J. Beel, B. Gipp, S. Lange, C. Breitinger // International Journal on Digital Libraries. – 2015. № 17 (4). – P. 305–338.
15. Besser H. The Past, Present, and Future of Digital Libraries / H. Besser // A Companion to Digital Humanities. – Blackwell Publishing Ltd, 2004. – P. 557–575.
16. Bini D. Numerical Methods for Structured Markov Chains / D. Bini, G. Latouche, B. Meini. – Oxford University Press ; New York, 2005. – 215 p.
17. Candela L. History, Evolution and Impact of Digital Libraries / L. Candela, D. Castelli, Pagano. – E-Publishing and Digital Libraries : Legal and Organizational Issues, IGI Global, 2011. – P. 1–30.
18. Ethier S. N., Kurtz T. G. Markov processes. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics / S. N. Ethier, T. G. Kurtz. – New York : John Wiley & Sons Inc., 1986. – P. 214–234.
19. Isah A. Digital Libraries: Analysis of Delos Reference Model and 5S Theory / A. Isah, B. C. Serema, A. Mutshewa, L. Kenosi // Journal of Information Science Theory and Practice. – 2013. – № 1 (4). – P. 38–47.
20. Kokoykina O. N. The development of the information society and issues of digital preservation. Formation of regional cultural policy in the context of the modernization of education / O. N. Kokoykina // International scientific-practical conf. – Orel, 2014. – P. 53–56.

21. Kotenko I. V. Analysis of the Sensitivity of Algorithms for Assessing the Harmful Information Indicators in the Interests of Cyber-Physical Security / I. V. Kotenko, I. B. Parashchuk // *Cyber-Physical Systems*. – 2019. – № 8 (3). – P. 1–15.
22. Lanagan J. Video digital libraries: contributive and decentralized / J. Lanagan, A. F. Smeaton // *International Journal on Digital Libraries*. – 2012. – № 12 (4). – P. 159–178.
23. Papy F. *Digital Libraries* / F. Papy. – 1st ed. – Amsterdam, Netherlands, 2016. – 152 p.
24. Pomerantz J., Marchionini G. The digital library as place / J. Pomerantz, G. Marchionini // *Journal of Documentation*. – 2007. – № 63 (4). – P. 505–533.
25. Reilly B. F. Jr. Rethinking the National Library / B. F. Reilly // *American Libraries*. – 2016. – Vol. 47, iss. 11/12. – 30 p.
26. Russian Association of Digital Libraries 2020. – Режим доступа: <http://www.aselib.ru/index.html>, свободный. – Заглавие с экрана. – Яз. рус. (дата обращения: 17.01.2020).
27. Trivedi M. Digital Libraries: Functionality, Usability, and Accessibility. *Library Philosophy and Practice (e-journal)* / M. Trivedi. – 2010. – P. 381. – Режим доступа: <http://digitalcommons.unl.edu/libphilprac/381>, свободный. – Заглавие с экрана. – Яз. рус. (дата обращения: 17.01.2020).
28. Van Le C. Opening the Doors to Digital Libraries: A Proposal to Exempt Digital Libraries From the Copyright Act / Van Le C. // *Case Western Reserve Journal of Law, Technology & The Internet*. – 2010. – № 1-2. – P. 135.
29. Witten I. H. *How to Build a Digital Library* / I. H. Witten, D. Bainbridge, D. M. Nichols. – 2nd ed. – Walham, US, 2009. – 656 p.
30. Yin G. G. *Continuous-Time Markov Chains and Applications: A Singular Perturbation Approach* / G. G. Yin, Q. Zhang. – Springer, 2012. – 351 p.

References

1. Bogdanova I. F., Bogdanova N. F. *Elektronnye biblioteki: istoriya i sovremennost* [Digital Libraries: History and Present]. *Informatsionnoe obshchestvo: obrazovanie, nauka, kultura i tekhnologii budushchego* [Information society: education, science, culture and future technologies], 2017, issue 1, pp. 133–154.
2. Burenin N. I., Rodimov A. P., Terentev V. M. *Metodologicheskie aspekty primeneniya metodov mnogokriterialnoy dinamicheskoy optimizatsii v zadachakh sinteza upravlyaemykh setey mnogokanalnoy radiosvyazi* [Methodological aspects of the application of multicriteria dynamic optimization methods in the synthesis of controlled multichannel radio communication networks]. Leningrad, VAS Publ., 1984. 92 p.
3. Zeyfman A. I., Bening V. E., Sokolov I. A. *Markovskie tsepi i modeli s nepreryvnym vremenem* [Markov chains and continuous time models]. Moscow, ELEKS-KM Publ., 2008. 167 p.
4. Zorin A. V., Zorin V. A., Proydakova E. V., Fedotkin M. A. *Vvedenie v obshhie cepi Markova* [Introduction to general Markov chains]. Nizhny Novgorod, Nizhny Novgorod State University, 2013. 51 p.
5. Zuykina K. L., Sokolova D. V., Skalaban A. V. *Elektronnye biblioteki v Rossii. Tekushchiy status i perspektivy razvitiya* [Digital Libraries in Russia. Current status and development prospects]. Moscow, Vash format Publ., 2017. 120 p.
6. Kalinin V. N. *Teoreticheskie osnovy sistemnykh issledovaniy* [Theoretical foundations of system research]. Saint Petersburg, 2013. 278 p.
7. Kryukova E. S., Parashchuk I. B. *Osobennosti razvitiya sovremennykh elektronnykh bibliotek i analiz podkhodov k otsenivaniyu ikh kachestva* [Features of the development of modern electronic libraries and analysis of approaches to assessing their quality]. *Sovremennye tekhnologii: aktualnye voprosy, dostizheniya i innovatsii : sbornik statey XXXI Mezhdunarodnoy nauchno-prakticheskoy konferentsii* [Modern technologies: current issues, achievements and innovations: collection of articles of the XXXI International scientific and practical conference]. Penza, Nauka i Prosveshchenie, 2019, pp. 34–36.
8. Matveev M. Yu. *Natsionalnye biblioteki: na puti k tsifrovomu budushchemu* [National Libraries: Towards a Digital Future]. *Vestnik Sankt-Peterburskogo gosudarstvennogo instituta kulturey* [Bulletin of Saint Petersburg Institute of Culture], 2017, no. 2 (31), pp. 88–95.
9. Nenadovich D. M., Parashchuk I. B. *Analiz chuvstvitelnosti protsessov filtratsii sostoyaniy upravlyaemoy radiotekhnicheskoy sistemy* [Sensitivity analysis of state filtration processes of a controlled radio engineering system]. *Radiotekhnika* [Radiotechnics], 1997, no. 4, pp. 23–25.
10. Petukhov G. B. *Osnovy teorii e ffektivnosti tselenapravlennykh protsessov* [Fundamentals of the theory of the effectiveness of targeted processes]. Moscow, Ministry of Defense of the USSR, 1989. 660 p.
11. Shakhnov I. F. (ed.) *Statisticheskie modeli i mnogokriterialnost zadachi prinyatiya resheniy* [Statistical models and multicriteria decision-making problems]. Moscow, Statistika, 1984. 563 p.
12. Terentev V. M., Parashchuk I. B. *Teoreticheskie osnovy upravleniya setyami mnogokanalnoy radiosvyazi* [Theoretical foundations of multichannel radio network management]. Saint Petersburg, VAS Publ., 1995. 195 p.
13. Shapovalova G. M. *Informatsionnoe obshchestvo: ot tsifrovyykh arkhivov k tsifrovomu kulturnomu naslediyu* [The Information Society: From Digital Archives to Digital Cultural Heritage]. *Mezhdunarodnyy nauchno-issledovatel'skiy zhurnal* [International Research Journal], 2016, no. 5-6, pp. 177–181.
14. Beel J., Gipp B., Lange S., Breitinger C. Research-paper recommender systems: a literature survey. *International Journal on Digital Libraries*, 2015, no. 17 (4), pp. 305–338.
15. Besser H. *The Past, Present, and Future of Digital Libraries. A Companion to Digital Humanities*. Blackwell Publishing Ltd, 2004, pp. 557–575.

16. Bini D., Latouche G., Meini B. *Numerical Methods for Structured Markov Chains*. Oxford University Press; New York, 2005. 215 p.
17. Candela L., Castelli D., Pagano. *History, Evolution and Impact of Digital Libraries. E-Publishing and Digital Libraries: Legal and Organizational Issues*. IGI Global, 2011, pp. 1–30.
18. Ethier S. N., Kurtz T. G. *Markov processes. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics*. New York, John Wiley & Sons Inc., 1986, pp. 214–234.
19. Isah A., Serema B. C., Mutsheba A., Kenosi L. Digital Libraries: Analysis of Delos Reference Model and 5S Theory. *Journal of Information Science Theory and Practice*, 2013, no. 1 (4), pp. 38–47.
20. Kokoykina O. N. The development of the information society and issues of digital preservation. Formation of regional cultural policy in the context of the modernization of education. *International scientific-practical conf.*, Orel, 2014, pp. 53–56.
21. Kotenko I. V., Parashchuk I. B. Analysis of the Sensitivity of Algorithms for Assessing the Harmful Information Indicators in the Interests of Cyber-Physical Security. *Cyber-Physical Systems*, 2019, no. 8 (3), pp. 1–15.
22. Lanagan J., Smeaton A. F. Video digital libraries: contributive and decentralized. *International Journal on Digital Libraries*, 2012, no. 12 (4), pp. 159–178.
23. Papy F. *Digital Libraries*. 1st ed. Amsterdam, Netherlands, 2016. 152 p.
24. Pomerantz J., Marchionini G. The digital library as place. *Journal of Documentation*, 2007, no. 63 (4), pp. 505–533.
25. Reilly B. F. Jr. Rethinking the National Library. *American Libraries*, 2016, vol. 47, iss. 11-12. 30 p.
26. *Russian Association of Digital Libraries 2020*. Available at: <http://www.aselib.ru/index.html> (accessed 17.01.2020).
27. Trivedi M. Digital Libraries: Functionality, Usability, and Accessibility. *Library Philosophy and Practice (e-journal)*, 2010, p. 381. Available at: <http://digitalcommons.unl.edu/libphilprac/381> (accessed 17.01.2020).
28. Van Le C. Opening the Doors to Digital Libraries: A Proposal to Exempt Digital Libraries From the Copyright Act. *Case Western Reserve Journal of Law, Technology & The Internet*, 2010, no 1-2, p. 135.
29. Witten I. H., Bainbridge D., Nichols D. M. *How to Build a Digital Library*. 2nd ed. Waltham, US, 2009. 656 p.
30. Yin G. G., Zhang Q. *Continuous-Time Markov Chains and Applications: A Singular Perturbation Approach*. Springer, 2012. 351 p.

DOI 10.21672/2074-1707.2020.49.4.131-143

УДК 519.6, 004.021

ПРИМЕНЕНИЕ РЕШЕНИЯ ОБРАТНОЙ ЗАДАЧИ МАТЕМАТИЧЕСКОГО МОДЕЛИРОВАНИЯ ГАЛЬВАНИЧЕСКОГО ПРОЦЕССА ДЛЯ ОПТИМИЗАЦИИ НЕРАВНОМЕРНОСТИ ТОЛЩИНЫ ПОКРЫТИЯ

Статья получена редакцией 20.02.2020, в окончательном варианте – 04.03.2020.

Соловьев Денис Сергеевич, Тамбовский государственный университет имени Г.Р. Державина, 392036, Российская Федерация, г. Тамбов, ул. Интернациональная, 33, кандидат технических наук, e-mail: solovjevdenis@mail.ru

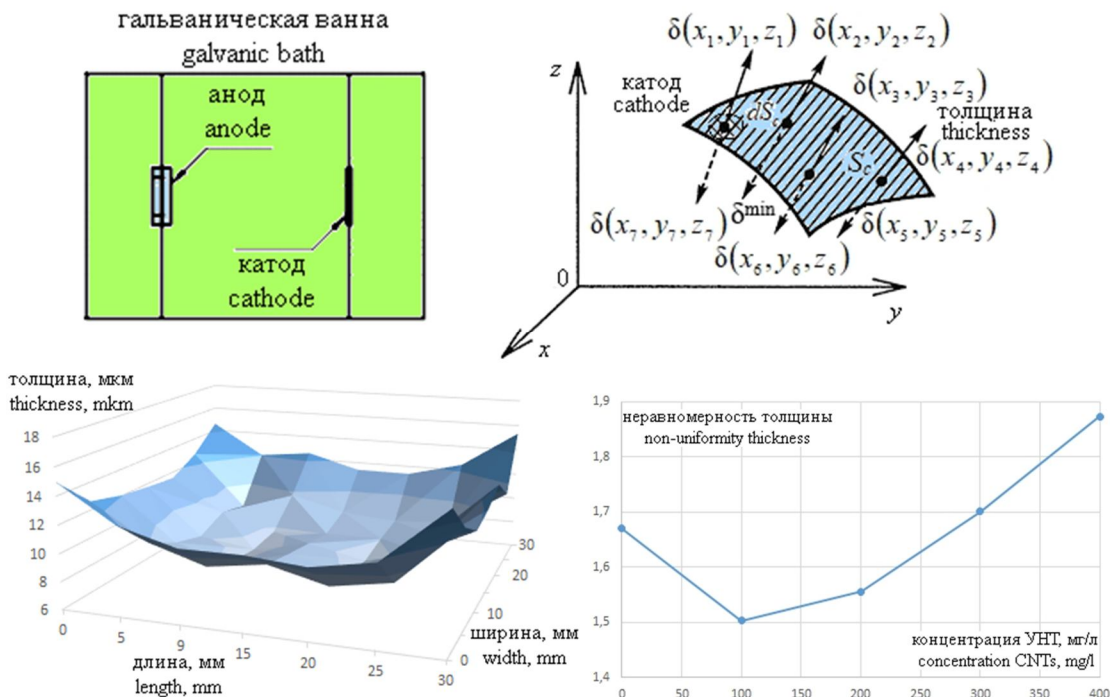
Соловьева Инна Александровна, Тамбовский государственный технический университет, 392000, Российская Федерация, г. Тамбов, ул. Советская, 106, аспирант, e-mail: good.win32@yandex.ru

Литовка Юрий Владимирович, Тамбовский государственный технический университет, Российская Федерация, 392000, г. Тамбов, ул. Советская, 106, доктор технических наук, профессор, e-mail: polychem@list.ru

В работе рассматривается применение углеродных нанотрубок для улучшения равномерности гальванических покрытий. Сформулирована задача поиска концентрации углеродных нанотрубок в электролите для улучшения равномерности толщины покрытия. Разработана математическая модель гальванического процесса, которая учитывает концентрацию углеродных нанотрубок в электролите. Для решения прямой задачи оптимизации (поиска оптимальной концентрации углеродных нанотрубок в электролите) поставлена обратная задача поиска функции катодной поляризации при добавлении в электролит углеродных нанотрубок. Увеличение скорости решения обратной задачи может быть достигнуто с применением технологии распараллеливания вычислений. Составлен алгоритм решения для прямой и обратной задач оптимизации. Для хранения требуемой информации разработаны базы данных. Проведена серия компьютерных и экспериментальных исследований по гальваническому нанесению цинкового покрытия с добавлением углеродных нанотрубок в электролит. Доказано, что присутствие углеродных нанотрубок в растворе электролита способно положительно влиять на равномерность получаемого покрытия.

Ключевые слова: прямая задача оптимизации, обратная задача оптимизации, математическое моделирование, эксперимент, углеродные нанотрубки, алгоритм, база данных, неравномерность покрытия

Графическая аннотация (Graphical annotation)



INVERSE PROBLEM SOLUTION OF THE MATHEMATICAL MODELING FOR GALVANIC PROCESS TO OPTIMIZE THE NON-UNIFORMITY OF THE COATING THICKNESS

The article was received by the editorial board on 20.02.2020, in the final version – 04.03.2020.

Solovjev Denis S., Tambov State University named after G.R. Derzhavin, 33 Internatsionalnaya St., Tambov, 392036, Russian Federation,

Cand. Sci. (Engineering), e-mail: solovjevdenis@mail.ru

Solovjeva Inna A., Tambov State Technical University, 106 Sovetskaya St., Tambov, 392000, Russian Federation,

post-graduate student, e-mail: good.win32@yandex.ru

Litovka Yuri V., Tambov State Technical University, 106, Sovetskaya St., Tambov, 392000, Russian Federation,

Doct. Sci. (Engineering), Professor, e-mail: polychem@list.ru

The article considers the use of carbon nanotubes to improve the uniformity of galvanic coatings. The problem of searching the concentration of carbon nanotubes in an electrolyte is formulated to improve the uniformity of the coating thickness. The mathematical model of the galvanic process has been developed that takes into account the concentration of carbon nanotubes in the electrolyte. The inverse problem of searching for the cathodic polarization function when carbon nanotubes are added to the electrolyte is posed to solve this (direct) optimization problem. The use of parallelization technology allows to increase the speed of solving the inverse problem. An algorithm for solving the direct and inverse optimization problems is compiled. Databases have been developed to store the required information. A series of computer and experimental studies on the galvanic deposition of a zinc coating using carbon nanotubes in an electrolyte was carried out. It is proved that the presence of carbon nanotubes in an electrolyte solution can have a positive effect on the uniformity of the resulting coating.

Key words: direct optimization problem, inverse optimization problem, mathematical modeling, experiment, carbon nanotubes, algorithm, database, coating non-uniformity

Введение. Исключительные свойства углеродных нанотрубок (УНТ), такие как жесткость, низкая плотность, высокая механическая прочность, электро- и теплопроводность, привлекли исследователей для использования их в различных структурах [1–11]. Не избежали модификации в виде использования УНТ и гальванические покрытия, применяемые для защиты деталей от коррозионного поражения и повышения их срока службы, придания декоративного вида и

улучшения показателей качественных характеристик. Среди них важное место занимает равномерность толщины покрытия. Обоснованием применения в составах электролитов наноматериалов являются результаты экспериментов по нанесению модифицированных УНТ гальванических покрытий с улучшенными показателями по сравнению с традиционными технологическими процессами их получения [12, 13]. Для отыскания оптимальной концентрации УНТ в электролите, обеспечивающей высокую равномерность покрытия, необходимо проводить многочисленные эксперименты, что связано со значительными материальными и временными затратами. Поэтому моделирование и оптимизация процессов нанесения модифицированных УНТ гальванических покрытий являются важными не только с технологических позиций и экономической точки зрения, но также и для уменьшения времени работы персонала в условиях воздействия агрессивных сред при проведении натурных экспериментов.

Целью данной работы является оптимизация равномерности модифицированных УНТ покрытий на основе математического моделирования гальванического процесса.

Постановка прямой и обратной задачи оптимизации. Для достижения поставленной цели сформулируем сначала задачу оптимизации. Определить степень неравномерности толщины покрытия можно согласно [14]:

$$R = \frac{1}{S_c} \int_{S_c} \frac{\delta(x, y, z) - \delta^{\min}}{\delta^{\min}} dS_c, \quad (1)$$

где δ – толщина покрытия в точке с координатами (x, y, z) на поверхности катода S_c ; δ^{\min} – минимальная толщина покрытия.

На рисунке 1 приведена графическая иллюстрация расчета неравномерности толщины покрытия по формуле (1) согласно принятым обозначениям.

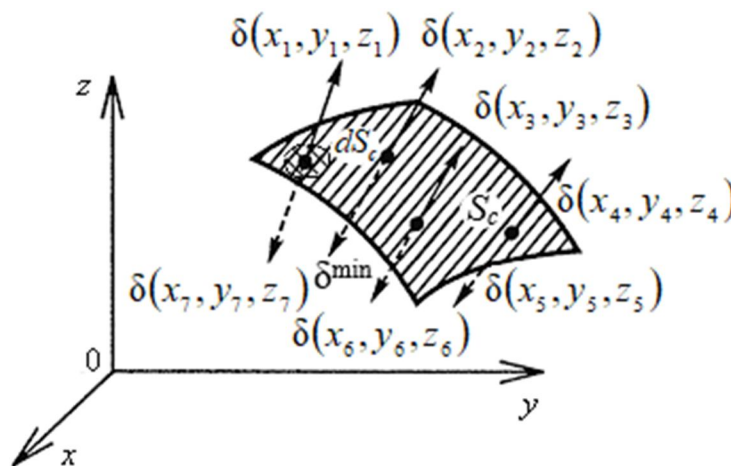


Рисунок 1 – Графическая иллюстрация расчета неравномерности толщины покрытия

Система уравнений математической модели гальванического процесса содержит уравнения Фарадея, Ома и Лапласа и имеет следующий вид:

$$\left\{ \begin{array}{l} \delta(x, y, z) = \frac{k\eta(t)T}{\rho} i_c(x, y, z) \\ i_a(x, y, z) = \chi \nabla \varphi(x, y, z) S_a \\ i_c(x, y, z) = -\chi \nabla \varphi(x, y, z) S_c \\ \Delta \varphi(x, y, z) = 0 \\ \varphi(x, y, z) + F_1(i_a) S_a = U \\ \varphi(x, y, z) + F_2(i_c, C_{УНТ}) S_c = 0 \\ \left. \frac{\partial \varphi(x, y, z)}{\partial \bar{n}} \right|_{S_{ms}} = 0 \end{array} \right., \quad (2)$$

где T – время; t – температура электролита; k – электрохимический эквивалент; η – выход металла по току; ρ – плотность металла покрытия; χ – электропроводность электролита; i_a, i_c – плотность тока на аноде и катоде; S_a, S_{ins} – площади поверхности анода и изолятора; \vec{n} – нормаль к S_{ins} ; φ – электрический потенциал поля; U – напряжение; $C_{УНТ}$ – концентрация УНТ в электролите; F_1, F_2 – функции анодной и катодной поляризации; L_x, L_y, L_z – длина, ширина и высота гальванической ванны.

На рисунке 2 показана схема расчета уравнений математической модели (2).

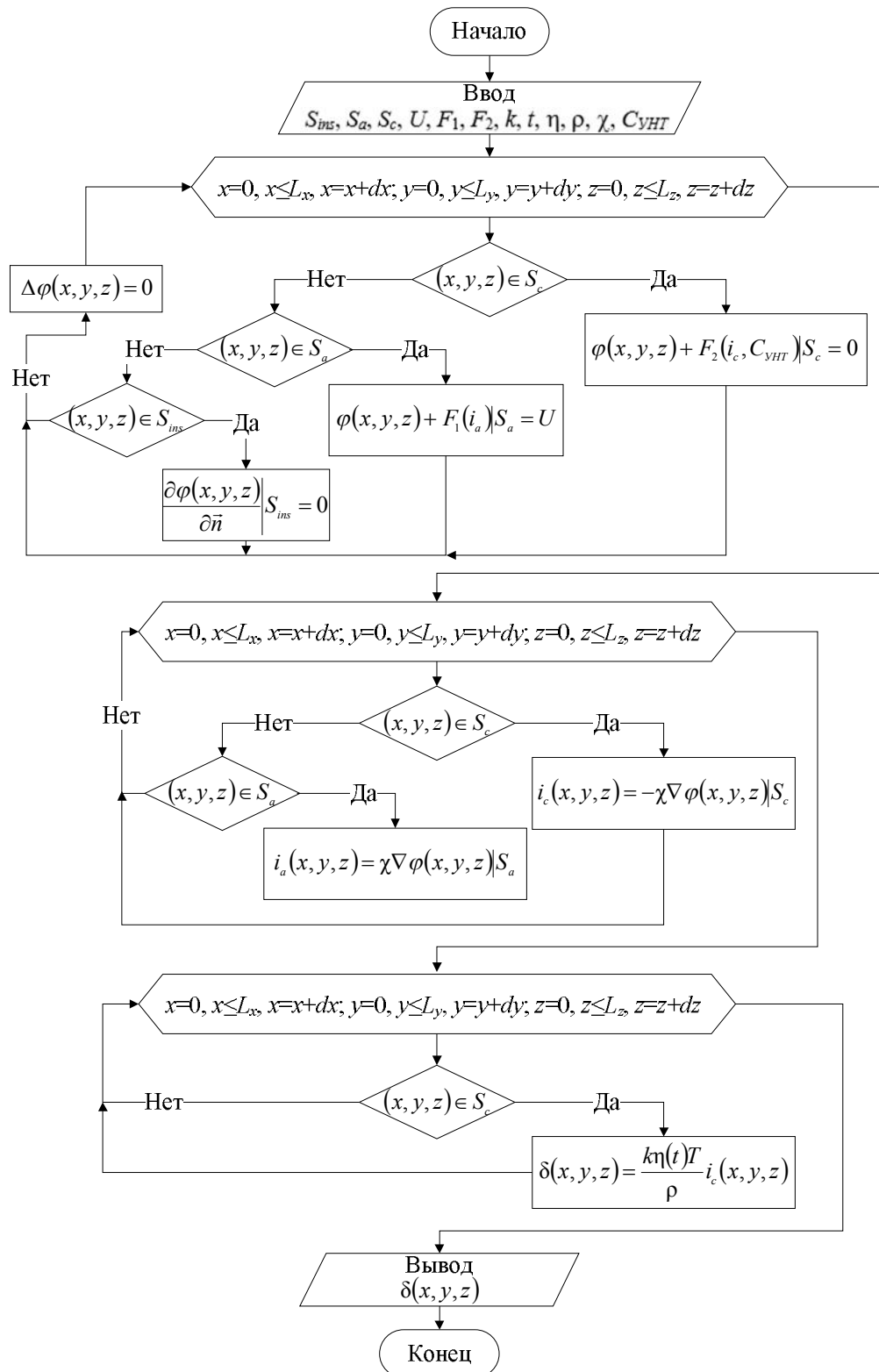


Рисунок 2 – Схема организации расчета по уравнениям математической модели

Отыскать значение концентрации $C_{УНТ}^*$, для которой коэффициент R в (1) минимален:

$$R \rightarrow \min, \tag{3}$$

при уравнениях связи (2) и следующих ограничениях:

$$\begin{cases} 0 \leq C_{УНТ} \leq C_{УНТ}^{\max} \\ \delta^{\min} \geq \delta^{\text{зад}} \end{cases}, \tag{4}$$

где \max , \min , зад – максимальное, минимальное и заданное значение.

Решение поставленной (прямой) задачи (3) может быть найдено с использованием методов нелинейного программирования [15]. Однако из-за того в присутствии УНТ в электролите функция F_2 изменяется, а известная аппроксимирующая зависимость для нее отсутствует. Поэтому сначала необходимо решить обратную задачу. Ее решение осложняется наличием в математической модели (2) уравнения в частных производных. Для отыскания аппроксимирующей зависимости воспользуемся методом Ритца [16] и представим функцию F_2 следующим образом:

$$F_2(i_c, C_{УНТ}) = F_2(i_c) + \sum_{j=0}^{N-1} a_j C_{УНТ}^j + a_N F_2(i_c) C_{УНТ}, \tag{5}$$

где $\vec{a} = (a_0, \dots, a_N)$ – неизвестные коэффициенты; $F_2(i_c)$ – функция поляризации без добавления УНТ.

Тогда обратную задачу можно сформулировать следующим образом.

Необходимо отыскать такие значения $\vec{a}^* = (a_0^*, \dots, a_N^*)$ в (5), для которых отклонение локальных значений толщины покрытия в выбранных точках поверхности для эксперимента и результатов моделирования являются минимальными:

$$\Lambda = \frac{1}{HS_c} \sum_{h=1}^H \int_{S_c} \frac{|\delta_h^\varepsilon(x, y, z) - \delta_h^p(x, y, z)|}{\delta_h^\varepsilon(x, y, z)} dS_c \rightarrow \min, \tag{6}$$

где H – количество экспериментов; индексы ε, p – соответственно экспериментальное и расчетное значения.

Для увеличения быстродействия нахождения решения задачи (6) возможно применение распараллеливания вычислений при определении критерия Λ ввиду независимости вычисления слагаемых H в нем. Пусть количество параллельных потоков вычислений G совпадает с количеством ядер процессора. Тогда разобьем диапазон $[1; H]$ на G равных частей, в каждой из которых вычисляется значение критерия Λ_g :

$$\Lambda_g = \frac{1}{S_c} \sum_{h=\frac{H}{G}(g-1)+1}^{\frac{H}{G}g} \int \frac{|\delta_h^\varepsilon(x, y, z) - \delta_h^p(x, y, z)|}{\delta_h^\varepsilon(x, y, z)} dS_c, \tag{7}$$

оформленное в отдельный поток. По завершении вычислений во всех G потоках определяется величина критерия Λ :

$$\Lambda = \sum_{g=1}^G \Lambda_g. \tag{8}$$

На рисунке 3 показана блок-схема основных шагов алгоритма решения прямой и обратной задач.

Поиск решения задачи оптимизации (6) предполагает значительный объем вычислений. Однако это не является недостатком, поскольку определение функции F_2 , согласно (5), на основе экспериментальных данных осуществляется однократно с последующим использованием в математической модели (2) для получения решения прямой задачи (3).

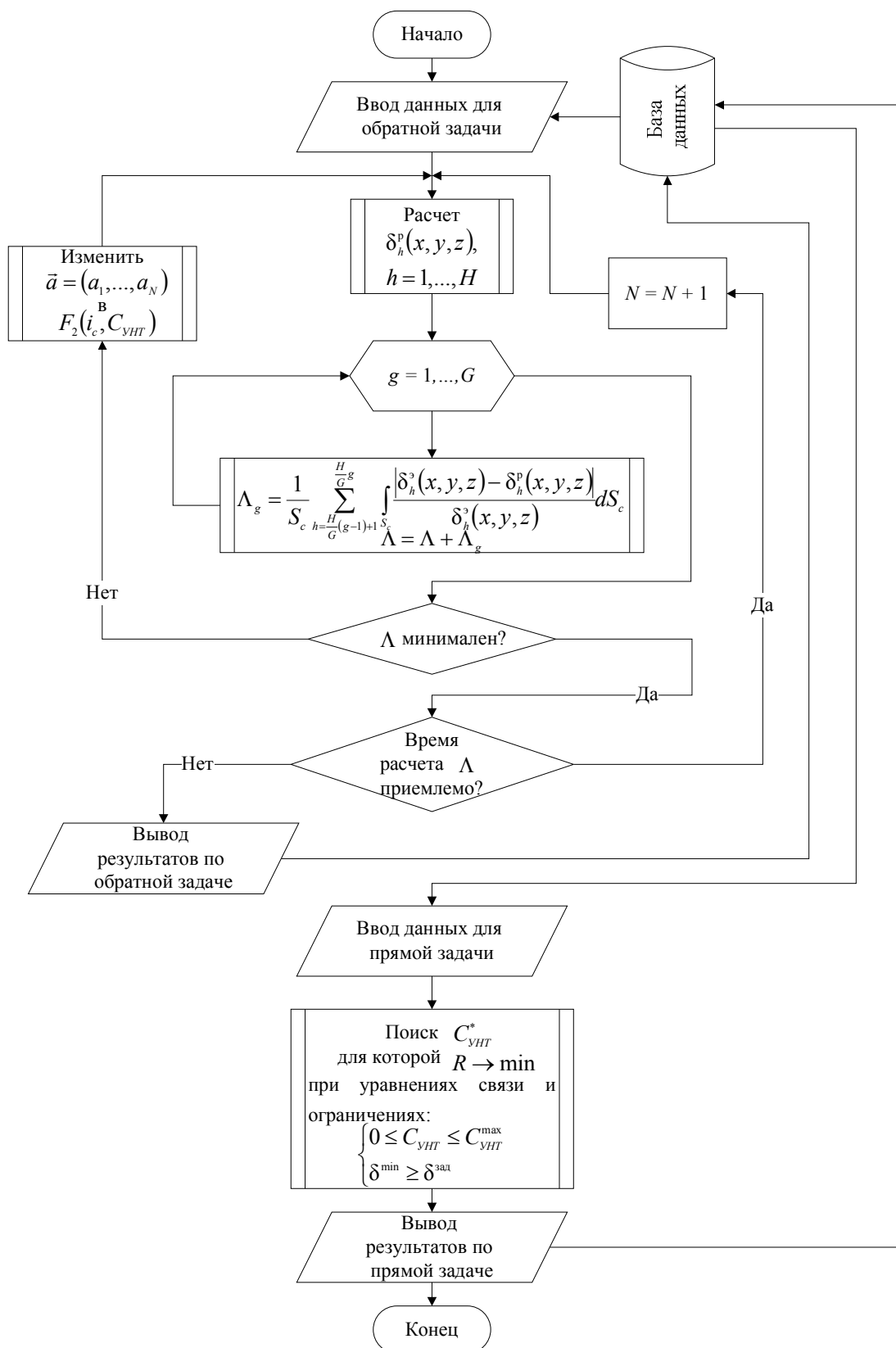


Рисунок 3 – Блок-схема алгоритма решения прямой и обратной задач

Даталогическая модель базы данных. Для реализации решения задач оптимизации была выбрана реляционная модель базы данных (БД). Она получила широкое распространение вследствие наиболее простой формы представления данных и является привычной для специалиста, пользующегося различной справочной литературой. В рамках этой модели все данные, необходимые в процессе решения задач, легко представляются в табличной форме. Кроме того, реляционная модель достаточно проста в реализации с помощью современных средств разработки приложений, она проста и в отношении ведения БД. Для хранения значительного объема требуемой информации в результате решения сформулированных прямой и обратной задач необходимы следующие БД [17]: готовых проектов, оборудования, электролитов и экспериментов.

Даталогическая модель базы данных приводится на рисунке 4.

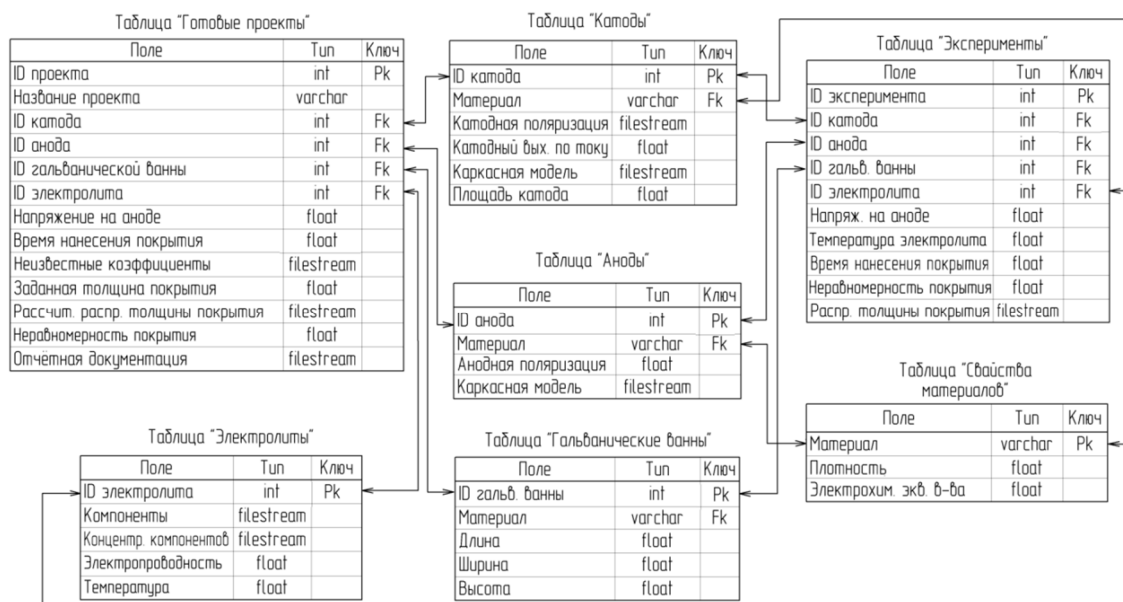


Рисунок 4 – Даталогическая модель базы данных

БД «Готовые проекты» содержит сведения об оптимизированных наномодифицированных гальванопокрытиях с найденной F_2 .

БД «Оборудование» состоит из таблиц «Гальванические ванны», «Катоды», «Аноды» и «Свойства материалов». Таблица «Гальванические ванны» содержит сведения об используемых ваннах: размерах и материалах, из которых они изготовлены. В таблице «Катоды» описаны применяемые катоды – в частности описывается материал катода, каркасная модель катода, катодная поляризация без учета $S_{УНТ}$. Таблица «Аноды» содержит сведения об используемых анодах и аналогична по составу таблице «Катоды». В таблице «Свойства материалов» описаны физико-химические характеристики применяемых материалов.

БД «Электролиты» включает одноименную таблицу, содержащую сведения о составе электролита и его физико-химических характеристиках.

В БД «Эксперименты» содержится таблица, поля которой отражают сведения о проводимых экспериментах: используемые ванны, катоды, аноды, электролиты; режимные параметры; распределение покрытия по катоду.

В связи с тем, что было отдано предпочтение реляционной модели данных, в качестве наиболее подходящей системы управления БД была выбрана программа Microsoft SQL Server 2019 15.0. Она имеет широкий спектр функций, а также используется для взаимодействия с БД размером от персональных до крупных, применяемых на предприятиях; является удобной в использовании.

Экспериментальная часть. Для решения сформулированных задач была проведена серия из $N = 20$ экспериментов по нанесению гальванического покрытия при различных концентрациях $S_{УНТ}$ из диапазона $[0; 400]$ мг/л на стальные образцы квадратной формы S_c размером $L_1 \times L_2 = 30 \times 30$ мм² с шероховатостью до $Rz = 10$ мкм и плотностью тока $i_c = 3$ А/дм² при времени покрытия $T = 20$ мин для получения $\delta^{зад} = 6$ мкм [18]. Подготовка поверхности катода включает в себя: очистку; обезжиривание; кислотное протравливание; промывку под проточной водой. В качестве металла покрытия выступал цинк, раствора - электролит для блестящего цинкования, УНТ «Таунит» – одномерные

наномасштабные нитевидные образования поликристаллического графита длиной более 2 мкм с наружными диаметрами от 15 до 40 нм в виде сыпучего порошка черного цвета [19]. Температура электролита $t = 20^\circ\text{C}$ при напряжении $U = 1\text{ В}$. Размеры ванны, в которой проходил процесс, представлены на рисунке 5 в миллиметрах.

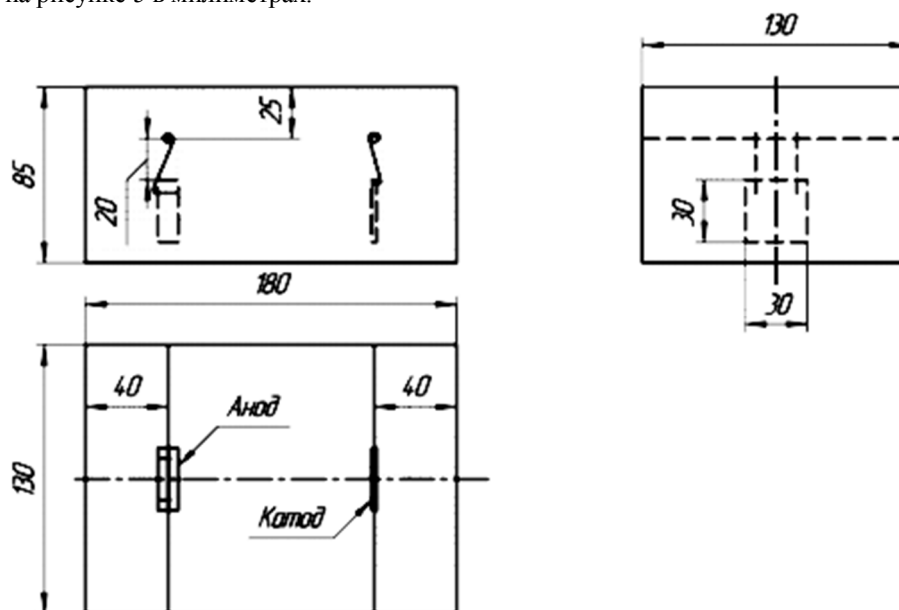


Рисунок 5 – Схема экспериментальной установки

Анод, изготовленный из цинка марки Ц0, при соотношении $S_a:S_c = 2:1$ с помощью двух токопроводящих крючков, закреплялся на медном проводнике, соединенном через клеммы с источником питания. Стенка катода, не обращенная к аноду, изолировалась. Перед началом нанесения гальванического покрытия в электролит вводился УНТ «Таунит». Дозирование исходного количества УНТ «Таунит» в электролит осуществлялось с использованием дозатора QUANTOS. При введении УНТ в виде порошка, последний подвергался диспергированию с использованием проточного ультразвукового диспергатора ИЛ 100-6/9. Для этого организовывался циркуляционный контур, по которому электролит перекачивался из ванны через проточный диспергатор и возвращался в ванну.

В результате проведенных экспериментов измерялась толщина покрытия $\delta_h^3(x, y, z)$ согласно [20] с использованием прибора «Константа К5». Данный прибор реализует использование комбинации импульсного индукционного, вихревого параметрического и фазового цифровых методов получения первичной информации, а также обладает возможностью контроля покрытий на деталях с шероховатостью до $Rz = 400\text{ мкм}$ на малоразмерных деталях и внутренних поверхностях полых изделий.

Решение прямой задачи оптимизации находилось с использованием метода золотого сечения [21]:

$$\begin{cases} \alpha_{k+1} = C_{УНТ1}^{(k+1)} \\ \beta_{k+1} = C_{УНТ2}^{(k+1)} \\ C_{УНТ1}^{(k+1)} = \alpha_{k+1} + \frac{3-\sqrt{5}}{2}(\beta_{k+1} - \alpha_{k+1}) \\ C_{УНТ2}^{(k+1)} = \beta_{k+1} - \frac{3-\sqrt{5}}{2}(\beta_{k+1} - \alpha_{k+1}) \end{cases} \quad (9)$$

Расчет уравнений распределений потенциала электрического поля в объеме электролита и плотности тока на поверхности катода в математической модели процесса осуществлялся с использованием методов конечных разностей и релаксации:

$$\varphi^{n+1}(x_i, y_j, z_k) = \omega \varphi^n(x_i, y_j, z_k) + (1 - \omega) \left[\frac{\varphi^n(x_{i-1}, y_j, z_k) + \varphi^n(x_{i+1}, y_j, z_k)}{2h_x^2 \left(\frac{1}{h_x^2} + \frac{1}{h_y^2} + \frac{1}{h_z^2} \right)} + \frac{\varphi^n(x_i, y_{j-1}, z_k) + \varphi^n(x_i, y_{j+1}, z_k)}{2h_y^2 \left(\frac{1}{h_x^2} + \frac{1}{h_y^2} + \frac{1}{h_z^2} \right)} + \frac{\varphi^n(x_i, y_j, z_{k-1}) + \varphi^n(x_i, y_j, z_{k+1})}{2h_z^2 \left(\frac{1}{h_x^2} + \frac{1}{h_y^2} + \frac{1}{h_z^2} \right)} - \varphi^n(x_i, y_j, z_k) \right] \quad (10)$$

$$i_c^{m+1}(x_i, y_j, z_k) = \omega i_c^m(x_i, y_j, z_k) + (1 - \omega) i_c^{m-1}(x_i, y_j, z_k) \quad (11)$$

где i, j, k – индексы дискретных значений координат (x, y, z) ; h_x, h_y, h_z – шаги сетки по соответствующим координатам; n, m – индексы итераций по потенциалу и плотности тока; ω – параметр релаксации, определяемый как:

$$\omega = \frac{2}{1 + \sin\left(\frac{\pi}{\max(N_x, N_y, N_z)}\right)}, \quad (12)$$

где N_x, N_y, N_z – количество узлов сетки по координатам (x, y, z) .

Решение обратной задачи оптимизации отыскивалось методом Пауэлла [22] с разбиением диапазона $[1; H]$ на $G = 4$ параллельных вычислительных потока с использованием механизма *Open MP* [23] на процессоре *AMD A10-5750M 2,5 GHz* с оперативной памятью 16 ГБ под управлением 64-разрядной операционной системы *Windows 10*.

Результаты и их обсуждение. При прохождении процесса электролиза на поверхности детали осаждается металл и частицы УНТ. Осаждение металла является первой фазой или матрицей, а осаждение УНТ, которые цементируются матрицей, – второй фазой.

Нанотрубки попадают на подложку в связи со следующими факторами. Во-первых, случайное нахождение вблизи покрываемой поверхности в начальный момент времени вследствие распределения УНТ в объеме электролита. Во-вторых, в процессе осаждения покрытия кристаллы металла механически «подводят» нанотрубки к поверхности. Добавление в гальванические электролиты УНТ способствует измельчению кристаллической структуры покрытия. Наночастицы, которые осаждаются вместе с металлом покрытия, выполняют роль дополнительных центров кристаллизации, оказывая влияние на неравномерность гальванического покрытия.

По результатам проведенных экспериментов на рисунке 6 представлена полученная зависимость критерия неравномерности R от $C_{УНТ}$.

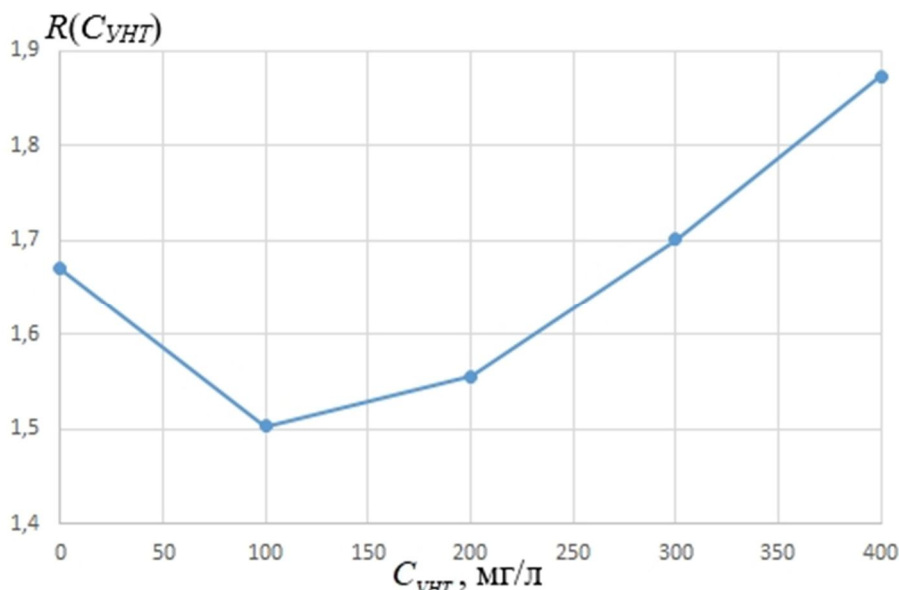


Рис. 6. Зависимость критерия неравномерности R от $C_{УНТ}$

Результаты определение количества неизвестных аппроксимирующих коэффициентов в (5) при решении обратной задачи (6) продемонстрированы на рисунке 7.

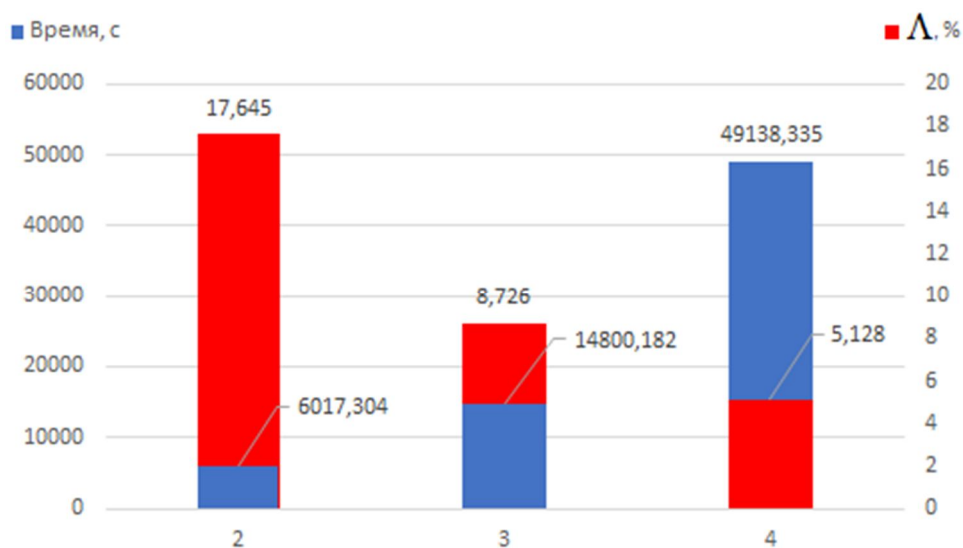


Рисунок 7 – Результаты решения обратной задачи для $N = 2$, $N = 3$ и $N = 4$

При увеличении количества элементов от $N = 2$ до $N = 3$ происходит снижение Λ на 8,91 % с увеличением времени решения задачи в 2,46 раза. При увеличении количества элементов от $N = 3$ до $N = 4$ происходит снижение Λ на 3,59 % с увеличением времени решения задачи в 3,32 раза. Поэтому в качестве количества неизвестных аппроксимирующих коэффициентов в (5) принимаем $N = 3$, которое обеспечивает $\Lambda = 8,73$ % соизмеримое отклонение результатов моделирования от точности измерительного прибора при допустимом времени решения обратной задачи.

На рисунке 8 приводится найденная функция поляризации F_2 (5) при $N = 3$, аналитическая запись которой имеет следующий вид:

$$F_2(i_c, C_{УНТ}) = F_2(i_c) + 0,9 \cdot 10^{-4} - 0,61 \cdot 10^{-2} C_{УНТ} + 0,92 \cdot 10^{-5} C_{УНТ}^2 - 0,43 \cdot 10^{-3} F_2(i_c) C_{УНТ},$$

где функция поляризации для электролита без добавления УНТ:

$$F_2(i_c) = 0,0087i_c^2 - 0,0873i_c - 0,36.$$

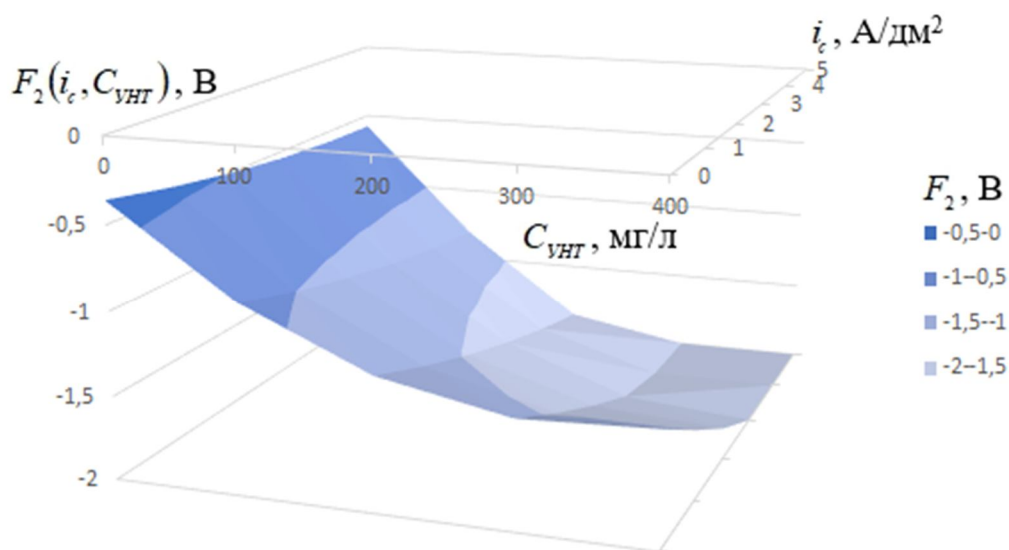


Рисунок 8 – Функция поляризации F_2

С использованием найденной $F_2(i_c, C_{УНТ})$ была решена прямая задача оптимизации (2)–(4) для рассматриваемой экспериментальной установки. На рисунке 9 демонстрируется толщина покрытия для расчета (а) и эксперимента (б) при найденном значении $C_{УНТ}^* = 179,5$ мг/л.

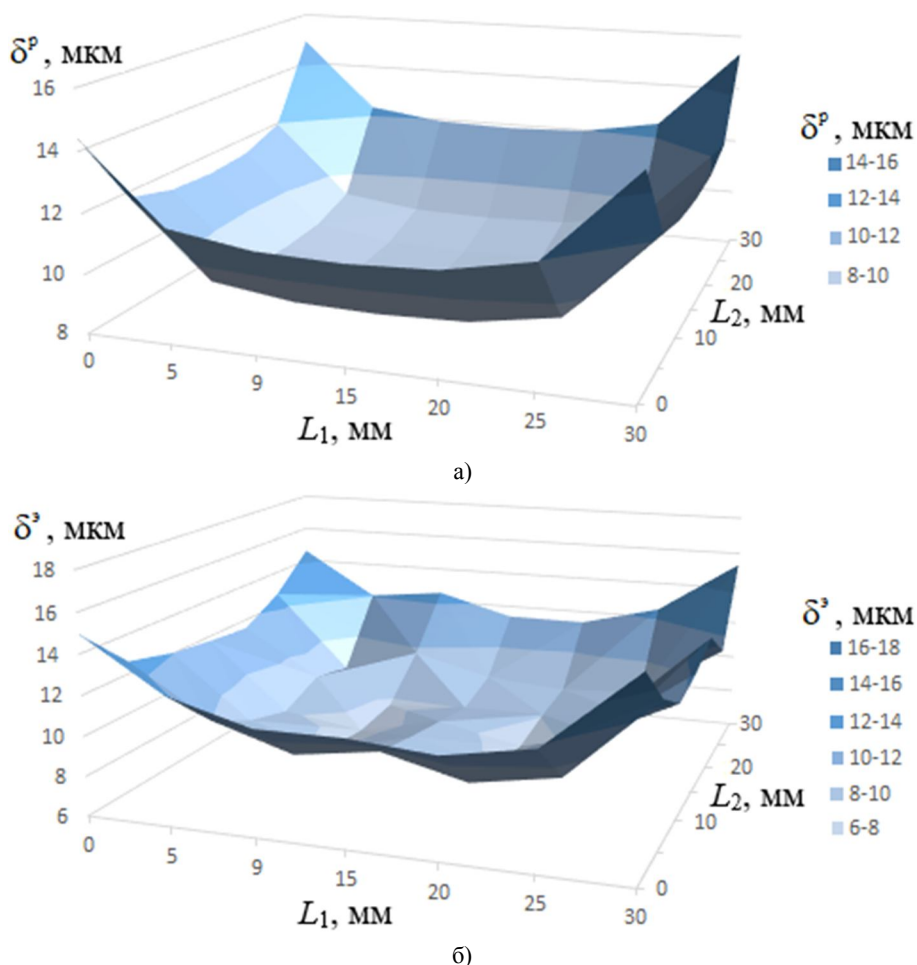


Рисунок 9 – Толщина покрытия для расчета (а) и эксперимента (б)

Расчетное и экспериментальное значение коэффициента R составляют соответственно 1,2726 и 1,3968. Для рассмотренного случая экспериментальное значение коэффициента R при добавлении УНТ снизилось на 16,7 % по сравнению с вариантом без добавления.

Заключение. Присутствие УНТ в растворе электролита способно оказывать положительный эффект на равномерность получаемого гальванического покрытия. Ввиду значительного многообразия металлов гальванических покрытий, составов электролитов и условий протекания процессов экспериментальный поиск оптимальной концентрации УНТ является неприемлемым решением. Применение методов математического моделирования, алгоритмов решения прямых и обратных задач оптимизации, реализованных с использованием современных технологий разработки программного обеспечения и баз данных, является наиболее подходящим способом поиска оптимальной концентрации УНТ для улучшения равномерности покрытия.

Полученные результаты представляют интерес для отдельных исследователей и организаций, занимающихся вопросами нанотехнологий, электрохимии, моделирования и оптимизации технологических процессов.

Библиографический список

1. Запороцкова И. В. Компьютерное моделирование взаимодействия тяжелых органических спиртов с однослойными углеродными нанотрубками / И. В. Запороцкова, Н. П. Поликарпова, Т. А. Ермакова, В. В. Яцышен // Прикаспийский журнал: управление и высокие технологии. – 2012. – № 3 (19). – С. 137–146. ([http://hi-tech.asu.edu.ru/files/3\(19\)/130-137.pdf](http://hi-tech.asu.edu.ru/files/3(19)/130-137.pdf))
2. Бадамшина Э. Р. Моделирование аномальных механических свойств полиуретана модифицированного углеродными однослойными нанотрубками / Э. Р. Бадамшина, Я. И. Эстрин, Г. С. Кулагина, С. А. Лурье, Ю. О. Соляев // Механика композиционных материалов и конструкций. – 2010. – Т. 16. – № 4. – С. 551–562.

3. Браже Р. А. Математическое моделирование спиральных супракристаллических нанотрубок / Р. А. Браже, А. Ф. Савин // Известия высших учебных заведений. Поволжский регион. Физико-математические науки. – 2015. – № 1 (33). – С. 120–129.
4. Созькин С. А. Выбор оптимальных параметров для моделирования атомной и электронной структуры углеродных нанотрубок в пакете SIESTA / С. А. Созькин, В. П. Бескачко, Г. П. Вяткин // Вестник Южно-Уральского государственного университета. Серия: Математика. Механика. Физика. – 2015. – Т. 7, № 3. – С. 78–85.
5. Нечаева Л. С. Компьютерное моделирование сорбции аминокислот на углеродных нанотрубках / Л. С. Нечаева, Е. В. Бутырская, С. А. Запрягаев // Журнал структурной химии. – 2017. – Т. 58, № 2. – С. 233–241.
6. Попов В. И. Компьютерное моделирование процессов распылительной сушки и искрового плазменного спекания керамоматричного и оксидного композиционного материала, армированного углеродными нанотрубками / В. И. Попов, М. В. Панченко, Н. В. Мамонова, Э. М. Кольцова // Успехи в химии и химической технологии. – 2018. – Т. 32, № 11 (207). – С. 53–55.
7. Леонтьев С. В. Многокритериальная оптимизация состава теплоизоляционного автоклавного газобетона, модифицированного дисперсией углеродных нанотрубок / С. В. Леонтьев, В. А. Шаманов, А. Д. Курзанов, Г. И. Яковлев // Строительные материалы. – 2017. – № 1-2. – С. 31–40.
8. Валуев И. А. Механизмы окисления бездефектных поверхностей углеродных наноструктур: влияние кривизны поверхности / И. А. Валуев, Г. Э. Норман, Б. Р. Шуб // Химическая физика. – 2011. – Т. 30, № 1. – С. 82–88.
9. Бандура А. В. Моделирование структуры и устойчивости углеродных нанолент / А. В. Бандура, В. А. Шур, Р. А. Эварестов // Журнал общей химии. – 2016. – Т. 86, № 8. – С. 1233–1243.
10. Кузнецов В. В. Конформационное поведение этана в нанотрубках / В. В. Кузнецов // Журнал органической химии. – 2013. – Т. 49, № 8. – С. 1245–1249.
11. Солтани М. Структура и свойства нанокompозита «магниевого сплава – углеродные нанотрубки» и их оптимизация с использованием метода планирования эксперимента / М. Солтани, М. Шаманьян, Б. Нируман // Металловедение и термическая обработка металлов. – 2017. – № 4 (742). – С. 59–65.
12. Хмыль А. А. Композиционные покрытия на основе проводящих металлов и углеродных кластеров для подвижных электрических контактов / А. А. Хмыль, В. Л. Ланин, Э. М. Шпилевский // Вестник Полоцкого государственного университета. Серия С: Фундаментальные науки. – 2006. – № 4. – С. 74–79.
13. Галперин В. А. Оптимизация процесса электрохимического осаждения меди в массивы углеродных нанотрубок / В. А. Галперин, П. К. Кондратьев, Д. Н. Зубов, Е. А. Кельм, А. А. Павлов // Нано- и микросистемная техника. – 2013. – № 4 (153). – С. 15–19.
14. Пэк В. В. Построение математической модели процесса нанесения наномодифицированного гальванического покрытия / В. В. Пэк, Ю. В. Литовка, Д. Ю. Попов, А. И. Хоботилина // Вестник Астраханского государственного технического университета. Серия: Управление, вычислительная техника и информатика. – 2010. – № 2. – С. 100–104.
15. Ширяев В. И. Исследование операций и численные методы оптимизации / В. И. Ширяев. – Москва : Ленанд, 2017. – 224 с.
16. Ritz W. Über eine neue methode zur lösung gewisser variationsprobleme der mathematischen physic / W. Ritz // Journal für die Reine und Angewandte Mathematik. – 1909. – Vol. 135. – pp. 1–61.
17. Голицына О. Л. Базы данных / О. Л. Голицына, Н. В. Максимов, И. И. Попов. – Москва : Форум, 2015. – 400 с.
18. Зажигаев Л. С. Методы планирования и обработки результатов физического эксперимента / Л. С. Зажигаев, А. А. Кишьян, Ю. И. Романиков. – Москва : Атомиздат, 2013. – 232 с.
19. Ткачев А. Г. Углеродный наноматериал Таунит – структура, свойства, производство и применение / А. Г. Ткачев // Перспективные материалы. – 2007. – № 3. – С. 5–9.
20. ГОСТ 9.302-88 Единая система защиты от коррозии и старения. Покрытия металлические и неметаллические неорганические. Методы контроля. – Москва : ИПК Издательство стандартов, 1990. – 40 с.
21. Kiefer J. Sequential minimax search for a maximum / J. Kiefer // Proceedings of the American Mathematical Society. – 1953. – Vol. 4 (3). – P. 502–506.
22. Powell M. J. D. An efficient method for finding the minimum of a function of several variables without calculating derivatives / M. J. D. Powell // Computer Journal. – 1964. Vol. 7 (2). – P. 155–162.
23. Антонов А. С. Параллельное программирование с использованием технологии OpenMP : учебное пособие / А. С. Антонов. – Москва : Изд-во МГУ, 2009. – 77 с.

References

1. Zaporotskova I. V., Polikarpova N. P., Ermakova T. A. Kompyuternoe modelirovanie vzaimodeystviya tyazhelykh organicheskikh spirtov s odnosloynymi uglerodnymi nanotrubbkami [Computer modeling of interaction of heavy organic alcohols with single-layer carbon nanotubes]. *Prikaspiyskiy zhurnal: upravlenie i vysokie tekhnologii* [Caspian Journal: Control and High Technologies], 2012, no. 3 (19), pp. 137–146. ([http://hi-tech.asu.edu.ru/files/3\(19\)/130-137.pdf](http://hi-tech.asu.edu.ru/files/3(19)/130-137.pdf))
2. Badamshina E. R., Estrin Ya. I., Kulaagina G. S., Lurie S. A., Solyaev Yu. O. Modelirovanie anomalnykh mekhanicheskikh svoystv poliuretana modifitsirovannogo uglerodnymi odnosloynymi nanotrubbkami [Scale modeling of abnormal mechanical properties of the polyurethane at modifying by the carbon single-walled nanotubes]. *Mekhanika kompozitsionnykh materialov i konstruksiy* [Journal on Composite Mechanics and Design], 2010, vol. 16, no. 4, pp. 551–562.

3. Brazhe R. A., Savin A. F. Matematicheskoe modelirovanie spiralnykh suprakristallicheskih nanotrubok [Mathematical modeling of the coiled supracrystalline nanotubes]. *Izvestiya vysshikh uchebnykh zavedeniy. Povolzhskiy region. Fiziko-matematicheskie nauki* [Physics and Mathematics Sciences. Physics], 2015, no. 1 (33), pp. 120–129.
4. Sozykin S. A., Beskachko V. P., Vyatkin G. P. Vybor optimalnykh parametrov dlya modelirovaniya atomnoy i elektronnoy struktury uglerodnykh nanotrubok v pakete SIESTA [Selection of the optimal parameters for simulation of atomic and electronic structure of carbon nanotubes by siesta package]. *Vestnik Yuzhno-Uralskogo gosudarstvennogo universiteta. Seriya: Matematika. Mekhanika. Fizika* [Bulletin of the South Ural State University. Series: Mathematics. Mechanics. Physics], 2015, vol. 7, no. 3, pp. 78–85.
5. Nechaeva L. S., Butyrskaya E. V., Zapryagaev S. A. Kompyuternoe modelirovanie sorbtсии aminokislot na uglerodnykh nanotrubkakh [Computer simulation of amino acid sorption on carbon nanotubes]. *Zhurnal strukturnoy himii* [Journal of Structural Chemistry], 2017, vol. 58, no. 2, pp. 217–225.
6. Popov V. I., Panchenko M. V., Mamonova N. V., Koltsova E. M. Kompyuternoe modelirovanie protsessov raspylitelnoyushki i iskrovogo plazmennogo spekaniya keramomatrichnogo i oksidnogo kompozitsionnogo materiala, armirovannogo uglerodnymi nanotrubkami [Computer simulation of the process of spray-drying and spark plasma sintering of ceramomatrix and oxide composites reinforced by carbon nanotubes]. *Uspekhi v khimii i khimicheskoy tekhnologii* [Advances in Chemistry and Technologies Used], 2018, vol. 32, no. 11 (207), pp. 53–55.
7. Leontev S. V., Shamanov V. A., Kurzanov A. D., Yakovlev G. I. Mnogokriterialnaya optimizatsiya sostava teploizolyacionnogo avtoklavnogo gazobetona, modifitsirovannogo dispersiyey uglerodnykh nanotrubok [Multiobjective Optimization of the Lightweight Autoclaved Aerated Concrete Modified with Carbon Nanotubes Dispersions]. *Stroitelnye materialy* [Construction Materials], 2017, no. 1–2, pp. 31–40.
8. Valuev I. A., Norman G. E., Shub B. R. Mekhanizmy okisleniya bezdefektnykh poverhnostey uglerodnykh nanostruktur: vliyaniye krivizny poverhnosti [Mechanisms of the oxidation of defect-free surfaces of carbon nanostructures: The influence of surface curvature]. *Khimicheskaya fizika* [Russian Journal of Physical Chemistry], 2011, vol. 5, no. 1, pp. 156–162.
9. Bandura A. V., Shur V. A., Evarestov R. A. Modelirovanie struktury i ustojchivosti uglerodnykh nanolent [Simulation of structure and stability of carbon nanoribbons]. *Zhurnal obshchey khimii* [Russian Journal of General Chemistry], 2016, vol. 86, no. 8, pp. 1777–1786.
10. Kuznetsov V. V. Konformacionnoye povedeniye etana v nanotrubkakh [Conformational behavior of ethane molecule encapsulated in a nanotube]. *Zhurnal organicheskoy khimii* [Russian Journal of Organic Chemistry], 2013, vol. 49, no. 8, pp. 1231–1235.
11. Soltani M., Shamanian M., Niroumand B. Struktura i svoystva nanokompozita "magnievyy splav – uglerodnye nanotrubki" i ikh optimizatsiya s ispolzovaniem metoda planirovaniya eksperimenta [Structure and properties of "magnesium alloy – carbon nanotubes" nanocomposite and their optimization using design of experiments]. *Metallovedeniye i termicheskaya obrabotka metallov* [Metal Science and Heat Treatment], 2017, no. 4 (742), pp. 59–65.
12. Khmyl A. A., Lanin V. L., Shpilevskiy E. M. Kompozitsionnye pokrytiya na osnove provodyashchikh metallov i uglerodnykh klasterov dlya podvizhnykh elektricheskikh kontaktov [Composite coatings based on conductive metals and carbon clusters for mobile electrical contacts]. *Vestnik Polotskogo gosudarstvennogo universiteta. Seriya S: Fundamentalnye nauki* [Bulletin of Polotsk State University. Series S: Basic Sciences], 2006, no. 4, pp. 74–79.
13. Galperin V. A., Kondratev P. K., Zubov D. N., Kelm E. A., Pavlov A. A. Optimizatsiya protsessa elektrokhimicheskogo osazhdeniya medi v massivy uglerodnykh nanotrubok [Process optimization for electrochemical deposition of copper into arrays of carbon nanotubes]. *Nano- i mikrosistemnaya tekhnika* [Nano and microsystem technology], 2013, no. 4 (153), pp. 15–19.
14. Pek V. V., Litovka Yu. V., Popov D. Yu., Khabotilina A. I. Postroeniye matematicheskoy modeli protsessa nanoseniya nanomodifitsirovannogo galvanicheskogo pokrytiya [Formation of mathematical model of the process of applying nanomodified galvanic coating]. *Vestnik Astrakhanskogo gosudarstvennogo tekhnicheskogo universiteta. Seriya: Upravleniye, vychislitel'naya tekhnika i informatika* [Vestnik of Astrakhan State Technical University. Series: Management, Computer Science and Informatics], 2010, no. 2, pp. 100–104.
15. Shiryayev V. I. *Issledovanie operatsiy i chislennyye metody optimizatsii* [Operations research and numerical optimization methods], Moscow, Lenand Publ., 2017. 224 p.
16. Ritz W. Über eine neue methode zur lösung gewisser variationsprobleme der mathematischen physic. *Journal für die Reine und Angewandte Mathematik*, 1909, vol. 135, pp. 1–61.
17. Golitsyna O. L., Maksimov N. V., Popov I. I. *Bazy dannykh* [Databases]. Moscow, Forum Publ., 2015. 400 p.
18. Zazhigaev L. S., Kishyan A. A., Romanikov Yu. I. *Metody planirovaniya i obrabotki rezultatov fizicheskogo eksperimenta* [Methods of planning and processing the results of a physical experiment]. Moscow, Atomizdat Publ., 2013. 232 p.
19. Tkachev A. G. Uglerodnyy nanomaterial Taunit – struktura, svoystva, proizvodstvo i primeniye [Carbon nanomaterial Taunit – structure, properties, production and application]. *Perspektivnye materialy* [Promising materials], 2007, no. 3, pp. 5–9.
20. *GOST 9.302-88 Edinaya sistema zashchity ot korrozii i stareniya. Pokrytiya metallicheskie i nemetallicheskie neorganicheskije. Metody kontrolya* [Unified system of corrosion and ageing protection. Metal and non-metal inorganic coatings. Control methods]. Moscow, IPK Izdatelstvo standartov, 1990. 40 p.
21. Kiefer J. Sequential minimax search for a maximum. *Proceedings of the American Mathematical Society*, 1953, vol. 4 (3), pp. 502–506.
22. Powell M. J. D. An efficient method for finding the minimum of a function of several variables without calculating derivatives. *Computer Journal*, 1964, vol. 7 (2), pp. 155–162.
23. Antonov A. S. *Parallelnoye programmirovaniye s ispolzovaniem tekhnologii OpenMP: uchebnoye posobie* [Parallel Programming Using OpenMP Technology: Tutorial]. Moscow, Lomonosov Moscow State University, 2009. 77 p.

ИНФОРМАЦИОННАЯ БЕЗОПАСНОСТЬ И ЗАЩИТА ИНФОРМАЦИИ

DOI 10.21672/2074-1707.2020.49.4.144-155
УДК 004.855.2, 004.056

ОБЗОР ИНСТРУМЕНТОВ МАШИННОГО ОБУЧЕНИЯ И ИХ ПРИМЕНЕНИЯ В ОБЛАСТИ КИБЕРБЕЗОПАСНОСТИ

Статья поступила в редакцию 23.01.2020, в окончательном варианте – 26.02.2020.

Власенко Александра Владимировна, Кубанский государственный технологический университет, 350072, Российская Федерация, г. Краснодар, ул. Московская, 2, кандидат технических наук, заведующая кафедрой компьютерных технологий и информационной безопасности Института компьютерных систем и информационной безопасности, e-mail: Vlasenko@kubstu.ru

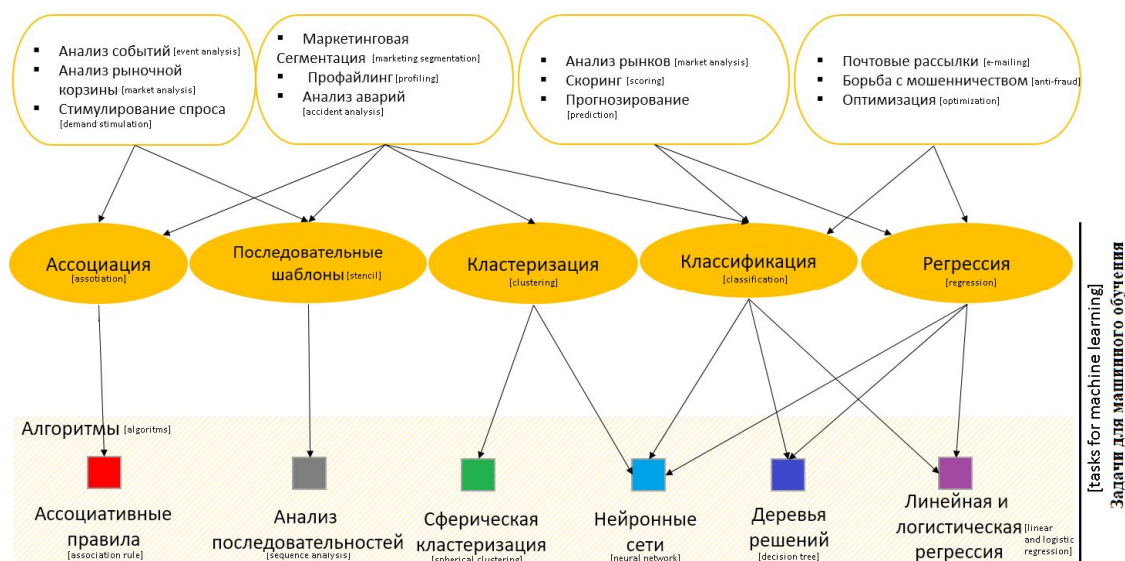
Дзьобан Павел Игоревич, Кубанский государственный технологический университет, 350072, Российская Федерация, г. Краснодар, ул. Московская, 2, кандидат технических наук, доцент кафедры компьютерных технологий и информационной безопасности Института компьютерных систем и информационной безопасности, e-mail: antiemoboy@mail.ru

Жук Роман Владимирович, Филиал «Макрорегион Юг» ООО ИК «СИБИНТЕК», 352800, Российская Федерация, г. Туапсе, ул. Карла Маркса, 36, главный специалист, e-mail: goonerkrd@gmail.com

Машинное обучение заслужено привлекает интерес специалистов в области кибербезопасности. Благодаря тому, что аппаратные и вычислительные мощности становятся все более доступными, методы машинного обучения могут использоваться для анализа и классификации природы возникновения аномалий, вредоносных активностей из агрегированных метаданных. Методы машинного обучения подразделяются на контролируемое (классификация, регрессия) и неконтролируемое обучение (кластеризация, сокращение количества измерений объектов). Оба эти подхода могут быть применены в области кибербезопасности для анализа вредоносных активностей в режиме реального времени, что устраняет недостатки традиционных методов обнаружения таких активностей. В данной статье для анализа активности хостов предлагается использовать данные с применением технологии экспорта потоков NetFlow. Также будут рассмотрены принципы обнаружения аномалий в сетевом трафике с применением различных инструментов машинного обучения (экстремальное машинное обучение, случайный лес, повышение градиента, логистическая регрессия), приведены примеры и успешные практики реализации методов обнаружения аномалий в сети.

Ключевые слова: кибербезопасность, атаки, сетевые аномалии, риски, мониторинг, сеть, машинное обучение, градиент, ботнет, агрегация, кластеризация, классификация, регрессия

Графическая аннотация (Graphical annotation)



**ANALYTICAL REVIEW OF MACHINE LEARNING TOOLS
AND THEIR APPLICATIONS IN THE FIELD OF CYBER SECURITY**

The article was received by the editorial board on 23.01.2020, in the final version – 26.02.2020.

Vlasenko Alexandra V., Kuban State Technological University, 2 Moskovskaya St., Krasnodar, 350072, Russian Federation,

Cand. Sci. (Engineering), Head of the Department of Computer Technologies and Information Security of the Institute of computer systems and information security, e-mail: Vlasenko@kubstu.ru

Dzoban Pavel I., Kuban State Technological University, 2 Moskovskaya St., Krasnodar, 350072, Russian Federation,

Cand. Sci. (Engineering), Associate Professor of the Department of Computer Technologies and Information Security of the Institute of Computer Systems and Information Security, e-mail: antimoboy@mail.ru

Zhuk Roman V., Branch «Macroregion South» Ltd Co IC «SIBINTEK», 36 Karl Marks St., Tuapse, 352800, Russian Federation,
chief specialist, e-mail: goonerkrd@gmail.com

Machine training deservedly attracts the interest of specialists in the field of cybersecurity. With the increasing availability of hardware and computing power, machine learning methods can be used to analyze and classify the nature of anomalies and malicious activities from aggregated metadata. Machine learning methods are divided into controlled (classification, regression) and uncontrolled learning (clustering, reducing the number of measurements of objects). Both of these approaches can be applied in the area of cybersecurity to analyse malicious activities in real time, thus eliminating the shortcomings of traditional methods of detecting such activities. This article proposes to use data using NetFlow flow export technology to analyze host activity, and also discusses the principles of detecting anomalies in network traffic using various machine learning tools (extreme machine learning, random forest, gradient increase, logistic regression), and provides examples and good practices of implementing anomaly detection methods in the network.

Key words: cybersecurity, attacks, network anomalies, risks, monitoring, network, machine learning, gradient, botnet, aggregation, clustering, classification, regression

Введение. В современном обществе Российской Федерации отрасль «информационная безопасность» (ИБ) постоянно развивается. Мотиваций и причин для развития этой отрасли достаточно много, но все они между собой тесно взаимосвязаны: начиная от постоянного роста компетенций злоумышленников и заканчивая требованиями правового поля, вынужденного «успевать» за тенденциями и современными реалиями. В данной статье рассматривается один из самых актуальных блоков отрасли – кибербезопасность. Сложные и все новые виды атак уже сегодня являются нормой; они становятся все более частыми и широко распространенными. Эта постоянная эволюция также требует инноваций в области кибербезопасности, занимающей особое место в отрасли.

Существуют решения и их комбинации (гибридизация), которые широко используются флагманами отрасли в области цифровизации и искусственного интеллекта. Системы обнаружения (далее – IDS) и предотвращения сетевых вторжений (далее – IPS) отслеживают вредоносные активности в сети и/или нарушения политики ИБ «владельца риска» на различных уровнях – пользователь, сегмент сети, организация, провайдер и т.д.

IDS и IPS на основе сигнатур опираются на известные сигнатуры и эффективно обнаруживают вредоносные активности, которые соответствуют этим сигнатурам. IDS и IPS на основе «поведенческого анализа», с другой стороны, аккумулируют данные, что является нормальным состоянием для системы, и сообщают о любом триггере, который отклоняется от принятого значения с выходом за границы заданного интервала. Оба типа систем довольно успешно применяются в лучших мировых практиках, хотя и не лишены недостатков.

Системы на основе сигнатур полагаются на сигнатуры известных угроз и поэтому неэффективны для атак «нулевого дня» или использования злоумышленниками новых образцов вредоносного программного обеспечения (ПО). Традиционные системы, использующие «анализ поведения», основаны на стандартном статичном шаблоне, который трудно динамически изменять в связи с растущей сложностью сетей и приложений. Следовательно, этот подход может быть неэффективен для обнаружения аномалий. Полный анализ пакетов данных – это еще один вариант, однако он требует значительных затрат вычислительных ресурсов и сопряжен с риском раскрытия конфиденциальной информации пользователей.

Машинное обучение заслужено привлекает интерес специалистов по теории и практике обеспечения ИБ, и в частности в области кибербезопасности. Благодаря тому, что аппаратные и вычислительные мощности становятся все более доступными, методы машинного обучения могут исполь-

зоваться для анализа и классификации природы возникновения аномалий, вредоносных активностей из агрегированных метаданных. Существуют сотни алгоритмов и подходов к машинному обучению, которые в целом подразделяются на контролируемое и неконтролируемое обучение. Подходы к обучению под наблюдением (контролируемые) выполняются в контексте классификации, (где ввод соответствует выводу) или регрессии (когда ввод отображается в непрерывный вывод). Самостоятельное обучение в основном достигается с помощью кластеризации и применяется для исследовательского анализа, сокращения количества измерений объектов. Оба эти подхода могут быть применены в области кибербезопасности для анализа вредоносных активностей в режиме реального времени, что устраняет недостатки традиционных методов обнаружения таких активностей.

Для защиты сети от угроз специалистам требуются интеллектуальные решения, которые носят всеобъемлющий характер, основаны на анализе поведения и дополняют существующие зонные средства обеспечения безопасности. Одно из таких решений заключается в использовании сетевой инфраструктуры в качестве датчика. Для анализа активности хостов можно использовать данные с применением технологии экспорта потоков sFlow и NetFlow. Изначально они разрабатывались для мониторинга и устранения неисправностей внутри сети [16].

Записи NetFlow предоставляют достаточно информации, чтобы однозначно идентифицировать трафик. NetFlow создавался как технология пакетной коммутации для маршрутизаторов Cisco. Положенная в основу NetFlow идея заключается в том, что первый пакет потока генерирует на коммутаторе или маршрутизаторе запись коммутации NetFlow. В дальнейшем эта запись используется при обработке пакетов из того же потока вплоть до его окончания. Поиск конкретного маршрута в таблице маршрутизации требуется только для первого пакета потока. NetFlow вместе с открытой версией стандарта IPFIX уже широко используется для мониторинга и управления сетью. Доступность данных NetFlow вместе с функциями конфиденциальности делает его эффективным средством вычисления сетевых аномалий.

Пример работы технологии NetFlow представлен на рисунке 1.

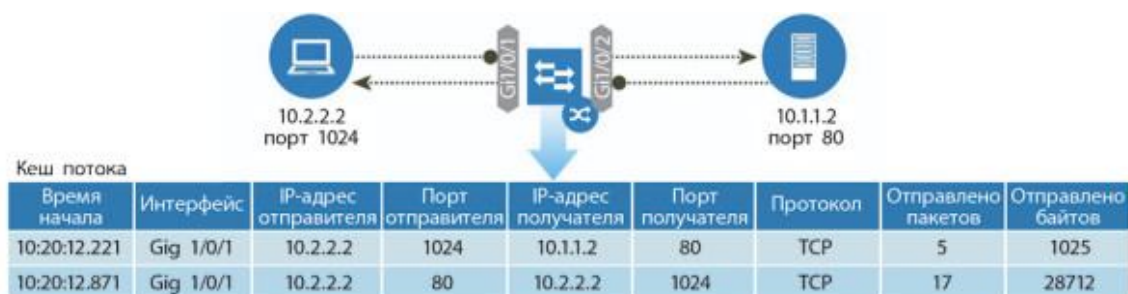


Рисунок 1 – Пример работы NetFlow

Стандарт sFlow, называемый также выборкой потоков, поддерживается на разных платформах [15]. Он представляет собой средство для экспорта усеченных пакетов, а также счетчиков интерфейса. Последовательное развитие sFlow сделало возможным получение им статуса отраслевого стандарта для экспорта пакетов на уровне 2 [16].

sFlow – это технология выборочного захвата пакетов. Из всего трафика через интерфейс выбирается «1 из n» пакетов (где n – частота выборки). Первые x байт (128 байт для sFlow версии 5) захваченного пакета копируются и экспортируются в пакеты UDP, называемые дейтаграммами sFlow. Первые x байт – это данные заголовка, необходимые для восстановления информации о трафике. Однако поскольку sFlow ориентирован на пакеты, то некоторые IP-потоки могут пропускаться. Когда пакеты захватываются для анализа, они не всегда представляют все IP-потоки (диалоги), проходящие через интерфейс. В результате IP-потоки, пакеты которых не собраны, не учитываются, что создает пробелы в анализе картины диалогов в сети.

Так как для агрегации метаданных, дальнейшего их парсинга, построения таблиц корреляции, вывода и подачи оператору отфильтрованной информации необходимы все пакеты, то нами в данной статье будет отдано предпочтение технологии NetFlow.

Детальная характеристика технологии NetFlow. С тех пор как NetFlow стал отраслевым стандартом для сбора данных сеансов, данные NetFlow предоставляют информацию, которая может быть использована для определения использования трафика в сети и состояния ресурсов. Таким образом, технология NetFlow позволяет обнаруживать сетевые аномалии и потенциальные кибератаки. Данный инструмент полезен для идентификации узловых устройств, которым требуется увеличить или уменьшить полосу пропускания, что скажется на повышении эффективности работы сети.

Такие инструменты, как NfSen/NfDump1, могут анализировать данные с NetFlow и отслеживать аномалии в сетевом трафике. Такой инструментарий успешно применяется для обеспечения мониторинга и управления сетью. Сегодня обеспечен широкий выбор инструментов анализа угроз и обнаружения аномалий, использующих трафик NetFlow [17].

NetFlow v5, NetFlow v9 и открытый стандарт IPFIX широко используются для решения различных задач построения защищенной ИТ-инфраструктуры. Записи NetFlow v5 включают данные, документирующие IP-адреса источника и назначения, порты источника и назначения и транспортный протокол. Процесс формирования отчетов о данных NetFlow включает захват потоков IP, агрегирование потоков на коммутаторе или маршрутизаторе и экспорт их в коллектор NetFlow. Процесс состоит из следующих этапов:

- а) конфигурация NetFlow настраивается для захвата потоков и помещения их в кеш NetFlow;
- б) экспорт NetFlow настраивается для отправки потоков коллектору;
- в) в кеше NetFlow ведется поиск устаревших потоков с помощью активных таймеров, неактивных таймеров и ограничений на кеш;
- г) одновременно в кеше NetFlow осуществляется поиск завершенных потоков – путем анализа флагов сброса TCP [TCP RST] и завершения [FIN];
- д) найденные в пунктах «в» и «г» потоки экспортируются на сервер коллектора NetFlow;
- е) 20–25 потоков объединяются в пакет и транспортируются на сервер коллектора NetFlow в формате User Datagram Protocol (UDP);
- ж) программное обеспечение коллектора NetFlow создает из полученных данных отчеты в реальном времени или исторические отчеты.

Инструменты NetFlow v9 и IPFIX обладают возможностью расширенной настройки. Это позволяет использовать дополнительные информативные поля, такие как имена пользователей, MAC-адреса и URL-адреса [16, 17].

Данные сетевого трафика. Сетевой трафик определяется как однонаправленная последовательность совокупности передаваемых пакетов с некоторыми общими свойствами, которые проходят через сетевое устройство. Записи трафика включают в себя различную информацию: IP-адреса, количество пакетов и байтов, временную метку, тип обслуживания, порты приложения, интерфейсы ввода и вывода и другую информацию [3,5]. Данные сеанса, содержащие IP-адреса клиента, номера порта клиента, IP-адреса сервера, номера порта сервера и протокола, включенного в данные потока, важны для идентификации соединения. Исследуя корреляцию запросов и ответов клиент-серверных диалогов, анализируя трафик, можно найти значимые траектории и связь сетевой аномалии с тем или иным событием/инцидентом ИБ [8, 11].

Выходные данные NetFlow включают, но не ограничиваются (в зависимости от версии инструмента), следующие атрибуты [16, 17]:

- время начала записанного потока;
- продолжительность потока;
- используемый протокол (TCP, UDP и т.д.);
- IP-адрес источника;
- исходный порт;
- направление общения;
- адрес назначения;
- порт назначения;
- протокол состояния;
- тип используемого сервиса;
- тип услуги назначения;
- общее количество обмененных пакетов;
- общее количество обмененных байтов;
- количество байтов, отправленных источником;
- метка, назначенная этому сетевому потоку.

Обнаружение аномалий в сетевом трафике. Традиционные методы обнаружения аномалий в сетевом трафике, такие как обнаружение вторжений и глубокая проверка пакетов, обычно требуют использования необработанных данных или подписей, опубликованных производителями [2].

Глубокая проверка пакетов предоставляет более точные данные, но требует проведения значительных объемов вычислений [4]. Данный инструмент «не работает» с зашифрованными данными. В современных реалиях это является негативным признаком, противоречит политике ИБ и стандартам охраны сведений конфиденциального характера, содержащих информацию о пользователях.

Данные Netflow, в сравнении с инструментами глубокой проверки пакетов, не содержат такой конфиденциальной информации и широко используются администраторами сетей. При использовании правильных методов анализа данные с NetFlow могут стать информативным источником для обнаружения аномалий. Один из основных недостатков NetFlow связан с объемом генерируемых данных. Это прямо влияет на точность результатов, количество информации и соответственно на вычислительные ресурсы, требуемые для обработки данных (хранения, накопления, изменения, передачи и пр.) [15].

Методы машинного обучения. Машинное обучение – это инструмент анализа данных, который используется для эффективного выполнения конкретных задач без применения явных инструкций, а опирается на шаблоны и умозаключения [2]. Возможности машинного обучения используются для решения различных проблем, в том числе и в области кибербезопасности. Постоянный анализ и обучение позволит генерировать различные прогнозы развития кибератак, моделировать различные ситуации и обнаруживать аномальные активности в сети.

Крупные отечественные провайдеры используют машинное обучение, интегрированное в облачные интеллектуальные системы, для выявления вредоносного и нелегитимного контента, изоляции зараженных хостов и вывода информативных графиков и дашбордов об общем состоянии того или иного сегмента [7]. Одной из основных трудностей в машинном обучении является создание интеллектуальных систем, способных изучать последовательные задачи, а затем передавать знания из ранее изученного «фундамента» для решения новых задач. Такая возможность называется непрерывным машинным обучением или интеллектуальными системами непрерывного обучения [9]. Применение машинного обучения для обнаружения ботнетов было широко исследовано, в том числе с использованием контролируемого машинного обучения [3, 7, 8, 11, 15, 16, 17]. Машинное обучение рассматривается как решение для аналитики метаданных, полученных с NetFlow. При этом основной проблемой будет выбор параметров для парсинга и корреляции, для обеспечения наилучших результатов при решении поставленных задач. Некоторые из распространенных методов машинного обучения подвергаются анализу в данной статье [10].

Экстремальное машинное обучение (Extreme Learning Machine, ELM) – это алгоритм обучения, который использует нейронные сети с прямой связью с одним или несколькими слоями скрытых узлов. Эти скрытые узлы настраиваются случайным образом, и соответствующие им выходные веса аналитически определяются алгоритмом. По словам создателей [18], этот алгоритм обучения может дать хорошую производительность обобщения и может учиться в тысячу раз быстрее, чем обычные алгоритмы обучения для нейронных сетей с прямой связью. Также стоит отметить, что аномалия характеризуется, как правило, не только экстремальными значениями отдельных признаков. Иллюстрация алгоритма экстремального машинного обучения (ELM) представлена на рисунке 2.

Extreme Learning Machine (ELM)

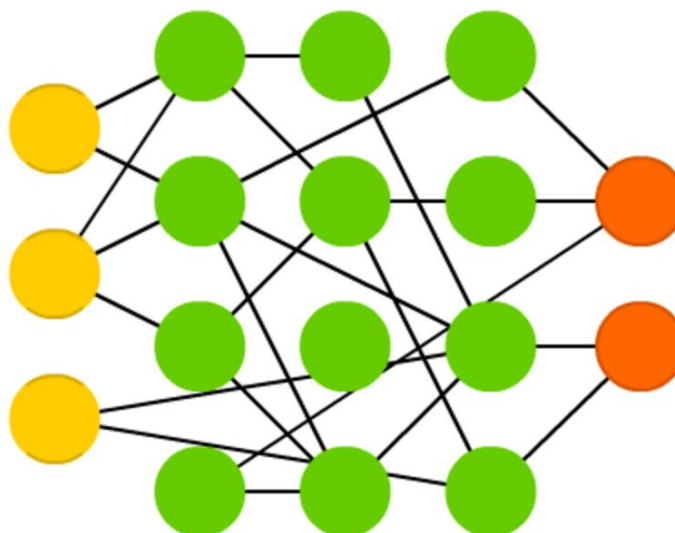


Рисунок 2 – Упрощенная иллюстрация алгоритма экстремального машинного обучения (ELM)

Случайный лес (Random forest, RF) – это контролируемый алгоритм машинного обучения, который предусматривает использование нескольких деревьев решений для выполнения задач классификации и регрессии. Алгоритм случайного леса считается ансамблевым алгоритмом машинного обучения, так как он включает в себя концепцию большинства рёбер нескольких деревьев. Выходные данные алгоритма, представленные как прогнозирование классов, определяются из совокупного результата всех классов, прогнозируемых отдельными деревьями. В последних исследованиях [7, 16, 18] изучались возможности использования алгоритма случайного леса в анализе кибератак, в частности, инъекционных атак; для фильтрации спама, для обнаружения вредоносных программ и многое другое.

Все деревья строятся независимо по следующему алгоритму:

1. Выбирается подвыборка обучающей выборки произвольного размера, по ней строится дерево (для каждого дерева – своя подвыборка).
2. Для построения каждого расщепления в дереве просматриваются максимальные значения случайных признаков (для каждого нового расщепления – свои случайные признаки).
3. Далее определяется наилучший признак и расщепление по нему (по заранее заданному критерию).

Дерево строится, как правило, до исчерпания выборки (пока в листьях не останутся представители только одного класса). Однако в современных реализациях алгоритма есть параметры, которые ограничивают высоту дерева, число объектов в листьях и число объектов в подвыборке, при котором проводится расщепление. Схематично, алгоритм случайного леса представлен на рисунке 3.

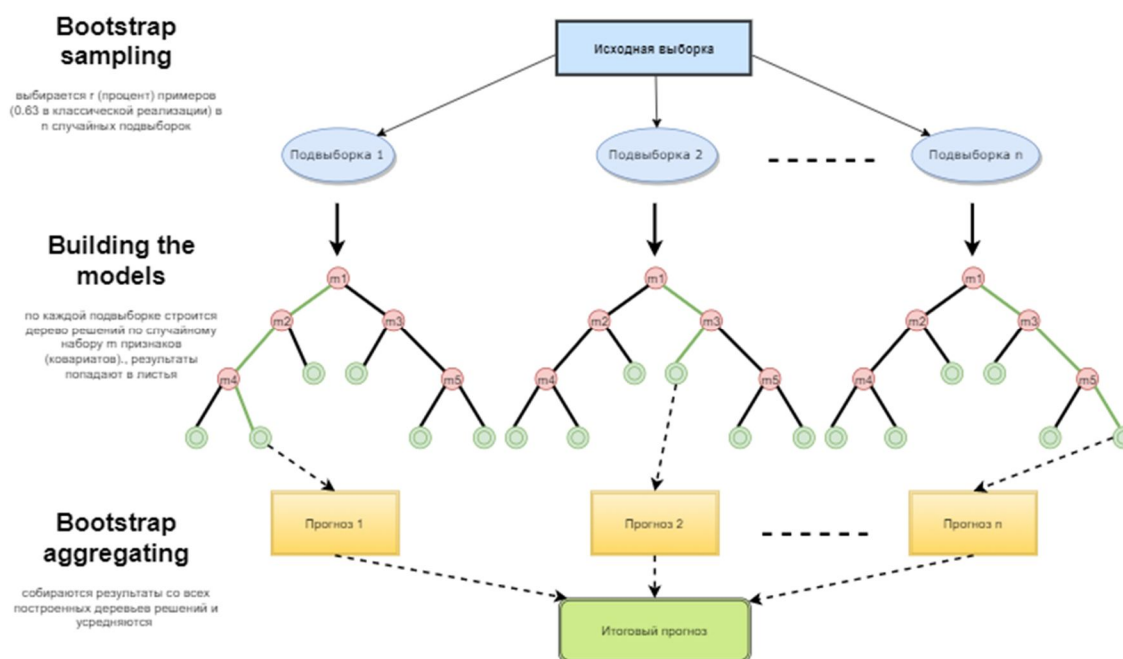


Рисунок 3 – Схема работы алгоритма случайного леса

Повышение градиента — это метод машинного обучения для задач регрессии и классификации, который создает модель прогнозирования в форме множества моделей слабого прогнозирования [15, 17], обычно деревьев решений. Этот метод использует идею о том, что следующая модель будет учиться на ошибках предыдущей. Модели имеют неравную вероятность появления в последующих моделях, и чаще появятся те, которые дают наибольшую ошибку. Предсказатели могут быть выбраны из широкого ассортимента моделей, например, деревья решений, регрессия, классификаторы и т.д. Из-за того, что предсказатели обучаются на ошибках, совершенных предыдущими, требуется меньше времени для того, чтобы добраться до реального ответа. Однако выбирать критерий остановки алгоритма следует с осторожностью, иначе это может привести к переобучению модели и искажению результатов. Цель любого алгоритма обучения – определить функцию потерь и минимизировать её. Когда алгоритм достигает стадии, на которой остатки не имеют какого-либо шаблона, который можно было бы смоделировать, то моделирование остатков будет остановлено – в противном случае это может привести к переобучению. Иными словами,

предсказания обновляются таким образом, чтобы сумма отклонений стремилась к нулю и предсказанные значения были близки к реальным. Математически это означает минимизацию функции потерь так, чтобы тестовые потери достигли своего минимума.

Исходя из вышеизложенного, за алгоритмом повышения градиента необходимо итеративно применять паттерны отклонений и улучшать предсказания, как показано на рисунке 4. Как только будет достигнут момент, когда отклонения не имеют никакого паттерна, стоит прекратить модернизацию модели (иначе это может привести к переобучению) [9]. Таким образом, для применения метода повышения градиента необходимо соблюдать следующую последовательность действий:

- построение простой модели;
- анализ ошибок;
- идентификация точек, которые не вписываются в простую модель;
- модификация моделями, обрабатывающими сложные случаи, которые были выявлены на начальной модели;
- агрегация и наложение построенных моделей, с определением весов каждого предсказателя [4, 17].

Пример работы алгоритма повышения градиента представлен на рисунке 4.

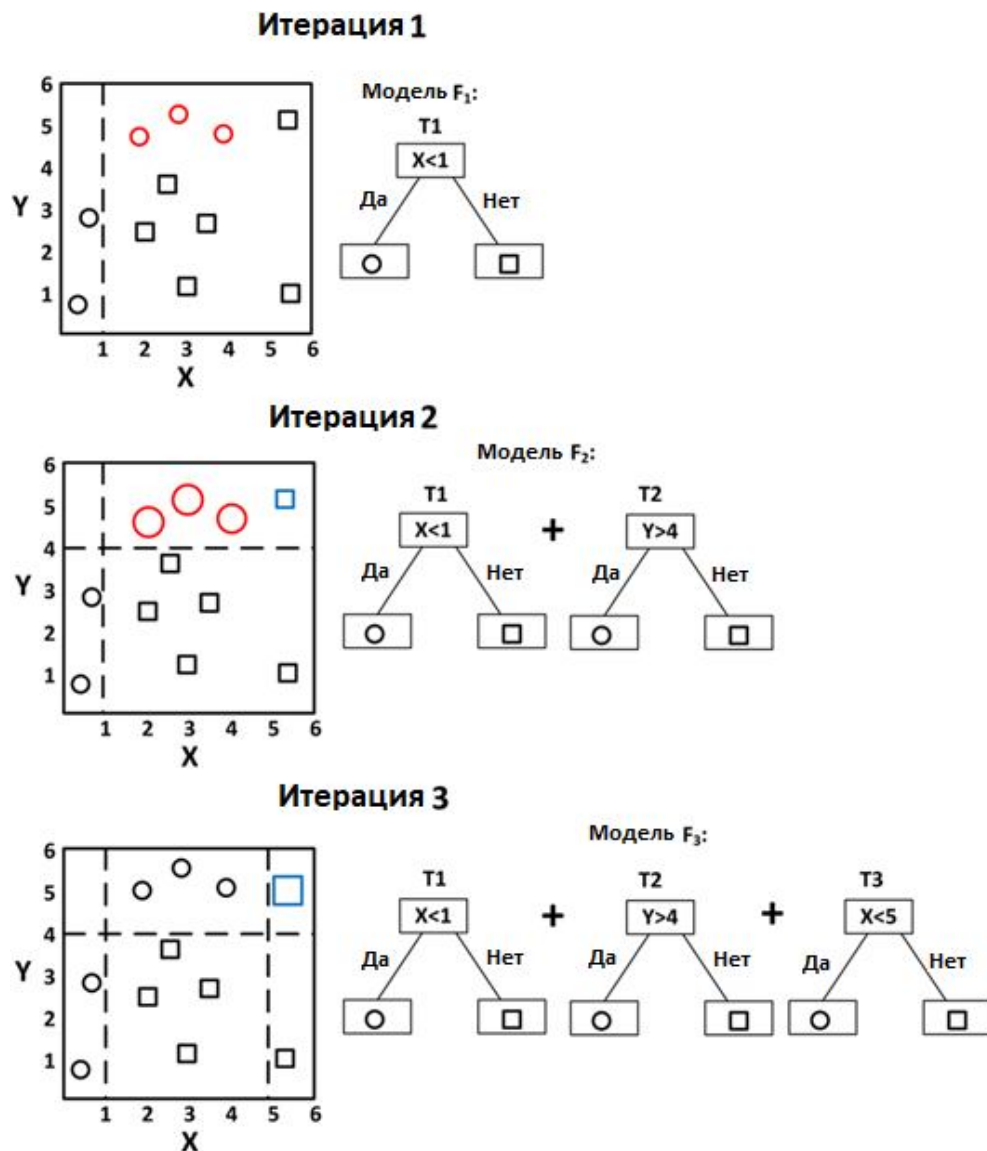


Рисунок 4 –Пример работы алгоритма повышения градиента

Логистическая регрессия является контролируемой моделью обучения, которая используется в качестве метода для двоичной классификации. Сам термин заимствован из статистики. В основе метода лежат логистические функции – сигмовидная кривая, которая полезна для ряда областей, включая нейронные сети. Логистическая регрессия моделирует вероятность проблем классификации с двумя возможными результатами и может использоваться для идентификации сетевого трафика как вредоносного (true) или нет (false).

Основная идея логистической регрессии заключается в том, что пространство исходных значений может быть разделено линейной границей (т.е. прямой) на две области, соответствующие классам. Под линейной границей подразумевается прямая линия – в случае двух измерений, в случае трех измерений – плоскость, и т.д. Эта граница задается в зависимости от имеющихся исходных данных и обучающего алгоритма. Для того чтобы алгоритм обрабатывал корректно, точки исходных данных должны разделяться линейной границей на две вышеупомянутых области. Если точки исходных данных удовлетворяют этому требованию, то их можно назвать линейно разделяемыми.

Пример работы метода логистической регрессии представлен на рисунке 5.

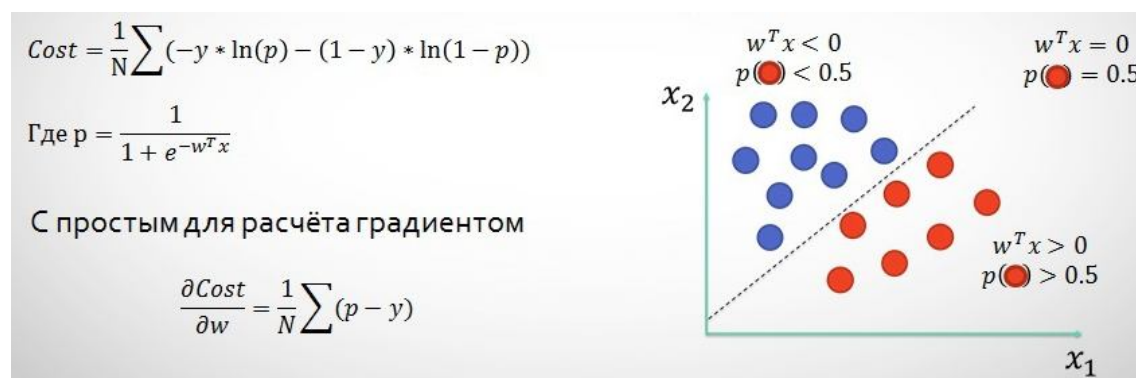


Рисунок 5 – Пример работы метода логистической регрессии

Методы обнаружения аномалий в сети. Аномалии — это объекты или инциденты, которые отличаются на более чем максимально допустимое отклонение (погрешность) от принятого значения. Таким образом, обнаружение сетевых аномалий сводится к их выявлению. Выявлять аномалии предлагается с использованием построенных триггеров на основе связи с редкими событиями при наблюдении за трафиком. Такие связи вызывают подозрение аналитиков, так как они существенно отличаются от большинства идентичных состояний легитимного сетевого трафика.

В машинном обучении обнаружение аномалий применяется в различных областях, включая обнаружение вторжений, обнаружение мошенничества и обнаружение нарушений в экосистеме.

Существует три широких категории обнаружения аномалий: неконтролируемые, контролируемые и полуправляемые [2, 13]. Некоторые из популярных методов обнаружения включают, среди прочего, основанный на плотности k-ближайший сосед, одноклассовый Support Vector Machine (SVM, набор схожих алгоритмов обучения с учителем, использующихся для задач классификации и регрессионного анализа), Байесовские сети, обнаружение выбросов на основе кластерного анализа [8,14]. Ряд систем анализа используют вышеуказанные методы обнаружения.

Кооперативный адаптивный механизм защиты сети (CAMNER) – это система обнаружения вторжений в сеть. Система CAMNER использует набор моделей обнаружения аномалий, которые поддерживают модель ожидаемого трафика в сети и сравнивают ее с реальным трафиком, чтобы выявить расхождения, которые определены как возможные атаки [1, 9]. Механизм имеет три основных уровня, которые оценивают трафик: детекторы аномалий, модели доверия и агрегаторы аномалий [3, 7, 16].

Устройства IDS/IPS применяются для оперативного реагирования на кибератаки. С учетом того факта, что легитимный трафик значительно превосходит по объему вредоносный, низкая частота ложных срабатываний делает систему непригодной для использования, так как нарушает отказоустойчивость и непрерывность ИТ-сервисов.

Корректность обрабатывания доверительной модели CAMNER заключается в перекрестной проверке аномалий в моделях триггеров. Каждая из этих моделей основана на различных характеристиках трафика. Для классификации кибератаки потоки из набора должны находиться вблизи центроидов [19]. На практике большинство атакующих потоков приходится на окрестности одного центроида, как показано на рисунке 6.

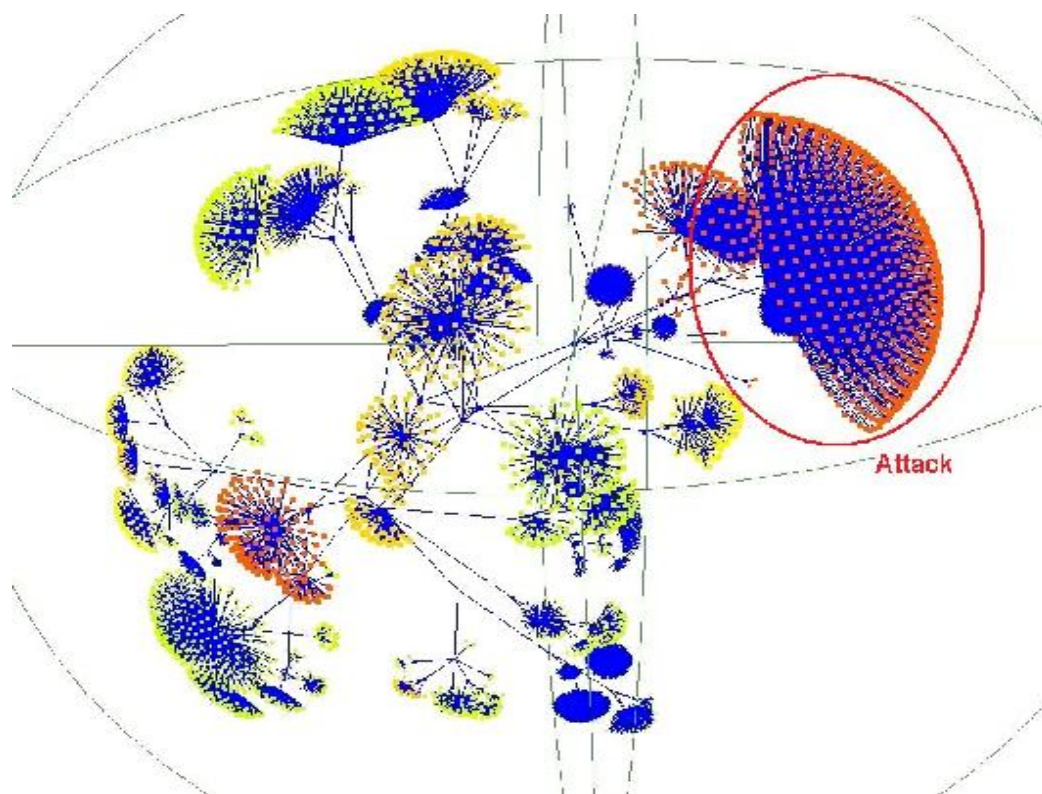


Рисунок 6 – Пример визуализации системы обнаружения вторжений CAMNEP

Уровень обнаружения аномалий анализирует NetFlow с использованием различных алгоритмов обнаружения аномалий, каждый из которых использует свой набор функций. Выходные данные агрегируются в события и отправляются в модели доверия. Модель доверия отображает NetFlow в транспортные кластеры. NetFlow с похожими поведенческими паттернами сгруппированы вместе. Уровень агрегатора создает композитный выход, который объединяет индивидуальное мнение нескольких детекторов аномалий [17].

Миннесотская система обнаружения вторжений (MINDS) использует набор методов интеллектуального анализа данных для автоматического обнаружения атак. Система MINDS создает контекстную информацию для каждого оцениваемого NetFlow с использованием следующих функций: количество сеансов NetFlow с исходного IP-адреса относительно других активных хостов из данного сегмента сети, то же по отношению к хосту назначения, то же по отношению к порту источника, активность NetFlow от одного хоста источника к тому же порту назначения. Значение аномалии для NetFlow основано на корреляции со значениями эталонной выборки [9, 16].

Методы кластеризации на основе поведения ботнета. Современный ботнет является флагманом среди наиболее эффективных методов атаки, доступных для современного киберпреступника. Ботнет – это совокупность зараженных рабочих станций (APM и прочих хостов) с сетевыми интерфейсами, вычислительные ресурсы которых задействованы в теневом режиме для выполнения вредоносного кода [5, 10]. Каждая из этих конечных точек или «ботов» регулярно взаимодействует с сервером контроля и управления (C & C), и весь ботнет может быть использован для управления гигантскими атаками DDoS (распределенного отказа в обслуживании), а также для нарушения конфиденциальности и целостности информации или распространения спама в массовом масштабе. Из-за размеров и сложностей их распознавания, ботнеты могут функционировать в дрейф-режиме (не проявляя активности) на протяжении долгого времени, аккумулируя нужную информацию. Ботнеты используются злоумышленниками для воровства финансовой информации и переводов денег в онлайн-банкинге, в расчетных палатах и системах начисления заработной платы пользователей сети [5, 11]. Например, ботнет «Zeus» проработал в дрейф-режиме более трех лет. Затем он был удаленно активирован злоумышленниками, когда было накоплено достаточное количество сведений, что принесло преступникам 70 миллионов долларов украденных средств [7, 15].

Метод BClus – это подход, использующий поведенческое обнаружение ботнета. BClus создает модель поведения с множеством триггеров известных ботнетов и использует их для обнаружения

аналогичного трафика в сети. Целью метода является кластеризация трафика, отправляемого каждым IP-адресом, и распознавание того, какие кластеры отличаются аномальной активностью, схожей с ботнетом [5, 6].

Обнаружение на стадии вредоносного заражения. Метод BotHunter полезен для обнаружения заражений и для координации передачи данных от ботнета к злоумышленникам. Он состоит из механизма корреляции, целью которого является выявление определенных этапов процесса заражения вредоносным программным обеспечением. Используется адаптивная версия Snort IDS с двумя запатентованными плагинами, которые обладают механизмом обнаружения аномалий статистического сканирования (SCADE) и механизмом обнаружения аномалий статистической нагрузки (SLADE) [15, 17, 18].

Система сбалансированного эффективного обучения (B-ELLA) является новым подходом к обнаружению кибератак на основе практического применения эффективной системы кибербезопасности с непрерывным обучением. Такое решение позволит оперативно решить текущие проблемы в области кибербезопасности, где каждая новая кибератака может рассматриваться как новый объект для анализа и изменения/формирования эталонного значения. Данный подход является расширением платформы ELLA. Он справляется с проблемой корреляции данных на основе уже сформированного набора данных по вредоносным активностям [14, 17].

Заключение. Итак, в связи с большим количеством метаданных, подлежащих анализу, применяемые сегодня методы обнаружения и предотвращения вторжений требуют модификации с использованием различных подходов машинного обучения для обнаружения вредоносных активностей в сетевом трафике или следов ботнетов из набора данных NetFlow в области кибербезопасности. Эталонное решение, предназначенное для обнаружения и предотвращения кибератак, должно анализировать поступающий набор данных любого объема, качества и глубины, и применяя методы искусственного интеллекта классифицировать трафик как легитимный или как вредоносный. В последнем случае ПО должно предоставить перечень угроз, с их классификацией, вероятностями и траекториями атак.

Обработка терабайтов метаданных для обеспечения кибербезопасности в вычислительном отношении является весьма трудоемкой. Во-первых, поступающие на анализ данные очень несбалансированные, то есть большая часть сетевого трафика безвредна, и только его меньшая часть является вредоносной. Это приводит к тому, что сложно сформировать корректную обучающую выборку на данный отрезок времени, с учетом того факта, что вектора кибератак все динамичны и постоянно изменяются. Более того, риск переобучения в процессе обучения высок, поскольку структура сети влияет на способ обучения модели, в то время как требуется независимый от сети алгоритм. Любая сетевая инфраструктура является динамической, постоянно поддается различным изменениям: связь узлов зависит от системного времени, и связи между серверами могут появляться и исчезать вместе с новыми запросами и новыми пользователями в сетевом сегменте.

Библиографический список

1. Ажмухамедов И. М. Методика формирования обучающего множества при использовании статических антивирусных методов эвристического анализа / И. М. Ажмухамедов, Р. Ю. Демина // Инженерный вестник Дона. – 2015. – № 37 (3). – С. 74.
2. Брумштейн Ю. М. Математические модели и методы решения задач информационного обеспечения, управления и оценки качества работы операторов в сложных человеко-машинных системах / Ю. М. Брумштейн, Д. А. Молимонов // Вестник Астраханского государственного технического университета. Серия: Управление, вычислительная техника и информатика. – 2019. – № 3. – С. 73–89.
3. Власенко А. В. Анализ уязвимостей и моделирование атак на данные трафика “https” / А. В. Власенко, П. И. Дзьобан // Вестник Адыгейского государственного университета. Серия 4: Естественно-математические и технические науки. – 2017. – № 2 (201). – С. 109–115.
4. Глухарев М. Л. Программа для автоматизированной верификации ограничений целостности баз данных / М. Л. Глухарев, А. Д. Хомоненко, А. П. Косаренко // Программные продукты и системы. – 2011. – № 1. – С. 91–95.
5. Дзьобан П. И. Свидетельство о государственной регистрации программы для ЭВМ № 2018618364. Программная среда криптографических преобразований / П. И. Дзьобан, А. В. Власенко, Б. В. Леваньков. – Зарегистрировано 11.06.2018.
6. Дзьобан П. И. Свидетельство о государственной регистрации изобретения и патент на изобретение № RU 0002699259 С1. Генератор псевдослучайных последовательностей / П. И. Дзьобан, А. В. Власенко, Б. В. Леваньков. – Зарегистрировано 04.09.2019.
7. Дзьобан П. И. Идентификация DDOS-атак на web-серверы / П. И. Дзьобан, А. В. Власенко // Прикаспийский журнал: управление и высокие технологии. – 2019. – № 1 (45). – С. 181–187.
8. Жук Р. В. Модель нарушителя комплексной системы обеспечения информационной безопасности объектов защиты / Р. В. Жук, А. С. Чебанов, А. В. Власенко, С. Ю. Сазонов // Известия Юго-Западного

государственного университета. Серия: управление, вычислительная техника, информатика, медицинское приборостроение. – 2013. – № 1. – С. 171–173.

9. Ключко В. И. Архитектуры систем поддержки принятия решений / В. И. Ключко, Е. А. Шумков, Власенко, Р. О. Карнизьян // Научный журнал КубГАУ. – 2013. – № 86. – С. 290–299.

10. Ключко В. И. Теория информации и сигналов : учеб. пособие / В. И. Ключко, А. В. Власенко, Н. В. Кушнир, А. В. Кушнир. – Краснодар : КубГТУ, 2011. – С. 132.

11. Корниенко А. А. Методика обнаружения и разрешения конфликтов программных средств защиты от кибератак на железнодорожном транспорте / А. А. Корниенко, М. А. Поляничко // Интеллектуальные технологии на транспорте. – 2015. – № 1. – С. 18–21.

12. Симанков В. С. Методологические аспекты поддержки принятия решений для организации функционирования интеллектуальной системы ситуационного центра / В. С. Симанков, А. Н. Черкасов // Глобальный научный потенциал. – 2015. – № 12 (45). – С. 114–122.

13. Хомоненко А. Д. Динамические модели отладки программ с вероятностным обнаружением ошибок и распределением Эрланга длительности их исправления / А. Д. Хомоненко, А. И. Данилов, А. А. Данилов // Научно-технический вестник информационных технологий, механики и оптики. – 2016. – № 16 (4). – С. 655–662.

14. Шарай В. А. Мониторинг состояния надежности и безопасности структурно-сложных систем на основе логико-числовых моделей / В. А. Шарай, О. С. Бурангулова, М. В. Андрица // Известия Южного федерального университета. Технические науки. – 2011. – № 125 (12). – С. 35–49.

15. Antoine Delplace Cyber Attack Detection thanks to Machine Learning Algorithms / Antoine Delplace, Sheryl Hermoso, Kristofer Anandita // University of Queensland / COMS7507: Advanced Security. – P. 3–15.

16. Cisco IOS NetFlow Command Reference. – Режим доступа: https://www.cisco.com/c/en/us/td/docs/ios/netflow/command/reference/nf_book/nf_01.html, свободный. – Заглавие с экрана. – Яз. англ. (дата обращения: 04.01.2020).

17. Kozik Rafal. Machine Learning Techniques for Cyber Attacks Detection / Kozik Rafal and Michal Choras // Image Processing and Communications Challenges. – 2014. – Vol. 5, № 233. – P. 391–398.

18. Tang J. Compressed-Domain Ship Detection on Spaceborne Optical Image Using Deep Neural Network and Extreme Learning Machine / J. Tang, C. Deng, G.-B. Huang & B. Zhao // IEEE Transactions on Geoscience and Remote Sensing. – 2015. – № 53 (3). – P. 1174–1185.

19. Reháč Martin. CAMNEP: Agent-Based Network Intrusion Detection System / Reháč Martin, Michal Pechoucek, Pavel Čeleda and Pavel Minárik. – 2008.

References

1. Azhmukhamedov I. M., Demina R. Yu. Metodika formirovaniya obuchayushchego mnozhestva pri ispolzovanii staticheskikh antivirusnykh metodov evristicheskogo analiza [The methodology for the formation of the training set using static antiviral methods of heuristic analysis]. *Inzhenernyy vestnik Dona* [Engineering Bulletin of the Don], 2015, no. 37 (3), p. 74.

2. Brumshteyn Yu. M., Malimono D. A. Matematicheskie modeli i metody resheniya zadach informatsionnogo obespecheniya, upravleniya i otsenki kachestva raboty operatorov v slozhnykh cheloveko-mashinnykh sistemakh [Mathematical models and methods for solving the problems of information support, management and quality assessment of operators in complex human-machine systems]. *Vestnik Astrakhanskogo gosudarstvennogo tekhnicheskogo universiteta. Seriya: Upravlenie, vychislitel'naya tekhnika i informatika* [Bulletin of Astrakhan State Technical University. Series: Management, computer engineering and informatics], 2019, no. 3, pp. 73–89.

3. Vlasenko A. V., Dzoban P. I. Analiz uyazvimostey i modelirovanie atak na dannye trafika “https” [Vulnerability analysis and attack modeling for “https” traffic data]. *Vestnik Adygeyskogo gosudarstvennogo universiteta. Seriya 4: Estestvenno-matematicheskie i tekhnicheskie nauki* [Vestnik of Adygea State University. Series 4: Natural-mathematical and technical sciences], 2017, no. 2 (201), pp. 109–115.

4. Glukharev M. L., Khomonenko A. D., Kosarenko A. P. Programma dlya avtomatizirovannoy verifikatsii ogranicheniy tselostnosti baz dannykh [Program for the automated verification of database integrity constraints]. *Programmnye produkty i sistemy* [Software products and systems], 2011, no. 1, pp. 91–95.

5. Dzoban P. I., Vlasenko A. V., Ivankov B. V. Svidetel'stvo o gosudarstvennoy registratsii programmy dlya EVM. Programmaya sreda kriptograficheskikh preobrazovaniy [The certificate of state registration of computer programs. Software environment for cryptographic transformations], no. 2018618364, registered 11.06.2018.

6. Dzoban P. I., Vlasenko A. V., Ivankov B. V. Svidetel'stvo o gosudarstvennoy registratsii izobreteniya i patent na izobretenie. Generator psevdosluchaynykh posledovatelnostey [The certificate on the state registration of the invention and the patent for the invention. Pseudorandom sequence generator], no. RU0002699259 C1, registered 04.09.2019.

7. Dzoban P. I., Vlasenko A. V. Identifikatsiya DDOS-atak na web-servery [Identification of DDOS attacks on web servers]. *Prikaspiyskiy zhurnal: upravlenie i vysokie tekhnologii* [Caspian Journal: Management and High Technologies], 2019, no. 1 (45), pp. 181–187.

8. Zhuk R. V., Chebanov A. S., Vlasenko A. V., Sazonov S. Yu. Model narushitelya kompleksnoy sistemy obespecheniya informatsionnoy bezopasnosti obektov zashchity [A model of an intruder of an integrated system for ensuring information security of objects of protection]. *Izvestiya Yugo-Zapadnogo gosudarstvennogo universiteta. Seriya: upravlenie, vychislitel'naya tekhnika, informatika, meditsinskoe priborostroenie* [Proceedings of the South-western State University. Series: control, computer engineering, informatics, medical instrument making], 2013, no. 1, pp. 171–173.

9. Klyuchko V. I., Shumkov E. A., Vlasenko A. V., Kernizan R. O. Arkhitektury sistem podderzhki prinyatiya resheniy [Architecture of decision support systems]. *Nauchnyy zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta* [Scientific Journal of the Kuban State Agrarian University], 2013, no. 86, pp. 290–299.
10. Klyuchko V. I., Vlasenko A. V., Kushnir N. V., Kushnir A. V. *Teoriya informatsii i signalov: uchebnoe posobie* [Theory of information and signals: textbook]. Krasnodar, Kuban State Technical University, 2011, p. 132.
11. Kornienko A. A., Polyanchko M. A. Metodika obnaruzheniya i razresheniya konfliktov programmykh sredstv zashhity ot kiberatak na zheleznodorozhnom transporte [A technique for detecting and resolving conflicts of software protection against cyber attacks in railway transport]. *Intellektualnye tekhnologii na transporte* [Intelligent Technologies in Transport], 2015, no. 1, pp. 18–21.
12. Simankov V. S., Cherkasov A. N. Metodologicheskie aspekty podderzhki prinyatiya resheniy dlya organizatsii funktsionirovaniya intellektualnoy sistemy situatsionnogo tsentra [Methodological aspects of decision support for organizing the functioning of the intellectual system of a situational center]. *Globalnyy nauchnyy potentsial* [Global Scientific Potential], 2015, no. 12 (45), pp. 114–122.
13. Khomonenko A. D., Danilov A. I., Danilov A. A. Dinamicheskie modeli otladki programm s veroyatnostnym obnaruzheniem oshibok i raspredeleniem Erlanga dlitelnosti ikh ispravleniya [Dynamic models of program debugging with probabilistic error detection and Erlang distribution of the duration of their correction]. *Nauchno-tekhnicheskiiy vestnik informatsionnykh tekhnologiy, mekhaniki i optiki* [Scientific and Technical Journal of Information Technologies, Mechanics and Optics], 2016, no. 16 (4), pp. 655–662.
14. Shari V. A., Burangulova O. S., Andriutsa M. V. Monitoring sostoyaniya nadezhnosti i bezopasnosti strukturno-slozhnykh sistem na osnove logiko-chislovykh modeley [Monitoring the state of reliability and safety of structurally complex systems based on logical-numerical models]. *Izvestiya Yuzhnogo federalnogo universiteta. Tekhnicheskiiye nauki* [News of Southern Federal University. Engineering], 2011, no. 125 (12), pp. 35–49.
15. Antoine Delplace, Sheryl Hermoso, Kristofer Anandita. Cyber Attack Detection thanks to Machine Learning Algorithms. *University of Queensland / COMS7507: Advanced Security*, pp. 3–15.
16. Cisco IOS NetFlow Command Reference. Available at: https://www.cisco.com/c/en/us/td/docs/ios/netflow/command/reference/nf_book/nf_01.html (accessed 04.01.2020).
17. Rafal Kozik and Michal Choras. Machine Learning Techniques for Cyber Attacks Detection. *Image Processing and Communications Challenges*, 2014, vol. 5, no. 233, pp. 391–398.
18. Tang J., Deng C., Huang G.-B. & Zhao B. Compressed-Domain Ship Detection on Spaceborne Optical Image Using Deep Neural Network and Extreme Learning Machine. *IEEE Transactions on Geoscience and Remote Sensing*, 2015, no. 53 (3), pp. 1174–1185.
19. Reháč, Martin, Michal Pechoucek, Pavel Čeleda and Pavel Minarik. *CAMNEP: Agent-Based Network Intrusion Detection System*, 2008.

DOI 10.21672/2074-1707.2020.49.4.155-161
УДК 004.77

СОВРЕМЕННЫЕ МЕТОДЫ АТАК ДЕАНОНИМИЗАЦИИ НА СЕТЬ TOR

Статья поступила в редакцию 05.12.2019, в окончательном варианте – 11.03.2020.

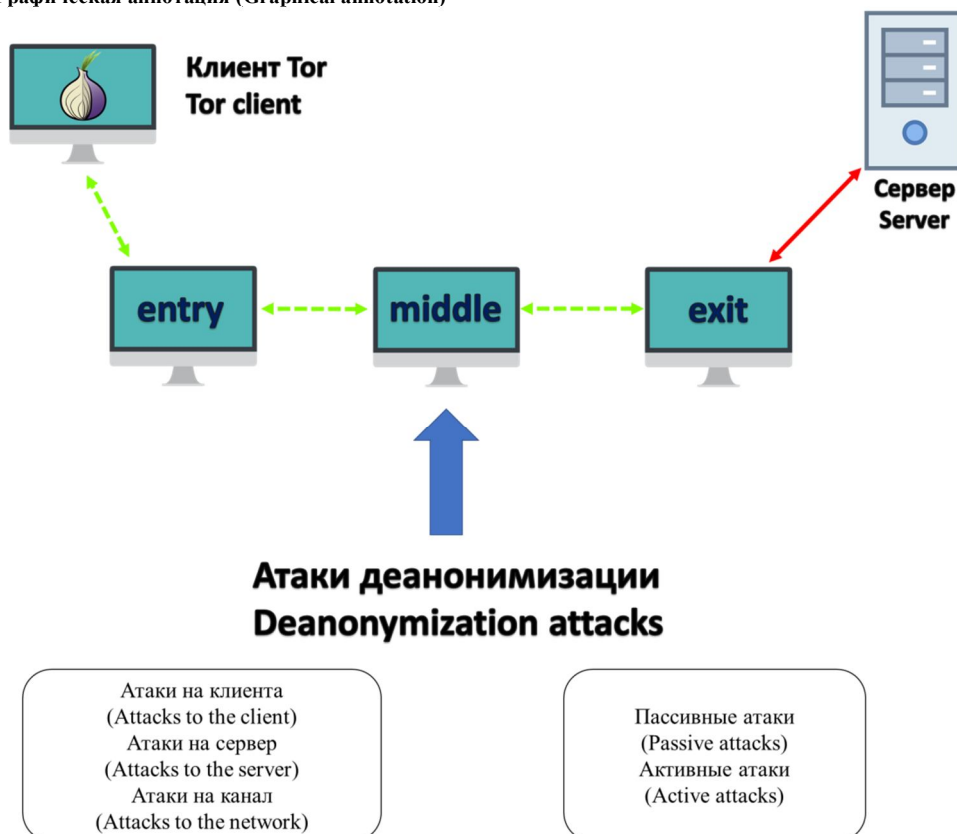
Новосельцева Алёна Вячеславовна, Краснодарский университет Министерства внутренних дел Российской Федерации, 350072, Российская Федерация, г. Краснодар, ул. Ярославская, 128, курсант, e-mail: AlenaNov98@mail.ru

Клюев Станислав Геннадиевич, Краснодарский университет Министерства внутренних дел Российской Федерации, 350072, Российская Федерация, г. Краснодар, ул. Ярославская, 128, кандидат технических наук, начальник кафедры информационной безопасности, e-mail: s.g.klyuev@mail.ru

В данной статье представлены современные методы атак деанонимизации на анонимную сеть Tor, а также предлагается их классификация. Изучены принципы взаимодействия узлов и построение цепочки луковой маршрутизации сети Tor. Рассмотрены более подробно атаки на клиентскую сторону сети (Raptor-атака, Torben-атака), атаки на сервер (атака с пометкой ячеек, Off-path MitM) и атаки на канал (Timing-атака, CellFlood DoS-attack). Также рассмотрены атаки в зависимости от воздействия на перехватываемый трафик (активные атаки, при которых происходит модификация трафика, и пассивные атаки, при которых трафик просто перехватывается и анализируется, но не модифицируется). Приведены примеры (прецеденты) некоторых атак и последствия. Сделан вывод о высокой значимости наличия больших вычислительных мощностей и ресурсов в осуществлении всех видов атак на Tor-сети.

Ключевые слова: анонимные сети, CellFlood DoS-атака, Tor, анализ трафика, деанонимизация, луковая маршрутизация, коррумпированные узлы, даркнет

Графическая аннотация (Graphical annotation)



MODERN METHODS OF ATTACKS OF DEANIMONIZATION ON THE TOR NETWORK

The article was received by the editorial board on 05.12.2019, in the final version – 11.03.2020.

Novoseltseva Alena V., Krasnodar University of the Ministry of Internal Affairs of the Russian Federation, Yaroslavskaya St., 128, Krasnodar, 350072, Russian Federation, cadet, e-mail: AlenaNov98@mail.ru

Klyuev Stanislav G., Krasnodar University of the Ministry of Internal Affairs of the Russian Federation, Yaroslavskaya St., 128, Krasnodar, 350072, Russian Federation, Cand. Sci. (Engineering), Head of Information Security Department, e-mail: s.g.klyuev@mail.ru

This article presents modern methods of deanonymization attacks on the anonymous Tor network, and also proposes their classification. The principles of node interaction and the construction of the onion routing chain of the Tor network are studied. Attacks on the client side of the network (Raptor attack, Torben attack), attacks on the server (attack marked with cells, Off-path MitM) and attacks on the channel (Timing-attack, CellFlood DoS-attack) are considered in more detail. Also considered attack depending on exposure to intercept traffic (active attacks, in which there is traffic modification and passive attacks, in which traffic is simply intercepted and analyzed, but not modified). Examples (precedents) of some attacks and consequences are given. The conclusion is drawn about the high importance of the availability of large computing power and resources in the implementation of all types of attacks.

Key words: anonymous networks, CellFlood DoS-attack, Tor, traffic analysis, deanonymization, onion routing, corrupt nodes, darknet

Введение. В эпоху коммуникаций глобальная сеть Интернет является главным источником информационных ресурсов в повседневной жизни, средством коммуникаций в электронной форме. Будучи важнейшим элементом в жизнедеятельности пользователей-физических лиц, а также правительств и систем критической инфраструктуры, сеть Интернет должна быть защищенной на достаточно высоком уровне.

В контексте обеспечения конфиденциальности важно обеспечить скрытие как контента, которым обмениваются объекты взаимодействия, так и идентичности самих объектов. Анонимные сети как раз и предназначены для сохранения конфиденциальности такого взаимодействия. Сама

анонимность достигается за счет шифрования пользовательских данных и пересылки трафика через ретрансляторы (узлы маршрутизации) или прокси.

Однако стоит отметить, что конфиденциальность взаимодействия пользователей, которая лежит в основе всех анонимных сетей, зачастую используется в преступных целях. Преимущества анонимизации трафика играют на руку террористам, продавцам оружия и наркотиков, мошенникам, экстремистам, сбытчикам запрещенных в обороте товаров и другому преступному контингенту пользователей таких сетей.

Целью работы является исследование существующих на данный момент типов атак деанонимизации на анонимную оверлейную сеть Tor; изучение технических особенностей таких атак, а также оценка возможностей по их реализации.

Анонимная сеть Tor. Существует множество анонимных сетей, включая I2P [6], Freenet [7], MorphMix [12] или Hownet [8]. Тем не менее в настоящее время наиболее популярной луковой сетью (технология анонимного обмена информацией, при котором сообщения неоднократно шифруются и потом отсылаются через несколько сетевых узлов, называемых луковыми маршрутизаторами [3]) является сеть Tor. Протокол Tor на данный момент считается одним из самых популярных сетевых протоколов для организации анонимного взаимодействия.

Tor – это анонимная сеть, которая полностью состоит из волонтерских серверов и программного обеспечения, которое позволяет получать доступ к этой сети. В ней каждый сервер является волонтерским, и чем больше сеть из пользователей – тем выше ее анонимность. Сеть Tor крайне популярна в России, и в 2019 г. страна стала мировым рекордсменом по числу ее пользователей [5] с посещаемостью 600 тысяч человек в день.

Для построения цепочки клиент обращается к серверу каталогов HSDir и загружает список доступных узлов в сети. Далее клиент выстраивает цепочку из абсолютно случайных узлов. Сама луковая цепочка, как правило, состоит из трех ретрансляторов – входной ноды, промежуточной ноды и выходной ноды (рис. 1). Эта цепь будет активной только 10 минут, далее она перестраивается таким же случайным образом. Такое перестроение также происходит при смене целевого сервера или при личном желании клиента.

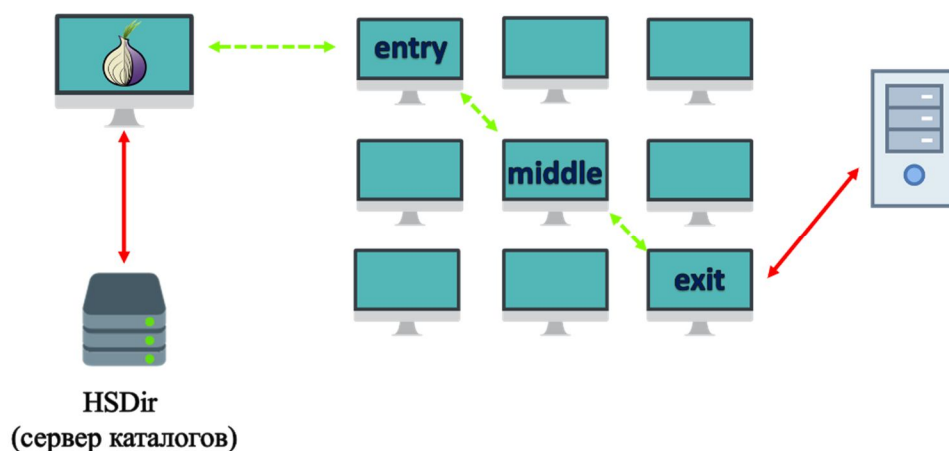


Рисунок 1 – Построение цепочки Tor

Атаки на Tor. Для организации и проведения атак анонимизации необходимо обладать определенными мощностями и ресурсами, например, коррумпированными узлами (узел, трафик которого может модифицировать и просматривать атакующий) или автономными системами.

Перед описанием особенностей атак необходимо отметить существующие на данный момент классификации атак анонимизации на Tor.

Так, в зависимости от воздействия на перехватываемый трафик выделяют активные и пассивные атаки. При активных атаках происходит модификация трафика злоумышленником. В то же время при пассивной атаке происходит лишь sniffing сетевого трафика (т.е. его перехват), как правило, с целью его дальнейшего анализа.

В зависимости от атакуемого объекта выделяются следующие виды:

- атаки на клиента;
- атаки на сервер;
- атаки на сеть.

Атаки на клиента. Наиболее эффективными атаками на клиента являются Torben-атака [10] и RAPTOR-атака [14].

Torben attack. В данной атаке предполагается, что злоумышленник может контролировать зашифрованную связь между клиентом и входной нодой, а также может внедрять специальные маркеры на интересующую клиента веб-страницу. Эти маркеры шаблонизируют трафик, и чаще всего их источником являются рекламные баннеры. Маркер генерирует характерный шаблон трафика, и он может быть обнаружен между клиентом и входной нодой (рис. 2). Это позволяет злоумышленнику деанонимизировать посетителей отмеченных веб-страниц.

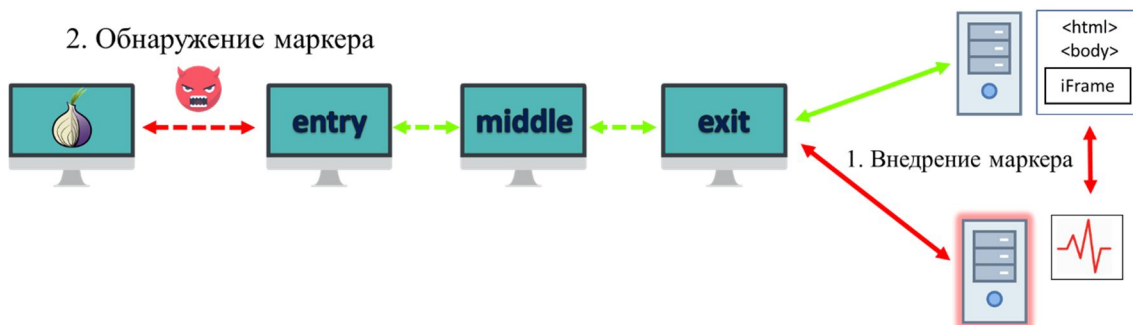


Рисунок 2 – Схема организации Torben-атаки

Маркер представляет собой 20-байтовое значение SHA-1 хэша отслеживаемого URL-адреса веб-страницы. Таким образом, существует соответствие между URL-адресами и их маркерами. Для обнаружения маркера существует многоклассовая машина опорных векторов (support vector machine – SVM) с вероятностными выводами последовательностей отдельных маркеров веб-страниц. Система использует вероятностное сравнение для выбора наиболее подходящего маркера.

По оценке исследователей, с 60 000 веб-страниц атака позволяет обнаруживать эти маркеры с точностью более 91 % без ложных срабатываний [10].

Входной узел Тог является единственным узлом, который напрямую связывается с клиентом. И чтобы побудить клиента использовать определенный «злонамеренный» узел входа, можно блокировать соединения с легальными узлами входа и предоставлять возможность соединения только с коррумпированными. На сетевом уровне это делается с помощью соответствующих политик, например, сетевыми администраторами или провайдерами.

Raptor. Routing attacks on privacy in Tor (RAPTOR) – атака на Tor, где в качестве атакующей стороны выступает автономная система. RAPTOR атака использует динамические аспекты протокола BGP (Border Gateway Protocol – протокол динамической маршрутизации).

Атака состоит из трех частей:

1) асимметричный анализ трафика. В нем анализируются поля TCP-заголовков для выявления номера TCP-последовательности и номер TCP-подтверждения доставки. Вычисляется корреляция между этими полями;

2) анализ натуральных перебоев. Путь между клиентом и входной нодой постоянно меняется, а это повышает вероятность попадания в коррумпированную автономную систему;

3) BGP-сниффинг – коррумпированная автономная система производит атаку «человек посередине» между клиентом и входным узлом. Это позволит автономной системе выполнять асимметричный анализ трафика.

По данной атаке был проведен ряд экспериментов, и атака в 90 % случаев была успешна [14].

Атаки на сервер. В этом типе атак целью являются скрытые сервисы. Рассмотрим некоторые из них.

Cell counting attack (атака с пометкой ячеек). Атакующий в данном случае должен имеет контроль над входным и выходным узлом клиента. Входной узел дублирует любое сообщение, при этом запоминает IP-адрес и время дублирования. Контролируя выходную ноду, атакующий должен засечь дубликат и получить IP-назначения и порт обращения. Таким образом, злоумышленник может вычислить пользователя и какие сервисы он посещает (рис. 3).

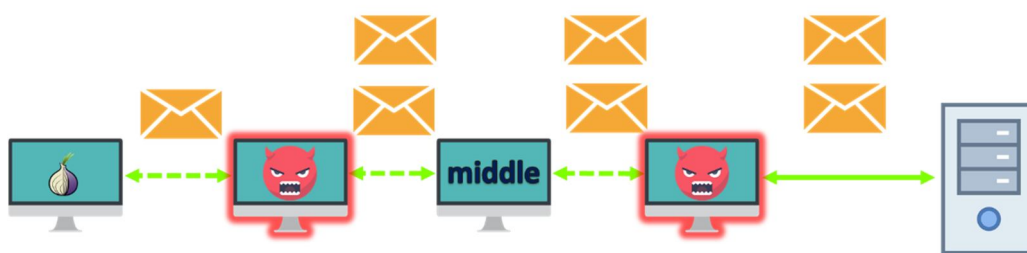


Рисунок 3 – Cell counting attack

Off-path MitM. Атака основана на слежении за метаданными скрытой службы, что позволяет атакующему узнать о ее существовании и доступности [13]. Центральным объектом атаки выступает HSDir.

HSDir – каталог скрытой службы (Hidden Service directory). Эти каталоги содержат информацию, позволяющую получать доступ к onion-доменам (псевдо-домены верхнего уровня, созданные для обеспечения доступа к анонимным или псевдо-анонимным адресам сети Tor [4]), не нарушая анонимности пользователя.

Злоумышленнику необходимо получить закрытый ключ скрытого сервиса. Также он должен создать клиента, который как будто бы обращается к скрытому сервису, и сервис для обманывания клиента.

Последовательность осуществления атаки показана на рисунке 4.

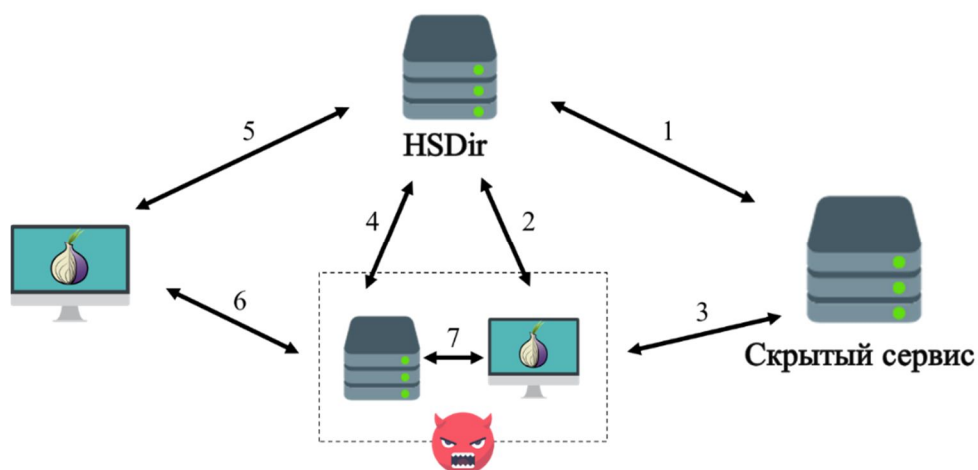


Рисунок 4 – Off-path MitM attack

Основные этапы:

1. Скрытый сервис обращается к каталогу и подгружает дескрипторы.
2. Злоумышленник извлекает информацию из необходимого каталога HSDir.
3. Имитируя работу клиента, злоумышленник устанавливает соединение со скрытым сервисом.
4. Имитируя работу скрытого сервиса, злоумышленник использует скомпрометированный закрытый ключ и загружает новые дескрипторы в HSDir.
5. Клиент, захотев подключиться к скрытому сервису, загружает дескрипторы с HSDir.
6. Клиент подключается к «скрытому сервису» (злоумышленнику).
7. Злоумышленник передает трафик клиента скрытой службе.
8. Злоумышленник находится между сервисом и выходной нодой, где трафик передается в открытом виде.

Таким образом, злоумышленник, который смог скомпрометировать закрытый ключ скрытой службы, может организовать атаку типа «человек посередине» на целевую скрытую службу.

Атаки на сеть (канал). В этом случае целью атаки является сама сеть Tor. Ориентируясь на всю сеть, важно учитывать, что в этом случае вредоносные действия могут влиять на несколько узлов.

Timing-атака. Если атакующий имеет возможность наблюдать трафик клиента и трафик на выходной нодой, то он может установить связь между ними. Данный метод построен на выстраивании корреляции временных шаблонов друг с другом и нахождении определенных зависимостей.

Идея основывается на том, что в Tor задержка по времени не может быть большой, т.е. временной шаблон пакетов данных должен сохраняться при продвижении через цепочку соединения [1].

Коррумпированный узел устанавливает соединение с другими узлами, чтобы измерить задержки соединений и замеряет эти значения. Так злоумышленник рассчитывает транспортную нагрузку на узлы в сети Tor. Когда атакующий получит шаблоны трафика всех узлов, он может воспроизвести атаку по сценарию атаки анализа трафика, высчитав корреляционные зависимости между данными о позиции ячейки, времени, IP-адресе, порте.

CellFlood DoS attack. В сети Tor существуют различные виды сообщений для обмена информацией между узлами. Именно обработка некоторых из них на узлах сети легла в основу этого вида атаки [11].

Атака использует запросы создания цепи (CREATE), которые быстро генерируются атакующим, однако они будут требовать большое количество вычислительных ресурсов от узла для ее обработки. Поэтому из-за криптографических операции обработка CREATE-сообщения занимает в 4 раза больше времени, чем его генерация.

Узел, получающий CREATE-сообщения быстрее, чем его процессор может обработать, отвечает на них, посылая DESTROY-сообщения в ответ. Следовательно, узел, находящийся под атакой, будет отклонять запросы от легитимных узлов (рис. 5).



Рисунок 5 – CellFlood DoS attack

Если атакующий заинтересован в том, чтобы исключить узел или набор узлов из сети Tor, ему выгоднее использовать данный вид атаки, нежели чем обычную и более требовательную по ресурсам классическую DDoS-атаку.

В течение нескольких лет этой уязвимостью активно пользуются злоумышленники. Сначала о подобного рода типе атак сообщали легитимные сайты даркнета (анонимная сеть не связанных между собой виртуальных туннелей, предоставляющая передачу данных в зашифрованном виде [2]), но в последнее время уязвимость используется преимущественно для атак на подпольные торговые площадки. В марте 2019 г. администрация одного из крупнейших черных рынков даркнета Dream Market объявила о его закрытии после серии мощных DDoS-атак [9]. Через месяц после закрытия Dream Market DDoS-атакам подверглись другие крупные торговые площадки, в том числе Empire Market и Nightmare Market. Постоянные DDoS-атаки вынуждают операторов onion-сайтов переходить с Tor на I2P. Действительно DDoS в обоих случаях.

Заключение. В статье рассмотрены современные методы атак деанонимизации на анонимную сеть Tor. Изучив особенности проведения этих атак, можно сделать вывод, что для их проведения требуются большие вычислительные мощности и ресурсы, автономные системы. Поэтому проведение вышеперечисленных типов атак возможно только государственными организациями или частными компаниями, обладающими соответствующим вычислительным потенциалом.

Библиографический список

2. Авдошин С. М. Технология анонимных сетей / С. М. Авдошин, А. В. Лазаренко // Информационные технологии. – 2016. – Т. 22, № 4. – С. 284–291.
3. Даркнет – интернет-энциклопедия «Википедия». – Режим доступа: <https://ru.wikipedia.org/wiki/Даркнет>, свободный. – Заглавие с экрана. – Яз. рус. (дата обращения: 26.12.2019).

4. Луковая маршрутизация – интернет-энциклопедия «Википедия». – Режим доступа: https://ru.wikipedia.org/wiki/Луковая_маршрутизация, свободный. – Заглавие с экрана. – Яз. рус. (дата обращения: 26.12.2019).
5. Onion – интернет-энциклопедия «Википедия». – Режим доступа: <https://ru.wikipedia.org/wiki/onion>, свободный. – Заглавие с экрана. – Яз. рус. (дата обращения: 26.12.2019).
6. Россия стала мировым рекордсменом по числу пользователей Tor. – Режим доступа: https://www.cnews.ru/news/top/2019-07-17_rossiya_ustanovila_mirovoj_rekord_po_chislu_polzovatelej, свободный. – Заглавие с экрана. – Яз. рус. (дата обращения: 29.12.2019).
7. I2P. – Режим доступа: <https://geti2p.net>, свободный. – Заглавие с экрана. – Яз. англ. (дата обращения: 29.12.2019).
8. Freenet. – Режим доступа: <https://freenetproject.org/>, свободный. – Заглавие с экрана. – Яз. англ. (дата обращения: 29.12.2019).
9. Hornet. – Режим доступа: <https://hornet.com/>, свободный. – Заглавие с экрана. – Яз. англ. (дата обращения: 29.12.2019).
10. Хозяева наркорынка в даркнете бежали, прихватив чужие биткойны на \$30 миллионов. – Режим доступа: https://safe.cnews.ru/news/top/2019-05-03_hozyaeva_narkorynka_v_darknete_bezhaliprihvativ, свободный. – Заглавие с экрана. – Яз. рус. (дата обращения: 13.12.2019).
11. Arp D., Torben: A practical side-channel attack for deanonymizing Tor communication / D. Arp, F. Yamaguchi, K. Rieck // Proc. 10th ACM Symp. Inf. Comput. Commun. Security (ASIA CCS). – 2015. – P. 597–602.
12. Ling Z. et al., A new cell-counting-based attack against tor / Z. Ling et al. // IEEE/ACM Trans. Netw. – Aug. 2012. – Vol. 20, no. 4. – P. 1245–1261.
13. Rennhard M. Introducing MorphMix: Peer-to-Peer based Anonymous Internet Usage with Collusion Detection / M. Rennhard, B. Plattner // Proceedings of the Workshop on Privacy in the Electronic Society (WPES 2002). – Washington, DC, 2002.
14. A. Sanatinia G. Off-path man-in-the-middle attack on tor hidden services / A. Sanatinia, G. Noubir. – New England Security Day, NESD, 2017.
15. Yixin Sun. Counter RAPTOR: Safeguarding Tor Against Active Routing Attacks / Yixin Sun, Anne Edmundson, Nick Feamster, Mung Chiang, and Prateek Mittal // Security and Privacy (SP) : IEEE Symposium. – 2017. – P. 977–992.

References

1. Avdoshin S. M., Lazarenko A. V. Tekhnologiya anonimnykh setey [Technology of anonymous networks]. *Informatsionnye tekhnologii* [Information Technologies], 2016, vol. 22, no. 4, pp. 284–291.
2. *Darknet – internet-entsiklopediya «Vikipediya»* [Darknet – internet encyclopedia «Wikipedia»]. Available at: <https://en.wikipedia.org/wiki/Darknet> (accessed 26.12.2019).
3. *Lukovaya marshrutizatsiya – internet entsiklopediya «Vikipediya»* [Onion Routing – internet encyclopedia «Wikipedia»]. Available at: https://en.wikipedia.org/wiki/Onion_routing (accessed 26.12.2019).
4. *Onion – internet-entsiklopediya «Vikipediya»* [Onion – internet encyclopedia «Wikipedia»]. Available at: <https://en.wikipedia.org/wiki/onion> (accessed 26.12.2019).
5. *Rossiya stala mirovym rekordsmenom po chislu polzovateley Tor* [Russia has become the world record holder in the number of Tor users]. Available at: https://www.cnews.ru/news/top/2019-07-17_rossiya_ustanovila_mirovoj_rekord_po_chislu_polzovatelej (accessed 29.12.2019).
6. *I2P*. Available at: <https://geti2p.net> (accessed 29.12.2019).
7. *Freenet*. Available at: <https://freenetproject.org/> (accessed 29.12.2019).
8. *Hornet*. Available at: <https://hornet.com/> (accessed 29.12.2019).
9. *Khozyaeva narkorynka v darknete bezhali, prikhvativ chuzhie bitkoiny na \$30 millionov* [The owners of the drug market fled on the darknet, grabbing other people's bitcoins for \$30 million]. Available at: https://safe.cnews.ru/news/top/2019-05-03_hozyaeva_narkorynka_v_darknete_bezhaliprihvativ (accessed 13.12.2019).
10. Arp D., Yamaguchi F., Rieck K. Torben: A practical side-channel attack for deanonymizing Tor communication. *Proc. 10th ACM Symp. Inf. Comput. Commun. Security (ASIA CCS)*, 2015, pp. 597–602.
11. Ling Z. et al. A new cell-counting-based attack against tor. *IEEE/ACM Trans. Netw.*, Aug. 2012, vol. 20, no. 4, pp. 1245–1261.
12. Rennhard M., Plattner B. Introducing MorphMix: Peer-to-Peer based Anonymous Internet Usage with Collusion Detection. *Proceedings of the Workshop on Privacy in the Electronic Society (WPES 2002)*. Washington, DC, 2002.
13. Sanatinia A., Noubir G. *Off-path man-in-the-middle attack on tor hidden services*. New England Security Day, NESD, 2017.
14. Yixin Sun, Anne Edmundson, Nick Feamster, Mung Chiang, and Prateek Mittal. Counter RAPTOR: Safeguarding Tor Against Active Routing Attacks. *Security and Privacy (SP): IEEE Symposium*, 2017, pp. 977–992.

DOI 10.21672/2074-1707.2020.49.4.162-169

УДК 004.056

ПОСТРОЕНИЕ ВЗАИМОСВЯЗИ МЕЖДУ НАРУШИТЕЛЕМ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ И УЯЗВИМОСТЯМИ ИНФОРМАЦИОННЫХ АКТИВОВ В ИНФОРМАЦИОННЫХ СИСТЕМАХ ОБРАБОТКИ ПЕРСОНАЛЬНЫХ ДАННЫХ

Статья поступила в редакцию 24.01.2020, в окончательном варианте – 05.03.2020.

Жук Роман Владимирович, Филиал «Макрорегион Юг» ООО ИК «СИБИНТЕК», 352800, Российская Федерация, г. Туапсе, ул. Карла Маркса, 36, главный специалист, e-mail: goonerkrd@gmail.com

Дзьобан Павел Игоревич, Кубанский государственный технологический университет, 350072, Российская Федерация, г. Краснодар, ул. Московская, 2,

кандидат технических наук, доцент кафедры компьютерных технологий и информационной безопасности Института компьютерных систем и информационной безопасности, e-mail: antiemoboy@mail.ru

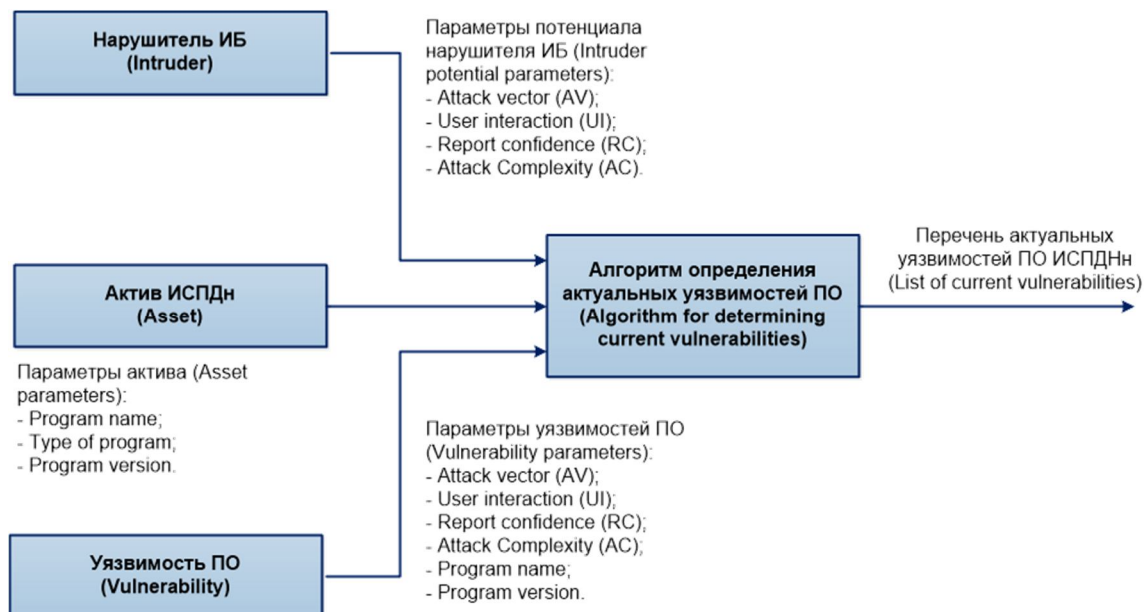
Власенко Александра Владимировна, Кубанский государственный технологический университет, 50072, г. Краснодар, ул. Московская, 2,

кандидат технических наук, доцент, заведующая кафедрой компьютерных технологий и информационной безопасности института компьютерных систем и информационной безопасности, e-mail: Alex_Vlasenko@list.ru

Рассмотрены методики определения активов в информационных системах, выбора уязвимостей программного обеспечения, определения нарушителей и угроз информационной безопасности. Проведен обзор методики оценки уязвимостей программного обеспечения и методики выбора потенциала для нарушителя информационной безопасности с определенными возможностями. Предложена унификация параметров актива информационной системы обработки персональных данных. Предложен способ количественной оценки потенциала нарушителя информационной безопасности, построена взаимосвязь между нарушителем информационной безопасности и уязвимостями программного обеспечения на основе проецирования и унификации метрик вектора оценки уязвимости программного обеспечения на параметры потенциала нарушителя информационной безопасности. Подготовлены производственные правила, позволяющие установить возможность реализации выявленных уязвимостей программного обеспечения выбранным нарушителем информационной безопасности.

Ключевые слова: актив, вектор, уязвимость программного обеспечения, метрика уязвимости, нарушитель информационной безопасности, производственная модель, угроза информационной безопасности

Графическая аннотация (Graphical annotation)



**BUILDING A RELATIONSHIP BETWEEN AN INFORMATION SECURITY
INTRUDER AND VULNERABILITIES OF INFORMATION ASSETS
IN INFORMATION SYSTEMS FOR PROCESSING PERSONAL DATA**

The article was received by the editorial board on 24.01.2020, in the final version – 05.03.2020.

Zhuk Roman V., Branch «Macroregion South» Ltd Co IC «SIBINTEK», 36 Karl Marks St., Tuapse, 352800, Russian Federation,
chief specialist, e-mail: goonerkrd@gmail.com

Dzoban Pavel I., Kuban State Technological University, 2 Moskovskaya St., Krasnodar, 350072, Russian Federation,

Cand. Sci. (Engineering), Associate Professor of the Department of Computer Technologies and Information Security of the Institute of Computer Systems and Information Security, e-mail: antiemoboy@mail.ru

Vlasenko Alexandra V., Kuban State Technological University, 2 Moskovskaya St., Krasnodar, 350072, Russian Federation,

Cand. Sci. (Engineering), Head of the Department of Computer Technologies and Information Security of the Institute of computer systems and information security, e-mail: Vlasenko@kubstu.ru

Methods for determining assets in information systems, selecting software vulnerabilities, and identifying information security intruder and threats to information security are considered. A review of the methodology for assessing software vulnerabilities and the methodology for selecting the potential for an information security intruder with certain capabilities is carried out. Unification of parameters of an asset of an information system for processing personal data is proposed. A method for quantifying the potential of an information security intruder is proposed, and the relationship between an information security intruder and software vulnerabilities is built on the basis of projecting and unifying metrics of the software vulnerability assessment vector on the parameters of the information security intruder's potential. We have prepared production rules that allow us to determine whether the identified software vulnerabilities can be implemented by the selected information security violator.

Key words: vector, software vulnerability, vulnerability metric, information security intruder, production model, information security threat

Введение. Как показывается практика применения международных стандартов по информационной безопасности (ИБ), существует различие между риском ИБ и актуальной угрозой ИБ. А именно, риск ИБ является мерой возможной реализации уязвимости актива определенной угрозой, в результате чего может быть нанесен определенный ущерб организации – владельцу актива [1]. Угроза ИБ в информационных системах обработки персональных данных (ИСПДн) является совокупностью несанкционированного, в том числе, доступа к информации, содержащей персональные данные (далее – ПДн). Результатом такого доступа может стать одно или несколько деструктивных действий (уничтожение, изменение, блокирование, копирование, распространение) [3]. Актуальной угрозой ИБ в ИСПДн считаются угрозы ИБ представляющие потенциальную опасность для ПДн и имеющие возможность реализации, оцененную по методике, представленной в [3].

В [6] представлена классификация угроз несанкционированного доступа в ИСПДн, а также типовые модели угроз ИБ в ИСПДн. Исходя из методик, приведенных в [3] и [6], угроза ИБ представлена как совокупность следующих факторов:

- источник угрозы ИБ;
- уязвимость ИСПДн;
- способ реализации;
- объект воздействия;
- деструктивное действие.

С появлением банка данных угроз безопасности [6] наблюдается тенденция к унификации методик определения угроз ИБ и формирования модели нарушителя ИБ и угроз ИБ в информационных системах (далее – ИС). Дополнительно, регулятором в 2015 г. для обсуждения был опубликован проект методического документа [6]. Однако данный документ на момент написания статьи не был введен в действие.

Рядом нормативных документов отечественных регуляторов в области ИБ для различных видов ИС предусмотрена разработка модели угроз ИБ. Например, данные требования аналогичны для государственных информационных систем [7], автоматизированных систем управления производственными и технологическими процессами [10].

Основным отличием при разработке модели угроз ИБ для ИСПДн от вышеперечисленных ИС является наличие утвержденных методик: [3] и [6].

В настоящее время для определения угроз ИБ в ИСПДн банк данных угроз безопасности представляет справочную информацию о существующих угрозах ИБ и уязвимостях программного обеспечения (далее – ПО). Определение перечня угроз ИБ осуществляется в соответствии с пользовательским фильтром, основными параметрами которого являются:

- тип нарушителя и его потенциал;
- нарушения свойств защищенности информации (конфиденциальность, целостность, доступность) по результатам реализации угрозы.

Определение перечня уязвимостей ИБ осуществляется аналогичным способом с использованием дополнительных параметров, применяемых в методиках оценки уязвимостей [7] и [11]:

- класс уязвимости;
- уровень опасности;
- базовый вектор (Access Vector, Attack Complexity, Authentication);
- идентификатор типа ошибки;
- наличие эксплойта;
- способ эксплуатации;
- способ устранения;
- операционная система.

На основании вышеизложенного разработка модели угроз ИБ посредством банка данных угроз безопасности для различных видов ИС существенным образом зависит от квалификации эксперта, либо группы экспертов, привлекаемых для выбора угроз ИБ. Применение банка данных угроз безопасности для разработки модели угроз ИБ в ИСПДн, совместно с методиками [3] и [6], также является непрактичным в виду отсутствия связи между используемыми банком и параметрами вышеуказанных методик.

Отсутствует взаимосвязь между угрозами ИБ и уязвимостями ПО, приведенными в банке данных угроз безопасности.

Целью данной работы является выбор параметров и организация взаимосвязи между активами ИСПДн, уязвимостями активов и возможностями нарушителя ИБ с целью формирования перечня уязвимостей ПО, которые могут быть реализованы в ИСПДн. Для достижения данной цели необходимо **решить следующие задачи**:

- классифицировать активы и подготовить перечень параметров для их определения;
- оптимизировать перечень нарушителей и разработать методику оценки их потенциала;
- разработать алгоритм выбора уязвимостей.

Подготовка перечня параметров активов ИСПДн. Классификация активов ИСПДн получена путем проецирования показателей исходной защищенности, приведенных в [3] и в проекте методического документа «Методика определения угроз безопасности информации в информационных системах», а также категорий нарушителя, представленных в [1], на перечень активов представленных в [1]. Учитывая необходимость унификации и автоматизации разрабатываемой методики, а также распределение уязвимостей по параметрам, представленное в банке данных угроз безопасности и общедоступных базах данных уязвимостей ПО, например [11], экспертным путем подготовлены следующие параметры активов:

- наименование ПО;
- тип ПО;
- версия ПО.

Унификация нарушителей информационной безопасности и количественная оценка из потенциала. Согласно классификации, представленной в [1], существует два типа нарушителей, имеющих различные категории в зависимости от наличия определенных возможностей:

- внешний (5 категорий);
- внутренний (8 категорий).

Банки данных угроз безопасности для определения угроз ИБ используют классификацию нарушителей ИБ, представленную в [6] и основанную на потенциале:

- внешний (3 типа потенциала);
- внутренний (3 типа потенциала).

Виды нарушителей ИБ представленные в [1] и [6] могут быть сопоставлены, а их перечень может быть проанализирован и оптимизирован экспертным путем. По результатам оптимизации перечень нарушителей ИБ с присвоенными потенциалами представлен на рисунке 1.

	Внешний	Внутренний
Высокий	Разведывательные службы государств	
Средний	Криминальные структуры, террористические, экстремистские группировки, конкурирующие организации	Системный администратор и администратор безопасности
	Разработчики, производители, поставщики программных, технических и программно-технических средств	
Низкий	Внешние субъекты (физические лица)	Лица, привлекаемые для установки, наладки, монтажа, пусконаладочных, а также для обеспечения функционирования ИСПДн
		Пользователи ИСПДн

Рисунок 1 – Оптимизированный перечень нарушителей ИБ

Согласно [6], потенциал нарушителя ИБ строго определен. Однако для новых нарушителей ИБ предусмотрен алгоритм и параметры определения потенциала:

- время, затрачиваемое нарушителем ИБ на поиск, определение и использование уязвимости;
- техническая компетентность;
- знание проекта ИС;
- оснащенность;
- возможности доступа к ИСПДн.

Присвоение значений параметрам определения потенциала нарушителя осуществляется экспертным путем. Основной проблемой при определении потенциала нарушителя ИБ по методике [6] является отсутствие связи нарушителя ИБ и уязвимостей ПО активов ИСПДн.

Одним из возможных способов построения связи между потенциалом нарушителя ИБ и уязвимостью ПО активов ИСПДн является проецирование базовых и временных метрик, используемых в общей системе оценки уязвимостей ПО [7] и [8], на параметры потенциала (табл. 1).

Таблица 1 – Метрики, сопоставленные с возможностями нарушителя ИБ

№ п/п	Наименование метрики	Возможность нарушителя ИБ	Метрика
1	Возможности физического доступа к активу:	Возможность доступа к ИС	Attack Vector (AV)
	Через сети общего доступа		Network (N)
	С помощью локально-вычислительной сети		Adjacent Network (A)
	Физический доступ		Physical (P)
2	Необходимость взаимодействия с пользователем	Знание проекта и информационной системы	User Interaction (UI)
	Необходимо		Required (R)
	Нет необходимости		None (N)
3	Наличие информации об уязвимости в общем доступе	Техническая компетентность нарушителя	Report Confidence (RC)
	Отсутствует описание		Unknown (U)
	Частично описана		Reasonable (R)
	Полностью описана		Confirmed (C)
4	Возможность применения специальных средств для эксплуатации уязвимости	Оснащенность нарушителя	Attack Complexity (AC)
	Применение специальных средств		High (H)
	Специальные средства не применяются		Low (L)

Для **разработки алгоритма выбора уязвимостей** проведен анализ различных источников информации об уязвимостях ПО, например [12], [13] и [14]. По результатам анализа установлены параметры, которые могут влиять на выбор уязвимостей для подготовленного перечня активов, в состав которого входит ПО:

- наименование ПО;
- версия ПО.

Таким образом, процесс определения уязвимостей ПО будет детализован на два этапа:

1. Определение перечня уязвимостей ПО для выбранных активов.
2. Сопоставление метрик, выбранных на «этапе 1» уязвимостей ПО с параметрами потенциала выбранного нарушителя ИБ.

Используя [4], спроецируем количественные значения на параметры потенциала (табл. 2).

Таблица 2 – Значения метрик уязвимости

№ п/п	Наименование метрики	Метрика	Числовое значение
1	Возможности физического доступа к активу:	Attack Vector (AV)	
	Через сети общего доступа	Network (N)	0,85
	С помощью локально-вычислительной сети	Adjacent Network (A)	0,62
	Физический доступ	Physical (P)	0,2
2	Необходимость взаимодействия с пользователем	User Interaction (UI)	
	Необходимо	Required (R)	0,62
	Нет необходимости	None (N)	0,85
3	Наличие информации об уязвимости в общем доступе	Report Confidence (RC)	
	Отсутствует описание	Unknown (U)	0,92
	Частично описана	Reasonable (R)	0,96
	Подтверждена производителем ПО	Confirmed (C)	1
4	Возможность применения специальных средств для эксплуатации уязвимости	Attack Complexity (AC)	
	Применение специальных средств	High (H)	0,44
	Специальные средства не применяются	Low (L)	0,77

Сопоставление параметров нарушителя с метриками вектора уязвимостей позволит разработать производционную модель вида:

$$R = \langle X_1A_1, X_2A_2, \dots, X_nA_n: YB \rangle.$$

Левая часть правила является антецедентом X и включает в себя перечень предпосылок для реализации правой части – консеквента Y. Между собой предпосылки связываются посредством операций «И», «ИЛИ».

Для реализации второго этапа создается следующее производционное правило:

- Vulnerability: если «AV уязвимости актива» <= «AV нарушителя ИБ» И «UI уязвимости актива» <= «UI нарушителя ИБ» И «RC уязвимости актива» <= «RC нарушителя ИБ» И «АС уязвимости актива» <= «АС нарушителя ИБ», то «Уязвимость №» = «Актуальная уязвимость».

Таким образом, для подготовки перечня уязвимостей ПО, которые могут быть реализованы существующим нарушителем ИБ, необходимым условием будет являться превосходство параметров потенциала нарушителя ИБ над метриками уязвимостей ПО из подготовленного списка для ИСПДн, либо их равенство.

По результатам сформированного перечня уязвимостей ПО, которые могут быть реализованы существующим нарушителем ИБ, возможно подготовить перечень угроз ИБ в ИСПДн.

Заключение. Используя параметры актива и проецируя метрики уязвимостей ПО на параметры потенциала нарушителя ИБ, может быть построена связь между нарушителем ИБ и существующими для активов ИСПДн уязвимостями ПО, а также создан объективный алгоритм подготовки перечня актуальных уязвимостей ПО, который может быть реализован в виде ПО или скриптов, для сокращения времени затраченного на разработку модели нарушителя ИБ и подготовку перечня уязвимостей ПО для активов ИСПДн.

Библиографический список

1. Методический документ ФСТЭК России «Базовая модель угроз безопасности персональных данных при их обработке в информационных системах персональных данных (выписка)» – 2008. – 69 с. – Режим доступа: <https://fstec.ru/component/attachments/download/289>, свободный. – Заглавие с экрана. – Яз. рус. (дата обращения: 20.01.2020).
2. ГОСТ Р ИСО/МЭК 27005-2010. Информационная технология. Методы и средства обеспечения безопасности. Менеджмент риска информационной безопасности. – Режим доступа: <http://docs.cntd.ru/document/gost-r-iso-mek-27005-2010>, свободный. – Заглавие с экрана. – Яз. рус. (дата обращения: 20.01.2020).
3. Информационное сообщение ФСТЭК России от 6 марта 2015 г. № 240/22/879 «О банке данных угроз безопасности информации». – Режим доступа: <https://fstec.ru/normotvorcheskaya/informatsionnye-i->

analiticheskie-materialy/956-informatsionnoe-soobshchenie-fstek-rossii-ot-6-marta-2015-g-240-22-879, свободный. – Заглавие с экрана. – Яз. рус. (дата обращения: 20.01.2020).

4. Калькулятор метрик уязвимостей. – Режим доступа: <https://www.first.org/cvss/calculator/cvsscalc30.js>, свободный. – Заглавие с экрана. – Яз. рус. (дата обращения: 20.01.2020).

5. Методический документ ФСТЭК России «Методика определения актуальных угроз безопасности персональных данных при их обработке в информационных системах персональных данных», – 2008. – 10 с. – Режим доступа: <https://fstec.ru/tekhnicheskaya-zashchita-informatsii/dokumenty/114-spetsialnye-normativnye-dokumenty/380>, свободный. – Заглавие с экрана. – Яз. рус. (дата обращения: 20.01.2020).

6. Методический документ ФСТЭК России «Методика определения угроз безопасности информации в информационных системах» (проект). – 2015. – 43 с. – Режим доступа: <https://fstec.ru/component/attachments/download/812>, свободный. – Заглавие с экрана. – Яз. рус. (дата обращения: 20.01.2020).

7. Общая система оценки уязвимостей, версия 2.0, июнь 2007 г., Петер Мелл, Карэн Шарфон, Национальный институт стандартизации и технологий, Саша Романовски, Университет Карнеги Меллон. – Режим доступа: <https://fstec.ru/tekhnicheskaya-zashchita-informatsii/dokumenty/114-spetsialnye-normativnye-dokumenty/380>, свободный. – Заглавие с экрана. – Яз. рус. (дата обращения: 17.01.2020).

8. Общая система оценки уязвимостей, версия 3.1, Технический документ, Редакция 1. – Режим доступа: <https://www.first.org/cvss/v3.1/specification-document>, свободный. – Заглавие с экрана. – Яз. рус. (дата обращения: 17.01.2020).

9. Приказ ФСТЭК России от 11.02.2013 г. № 17 «Об утверждении требований о защите информации, не составляющей государственную тайну, содержащейся в государственных информационных системах». – Режим доступа: <https://fstec.ru/normotvorcheskaya/akty/53-prikazy/702>, свободный. – Заглавие с экрана. – Яз. рус. (дата обращения: 19.01.2020).

10. Приказ ФСТЭК России от 14.03.2014 г. № 31 «Об утверждении требований к обеспечению защиты информации в автоматизированных системах управления производственными и технологическими процессами на критически важных объектах, потенциально опасных объектах, а также объектах, представляющих повышенную опасность для жизни и здоровья людей и для окружающей природной среды». – Режим доступа: <https://fstec.ru/index?id=868:prikaz-fstek-rossii-ot-14-marta-2014-g-n-31>, свободный. – Заглавие с экрана. – Яз. рус. (дата обращения: 19.01.2020).

11. Сайт базы данных уязвимостей. – Режим доступа: <https://nvd.nist.gov/vuln-metrics/cvss>, свободный. – Заглавие с экрана. – Яз. рус. (дата обращения: 20.01.2020).

12. Сайт базы данных уязвимостей. – Режим доступа: <https://nvd.nist.gov/vuln/>, свободный. – Заглавие с экрана. – Яз. рус. (дата обращения: 20.01.2020).

13. Сайт базы данных уязвимостей. – Режим доступа: <http://www.cvedetails.com>, свободный. – Заглавие с экрана. – Яз. рус. (дата обращения: 20.01.2020).

14. Сайт базы данных уязвимостей. – Режим доступа: <http://www.securityfocus.com>, свободный. – Заглавие с экрана. – Яз. рус. (дата обращения: 20.01.2020).

References

1. *Metodicheskiy dokument FSTEK Rossii "Bazovaya model ugroz bezopasnosti personalnykh dannykh pri ikh obrabotke v informatsionnykh sistemakh personalnykh dannykh (vypiska)"* [Guidance document of the FSTEC of Russia "The basic model of personal data security threats when they are processed in personal data information systems (extract)"], 2008. 69 p. Available at: <https://fstec.ru/component/attachments/download/289> (accessed 20.01.2020).

2. *GOST R ISO/ MEK 27005-2010. Informatsionnaya tehnologiya. Metodi i sredstva obespecheniya informacionnoi bezopashnosti* [GOST R ISO / IEC 27005-2010. Information technology. Methods and means of ensuring security. Information security risk management]. Available at: <http://docs.cntd.ru/document/gost-r-iso-mek-27005-2010> (accessed 20.01.2020).

3. *Informatsionnoe pismo FSTEK Rossii ot 6 marta 2015 g. №240/22/879 "O banke dannykh ugroz bezopasnosti informatsii"* [FSTEC information message no. 240/22/879 of March 6, 2015 «On the data Bank of information security threats.».] Available at: <https://fstec.ru/normotvorcheskaya/informatsionnye-i-analiticheskie-materialy/956-informatsionnoe-soobshchenie-fstek-rossii-ot-6-marta-2015-g-240-22-879> (accessed 20.01.2020).

4. *Kalkulyator metrik uyazvimostey* [Calculator metrics of the vulnerabilities.] Available at: <https://www.first.org/cvss/calculator/cvsscalc30.js> (accessed 20.01.2020).

5. *Metodicheskiy dokument FSTEK Rossii "Methodika opredeleniya aktualnykh ugroz bezopasnosti personalnykh dannykh v informatsionnykh sistemakh personalnykh dannykh"* [Guidance document of the FSTEC of Russia "Methodology for determining current threats to the security of personal data when they are processed in personal data information systems"], 2008. 10 p. Available at: <https://fstec.ru/tekhnicheskaya-zashchita-informatsii/dokumenty/114-spetsialnye-normativnye-dokumenty/380> (accessed 20.01.2020).

6. *Metodicheskiy dokument FSTEK Rossii "Methodika opredeleniya ugroz bezopasnosti informacii v informacionnih sistemakh" (proekt)* [Methods for determining information security threats in information systems (project)]. Available at: <https://fstec.ru/component/attachments/download/812> (accessed 20.01.2020).

7. *Obshchaya otsenka uyazvimostey* [General vulnerability assessment system] version 2.0, June 2007, Peter Mell, Karen Sharfon, national Institute of standardization and technology, Sasha Romanovsky, Carnegie Mellon University]. Available at: <https://fstec.ru/tekhnicheskaya-zashchita-informatsii/dokumenty/114-spetsialnye-normativnye-dokumenty/380> (accessed 17.01.2020).

8. *General vulnerability assessment system, version 3.1, Technical document, Revision 1.* Available at: <https://www.first.org/cvss/v3.1/specification-document> (accessed 19.01.2020).

9. *Prikaz FSTEC Rossii ot 11.02.2013 №17 "Ob utverzhenii trebovaniy o zashchite informatsii, ne sostavlyayushchey gosudarstvennyuyu tayny, sodergashcheysya v gosydarstvennykh informatsionnykh sistemakh"* [FSTEC order no. 17.11.2013 «On approval of requirements for the protection of information that does not constitute a state secret contained in state information systems»]. Available at: <https://fstec.ru/normotvorcheskaya/akty/53-prikazy/702> (accessed 19.01.2020).

10. *Prikaz FSTEC Rossii ot 31.03.2014 № 31 "Ob utverzhenii trebovaniy o zashchite informatsii v avtomatizirovannykh sistemakh upravleniya proizvodstvennyymi i tekhnologicheskimi protsessami na kriticheskikh vazhnykh obektakh, potentsialno opasnykh obektakh, a takzhe obektakh predstavlyayushikh povyshennuyu opasnost dlya zhizni izdorovya ludey i dlya okruzhayushchey prirodnoy sredy"* [FSTEC order no. 31 of March 14, 2014 «On approval of requirements for ensuring information security in automated production and process control systems at critical facilities, potentially dangerous facilities, as well as objects that pose an increased risk to human life and health and the environment»]. Available at: <https://fstec.ru/index?id=868:prikaz-fstek-rossii-ot-14-marta-2014-g-n-31> (accessed 19.01.2020).

11. *Sayt bazy dannykh uyazvimostey* [Vulnerability database site]. Available at: <https://nvd.nist.gov/vuln-metrics/cvss> (accessed 20.01.2020).

12. *Sayt bazy dannykh uyazvimostey* [Vulnerability database site]. Available at: <https://nvd.nist.gov/vuln/> (accessed 20.01.2020).

13. *Sayt bazy dannykh uyazvimostey* [Vulnerability database site]. Available at: <http://www.cvedetails.com> (accessed 20.01.2020).

14. *Sayt bazy dannykh uyazvimostey* [Vulnerability database site]. Available at: <http://www.securityfocus.com> (accessed 20.01.2020).

DOI 10.21672/2074-1707.2020.49.4.169-178

УДК 004.056

ОПРЕДЕЛЕНИЕ АКТУАЛЬНОСТИ УГРОЗ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ В ИНФОРМАЦИОННЫХ СИСТЕМАХ ОБРАБОТКИ ПЕРСОНАЛЬНЫХ ДАННЫХ С ИСПОЛЬЗОВАНИЕМ МАТЕМАТИЧЕСКОГО АППАРАТА НЕЙРОННЫХ СЕТЕЙ

Статья поступила в редакцию 04.02.2020, в окончательном варианте – 09.03.2020.

Жук Роман Владимирович, Филиал «Макрорегион Юг» ООО ИК «СИБИНТЕК», 352800, Российская Федерация, г. Туапсе, ул. Карла Маркса, 36, главный специалист, e-mail: goonerkrd@gmail.com

Дзьобан Павел Игоревич, Кубанский государственный технологический университет, 350072, Российская Федерация, г. Краснодар, ул. Московская, 2, кандидат технических наук, доцент кафедры компьютерных технологий и информационной безопасности Института компьютерных систем и информационной безопасности, e-mail: antimoboy@mail.ru

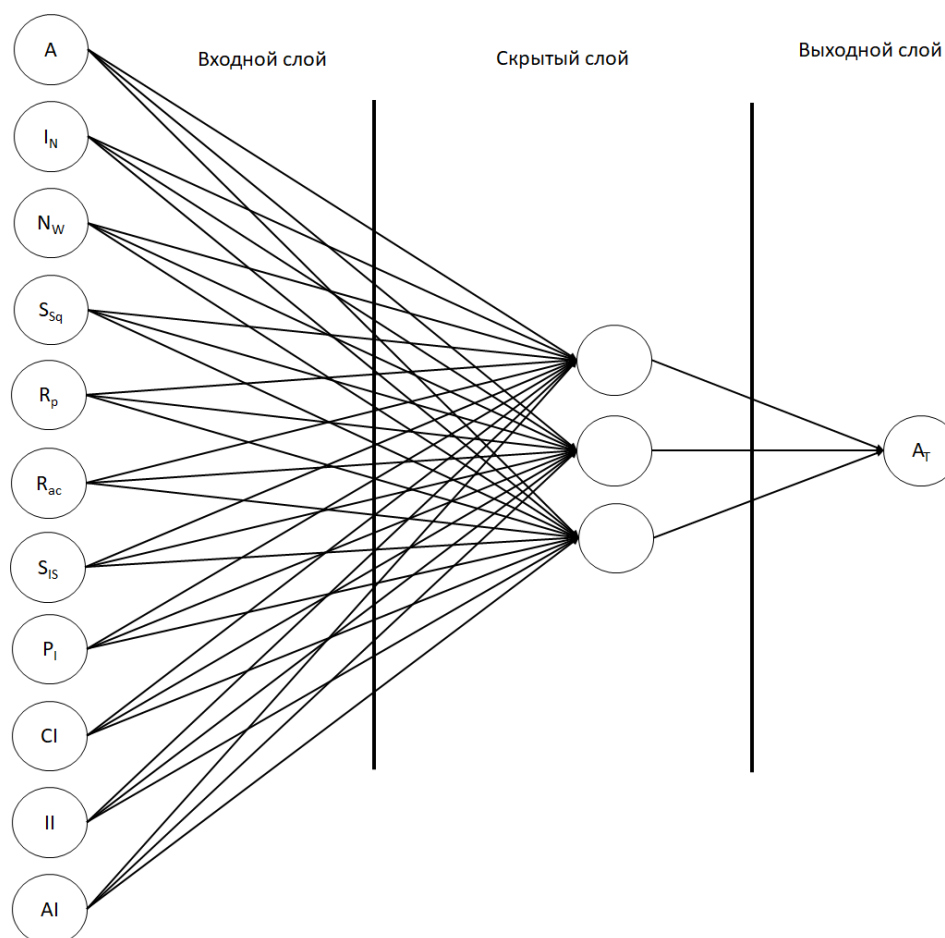
Власенко Александра Владимировна, Кубанский государственный технологический университет, 50072, г. Краснодар, ул. Московская, 2,

кандидат технических наук, доцент, заведующая кафедрой компьютерных технологий и информационной безопасности института компьютерных систем и информационной безопасности, e-mail: Alex_Vlasenko@list.ru

Рассмотрены методики определения угроз информационной безопасности в информационных системах обработки персональных данных. В связи с отсутствием согласованности существующей утвержденной методики с применяемой базой данных угроз информационной безопасности (<https://bdu.fstec.ru/>) был проведен анализ параметров актуальности угроз информационной безопасности и предложен способ определения актуальности угроз информационной безопасности с использованием математического аппарата искусственных нейронных сетей. Для реализации этого способа был выполнен анализ топологий искусственных нейронных сетей и методов вычисления ошибок в них. Разработана искусственная нейронная сеть на основании топологии многослойного перцептрона с обратным распространением ошибки. Проведено обучение разработанной искусственной нейронной сети путем подготовки и использования обучающей выборки. Осуществлено сравнение быстродействия функционирования разработанной искусственной нейронной сети с быстродействием привлеченной группы экспертов, действующих по существующей утвержденной методике определения актуальности угроз информационной безопасности в информационных системах обработки персональных данных.

Ключевые слова: показатель исходной защищенности, потенциал нарушителя информационной безопасности, искусственная нейронная сеть, перцептрон, искусственный нейрон, слой, входной сигнал, выходной сигнал, функция активации, сигмоидальная функция, обучающая выборка, угроза информационной безопасности

Графическая аннотация (Graphical annotation)



**DETERMINING THE RELEVANCE OF INFORMATION SECURITY THREATS
IN INFORMATION SYSTEMS FOR PROCESSING PERSONAL DATA
USING THE MATHEMATICAL APPARATUS OF NEURAL NETWORKS**

The article was received by the editorial board on 04.02.2020, in the final version – 09.03.2020.

Zhuk Roman V., Branch «Macroregion South» Ltd Co IC «SIBINTEK», 36 Karl Marks St., Tuapse, 352800, Russian Federation,
chief specialist, e-mail: goonerkrd@gmail.com

Dzoban Pavel I., Kuban State Technological University, 2 Moskovskaya St., Krasnodar, 350072, Russian Federation,

Cand. Sci. (Engineering), Associate Professor of the Department of Computer Technologies and Information Security of the Institute of Computer Systems and Information Security, e-mail: antiemoboy@mail.ru

Vlasenko Alexandra V., Kuban State Technological University, 2 Moskovskaya St., Krasnodar, 350072, Russian Federation,

Cand. Sci. (Engineering), Head of the Department of Computer Technologies and Information Security of the Institute of computer systems and information security, e-mail: Vlasenko@kubstu.ru

Describes methods of determining threats to the information security in information systems of personal data processing. Due to the lack of consistency of the existing approved methodology with the one used by the information security threat data Bank (<https://bdu.fstec.ru/>) the analysis of the parameters of the relevance of threats is carried out and a method for determining the relevance of is threats using the mathematical apparatus of artificial neural networks is proposed. To implement the method, the analysis of artificial neural networks topologies and methods for calculating errors in artificial neural networks is performed. The artificial neural network was developed based on the topology of a multi-layer perceptron with reverse error propagation. Training of the developed information systems was conducted by preparing a training sample. The performance of the developed information

systems was compared with the performance of the involved group of experts acting on the existing approved methodology for determining the relevance of is threats in the personal data processing.

Key words: initial security indicator, potential of information security intruder, artificial neural network, perceptron, artificial neuron, layer, input signal, output signal, activation function, sigmoid function, training sample, information security threat

Введение. В настоящее время математический аппарат искусственных нейронных сетей (ИНС) применяется в различных областях науки для решения широкого спектра задач, оперирующих со значительными объемами данных. В частности, широкое распространение ИНС получили в рамках решения задач машинного обучения. Основными направлениями аппарата ИНС сейчас являются распознавание образов в изображениях и классификация объектов.

Также ИНС активно применяются в различных областях деятельности, связанных с информационной безопасностью (ИБ) [1].

- идентификация и аутентификация объектов (в частном случае – биометрия);
- методы обнаружения и предотвращения вторжений;
- механизмы обнаружения и классификации вирусной активности;
- анализ защищенности информационных систем (ИС);
- выявление и классификация уязвимостей программного обеспечения (далее – уязвимости ПО);
- оценка рисков ИБ, направленных на информационные активы

К последнему пункту также может быть отнесена задача подготовка модели угроз ИБ для ИС. Частной подзадачей такой задачи является подготовка перечня угроз ИБ для ИС обработки персональных данных (далее – ИСПДн).

Многие эксперты в области ИБ при разработке модели угроз ИБ для ИСПДн сталкиваются с рядом проблем, решение которых отсутствует на протяжении долгого времени.

1. Существующие методики определения актуальных угроз ИБ в ИСПДн [4, 6] морально устарели.

2. Использование разработанного регуляторами в области ИБ банка угроз ИБ [5] противоречит утвержденным методикам [4, 6].

3. В банке данных угроз ИБ (<https://bdu.fstec.ru/>) отсутствует утвержденная методика определения потенциала нарушителя ИБ.

4. В банке данных угроз ИБ (<https://bdu.fstec.ru/>) отсутствует взаимосвязь между уязвимостями ПО и угрозами ИБ.

Целью данной работы является анализ и выбор параметров актуальности угрозы ИБ в ИСПДн, а также разработка механизма определения актуальности угроз ИБ посредством применения математического аппарата ИНС. Для достижения данной цели необходимо **решить следующие задачи:**

- произвести анализ и подготовить перечень параметров актуальности угрозы ИБ ИСПДн;
- разработать модель ИНС и осуществить ее обучение.

Анализ существующих методик определения актуальных угроз ИБ в ИСПДн. Проблема построения взаимосвязи между уязвимостями ПО и угрозами ИБ в «банке данных угроз ИБ» является фундаментальной – ранее на нее было обращено внимание множества экспертов в области ИБ. Один из способов решения данной проблемы был предложен в [2] и заключается в применении математического аппарата ИНС для определения вероятности угрозы ИБ. При этом входными сигналами для создаваемой ИНС являются формализованные параметры записи об уязвимости ПО в «банке данных угроз ИБ», например:

- класс уязвимости ПО;
- тип ошибки ПО;
- тип ПО.

Веса для нейронов входного слоя ИНС назначаются в случайном порядке. Выходными параметрами данной ИНС являются числовые значения вероятности реализации угрозы ИБ в интервале от 0 до 1. Способ обучения разработанной ИНС предложен на примере угроз ИБ из «банка данных угроз ИБ», входящих в перечень OWASP Top 10 [9]. Для каждой выбранной угрозы ИБ формируется отдельная ИНС. Обучение ИНС проводится на основе выборки параметров из записей уязвимостей ПО в банке угроз ИБ. Для проведения обучения ИНС формируются две обучающие выборки для уязвимостей ПО:

- эксплуатация которых может повлечь реализацию угрозы ИБ;
- эксплуатация которых не может повлечь реализацию угрозы ИБ.

Недостатком данного подхода является отсутствие алгоритма подготовки перечня уязвимостей ПО и угроз ИБ для ИСПДн. Первоначально перечень угроз ИБ готовится на основании выбранного потенциала нарушителя ИБ и статистической информации из открытых источников интернета. Дополнительно в [2] не рассмотрено отсутствие статистической информации при обучении ИНС. На

следующем этапе выбранные угрозы ИБ экспертным путем сопоставляются с уязвимостями ПО. Затем осуществляется выборка актуальных угроз ИБ посредством ИНС, на вход которой подаются параметры выбранных экспертом уязвимостей ПО. Таким образом, мерой выбора угрозы ИБ является потенциал нарушителя ИБ. При этом выбор уязвимостей ПО полностью зависит от квалификации привлекаемого эксперта.

В [4] и [6] актуальность угрозы ИБ в ИСПДн представлена как соотношение «вероятности или возможности ее реализации нарушителем ИБ» со «степенью возможного ущерба от реализации данной угрозы ИБ». Возможность реализации угрозы ИБ определяется в зависимости от уровня защищенности для двух видов ИСПДн:

- проектируемых;
- введенных в эксплуатацию (эксплуатируемых).

В связи с отсутствием статистической информации по введенным в эксплуатацию ИСПДн, предлагается использовать структурно-функциональные характеристики ИСПДн уровня проектной защищенности [7]. Для подготовки исчерпывающего перечня данных характеристик осуществим сравнение предлагаемых в [7] параметров с аналогичными показателями уровня исходной защищенности [4]. Результат сравнения и оптимизации данных параметров представлен в таблице.

Таблица – Показатели исходной защищенности ИСПДн

№ п/п	Показатель	Влияние на уровень исходной защищенности		
		Высокий	Средний	Низкий
1	Архитектура ИС:			
	Тонкий клиент	+		
	Одноранговая сеть		+	
	Файл-серверная			+
	ЦОД			+
	Удаленный доступ			+
	Разные типы ОС (гетерогенность среды)		+	
	ППО, независимое от ОС		+	
2	Взаимодействие ИСПДн со сторонними ИСПДн:			
	Взаимодействует			+
	Не взаимодействует		+	
3	Взаимодействие ИСПДн с сетями общего пользования:			
	Подключена;			+
	Подключена через выделенную инфраструктуру		+	
	Не подключена	+		
4	Территориальное размещение ИСПДн:			
	Распределенная			+
	Локальная в пределах одной КЗ		+	
	Локальная на одном АРМ, не подключенном к сети	+		
5	Режим обработки информации в ИСПДн:			
	Многопользовательский			+
	Однопользовательский	+		

Продолжение таблицы

6	Разграничение прав доступа в ИСПДн:			
	Без разграничения			+
	С разграничением		+	
7	Наличие сегментирования ИС			
	Отсутствует			+
	Имеется		+	

Количественно потенциал нарушителя ИБ также может быть представлен путем проецирования метрик уязвимостей ПО [10] на возможности нарушителя ИБ [6].

ИНС является одной из реализаций систем искусственного интеллекта и представляет собой совокупность искусственных нейронов (ИН), объединённых синапсами (имеющими определенные «веса») и размещенных на определенном количестве слоев (входной, скрытый, выходной). Необходимо отметить, что количество скрытых слоев является произвольным и может быть выбрано в соответствии с решаемой задачей.

Основной теорией построения ИНС является представление об ИН как о логическом элементе, имеющим множество входных сигналов $x(t_i)$ и выходных сигналов $y(t_i+1)$ в момент времени, который работает по принципу «все или ничего» [1]. Данные сигналы аппроксимируются единичными импульсами прямоугольной формы или единичными потенциалами, которые могут быть представлены булевыми переменными $x(t_i)$, $y(t_i+1)$ в интервале от 0 до 1.

По результатам анализа для определения актуальных угроз ИБ с использованием математического аппарата ИНС предлагается использовать в качестве искусственных нейронов (ИН) входного слоя следующие показатели актуальности угрозы ИБ:

- уровень проектной защищенности;
- потенциал нарушителя ИБ;
- нарушение свойств защищенности, используемые при расчете возможного ущерба от реализации угрозы ИБ для ИСПДн.

На основании вышеизложенного параметры показателей актуальности угрозы ИБ на входном слое ИНС представляются как перечень следующих ИН:

- архитектура ИС (A);
- взаимодействие со сторонними ИСПДн (I_N);
- взаимодействие с сетями общего пользования (N_W);
- территориальное размещение ИСПДн (S_{Sq});
- режим обработки информации в ИСПДн (R_p);
- разграничение прав доступа (R_{ac});
- наличие сегментирования ИС (S_{IS});
- потенциал нарушителя (P_i);
- нарушение конфиденциальности (CI);
- нарушение целостности (II);
- нарушение доступности (AI).

Таким образом, общее количество ИН во входном слое ИНС составляет «11». На основании методик, приведенных в [4, 6], можно считать, что данный перечень является необходимым и достаточным для определения актуальности угроз ИБ в ИСПДн.

Выходным ИН будет являться величина актуальности угрозы ИБ (A_T).

Все значения для входных и выходных ИН будут принадлежать интервалу [0, 1].

В процессе выбора топологии был проведен анализ следующих типов ИНС:

- однослойная сеть;
- многослойная сеть;
- рекуррентная сеть;
- сеть латеральными связями;
- сеть с локальными связями.

Применение однослойной ИНС является невозможным в связи с представлением весовых коэффициентов ИН в виде матрицы, что ограничивает возможности ее обучения [1]. Также задача определения актуальности угроз ИБ в ИСПДн не является линейной. ИН выходного слоя рекуррентной ИНС подаются на ИН входного слоя и участвуют в обработке следующего входного век-

тора. Процесс определения угроз ИБ носит постоянный, но не динамический характер – поэтому применение рекуррентных ИНС нецелесообразно. ИНС с латеральными связями используют боковые связи ИН в слое и чаще всего применяются для построения сетей распознавания графических объектов [1]. ИНС с локальными связями ориентированы на обработку и восстановление изображений. По результатам выполненного обзора было принято решение использовать модель многослойной ИНС обратного распространения ошибки [1].

Для сравнения и выбора количества скрытых слоев ИНС экспериментальным путем созданы две ИНС. Первая – с учетом количества скрытых слоев ИНС представлена в [2], т.е. входным, двумя скрытыми и выходным слоем. Вторая ИНС – с одним скрытым слоем.

Количество ИН на каждом слое определяются по эвристическому правилу геометрической пирамиды. Суть данного правила заключается в сопоставлении слоев ИНС с формой пирамиды, т.е. количество ИН от входного к выходному слою должно уменьшаться в геометрической прогрессии.

Расчет количества ИН на скрытых слоях для первой ИНС осуществлялся по следующим формулам [11]:

$$r = \sqrt{\frac{n}{m}}, \quad (1)$$

$$k_1 = mr^2, \quad (2)$$

$$k_2 = mr, \quad (3)$$

где n – число ИН входного слоя; m – число ИН выходного слоя; k_1 – число ИН в первом скрытом слое; k_2 – число ИН во втором скрытом слое. Исходя из количества ИН входного слоя $n = 11$ и выходного слоя $m = 1$, $k_1 = 4,9$, округляется в большую сторону до целого числа 5, $k_2 = 2,22$, округляется в меньшую сторону до целого числа 2.

Расчет количества ИН в скрытом слое для второй ИНС осуществляется по формуле для трехслойной сети [11]:

$$r = \sqrt{nm}, \quad (4)$$

вследствие чего количество нейронов на скрытом слое $r = 3,31$, округляется в меньшую сторону до целого числа 3.

Для осуществления выбора топологии была разработана обучающая выборка, параметры уровня проектной защищенности и потенциал нарушителя ИБ (8 «единиц»), выбираются с привлечением эксперта. В связи с этим количество комбинаций ограничено. Параметры нарушения свойств защищенности (конфиденциальности, целостности, доступности) для подготовки обучающей выборки были выбраны из ранее подготовленного экспертным путем перечня угроз ИБ.

Проблемным вопросом при формировании обучающей выборки является отсутствие статистической информации в банке данных угроз ИБ.

Для решения данной проблемы проведен анализ статистической информации из реестра операторов, осуществляющих обработку персональных данных [8]. Установлено, что юридические лица в среднем регистрируют порядка трех ИСПДн в реестре, в то время как государственные и муниципальные учреждения – от 5 до 15 (с учетом учреждений здравоохранения).

На основании вышеизложенного была подготовлена обучающая выборка для усредненного количества ИСПДн – 5 шт. В нее была включена информация о параметрах нарушения свойств защищенности 8-ми различных угроз ИБ (рис. 1).

Тренировочный сет.

```
training_set_inputs = array([[0.2, 0.2, 0.6, 0.6, 0.6, 0.6, 0.6, 0.25, 0.9, 0.6, 0.6],
                             [0.6, 0.6, 0.6, 0.6, 0.6, 0.6, 0.6, 0.25, 0.9, 0.6, 0.6],
                             [0.2, 0.2, 0.6, 0.6, 0.6, 0.6, 0.6, 0.25, 0.2, 0.3, 0.4],
                             [0.8, 0.2, 0.8, 0.2, 0.6, 0.6, 0.6, 0.25, 0.4, 0.4, 0.2],
                             [0.2, 0.6, 0.8, 0.8, 0.6, 0.6, 0.6, 0.25, 0.9, 0.6, 0.6],
                             [0.2, 0.2, 0.6, 0.6, 0.6, 0.6, 0.6, 0.25, 0.4, 0.8, 0.3],
                             [0.6, 0.6, 0.6, 0.6, 0.6, 0.6, 0.6, 0.25, 0.2, 0.8, 0.3],
                             [0.2, 0.2, 0.6, 0.6, 0.6, 0.6, 0.6, 0.25, 0.9, 0.9, 0.4],
                             [0.8, 0.2, 0.8, 0.2, 0.6, 0.6, 0.6, 0.25, 0.2, 0.6, 0.1],
                             [0.2, 0.6, 0.8, 0.8, 0.6, 0.6, 0.6, 0.25, 0.3, 0.2, 0.3]])
training_set_outputs = array([[1, 1, 0, 0, 1, 1, 1, 1, 1, 0]])
```

Рисунок 1 – Обучающая выборка для ИНС

Для обучения ИНС используется скрипт, написанный на языке программирования Python (<https://github.com/miloharper/simple-neural-network/blob/master/main.py>).

При запуске скрипта через среду разработки Python – IDLE было установлено, что разница в вычислении ошибок между ИНС с одним скрытым слоем и ИНС с двумя скрытыми слоями, составляет менее 1 %.

Однако ИНС с одним скрытым слоем функционирует на 1 (одну) секунду быстрее, что является преимуществом при дальнейшем увеличении объемов обучающей выборки. В связи с отсутствием статистической информации, каждый новый случай предлагается добавлять в обучающую выборку и проводить переобучение ИНС.

Для упрощения обучения ИНС, а также осуществления вычислений, было принято решение использовать ИНС с одним скрытым слоем, состоящим из 3-х ИН.

Все значения ИН входного слоя предлагается использовать на интервале от 0 до 1 в соответствии с предлагаемой в [10] градацией.

После обучения ИНС с помощью вышеуказанного скрипта было проведено определение актуальности выбранной тестовой угрозы ИБ для смоделированной ИСПДн.

Проблемами, вытекающими из результатов функционирования ИНС, является зависимость результата ИНС от следующих факторов: 1) объема обучающей выборки; 2) от квалификации учителя, либо источника, который используется для сопоставления входных и выходных сигналов.

Решение вышеуказанных проблем может быть достигнуто посредством создания дополнительной площадки для банка данных угроз ИБ. Целью ее создания будет являться сбор и накопление статистической информации об угрозах ИБ в ИСПДн операторов и калькуляция актуальности угроз ИБ. Удобство данного способа будет заключаться в обезличенной передаче значений входных ИН для общего перечня угроз ИБ и получении выгрузки по актуальности выбранных оператором угроз ИБ также в обезличенной виде.

Визуализация функционирования ИНС при определении актуальности выбранной тестовой угрозы ИБ для смоделированной ИСПДн на основании представленной на рисунке 1 обучающей выборки представлена на рисунках 2 и 3. Точность функционирования ИНС вычисляется (оценивается) по результатам прохождения каждой итерации, и включает в себя: вычисление погрешности методом обратного распространения ошибки текущей итерации и корректировку весов слоев ИНС после прохождения итерации.

Для визуализации последовательного увеличения точности определения актуальности угроз ИБ в ИСПДн посредством ИНС на рисунках 2 и 3 на основании проведенного эксперимента, было принято решение отобразить 5950 итерацией (шагов). Данное количество итераций позволяет визуализировать вычисление ошибки ИНС. Однако дальнейшее отображение итераций делает график точности ИНС не информативным.

Обучение ИНС проводилось с помощью одной эпохи, состоящей из 10000 итераций. Данное количество было установлено экспериментально и является достаточным для достижения результата с приемлемой ошибкой. Также экспертным путем для контроля результатов функционирования ИНС была установлена градация актуальности угрозы ИБ в ИСПДн на уровне значения ИН выходного слоя равного 0,51. В рамках дальнейшей работы планируется провести декомпозицию ИН выходного слоя в соответствии с таблицей оценки уязвимостей ПО [10].

Скорость функционирования скрипта при определении актуальности тестовой угрозы ИБ для смоделированной ИСПДн на основании подготовленной обучающей выборки, составляет 2 (две) секунды. Наряду с этим, определение актуальности угрозы ИБ в ИСПДн по методикам, представленным в [6], а также с применением банка данных угроз ИБ требует создания рабочей группы и проведения анкетирования экспертов, что является достаточно трудоемким и продолжительным по времени процессом. По результатам проведенного эксперимента с привлечением группы экспертов в составе 5-ти человек, время, затраченное на определение актуальности тестовой угроз ИБ в смоделированной ИСПДн, равняется приблизительно 30 минутам.

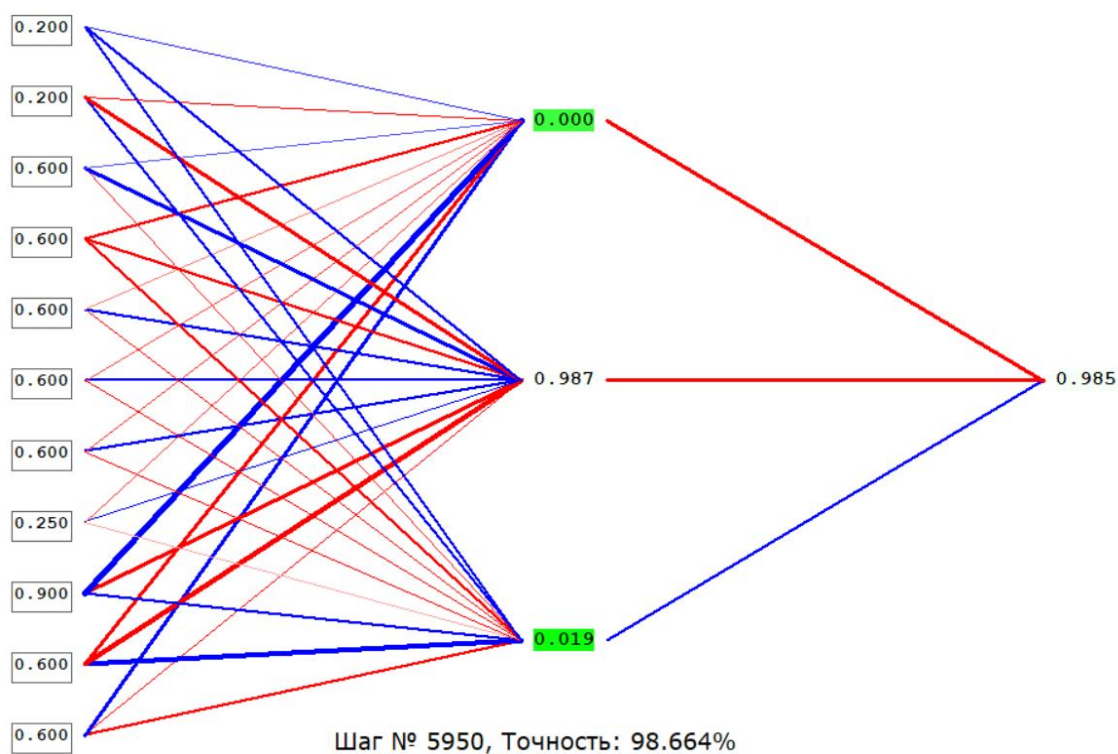


Рисунок 2 – Визуализация определения актуальности тестовой угрозы ИБ для смоделированной ИСПДн на основании обучающей выборки

Шаг № 5950, Точность: 98.664%



Рисунок 3 – Изменение точности применяемой ИНС в зависимости от количества шагов

На основании вышеизложенного делаем вывод, что применение математического аппарата ИНС позволит существенно сократить время, затрачиваемое на определение актуальности угрозы ИБ в ИСПДн по сравнению со способом основанном на экспертной оценке. Для расчета сравнительного эффекта сокращения времени используем следующую формулу:

$$T = T_{\Sigma} / T_{\text{инс}}, \quad (5)$$

где T_{Σ} – время, затраченное на определение актуальности угрозы ИБ в ИСПДн группой экспертов; а $T_{\text{инс}}$ – время, затраченное на определение актуальности угрозы ИБ в ИСПДн с помощью ИНС. Таким образом, в рамках конкретного эксперимента на основании времени функционирования скрипта, указанного выше, и времени, затрачиваемого группой экспертов, эффект применения ИНС представляет собой сокращение времени, затрачиваемого на определение актуальности тестовой угрозы ИБ в смоделированной ИСПДн и составляет 900 раз.

Учитывая предложенный подход к решению проблем накопления статистической информации об актуальных угрозах ИБ в ИСПДн и дополнению обучающей выборки с переобучением, математический аппарат ИНС позволит определять актуальность угроз ИБ в динамике изменения входных параметров, например, показателей защищенности.

Научной новизной представленного подхода, в том числе в сравнении с [2], является унификация состава ИН входного слоя; организация связи угроз ИБ с уязвимостями ИБ через значения нарушения свойств защищенности. Существенным отличием представленного метода от [2] является возможность применения одной ИНС для определения актуальности множества угроз ИБ в отличие от подхода, требующего формирования отдельных ИНС для отдельных угроз ИБ.

Заключение. Таким образом, используя параметры уровня проектной защищенности, потенциал нарушителя ИБ и нарушение свойств защищенности, можно разработать простую для реализации ИНС, с минимальным значением ошибки сети. Данная ИНС позволит минимизировать участие эксперта в процессе определения угроз ИБ в ИСПДн.

Библиографический список

1. Брюхомицкий Ю. А. Нейросетевые модели для систем информационной безопасности : учебное пособие / Ю. А. Брюхомицкий. – Таганрог : Изд-во ТРТУ, 2005. – 160 с.
2. Соловьев С. В. Применение аппарата нейронных сетей для определения актуальных угроз безопасности информации информационных систем / С. В. Соловьев, В. В. Мамута // Научные технологии в космических исследованиях Земли. – 2016 – Т. 8. № 5 – С.78-82.
3. Цветкова О. Л. О применении теории искусственных нейронных сетей в решении задач обеспечения информационной безопасности / О. Л. Цветкова, А. И. Крепер // Символ науки. – 2017. – № 04-2 – С. 105–107.
4. Методический документ ФСТЭК России «Базовая модель угроз безопасности персональных данных при их обработке в информационных системах персональных данных (выписка)» – 2008. – 69 с. – Режим доступа: <https://fstec.ru/component/attachments/download/289>, свободный. – Заглавие с экрана. – Яз. рус. (дата обращения: 27.12.2019).
5. Информационное сообщение ФСТЭК России от 6 марта 2015 г. № 240/22/879 «О банке данных угроз безопасности информации». – Режим доступа: <https://fstec.ru/normotvorcheskaya/informatsionnye-i-analiticheskie-materialy/956-informatsionnoe-soobshchenie-fstek-rossii-ot-6-marta-2015-g-240-22-879>, свободный. – Заглавие с экрана. – Яз. рус. (дата обращения: 24.12.2019).
6. Методический документ ФСТЭК России «Методика определения актуальных угроз безопасности персональных данных при их обработке в информационных системах персональных данных», – 2008. – 10 с. – Режим доступа: <https://fstec.ru/tekhnicheskaya-zashchita-informatsii/dokumenty/114-spetsialnye-normativnye-dokumenty/380>, свободный. – Заглавие с экрана. – Яз. рус. (дата обращения: 27.12.2019).
7. Методический документ ФСТЭК России «Методика определения угроз безопасности информации в информационных системах» (проект). – 2015. – 43 с. – Режим доступа: <https://fstec.ru/component/attachments/download/812>, свободный. – Заглавие с экрана. – Яз. рус. (дата обращения: 20.12.2019).
8. Реестр операторов, осуществляющих обработку персональных данных. – Режим доступа: <http://rkn.gov.ru/personal-data/register/>, свободный. – Заглавие с экрана. – Яз. рус. (дата обращения: 25.01.2020).
9. Сайт OWASP – Режим доступа: <https://owasp.org/www-project-top-ten/>, свободный. – Заглавие с экрана. – Яз. англ. (дата обращения: 25.01.2020).
10. Common Vulnerability Scoring System v3.0: Specification Document. – Режим доступа: <https://www.first.org/cvss/specification-document> свободный. – Заглавие с экрана. – Яз. англ. (дата обращения: 20.01.2020).
11. Masters T. Practical Neural Network Recipes in C++ / T. Masters. – Academic Press, 1993. – 504 p.

References

1. Bryukhomitskiy Yu. A. *Neurosetevye modeli dlya sistem informatsionnoy bezopasnosti* [Neural network models for information security systems]. Taganrog, Taganrog State Radioengineering University, 2005. 160 p.
2. Soloviev S. V., Mamuta V. V. *Primenenie apparata neyronnykh setey dlya opredeleniya aktualnykh ugroz bezopasnosti informatsii informatsionnykh sistem* [The use of the apparatus of neural networks to determine the current threats to the security of information of information systems]. *Nauchnye tekhnologii v kosmicheskikh issledovaniyakh zemli* [High Technology in Space Research of the Earth], 2016, vol. 8, issue 5, pp. 78–82.
3. Tsvetkova O. L., Kreper A. I. *O primenenii teorii neyronnykh setey v reshenii zadach obespecheniya informatsionnoy bezopasnosti* [On the application of the theory of artificial neural networks in solving problems of ensuring information security]. *Simvol nauki* [Symbol of Science], 2017, issue 04-2, pp. 105–107.
4. *Metodicheskiy dokument FSTEC Rossii "Bazovaya model ugroz bezopasnosti personalnykh dannykh pri ikh obrabotke v informatsionnykh sistemakh personalnykh dannykh (vypiska)"* [Guidance document of the FSTEC of Russia "The basic model of personal data security threats when they are processed in personal data information systems (extract)"], 2008. 69 p. Available at: <https://fstec.ru/component/attachments/download/289> (accessed 27.12.2019).
5. *Informatsionnoe pismo FSTEC Rossii ot 6 marta 2015 g. № 240/22/879 "O banke dannykh ugroz bezopasnosti informatsii"* [FSTEC information message no. 240/22/879 of March 6, 2015 «On the data bank of information security threats»]. Available at: <https://fstec.ru/normotvorcheskaya/informatsionnye-i-analiticheskie-materialy/956-informatsionnoe-soobshchenie-fstek-rossii-ot-6-marta-2015-g-240-22-879> (accessed 24.12.2019).
6. *Metodicheskiy dokument FSTEC Rossii "Metodika opredeleniya aktualnykh ugroz bezopasnosti personalnykh dannykh v informatsionnykh sistemakh personalnykh dannykh"* [Guidance document of the FSTEC of Russia "Methodology for determining current threats to the security of personal data when they are processed in personal data information systems"], 2008. 10 p. Available at: <https://fstec.ru/tekhnicheskaya-zashchita-informatsii/dokumenty/114-spetsialnye-normativnye-dokumenty/380> (accessed 27.12.2019).

7. *Metodicheskiy dokument FSTEC Rossii "Methodika opredeleniya ugroz bezopasnosti informatsii v informatsionnykh sistemakh" (proekt)* [Methods for determining information security threats in information systems (project)]. Available at: <https://fstec.ru/component/attachments/download/812> (accessed 20.12.2019).
8. *Reestr operatorov personalnykh dannykh* [Register of operators that process personal data]. Available at: <http://rkn.gov.ru/personal-data/register/> (accessed: 01/25/2020).
9. *OWASP website*. Available at: <https://owasp.org/www-project-top-ten/> (accessed: 01/25/2020).
10. *Common Vulnerability Scoring System v3.0: Specification Document*. Available at: <https://www.first.org/cvss/specification-document> (accessed 01.20.2020).
11. Masters T. *Practical Neural Network Recipes in C++*. Academic Press, 1993. 504 p.

ИНФОРМАЦИЯ О ПРЕДСТОЯЩИХ НАУЧНЫХ МЕРОПРИЯТИЯХ

Приведена информация о предстоящих научных мероприятиях: Международная научная конференция «**Математические методы в технике и технологиях – ММТТ-33**» (июнь-ноябрь 2020 г., Санкт-Петербург, Саратов, Минск); XXV Международный симпозиум «**Надежность и качество-2020**» (г. Пенза), май 2020 г.; **Метананобио-2020** (г. Саратов, май 2020 г.); **1st International conference on innovative informational and engineering technologies (ИИТ-2020)** – Ставрополь, 2019 г. и др.

Ключевые слова: научные мероприятия, тематика, направления, объявления, информационные технологии

1. XXXIII Международная научная конференция «**Математические методы в технике и технологиях – ММТТ-33**». 1–5 июня 2020 г., Казанский национальный исследовательский технологический университет (КНИТУ), Казань, Россия; 5–10 октября 2020 г., Калининградский государственный технический университет (КГТУ), Калининград, Россия; 14–16 октября 2020 г., Саратовский государственный технический университет имени Гагарина Ю.А., Саратов, Россия; 26–30 октября 2020 г., Белорусский национальный технический университет (БНТУ), Минск, Беларусь.

По рекомендациям рецензентов расширенные варианты поданных материалов с видоизменными наименованиями могут быть опубликованы в журналах из Перечня ВАК; в коллективной научной монографии с последующей индексацией ее в Scopus.

Подробная информация о конференции и условиях участия в ней размещается на сайте <http://mmtt.sstu.ru/>

2. XXV Международный симпозиум «**Надежность и качество-2020**», 25–31 мая 2020 г., г. Пенза, Пензенский государственный университет, e-mail: kipra@pnzgu.ru, nika@nika-penza.ru, тел./факс +7(841-2) 56-43-46, тел. +7(841-2) 36-82-12.

3. Международная заочная научно-техническая конференция молодых учёных, специалистов и студентов «**Инновационные технологии в проектировании и производстве**» с включением материалов в книгу «Труды международного симпозиума «**Надежность и качество**»». Сроки проведения: с 10 по 15 мая 2020 г. Организатор: Пензенский государственный университет. Обращаться: Николай Кондратьевич Юрков, www.dpr_kipra.pnzgu.ru, www.nika-penza.ru.

4. **Метананобио-2020**. Седьмая Всероссийская научная школа-семинар «**Взаимодействие сверхвысокочастотного, терагерцового и оптического излучения с полупроводниковыми микро- и наноструктурами, метаматериалами и биообъектами**». Сроки проведения: – с 19 по 20 мая 2020 г. на базе Саратовского государственного университета имени Н.Г. Чернышевского. Сайт: <http://www.metananobio.ru/>. Страница на сайте Саратовского государственного университета: <https://www.sgu.ru/structure/fnbmt/news/2020-03-07/metananobio-2020>. Наименования тематических секций: полупроводниковая микро- и наноэлектроника; Фотонные кристаллы и метаматериалы; медико-биологические системы и технологии; твердотельная СВЧ-электроника; терагерцовые технологии. Образец оформления представляемых материалов по адресу: <http://www.metananobio.ru/docs/forma.doc>.

ПРАВИЛА ДЛЯ АВТОРОВ

1. В журнале публикуются материалы на английском и русском языках по тематике, соответствующей утвержденным для журнала отраслям наук, группам специальностей.

2. В список соавторов работ включаются только те лица, которые внесли творческий вклад в подготовку представленных материалов. Лицам, оказавшим только техническую помощь, можно выразить благодарность в конце статьи. Один человек может быть автором (соавтором) не более чем двух статей в одном номере журнала, причем единственным автором он может быть только в одной статье.

3. Объем публикаций для научных статей должен быть не менее 8 страниц, а количество источников в библиографическом списке (списке литературы) – не менее 15 позиций.

4. **Содержание** каждой статьи должно включать следующие элементы: УДК; название статьи; сведения об авторах, включая их место работы, должность, адрес электронной почты; аннотацию объемом от 100 до 250 слов, ключевые слова (от 9 до 13); графическую аннотацию, отражающую содержание статьи; название статьи, сведения об авторах, аннотацию и ключевые слова на английском языке (для англоязычных статей – на русском языке); введение – оно должно заканчиваться формулировкой цели работы в явной форме; собственно текст статьи – очень желательна его сегментация на разделы, имеющие содержательные заголовки; выводы или заключение (должны соответствовать формулировке цели статьи).

5. Для русскоязычных статей приводится два библиографических списка: на языке оригинала статьи; список с транслитерацией русскоязычных источников на латиницу и (дополнительно) приведением в квадратных скобках переводов названий статей и названий источников на английский язык.

В «русскоязычном» библиографическом списке (списке литературы) порядок следования источников – по алфавиту фамилий авторов (сначала русскоязычные источники, потом иноязычные). На все источники, включенные в библиографический список, должны быть даны ссылки в тексте статьи в квадратных скобках. При необходимости авторы могут указывать номера страниц в источниках, на которые даются ссылки. Приветствуются ссылки на иноязычные источники, а также на материалы, опубликованные ранее в журнале «Прикаспийский журнал: управление и высокие технологии». Однако в последнем случае количество таких ссылок не должно превышать 20 % от общего количества источников, включенных в библиографический список. Для источников, имеющих DOI, целесообразно его указывать. При ссылках на статьи, опубликованные в журнале «Прикаспийский журнал: управление и высокие технологии», целесообразно в конце библиографического описания источника в круглых скобках указывать гиперссылку, указывающую на место размещения статьи на странице сайта Астраханского государственного университета.

Ссылки в библиографическом списке на материалы, размещенные в интернете, допускаются при соблюдении следующих условий: если у материала, на который дается ссылка, имеется автор и/или название, то они должны быть указаны для этого источника; должен быть приведен полный маршрут доступа к источнику в интернете; должна быть указана дата обращения (доступа) к источнику.

Ограничения по списку литературы: доля самоцитирований для любого из авторов статьи, а также по совокупности всех авторов статьи, не должна превышать 25 %; доля ссылок на статьи с участием одного автора, не являющегося автором (соавтором) статьи, не должна превышать 25 %.

6. Суммарная доля таблиц и иллюстраций в общем объеме представляемой статьи не должна превышать 40 %. Под иллюстрациями понимаются следующие объекты: диаграммы; графики; рисунки; эскизы; фотографии; карты и т.п.

7. Доля оригинального текста в статьях (оцениваемого через систему «Антиплагиат» на сайте www.antiplagiat.ru) должна быть не менее 80 %.

8. Указание на то, что работа финансируется по какому-либо гранту, в рамках Федеральной целевой программы, государственного заказа и пр. дается в виде постраничной сноски после заголовка (названия) работы.

9. В сведения об авторах работ помимо места работы и должности целесообразно включать ORCID автора и гиперссылку на страничку с его личными наукометрическими показателями на сайте www.elibrary.ru. По желанию можно привести также ссылки на странички с наукометрическими показателями на Scopus, в ResearchGate; на личную страничку, размещенную на сайте организации.

10. Основные технические требования к оформлению статей (материалов):

10.1. Текст должен быть расположен по ширине страницы формата А4 с учётом полей (все поля по 2,5 см), набран шрифтом Times New Roman, кегль 14, межстрочный интервал 1.0. В таблицах, подрисуночных надписях допускается уменьшенный шрифт – вплоть до 10 кегля. Альбомная ориентация страниц допускается только в порядке исключения для следующих случаев: широкоформатные таблицы с большим количеством колонок; иллюстрации большого размера, которые не умещаются на странице с книжной ориентацией.

Абзацные отступы одинаковы по всему тексту – 1,25 см. Кавычки («»»), скобки ([], ()), маркеры и другие знаки должны быть аналогичными на протяжении всего предоставляемого для публикации материала.

10.2. Все таблицы, рисунки, формулы должны иметь сквозную нумерацию в пределах текста статьи. Заголовки таблиц пишутся над ними и должны включать в себя номер таблицы и ее содержательное наименование.

10.3. Формулы должны иметь сквозную нумерацию в пределах текста статьи. Для формул желательно избегать «многоэтажных конструкций». Нумероваться могут только те формулы, на которые есть ссылки в статье. Размеры шрифтов для формул в MS Equation Editor: основные символы – 14 пт.; подстрочные и надстрочные индексы – 10 пт.; дополнительные индексы для подстрочных и надстрочных индексов – 8 пунктов. Константы в формулах записываются прямым шрифтом, переменные – курсивом (наклонным шрифтом); векторные и матричные величины – полужирным шрифтом. Матричные величины могут указываться в квадратных скобках, векторные величины (наборы скаляров) – в фигурных скобках.

Дополнительная информация для авторов статей

Основная тематика журнала и требования к научному уровню представляемых статей – <http://hi-tech.asu.edu.ru/docs/ru-rules/trebvaniya.pdf>.

Адрес и телефоны редакционно-издательского дома – <http://hi-tech.asu.edu.ru/docs/ru-rules/address.pdf>.

Пример оформления текста статьи – <http://hi-tech.asu.edu.ru/docs/ru-rules/primer.pdf>.

Рекомендации по оформлению графических аннотаций – http://hi-tech.asu.edu.ru/docs/ru-rules/graf_ann.pdf, дополнительных материалов к статьям – http://hi-tech.asu.edu.ru/docs/ru-rules/dop_mat.pdf.

Подробные правила оформления таблиц – <http://hi-tech.asu.edu.ru/docs/ru-rules/tables.pdf>, оформления формул – <http://hi-tech.asu.edu.ru/docs/ru-rules/formul.pdf>, оформления иллюстраций – <http://hi-tech.asu.edu.ru/docs/ru-rules/ilustr.pdf>.

Пример оформления основного списка литературы к статье – http://hi-tech.asu.edu.ru/docs/ru-rules/spisok_liter.pdf, транслитерированного списка – http://hi-tech.asu.edu.ru/docs/ru-rules/trans_spisok_liter.pdf.

Дополнительные документы к статьям, представляемые авторами – http://hi-tech.asu.edu.ru/docs/ru-rules/dop_doc.pdf.

Переписка с авторами и порядок рецензирования статей – http://hi-tech.asu.edu.ru/docs/ru-rules/poryadok_recenz.pdf.

Условия и порядок оплаты статей – http://hi-tech.asu.edu.ru/docs/ru-rules/poryadok_oplaty.pdf.

RULES FOR THE AUTHORS – <http://hi-tech.asu.edu.ru/docs/en-rules/rules-for-authors.pdf>.

**Подписка на наши издания осуществляется
по Объединённому каталогу «Пресса России»**

Журнал фундаментальных и прикладных исследований «Гуманитарные исследования»

Подписной индекс – 11109

В журнале публикуются статьи по широкому спектру проблем гуманитарного знания. Ведущие направления публикаций отражены в следующих рубриках: «Языкознание», «Литературоведение».

Периодичность издания – 4 раза в год.

Ориентировочная стоимость одного номера – 3762 руб.

Телефон: (8512) 24-64-95. E-mail: asupress@yandex.ru

Научно-технический журнал «Геология, география и глобальная энергия»

Подписной индекс – 11173

Редколлегия журнала принимает к рассмотрению статьи по проблемам геологии, нефтегазоносности различных регионов, охватывающие важнейшие и крайне полезные для науки и производства, а также для обучения студентов естественного направления.

Периодичность издания – 4 раза в год.

Ориентировочная стоимость одного номера – 4500 руб.

Телефон: (8512) 24-64-95. E-mail: asupress@yandex.ru

Научный журнал «Каспийский регион: политика, экономика, культура»

Подписной индекс – 11170

Профиль журнала – анализ проблем настоящего, прошлого и будущего Каспийского региона в их взаимосвязи с современным развитием мира.

Издание имеет многоплановый, междисциплинарный характер, знакомит читателя с исследованиями и дискуссиями во всех областях социальных и гуманитарных знаний по проблемам Каспийского региона.

Периодичность издания – 4 раза в год.

Ориентировочная стоимость одного номера – 3500 руб.

Телефон: (8512) 24-64-95. E-mail: asupress@yandex.ru

Научно-технический журнал «Прикаспийский журнал: управление и высокие технологии»

Подписной индекс – 73313

На страницах журнала представлены результаты исследований и новейшие разработки в области технических наук.

Ведущие направления публикаций отражены в следующих рубриках: «Управление в социальных и экономических системах», «Системный анализ, управление и обработка информации», «Математическое моделирование, численные методы и комплексы программ», «Информационная безопасность и защита информации», «Информационно-измерительные и управляющие системы».

Периодичность издания – 4 раза в год.

Ориентировочная стоимость одного номера – 9000 руб.

Телефон: (8512) 24-64-95. E-mail: asupress@yandex.ru

Предлагаем всем желающим разместить в наших изданиях рекламу.
Адрес Издательского дома «Астраханский университет»:
414056, г. Астрахань, ул. Татищева, 20а; тел. (8512) 24-64-95, 24-68-37
e-mail: asupress@yandex.ru

ПРИКАСПИЙСКИЙ ЖУРНАЛ: управление и высокие технологии

НАУЧНО-ТЕХНИЧЕСКИЙ ЖУРНАЛ

**2020
№ 1 (49)**

Свидетельство о регистрации средства массовой информации
Федеральной службы по надзору в сфере массовых коммуникаций,
связи и охраны культурного наследия
ПИ № ФС77-31932 от 16 мая 2008 г.

Учредитель
Астраханский государственный университет
Российская Федерация, 414056, г. Астрахань, ул. Татищева, 20а

Адрес редакции:
Российская Федерация, 414056, г. Астрахань, ул. Татищева, 20

Адрес издателя:
Российская Федерация, 414056, г. Астрахань, ул. Татищева, 20а

Подписной индекс – 73313
По Объединенному каталогу «Пресса России»

Техническое редактирование,
компьютерная правка, верстка *Н.Н. Сахно*

Дата выхода в свет 03.04.2020 г.

Цена свободная
Уч.-изд. 14,3. Усл. печ. л. 20,0.
Заказ № 4146. Тираж 500 экз. (первый завод – 22 экз.)

Оттиражировано на базе Издательского дома «Астраханский университет»
414056, г. Астрахань, ул. Татищева, 20а
Тел. (8512) 24-64-95, тел./факс (8512) 24-68-37
E-mail: asupress@yandex.ru