

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

ИНФОРМАЦИОННО- УПРАВЛЯЮЩИЕ СИСТЕМЫ

НАУЧНЫЙ ЖУРНАЛ

5(132)/2024

5(132)/2024

PEER REVIEWED JOURNAL

INFORMATSIONNO- UPRAVLIAYUSHCHIE SISTEMY (INFORMATION AND CONTROL SYSTEMS)

Founder

A. Vostrikov

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

E. Krouk

Dr. Sc., Professor, Moscow, Russia

Executive secretary

O. Muravtsova

Editorial Board

S. Andreev

Dr. Sc., Tampere, Finland

V. Anisimov

Dr. Sc., Professor, Saint Petersburg, Russia

B. Bezruchko

Dr. Sc., Professor, Saratov, Russia

N. Blaunstein

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

M. Buzdalov,

PhD, Researcher, Saint Petersburg, Russia

C. Christodoulou

PhD, Professor, Albuquerque, New Mexico, USA

A. Dudin

Dr. Sc., Professor, Minsk, Belarus

I. Dumer

PhD, Professor, Riverside, USA

M. Favorskaya

Dr. Sc., Professor, Krasnoyarsk, Russia

L. Fortuna

PhD, Professor, Catania, Italy

A. Fradkov

Dr. Sc., Professor, Saint Petersburg, Russia

A. Hramov

Dr. Sc., Professor, Innopolis, Russia

L. Jain

PhD, Professor, Canberra, Australia

G. Matvienko

Dr. Sc., Professor, Tomsk, Russia

A. Myllari

PhD, Professor, Grenada, West Indies

K. Samouylov

Dr. Sc., Professor, Moscow, Russia

J. Seberry

PhD, Professor, Wollongong, Australia

M. Sergeev

Dr. Sc., Professor, Saint Petersburg, Russia

A. Shalyto

Dr. Sc., Professor, Saint Petersburg, Russia

A. Shepeta

Dr. Sc., Professor, Saint Petersburg, Russia

Yu. Shokin

RAS Academician, Dr. Sc., Novosibirsk, Russia

A. Smirnov

Dr. Sc., Professor, Saint Petersburg, Russia

T. Sutikno

PhD, Associate Professor, Yogyakarta, Indonesia

Z. Yuldashev

Dr. Sc., Professor, Saint Petersburg, Russia

A. Zeifman

Dr. Sc., Professor, Vologda, Russia

Editor: A. Larionova**Proofreader:** T. Zvertanovskaia**Design:** M. Chernenko, Yu. Umnitsyna**Layout and composition:** Yu. Umnitsyna**Contact information**The Editorial and Publishing Center, SUAI
67A, Bol'shaya Morskaya, 190000, Saint Petersburg, Russia
Website: <http://i-us.ru/en>, e-mail: i-us.spb@gmail.com
Tel.: +7 - 812 494 70 02**INFORMATION PROCESSING AND CONTROL****Kapusta K. L., Kipyatkova I. S., Kagirov I. A.** Analytical survey
of transformer-based end-to-end speech recognition models
and strategies 2**SYSTEM AND PROCESS MODELING****Akimenko T. A., Larkin E. V., Bogomolov A. V., Privalov A. N.** Model of
scene image formation in the television module of a mobile robot
moving over rough terrain 16**Kolesnikova S. I., Fomenkova A. A.** Model of the stabilization system
of a stochastic bioengineering object on a set of target states
and synergetic control theory 25**INFORMATION SECURITY****Bashmakov N. M., Vasilyev V. I., Vulfin A. M., Kartak V. M., Kirillova A. D.**
Detection of network botnet attacks based on machine learning
and knowledge transfer technologies 41**Novikova E. S., Kuznetsova E. O., Golubev S. A.** Network intrusion
detection based on convolutional neural networks in industrial
cyber-physical systems 57**INFORMATION CODING AND TRANSMISSION****Taubin F. A., Trofimov A. N.** Product codes with semi-blind decoding
for single carrier transmission over multipath fading channel 68**INFORMATION ABOUT THE AUTHORS**

81

5(132)/2024

РЕЦЕНЗИРУЕМОЕ ИЗДАНИЕ

ИНФОРМАЦИОННО-
УПРАВЛЯЮЩИЕ
СИСТЕМЫ

Учредитель

А. А. Востриков

Издатель

Санкт-Петербургский государственный университет
аэрокосмического приборостроения

Главный редактор

Е. А. Крук,

д-р техн. наук, проф., Москва, РФ

Ответственный секретарь

О. В. Муравцова

Редакционная коллегия:

С. Д. Андреев,

д-р техн. наук, Тампере, Финляндия

В. Г. Анисимов,

д-р техн. наук, проф., Санкт-Петербург, РФ

Б. П. Безручко,

д-р физ.-мат. наук, проф., Саратов, РФ

Н. Блаунштейн,

д-р физ.-мат. наук, проф., Беэр-Шева, Израиль

М. В. Буэдалов,

канд. техн. наук, научный сотрудник, Санкт-Петербург, РФ

Л. С. Джайн,

д-р наук, проф., Канберра, Австралия

А. Н. Дудин,

д-р физ.-мат. наук, проф., Минск, Беларусь

И. И. Думер,

д-р наук, проф., Риверсайд, США

А. И. Зейфман,

д-р физ.-мат. наук, проф., Вологда, РФ

К. Кристоделу,

д-р наук, проф., Альбукерке, Нью-Мексико, США

Г. Г. Матвиенко,

д-р физ.-мат. наук, проф., Томск, РФ

А. А. Мюллери,

д-р наук, профессор, Гренада, Вест-Индия

К. Е. Самуилов,

д-р техн. наук, проф., Москва, РФ

Д. Себерри,

д-р наук, проф., Волонгонг, Австралия

М. Б. Сергеев,

д-р техн. наук, проф., Санкт-Петербург, РФ

А. В. Смирнов,

д-р техн. наук, проф., Санкт-Петербург, РФ

Т. Сутикну,

д-р наук, доцент, Джокьякарта, Индонезия

М. Н. Фаворская,

д-р техн. наук, проф., Красноярск, РФ

Л. Фортуна,

д-р наук, проф., Катания, Италия

А. Л. Фрадков,

д-р техн. наук, проф., Санкт-Петербург, РФ

А. Е. Храмов,

д-р физ.-мат. наук, Иннополис, РФ

А. А. Шальто,

д-р техн. наук, проф., Санкт-Петербург, РФ

А. П. Шепета,

д-р техн. наук, проф., Санкт-Петербург, РФ

Ю. И. Шокин,

акад. РАН, д-р физ.-мат. наук, проф., Новосибирск, РФ

З. М. Юлдашев,

д-р техн. наук, проф., Санкт-Петербург, РФ

Редактор: А. Г. Ларионова

Корректор: Т. В. Звертановская

Дизайн: М. Л. Черненко, Ю. В. Умницына

Компьютерная верстка: Ю. В. Умницына

Адрес редакции: 190000, г. Санкт-Петербург,

ул. Большая Морская, д. 67, лит. А, ГУАП, РИЦ

Тел.: (812) 494-70-02, эл. адрес: ius.spb@gmail.com,

сайт: <http://i-us.ru>

ОБРАБОТКА ИНФОРМАЦИИ И УПРАВЛЕНИЕ

Капуста К. Л., Кипяtkова И. С., Кагиров И. А. Аналитический обзор интегральных моделей и стратегий распознавания речи на основе архитектуры трансформера

2

МОДЕЛИРОВАНИЕ СИСТЕМ И ПРОЦЕССОВ

Акименко Т. А., Ларкин Е. В., Богомолов А. В., Привалов А. Н. Модель формирования образа сцены в телевизионном модуле мобильного робота, движущегося по пересеченной местности

16

Колесникова С. И., Фоменкова А. А. Модель системы стабилизации стохастического биоинженерного объекта на множестве целевых состояний и синергетической теории управления

25

ЗАЩИТА ИНФОРМАЦИИ

Башмаков Н. М., Васильев В. И., Вульфин А. М., Картак В. М., Кириллова А. Д. Обнаружение сетевых атак ботнетов на основе технологий машинного обучения и переноса знаний

41

Новикова Е. С., Кузнецова Е. О., Голубев С. А. Выявление сетевых вторжений в промышленных киберфизических системах на основе сверточных нейронных сетей

57

КОДИРОВАНИЕ И ПЕРЕДАЧА ИНФОРМАЦИИ

Таубин Ф. А., Трофимов А. Н. Коды-произведения с полуслепым декодированием для одночастотной передачи в многолучевом канале с замираниями

68

СВЕДЕНИЯ ОБ АВТОРАХ

81

Журнал входит в БД Scopus и в Перечень рецензируемых научных изданий, в которых должны быть опубликованы основные научные результаты диссертаций на соискание ученой степени кандидата наук, на соискание ученой степени доктора наук.

Сдано в набор 10.09.24. Подписано в печать 12.11.24. Дата выхода в свет: 15.11.2024.

Формат 60×841/8. Гарнитура CentSchbkCyrill BT. Печать цифровая.

Усл. печ. л. 10.0. Уч.-изд. л. 13.5. Тираж 1000 экз (1-й завод 50 экз.). Заказ № 414.

Оригинал-макет изготовлен в редакционно-издательском центре ГУАП.

190000, г. Санкт-Петербург, ул. Большая Морская, д. 67, лит. А.

Отпечатано в редакционно-издательском центре ГУАП.

190000, г. Санкт-Петербург, ул. Большая Морская, д. 67, лит. А.

Распространяется бесплатно.

Журнал зарегистрирован в Министерстве РФ по делам печати, телерадиовещания и средств массовых коммуникаций. Свидетельство о регистрации ПИ № 77-12412 от 19 апреля 2002 г. Перерегистрирован в Роскомнадзоре. Свидетельство о регистрации ПИ № ФС77-82226 от 23 ноября 2021 г.

© А. А. Востриков, 2024



Аналитический обзор интегральных моделей и стратегий распознавания речи на основе архитектуры трансформер

К. Л. Капуста^а, программист, orcid.org/0009-0008-8623-0101

И. С. Кипяткова^а, канд. техн. наук, доцент, старший научный сотрудник, orcid.org/0000-0002-1264-4458, kipyatkova@iiias.spb.su

И. А. Кагиров^а, научный сотрудник, orcid.org/0000-0003-1196-1117

^аСанкт-Петербургский Федеральный исследовательский центр РАН, 14-я линия В. О., 39, Санкт-Петербург, 199178, РФ

Введение: одной из тенденций в области распознавания естественных языков является переход от модульных архитектур к интегральным моделям. Эти системы объединяют различные этапы обработки, такие как акустическое, языковое и лексическое моделирование и декодирование, в единую архитектуру. Среди современных архитектур, наиболее часто используемых для интегрального распознавания речи, находится архитектура трансформер, а также ее модификации. **Цель:** выполнить подробный обзор моделей интегрального распознавания речи на базе архитектуры трансформер. **Результаты:** анализ различных стратегий декодирования позволил сделать ряд выводов. Так, коннекционная временная классификация эффективна при отсутствии выравнивания между речевым сигналом и текстовыми транскрипциями, но ее применение не рационально, если длина входных данных меньше длины выходных. Основным недостатком моделей, работающих по стратегии коннекционной временной классификации, является предположение о независимости выходных символов. Гораздо перспективнее оказываются трансдюсеры, учитывающие предшествующий контекст для каждого выходного символа, и шифраторы-дешифраторы с механизмом внимания, позволяющие учитывать долгосрочные зависимости и контекст. Обратной стороной последней стратегии является невысокая скорость, что ограничивает ее использование в реальном времени. Каждая из рассмотренных в статье стратегий, таким образом, имеет свои достоинства, но лучше всего проявляет себя с задачами конкретного типа. **Практическая значимость:** представленный обзор рассматривается как вклад в изучение быстроразвивающейся области интегрального распознавания речи независимо от конкретных естественных языков. Полученные выводы могут найти практическое применение при создании систем автоматического распознавания речи на естественных языках, в том числе и на малоресурсных языках. **Обсуждение:** существующая тенденция к увеличению размера моделей делает наиболее перспективными гибридные решения, учитывающие необходимость использования систем распознавания речи в режиме реального времени и требующие меньших вычислительных ресурсов.

Ключевые слова – интегральные модели, трансформер, трансдюсер, декодирование, автоматическое распознавание речи.

Для цитирования: Капуста К. Л., Кипяткова И. С., Кагиров И. А. Аналитический обзор интегральных моделей и стратегий распознавания речи на основе архитектуры трансформер. *Информационно-управляющие системы*, 2024, № 5, с. 2–15. doi:10.31799/1684-8853-2024-5-2-15, EDN: MWTGXE

For citation: Kapusta K. L., Kipyatkova I. S., Kagirov I. A. Analytical survey of transformer-based end-to-end speech recognition models and strategies. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2024, no. 5, pp. 2–15 (In Russian). doi:10.31799/1684-8853-2024-5-2-15, EDN: MWTGXE

Введение

За прошедшее десятилетие системы автоматического распознавания речи (САРР) прошли большой путь от классических архитектур до современного интегрального (англ. *end-to-end*) подхода [1]. Внедрение глубоких искусственных нейронных сетей в САРР привело к значительному приросту точности распознавания [2]. По результатам сравнительных исследований [3], применение таких сетей улучшило качество распознавания более чем на 50 % по метрике Word Error Rate (WER). Следующим этапом стало появление САРР, реализованных в рамках одной глубокой нейронной сети, выполняющей роль как акустической, так и языковой модели. Именно такие системы принято называть интегральными [4].

Следует заметить, что, несмотря на общий прирост точности, результаты интегрального распознавания могут значительно варьироваться в зависимости от условий работы САРР и особенностей целевого языка. Например, системы, обученные на ограниченном объеме данных, могут показывать недостаточную точность при распознавании речи в шумных средах, а также при наличии у диктора нестандартного выговора или акцента [5, 6]. Важно понимать, что и лингвистические особенности целевого языка тоже оказывают прямое влияние на работу САРР. Так, результаты работы одной и той же мультязычной системы, обученной на сопоставимом объеме данных для двух разных языков, могут различаться по точности в несколько раз в зависимости от исходного языка. Например, в работе [7]

при тестировании мультязычной модели на корпусе Fleurs [8] был получен результат 10,3 % WER для вьетнамского языка и 27,1 % для иврита при том, что объем обучающих корпусов составил 691 и 688 часов речи соответственно, т. е. ключевым фактором, повлиявшим на точность системы распознавания, оказались лингвистические особенности иврита (при практически аналогичном объеме вьетнамского набора данных точность его распознавания гораздо выше).

В настоящее время одной из наиболее широко используемых архитектур, применяемых для интегрального распознавания речи, является архитектура трансформер. В этой статье представлен обзор основных разновидностей трансформеров, а также связанных с ними различных стратегий декодирования, применяемых для оптимизации работы САРР.

Основные принципы интегрального распознавания речи

Задачу автоматического распознавания речи можно сформулировать как поиск наиболее вероятной последовательности слов по входному звуковому сигналу [9]. Основными составными частями «традиционной» САРР являются блоки, отвечающие за выделение признаков, акустическое моделирование, языковое моделирование, лексикон и декодирование. При интегральном подходе все вышеперечисленные модули представлены в рамках одной искусственной нейронной сети, которая реализует их совместное обучение [10].

Важной особенностью интегральных САРР является наличие нейросетевых механизмов кодера (или шифратора; англ. *encoder*) и декодера (или дешифратора; англ. *decoder*) [4]. Задача кодера в САРР состоит в преобразовании входной последовательности речевых данных в промежуточное представление признаков. Декодер в свою очередь преобразует промежуточное представление в выходную последовательность графем или лексем (слов) [4, 11].

Интегральная система автоматического распознавания речи работает по следующему принципу: на вход подается необработанный звуковой сигнал, который далее подвергается обработке для извлечения акустических признаков, например признаков, называемых *filterbank* (т. е. признаков, полученных с помощью гребенки полосовых фильтров) [12] или мел-кепстральных частотных коэффициентов (англ. MFCC) [13]. Также на вход модели могут подаваться заранее извлеченные акустические признаки. На этапе шифрования модель получает промежуточное представление признаков из входной последова-

тельности. На этапе декодирования из промежуточного представления строится выходная последовательность в виде готовой транскрипции. Опционально при декодировании может быть интегрирована внешняя языковая модель для улучшения результатов декодирования за счет охвата большего языкового контекста [14].

Трансформеры и их применение в САРР

На сегодня архитектура трансформер является одной из самых распространенных и эффективных архитектур для задач обработки естественного языка, в том числе и для задачи распознавания речи. Впервые она была представлена в статье “Attention is all you need” в 2017 г. [15].

Архитектура трансформер состоит из двух основных блоков: кодера и декодера. Блок кодера преобразует входную последовательность в скрытое представление, а декодер генерирует выходную последовательность на основе скрытого представления и предыдущих элементов выходной последовательности (рис. 1).

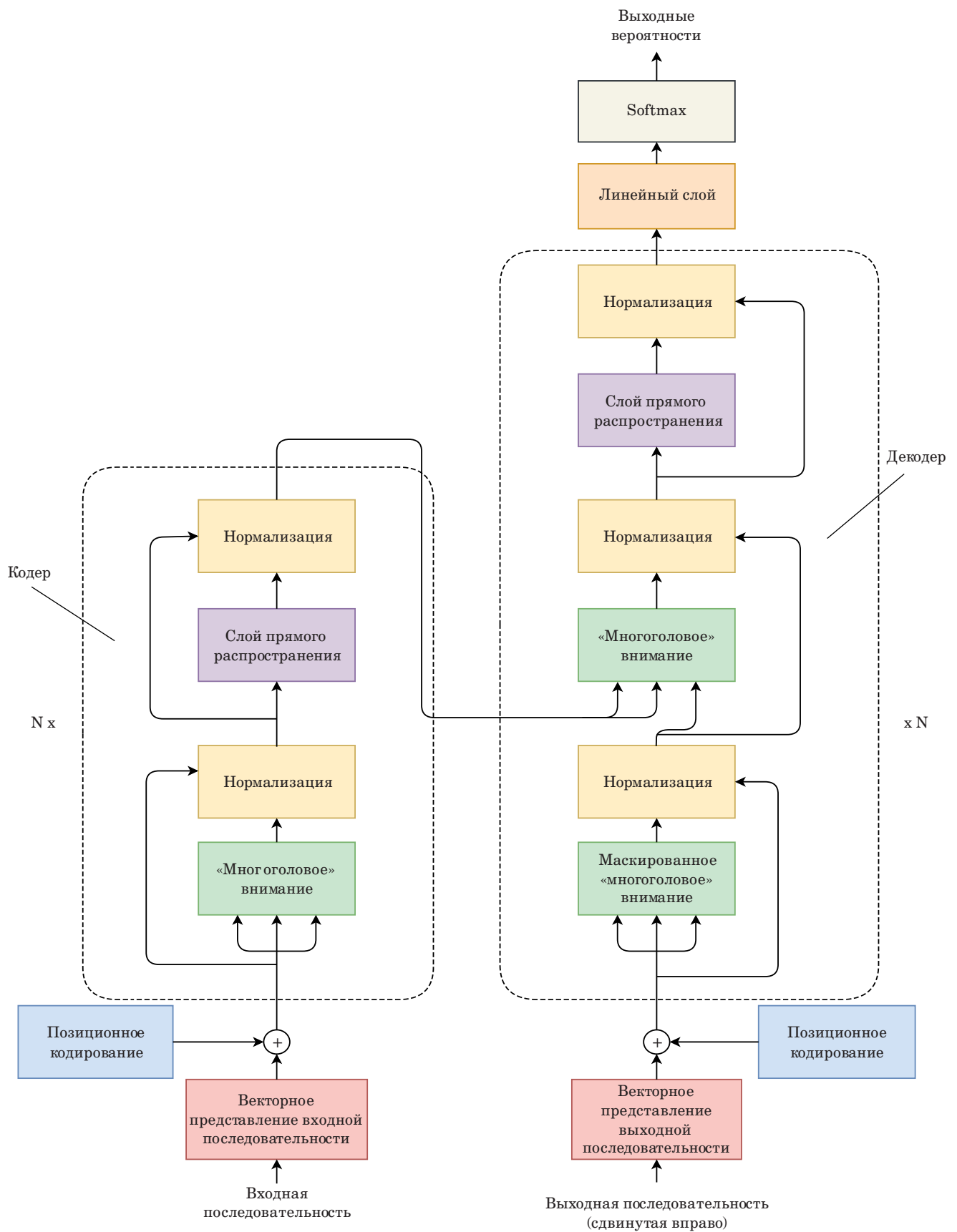
Основным компонентом архитектуры трансформер является механизм внимания (англ. *attention mechanism*), который позволяет модели учитывать все элементы входной последовательности при вычислении каждого элемента выходной последовательности. В отличие от рекуррентных нейронных сетей (англ. *recurrent neural network*, RNN), которые обрабатывают элементы последовательности одну за другой, трансформер может параллельно обрабатывать сразу всю последовательность. В работе [15] механизм внимания определяется как сопоставление вектора запроса (англ. *query vector*), вектора ключа (англ. *key vector*) и вектора значения (англ. *value vector*) с выходными данными. Механизм внимания реализуется следующим образом. Входные данные (векторные представления слов) преобразуются в три вектора — запроса, ключа и значения:

$$\mathbf{q}_i = \mathbf{W}_q \mathbf{x}_i, \quad \mathbf{k}_i = \mathbf{W}_k \mathbf{x}_i, \quad \mathbf{v}_i = \mathbf{W}_v \mathbf{x}_i,$$

где \mathbf{q}_i , \mathbf{k}_i , \mathbf{v}_i — векторы запросов, ключей и значений соответственно; \mathbf{W}_q , \mathbf{W}_k , \mathbf{W}_v — весовые матрицы, которые определяются в ходе обучения; \mathbf{x}_i — векторное представление i -го слова. Внимание, называемое вниманием на основе масштабированного скалярного произведения, вычисляется следующим образом:

$$\alpha = \text{softmax} \left(\frac{\mathbf{QK}^T}{\sqrt{d_k}} \right) \mathbf{V},$$

где d_k — размерность ключа.



■ **Рис. 1.** Архитектура трансформер [15]
 ■ **Fig. 1.** Transformer architecture [15]

В архитектуре трансформера используется множество наборов матриц запроса, ключа и значения. Каждый набор матриц называется «головой» (англ. *head*), поэтому такой тип внимания получил название «многоголовый» (англ. *multi-head attention*). Каждая голова может фокусироваться на разных аспектах входной последовательности, позволяя модели захватывать более сложные зависимости между элементами данных. Для каждой головы вычисляются свои векторы запроса, ключа и значения, на основе которых определяется внимание. Выходы всех голов объединяются путем конкатенации в одну матрицу:

$$\alpha_{MH} = \text{Concat}(\alpha_1, \alpha_2, \dots, \alpha_M) \mathbf{W}_o,$$

где M — число векторов внимания; \mathbf{W}_o — матрица весов.

В трансформере используются три типа «многоголового» внимания.

1. Кросс-внимание (англ. *cross-attention*). Данный тип внимания используется в декодере, при этом запросы поступают из декодера, а ключи и значения — из кодера. Это позволяет декодеру фокусироваться на релевантных элементах входной последовательности для генерации элемента выходной последовательности.

2. Самовнимание. Используется в кодере, при этом ключи, значение и запросы поступают из предыдущего слоя кодера. Самовнимание позволяет модели учитывать зависимости между всеми элементами входной последовательности.

3. Маскированное самовнимание. Применяется в декодере для предотвращения использования будущих элементов последовательности при генерации текущего элемента. Маска блокирует доступ к будущим элементам при вычислении внимания.

Кроме того, в трансформерах используется позиционное кодирование, которое добавляет к векторному представлению элемента входной последовательности информацию о его порядке в последовательности.

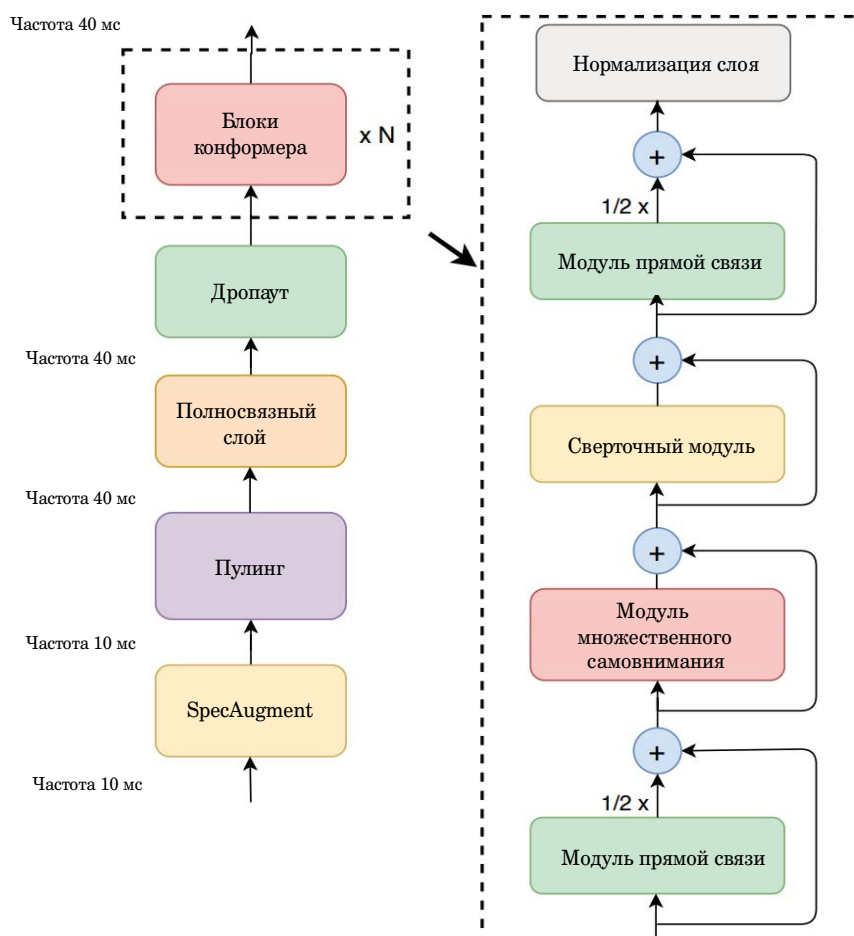
Трансформер в своем изначальном варианте подразумевает отсутствие любых рекуррентных слоев, вместо них применяются различные варианты механизма внимания. Это позволяет решить проблему затухающего градиента, возникающую при обработке последовательностей при помощи рекуррентных нейронных сетей [16], а также дает возможность обработки более длинных входных последовательностей, позволяя более эффективно распараллелить вычисления [17].

Во многих современных CARP архитектура трансформера по-прежнему применяется без существенных изменений [18, 19]. Так, в [19]

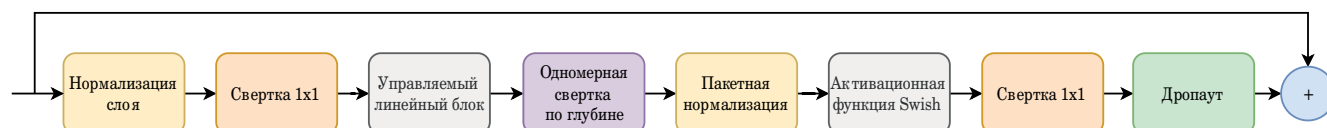
авторам удалось добиться лучших результатов по сравнению с архитектурами с долгой краткосрочной памятью (англ. *long short-term memory*, LSTM). За счет применения трансформера с восемью «головами» внимания и внешней языковой моделью ошибка распознавания составила 2,81 % против 2,98 % WER для базовой архитектуры на данных корпуса Librispeech Test Clean. В работе [20] была применена архитектура, которая представляет из себя кодер, состоящий из 18 модулей, и декодер, состоящий из шести модулей. Это позволило улучшить метрику WER до 7,83 % против 9,25 % у базовой модели, построенной на рекуррентной нейронной сети на основе трансдюсера (англ. *recurrent neural network transducer*, RNN-T), и до 8,05% у модели на основе шифратора и дешифратора с механизмом внимания (англ. *attention-based encoder-decoder*, AED). В более поздних работах представлены системы на основе архитектуры трансформера, адаптированной для задачи распознавания речи в реальном времени. Например, авторы [21] предлагают решение в виде классического трансформера с восемью «головами» внимания и дополнительной памятью. На корпусе Librispeech Clean предложенная модель набирает 2,8 % против 3,3 % WER по сравнению с рекуррентными сетями типа *latency-controlled bidirectional long short-term memory* (LC-BLSTM) [22].

Существенным недостатком трансформеров является неэффективная работа с локальным контекстом (информацией, содержащейся непосредственно рядом с обрабатываемым токеном) по сравнению с глобальным (т. е. соседними предложениями или даже другими абзацами). Эту проблему решает архитектура, названная «конформер» (англ. *conformer*) [23], которая внедряет в базовую архитектуру трансформера сверточные слои, что позволяет модели эффективно фокусироваться на локальных контекстах.

Общая схема архитектуры конформера (рис. 2) имеет следующий вид: в начале идет блок спектральной аугментации (SpecAugment) [24], который искусственно увеличивает разнообразие обучающих данных путем внесения случайных искажений в спектрограммы аудиозаписей и при этом сохраняет семантическую информацию, необходимую для распознавания речи. Далее — слой пулинга, полносвязный слой, а также слой прореживания нейронов (англ. *dropout*), за которыми следует последовательность из конформер-блоков, представляющих собой группу из модулей прямой связи, внимания и сверточных модулей. Схема сверточного модуля представлена на рис. 3. Вначале идет сверточный слой с ядром 1x1 (англ. *pointwise convolution*), затем — слой



■ **Рис. 2.** Схема модели конформер [23]
 ■ **Fig. 2.** Architecture of the conformer model [23]



■ **Рис. 3.** Схема сверточного блока модели конформер [23]
 ■ **Fig. 3.** Architecture of the convolutional module of the conformer model [23]

управляемого линейного блока (англ. *gated linear unit*; GLU), следом расположен слой, выполняющий одномерную свертку по глубине. После свертки выполняется пакетная нормализация. Далее идет слой с активационной функцией swish.

Математически работу конформер-блока можно записать следующим образом [23]:

$$\tilde{x}_i = x_i + \frac{1}{2} \text{FFN}(x_i);$$

$$x'_i = \tilde{x}_i + \text{MHSA}(\tilde{x}_i);$$

$$x''_i = x'_i + \text{Conv}(x'_i);$$

$$y_i = \text{Layernorm}\left(x''_i + \frac{1}{2} \text{FFN}(x''_i)\right),$$

где x_i – признак, поступающий на вход конформер-блока; y_i – выход конформер-блока; FFN означает модуль прямого распространения; MHSA – модуль «многоголового» самовнимания; Conv – модуль свертки.

В качестве декодера в исходной статье [23], описывающей архитектуру конформер, использовалась однослойная LSTM.

Модель конформер объединяет преимущества механизма внимания и свертки. «Многоголовое» самовнимание позволяет захватывать глобальный контекст, а сверточные слои позволяют конформеру эффективно фокусироваться на локальном контексте, что является преимуществом модели конформер по сравнению с моделью трансформер.

Как показывает исследование [25], конформер можно оптимизировать для работы в режиме реального времени. Авторы [25] предложили удалить блок самовнимания, что позволило достигнуть снижения вычислительной сложности на 10 % и применить модель в работе голосового помощника Alexa; при этом было показано, что упрощение архитектуры не снижает точности распознавания коротких фраз, которые характерны при обращении к голосовым помощникам. Авторы назвали данную архитектуру *Commer*.

В работе [26] предложена модель *HyperConformer*, требующая меньших вычислительных ресурсов и памяти. *HyperConformer* вместо «многоголового» самовнимания использует модель *HyperMixer*, предложенную в работе [27] для динамического формирования многослойного перцептрона для смешивания признаков. *HyperConformer* позволяет моделировать как локальное, так и глобальное взаимодействие между токенами.

В работе [28] между кодером и декодером трансформера добавлен дополнительный блок памяти, аналогичный нейронной машине Тьюринга, что позволяет лучше распознавать длинные фразы.

Уже из представленного в текущем разделе обзора архитектур можно сделать вывод, что, несмотря на общую эффективность трансформеров (и в том числе модификаций этой архитектуры), их применение связано с рядом ограничений. Так, трансформеры часто сталкиваются с проблемами при обработке длинных последовательностей (и нередко длина аудиозаписи может превышать оптимальную длину входных данных). Кроме того, обучение моделей на основе трансформеров требует больших объемов данных и времени, что может вызывать трудности в условиях ограниченного доступа к качественным и размеченным аудиоданным. Наконец, чрезвычайно важным аспектом успешного применения архитектуры трансформер в распознавании речи является выбор правильной стратегии декодирования. Даже если сама архитектура эффективна, без оптимальной стратегии декодирования качество результатов может значительно пострадать. В следующем разделе рассмотрены различные стратегии декодирования, которые могут применяться совместно с моделями трансформер.

Стратегии декодирования в интегральных системах распознавания речи

Декодер в интегральной системе распознавания речи играет фундаментальную роль в преобразовании выходных данных модели, полученных после анализа речевых данных, в текстовое представление, доступное для человека. Его функция заключается в том, чтобы эффективно переводить последовательность признаков или вероятностей, полученных от модели, в читаемый или последовательность символов. При этом декодер должен учитывать контекст и долгосрочные зависимости в речевой информации, чтобы генерировать наиболее вероятное текстовое представление, а также управлять переменной длиной входных и выходных последовательностей, обеспечивая корректное соответствие между аудио- и текстовыми данными различной длины. Более того, декодер должен быть способен обрабатывать неоднозначности и ошибки, возникающие в процессе распознавания речи, например, выбирая наиболее вероятные интерпретации или исправляя ошибки в распознанном тексте, чтобы обеспечить высокую точность и понятность результатов. По большому счету, функционально декодер выполняет роль языковой модели, при этом единицами словаря могут быть слова, графемы или отрезки слов (англ. *pieces of words* — «кусочки слов», *subword units* — «подсловные единицы»), не поддающиеся обобщению с лингвистической точки зрения. Процедура разбиения текста на такие единицы называется токенизацией. Большое количество алгоритмов токенизации, например *Byte Pair Encoding*, *WordPiece* или *Unigram Language Model*, позволяет системе эффективно разбивать слова на подсловные единицы, решая, в частности, проблему слов, не вошедших в обучающий словарь.

Несколько стратегий декодирования получили наибольшее распространение при создании интегральных систем распознавания речи в наши дни.

Коннекционная временная классификация. Коннекционная временная классификация (англ. *connectionist temporal classification*, СТС) [29] используется для обучения моделей в том случае, если необходимо учитывать выравнивание входных и выходных последовательностей. Это позволяет модели пропускать некоторые временные шаги. Модель СТС преобразует звуковой сигнал в последовательность символов, а затем удаляет из нее ненужные символы, пробелы и повторы. Выходной слой нейронной сети содержит по одному блоку для каждого символа выходной последовательности (графем, фонем,

знаков препинания) и еще один для дополнительного символа «пропуск» (“blank”), соответствующего пустому выходному символу. Более подробно модель CTC описана в [4, 29].

Стратегия CTC может быть применена к архитектурам RNN (LSTM, GRU) и трансформерам (с модификациями для поддержки CTC). Одним из главных преимуществ CTC является гибкость, поскольку эта стратегия не требует точного выравнивания входных и выходных последовательностей, что полезно для задач с изменчивыми длительностями событий. Также CTC хорошо поддерживает длинные последовательности, в то же время позволяя моделям работать без явного выравнивания между входными и выходными данными. Последнее особенно полезно в рамках распознавания речи, так как не всегда длина последовательности фонем соответствует длине слова, записанного в соответствии с правилами графики исходного языка.

Тем не менее применение этой стратегии оправдано только в том случае, если длина входной последовательности больше, чем длина выходной. В частности, это ограничивает возможность применения техник понижения размерности. Помимо этого, в CTC элементы выходной последовательности априорно считаются независимыми, что иногда приводит к ошибкам в итоговой гипотезе распознавания.

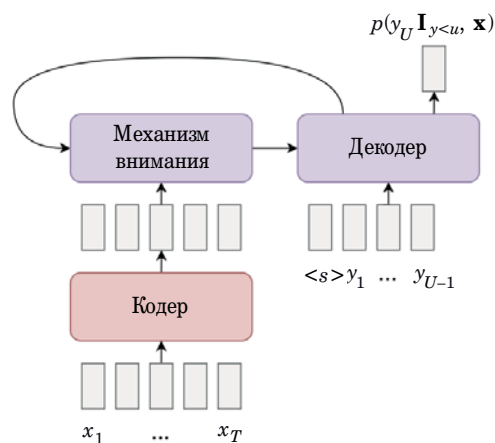
В статье [30] предлагается смешанная архитектура нейросетевой модели CTC в сочетании с механизмом внимания, ограниченным по времени. Это решение показывает, что можно создать систему с использованием механизма внимания, которая будет работать в режиме реального времени, однако такая система имеет все недостатки подхода CTC и лишь частично реализует преимущества механизма внимания из-за ограничения по времени.

Совместное использование конформера и CTC сообщает моделям большую устойчивость к вариациям по времени и дает более эффективное обучение на небольших объемах данных. Успешные реализации такого подхода можно найти в [31, 32].

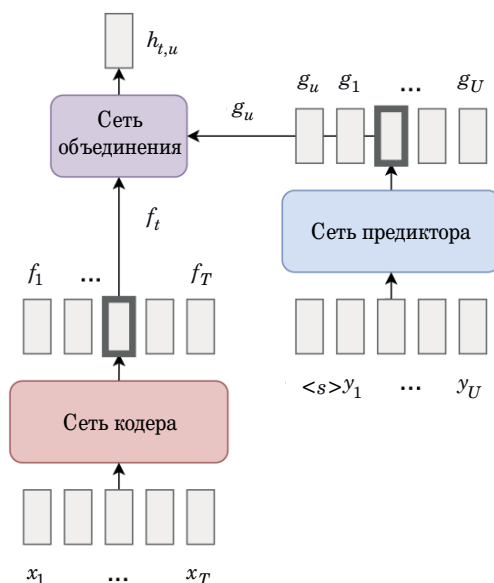
AED. Шифратор и дешифратор с механизмом внимания (AED) [33, 34] — это стратегия, основанная на механизме внимания, который позволяет модели обрабатывать только определенные части входных данных при генерации выходной последовательности. Эта стратегия наиболее эффективно применяется в связке с трансформерами для точного моделирования сложных зависимостей, содержащихся в речевом сигнале, и учета контекста на длительных временных промежутках. Это выгодно отличает AED от CTC и в конечном итоге приводит к улучшению точности.

Однако у моделей, использующих эту стратегию декодирования, есть недостаток — они серьезно уступают другим стратегиям по части скорости распознавания. Это связано с тем, что механизм внимания ожидает сразу всю входную последовательность для генерации транскрипции на выходе. В большинстве случаев это делает невозможным использование этой стратегии в режиме реального времени, что является крупным недостатком для современных CAPT. Несмотря на то, что существуют некоторые способы адаптировать AED для использования в режиме реального времени [30, 35, 36], такие модели все равно сильно проигрывают по скорости моделям, использующим CTC и RNN-T. Схема AED представлена на рис. 4 (x, y — значения элементов входной последовательности признаков кодера и декодера; $T, U - 1$ — размеры входных последовательностей признаков кодера и декодера соответственно).

RNN-T. Архитектура трансдюсер (RNN-T) была впервые представлена в 2012 г. [38]. RNN-T решает основные проблемы CTC, связанные с выравниванием длины последовательности, при помощи генерации сразу целого набора из нескольких выходных последовательностей для каждой входной последовательности. Наличие двух отдельных сетей: кодера и предсказательной сети, — а также объединяющей их полносвязной сети позволяет таким моделям лучше фокусироваться на контексте. При всех достоинствах данный подход не свободен от недостатков. Ввиду необходимости построения целого набора выравниваний для каждой входной последовательности, модели, построенные на трансдюсере, требуют большого количества памяти для обучения. Схема RNN-модели представлена на рис. 5 (x, y — значения элементов входной последовательности признаков кодера и предиктора; $T,$



■ **Рис. 4.** Схема AED-модели [37]
 ■ **Fig. 4.** AED model diagram [37]



■ **Рис. 5.** Схема модели RNN-T [30]
 ■ **Fig. 5.** RNN-T model diagram [30]

U — размер входных последовательностей сетей кодера и предиктора соответственно; индексы f, g — значения выходных последовательностей признаков кодера и предиктора соответственно).

Ряд исследований [3, 18, 39] показывает, что обычный RNN-T используется для построения SAPR реже, чем трансформер. К тому же эти модели недостаточно эффективны при распознавании данных, не соответствующих обучающей выборке [40]. В работах, выдаваемых в системе Google Scholar по запросу «распознавание речи», трансформеры используются примерно в 10 раз чаще, чем RNN-T. Как уже сказано, трансдюсер хорошо подходит для распознавания речи в режиме реального времени, однако трансформеры также эффективно справляются с подобными задачами [21, 31, 41, 42]. Собственно, немодифицированные трансдюсеры на сегодня достаточно редки. Более актуальными являются комбинированные модели, сочетающие в себе одновременно элементы как трансдюсера, так и других архитектур. Как пример можно привести сочетание трансформера и трансдюсера (англ. *transformer-transducer*, или Т-Т) [43–46]. Эта модель поддерживает работу с последовательностями переменной длины и обеспечивает улучшение точности распознавания речи, устраняя недостатки предыдущих моделей [47]. В этой связке трансформер играет роль кодера, который генерирует высокоуровневые признаки, основанные на глобальном контексте, а использование трансдюсера подразумевает дальнейшую интерпретацию выходных данных с трансформера для построения выравниваний, как в клас-

сическом варианте с СТС или обычным RNN-T. Это позволяет совместить преимущества обоих подходов [43].

Одной из самых эффективных является модель конформер-трансдюсер, в которой LSTM-слой декодера, применяемый в базовом конформере, заменен на трансдюсер [48]. Данная архитектура обладает достоинствами архитектуры конформера, а именно способностью эффективно моделировать как долгосрочные временные зависимости при помощи механизма самовнимания, так и локальные зависимости благодаря сверточным ядрам. При этом кодер трансформера и декодер трансдюсера объединены в единую структуру, что повышает эффективность вычислений, точность, а также дает способность выравнивания входных и выходных данных переменной длины. Одной из самых последних модификаций такой модели является Fast Conformer, представленный Nvidia [49]. Скорость этой модели в 2,8 раза выше, чем у оригинальной модели конформер-трансдюсер, за счет дополнительного уменьшения частоты дискретизации входной аудиоспектрограммы в два раза, а также превосходит базовую на 0,2 % WER на части Test Other корпуса Librispeech.

Обсуждение

На основе описанных моделей с архитектурой трансформер и их модификаций, а также стратегий декодирования можно сделать ряд выводов. Трансформеры обеспечивают эффективное параллельное вычисление и способны обрабатывать длинные последовательности благодаря механизму внимания. Основным недостатком в этом случае заключается в высоких вычислительных затратах, особенно при работе с большими моделями и длинными последовательностями. При этом трансформеры работают лучше в сценариях, где требуется учесть глобальный контекст всей последовательности, таких как длинные предложения или монологи, что позволяет учитывать взаимосвязи между удаленными элементами входных данных.

Конформеры, дополняя трансформеры сверточными слоями, улучшают обработку локальных контекстов, что особенно полезно для обработки непрерывной речи с высокой плотностью информации или шумовых данных. Конформеры объединяют преимущества трансформеров в захвате глобальных зависимостей с возможностями сверточных нейронных сетей в захвате локальных зависимостей. Это делает их пригодными для задач, в которых важна высокая точность распознавания в условиях высокого уровня шума и (или) вариативности речи.

В то же время конформеры зачастую требуют больших вычислительных ресурсов из-за большого количества параметров и сложности архитектуры, что влечет за собой большие временные затраты на обучение. Кроме того, конформерные модели могут требовать большого количества данных для эффективного обучения, особенно если они используются для работы с разнообразными типами данных.

Итак, рассмотрено несколько стратегий декодирования: CTC, AED и трансдюсер (RNN-T). CTC подходит для задач, где важно учитывать выравнивание входных и выходных последовательностей, но плохо применима в тех случаях, когда длина входной последовательности меньше выходной. В качестве недостатка данных моделей можно отметить то, что эти модели предполагают независимость выходных символов. В трансдюсерах, напротив, выходной символ зависит от предшествующего контекста. AED позволяет учитывать долгосрочные зависимости и контекст, способствуя повышению точности, но снижает скорость распознавания, ограничивая ее работу в режиме реального времени. У механизма внимания есть недостаток, который заключается в том, что модели, его использующие, требуют на вход подачу сразу всей обрабатываемой последовательности. Это делает их медленными и непригодными для систем потокового распознавания речи. Основные достоинства и недостатки различных стратегий декодирования представлены в таблице.

Другим важным наблюдением является то, что точность работы SAPP сильно зависит от условий применения и специфики целевого языка. Модели, обученные на ограниченном объеме дан-

ных, могут показывать низкую точность в сложных условиях, таких как шумная среда или нестандартное произношение. Лингвистические особенности целевого языка также играют важную роль, что демонстрируют расхождения в показателях точности для различных языков при аналогичном объеме обучающих данных.

Качество и объем обучающих данных также остаются критически важными для успешного обучения моделей SAPP. Модели, обученные на большом объеме данных, показывают лучшую производительность, особенно в условиях шумной среды или нестандартного произношения.

В последнее время тенденцией является увеличение размера моделей [18]. Очевидно, что для обучения особенно больших моделей требуется большое количество как вычислительных ресурсов, так и обучающих данных, что усложняет процесс обучения таких моделей. Подход, используемый для моделей wav2vec [51], призван решить эту проблему. Он подразумевает предварительное обучение модели на большом количестве неразмеченных аудиозаписей, что дает возможность выделить необходимый контекст для условной акустической модели, а затем уже дообучить ее на относительно небольших по размеру данных в виде транскрибированной речи, что, в свою очередь, позволяет условной языковой модели выделить языковой контекст [51–53]. Подход хорошо себя показал в задаче малоресурсного распознавания речи. В статье [54] достигнут результат по метрике WER 26,5 % для языка суахили при размере набора данных всего 30 часов и 18,7 % CER для чукотского языка при всего трех часах речевых данных.

- Характеристики основных стратегий декодирования в контексте интегрального подхода
- Features of the main decoding strategies within the framework of end-to-end

Характеристики и метаданные	Стратегия декодирования						
	CTC			AED	Трансдюсер		
Достоинства	Гибкость выравнивания, хорошая работа с длинными временными последовательностями			Высокая точность, гибкость, эффективная работа с длинными последовательностями	Учитывает предшествующий контекст		
Недостатки	Элементы выходной последовательности априорно считаются независимыми			Высокая вычислительная сложность, сложность настройки	Требует больших вычислительных ресурсов		
Пример	[50]	[49]	[49]	[20]	[43]	[49]	[49]
Кодек	T	C	FC	T	T	C	FC
Тестовый корпус	TED-LIUM 2	LibriSpeech test other		Microsoft anonymized training data	LibriSpeech test other		
WER, %	14,2	4,50	4,19	7,83	5,6	3,74	3,79

Примечание: T – трансформер, C – конформер, FC – Fast Conformer.

Заключение

В настоящей статье были рассмотрены модели интегрального распознавания речи на основе архитектуры трансформер и основные стратегии декодирования, которые могут применяться совместно с трансформерами. Использование интегральных систем, объединяющих акустическую и языковую модели в рамках одной глубокой нейронной сети, зарекомендовало себя как эффективное решение по сравнению с традиционными подходами. Тем не менее различные стратегии реализации имеют свои сильные и слабые стороны, что заставляет адаптировать их в зависимости от конкретных задач. Особенно это важно при решении задач распознавания естественных языков.

С учетом интеграции различных технологий дальнейшее развитие САРР может быть связано с мультимодальным обучением, где модели обучаются на данных разных типов, таких как текст, аудио и видео. Это может значительно улучшить понимание контекста и повысить точность распознавания речи в сложных условиях. Дальнейшие исследования авторов могут быть связаны именно с обоснованием методов работы с различными модальностями в рамках интегрального подхода.

Финансовая поддержка

Исследование выполнено в рамках бюджетной темы СПб ФИЦ РАН (№ FFZF-2022-0005).

Литература

1. Malik M., Malik M. K., Mehmood K., Makhdoom I. Automatic speech recognition: A survey. *Multimedia Tools and Applications*, 2021, no. 80, pp. 9411–9457. doi:10.1007/s11042-020-10073-7
2. Adolfi F., Bowers J. S., Poeppel D. Successes and critical failures of neural networks in capturing human-like speech recognition. *Neural Networks*, 2023, vol. 162, pp. 199–211. doi:10.1016/j.neunet.2023.02.032
3. Prabhavalkar R., Hori T., Sainath T. N., Schlüter R., Watanabe S. End-to-end speech recognition: A survey. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 2024, no. 32, pp. 325–351. doi:10.1109/TASLP.2023.3328283
4. Марковников Н. М., Кипяткова И. С. Аналитический обзор интегральных систем распознавания речи. *Труды СПИИРАН*, 2018, т. 58, № 3, с. 77–110. doi:10.15622/sp.58.4
5. Zhang Z., Geiger J., Pohjalainen J., Mousa A. E.-D., Jin W., Schuller B. Deep learning for environmentally robust speech recognition: An overview of recent developments. *ACM Transactions on Intelligent Systems and Technology*, 2018, vol. 9, no. 5, Article 49. doi:10.1145/31781
6. Najafian M., Russell M. Automatic accent identification as an analytical tool for accent robust automatic speech recognition. *Speech Communication*, 2020, no. 122, pp. 44–55. doi:10.1016/j.specom.2020.05.003
7. Radford A., Kim J. W., Xu T., Brockman G., McLeavey C., Sutskever I. Robust speech recognition via large-scale weak supervision. *Proc. of the 40th Intern. Conf. on Machine Learning Intern. Conf. on Machine Learning (PMLR-2023)*, 2023, pp. 28492–28518.
8. Conneau A., Ma M., Khanuja S., Zhang Y., Axelrod V., Dalmia S., Riesa J., Rivera C., Bapna A. FLEURS: Few-shot learning evaluation of universal representations of speech. *Proc. of 2022 IEEE Spoken Language Technology Workshop (SLT-2022)*, 2022, pp. 798–805. doi:10.1109/SLT54892.2023.10023141
9. Jelinek F., Bahl L., Mercer R. Design of a linguistic statistical decoder for the recognition of continuous speech. *IEEE Transactions on Information Theory*, 1975, vol. 21, no. 3, pp. 250–256.
10. Wang S., Li G. Overview of end-to-end speech recognition. *Journal of Physics: Conference Series*, 2019, vol. 1187, Article 052068. doi:10.1088/1742-6596/1187/5/052068
11. Чучупал В. Я. Нейросетевые модели языка для систем распознавания речи. *Речевые технологии*, 2020, № 1, с. 27–47.
12. Seki H., Yamamoto K., Nakagawa S. A deep neural network integrated with filterbank learning for speech recognition. *Proc. of 2017 IEEE Intern. Conf. on Acoustics, Speech and Signal Processing (ICASSP-2017)*, 2017, pp. 5480–5484. doi:10.1109/ICASSP.2017.7953204
13. Ittichaichareon C., Suksri S., Yingthawornsuk T. Speech recognition using MFCC. *Proc. of Intern. Conf. on Computer Graphics, Simulation and Modeling*, 2012, pp. 135–138.
14. Hori T., Cho J., Watanabe S. End-to-end speech recognition with word-based RNN language models. *Proc. of 2018 IEEE Spoken Language Technology Workshop (SLT-2018)*, 2018, pp. 389–396. doi:10.1109/SLT.2018.8639693
15. Vaswani A., Shazeer N., Parmar N., Uszkoreit J., Jones L., Gomez A. N., Kaiser Ł., Polosukhin I. Attention is all you need. *Proc. of the 31st Intern. Conf. on Neural Information Processing Systems (NIPS-2017)*, 2017, pp. 6000–6010.
16. Hochreiter S. The vanishing gradient problem during learning recurrent neural nets and problem solutions. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 1998, vol. 06, no. 02, pp. 107–116.

17. Karita S., Chen N., Hayashi T., Hori T., Inaguma H., Jiang Z., Someki M., Soplin N. E. Y., Yamamoto R., Wang X., Watanabe S., Yoshimura T., Zhang W. A comparative study on transformer vs RNN in speech applications. *Proc. of 2019 IEEE Automatic Speech Recognition and Understanding Workshop (ASRU-2019)*, 2019, pp. 449–456. doi:10.1109/ASRU46091.2019.9003750
18. Latif S., Zaidi A., Cuayahuitl H., Shamshad F., Shoukat M., Qadir J. Transformers in speech processing: A survey. *arXiv preprint*, 2023. arXiv:2303.11607. doi:10.48550/arXiv.2303.11607
19. Zeyer A., Bahar P., Irie K., Schluter R., Ney H. A comparison of transformer and LSTM encoder decoder models for ASR. *Proc. of 2019 IEEE Automatic Speech Recognition and Understanding Workshop (ASRU-2019)*, 2019, pp. 8–15. doi:10.1109/ASRU46091.2019.9004025
20. Li J., Wu Y., Gaur Y., Wang C., Zhao R., Liu S. On the comparison of popular end-to-end models for large scale speech recognition. *Proc. of the 21st Annual Conf. of the International Speech Communication Association (Interspeech-2020)*, 2020, pp. 1–5. doi:10.21437/Interspeech.2020-2846
21. Wu C., Wang Y., Shi Y., Yeh C.-F., Zhang F. Streaming Transformer-based acoustic models using self-attention with augmented memory. *Proc. of the 21st Annual Conf. of the International Speech Communication Association (Interspeech-2020)*, 2020, pp. 2132–2136. doi:10.21437/Interspeech.2020-2079
22. Zhang Y., Chen G., Yu D., Yao K., Khudanpur S., Glass J. Highway long short-term memory RNNs for distant speech recognition. *Proc. of 2016 IEEE Intern. Conf. on Acoustics, Speech and Signal Processing (ICASSP-2016)*, 2016, pp. 5755–5759. doi:10.1109/ICASSP.2016.7472780
23. Gulati A., Qin J., Chiu C.-C., Parmar N., Zhang Y., Yu J., Han W., Wang S., Zhang Z., Wu Y., Pang R. Conformer: Convolution-augmented Transformer for speech recognition. *Proc. of the 21st Annual Conf. of the International Speech Communication Association (Interspeech-2020)*, 2020, pp. 5036–5040. doi:10.21437/Interspeech.2020-3015
24. Park D. S., Chan W., Zhang Y., Chiu C.-C., Zoph B., Cubuk E. D., Le Q. V. SpecAugment: A simple data augmentation method for automatic speech recognition. *Proc. of the 20th Annual Conf. of the International Speech Communication Association (Interspeech-2019)*, 2019, pp. 2613–2617. doi:10.21437/Interspeech.2019-2680
25. Radfar M., Lyskawa P., Trujillo B., Xie Y., Zhen K., Heymann J., Filimonov D., Strimel G., Susanj N., Mouchtaris T. Conmer: Streaming conformer without self-attention for interactive voice assistants. *Proc. of the 24th Annual Conf. of the International Speech Communication Association (Interspeech-2023)*, 2023, pp. 2198–2202. doi:10.21437/Interspeech.2023-2228
26. Mai F., Zuluaga-Gomez J., Parcollet T., Motlicek P. HyperConformer: Multi-head HyperMixer for efficient speech recognition. *Proc. of the 24th Annual Conf. of the International Speech Communication Association (Interspeech-2023)*, 2023, pp. 2213–2217. doi:10.21437/Interspeech.2023-1611
27. Mai F., Pannatier A., Fehr F., Chen H., Marelli F., Fleuret F., Henderson J. Hypermixer: An MLP-based low cost alternative to transformers. *Proc. of the 61st Annual Meeting of the Association for Computational Linguistics*, 2023, vol. 1: Long Papers, pp. 15632–15654. doi:10.18653/v1/2023.acl-long.871
28. Carvalho C., Abad A. Memory-augmented conformer for improved end-to-end long-form ASR. *Proc. of the 24th Annual Conf. of the International Speech Communication Association (Interspeech-2023)*, 2023, pp. 2218–2222. doi:10.21437/Interspeech.2023-893
29. Graves A., Fernández S., Gomez F., Schmidhuber J. Connectionist temporal classification: labelling unsegmented sequence data with recurrent neural networks. *Proc. of the 23rd Intern. Conf. on Machine Learning (ICML-2006)*, 2006, pp. 369–376. doi:10.1145/1143844.1143891
30. Moritz N., Hori T., Le J. Streaming automatic speech recognition with the Transformer model. *Proc. of 2020 IEEE Intern. Conf. on Acoustics, Speech and Signal Processing (ICASSP-2020)*, 2020, pp. 6074–6078. doi:10.1109/ICASSP40776.2020.9054476
31. Burchi M., Vielzeuf V. Efficient Conformer: Progressive downsampling and grouped attention for automatic speech recognition. *Proc. of 2021 IEEE Automatic Speech Recognition and Understanding Workshop (ASRU-2021)*, 2021, pp. 8–15. doi:10.1109/ASRU51503.2021.9687874
32. Guo H., Chen Y., Xie X., Xu G., Guo W. Efficient Conformer-based CTC model for intelligent cockpit speech recognition. *Proc. of 13th Intern. Symp. on Chinese Spoken Language Processing (ISCSLP-2022)*, 2022, pp. 522–526. doi:10.1109/ISCSLP57327.2022.10037993
33. Chan W., Jaitly N., Le Q. V., Vinyals O. Listen, attend and spell: A neural network for large vocabulary conversational speech recognition. *Proc. of 2016 IEEE Intern. Conf. on Acoustics, Speech and Signal Processing (ICASSP-2016)*, 2016, pp. 4960–4964. doi:10.1109/ICASSP.2016.7472621
34. Chorowski J. K., Bahdanau D., Serdyuk D., Cho K., Bengio Y. Attention-based models for speech recognition. *Proc. of the 28th Intern. Conf. on Neural Information Processing Systems (NIPS-2015)*, 2015, vol. 1, pp. 577–585.
35. Chiu C.-C., Raffel C. Monotonic chunkwise attention. *arXiv preprint*, 2018. arXiv:1712.05382. doi:10.48550/arXiv.1712.05382
36. Inaguma H., Kawahara T. Alignment knowledge distillation for online streaming attention-based speech recognition. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 2021, vol. 31, pp. 1371–1385. doi:10.1109/TASLP.2021.31332

37. **Lugosh L.** *Sequence-to-sequence learning with Transducers*. <https://lorenlugosch.github.io/posts/2020/11/transducer> (дата обращения: 08.08.2024).
38. **Graves A.** Sequence transduction with recurrent neural networks. *arXiv preprint*, 2012. arXiv:1211.3711. doi:10.48550/arXiv.1211.3711
39. **Karmakar P., Teng S. W., Lu G.** Thank you for attention: A survey on attention-based artificial neural networks for automatic speech recognition. *Intelligent Systems with Applications*, 2024, vol. 23, Article 200406. doi:10.1016/j.iswa.2024.200406
40. **Chiu C.-C., Narayanan A., Han W., Prabhavalkar R., Zhang Y., Jaitly N., Pang R., Sainath T. N., Nguyen P., Cao L., Wu Y.** RNN-T models fail to generalize to out-of-domain audio: Causes and solutions. *Proc. of 2021 IEEE Spoken Language Technology Workshop (SLT-2021)*, 2021, pp. 873–880. doi:10.1109/SLT48900.2021.9383518
41. **Shi Y., Wang Y., Wu C., Yeh C.-F., Chan J., Zhang F., Le D., Seltzer M.** Emformer: Efficient memory Transformer based acoustic model for low latency streaming speech recognition. *Proc. of 2021 IEEE Intern. Conf. on Acoustics, Speech and Signal Processing (ICASSP-2021)*, 2021, pp. 6783–6787. doi:10.1109/ICASSP39728.2021.9414560
42. **Tsunoo E., Kashiwagi Y., Watanabe S.** Streaming Transformer ASR with blockwise synchronous beam search. *Proc. of 2021 IEEE Spoken Language Technology Workshop (SLT-2021)*, 2021, pp. 22–29. doi:10.1109/SLT48900.2021.9383517
43. **Zhang Q., Lu H., Sak H., Tripathi A., McDermott E., Koo S., Kumar S.** Transformer Transducer: A streamable speech recognition model with Transformer Encoders and RNN-T loss. *Proc. of 2020 IEEE Intern. Conf. on Acoustics, Speech and Signal Processing (ICASSP-2020)*, 2020, pp. 7829–7833. doi:10.1109/ICASSP40776.2020.9053896
44. **Liu C., Zhang F., Le D., Kim S., Saraf Y., Zweig G.** Improving RNN Transducer based ASR with auxiliary tasks. *Proc. of 2021 IEEE Spoken Language Technology Workshop (SLT-2021)*, 2021, pp. 172–179. doi:10.1109/SLT48900.2021.9383548
45. **Tripathi A., Kim J., Zhang Q., Lu H., Sak H.** Transformer Transducer: One model unifying streaming and non-streaming speech recognition. *arXiv preprint*, 2020. arXiv:2010.03192. doi:10.48550/arXiv.2010.03192
46. **Cui M., Kang J., Deng J., Yin X., Xie Y., Chen X., Liu X.** Towards effective and compact contextual representation for Conformer Transducer speech recognition systems. *Proc. of the 24th Annual Conf. of the International Speech Communication Association (Interspeech-2023)*, 2023, pp. 2223–2227. doi:10.21437/Interspeech.2023-552
47. **Chen X., Wu Y., Wang Z., Liu S., Li J.** Developing real-time streaming Transformer Transducer for speech recognition on large-scale dataset. *Proc. of 2021 IEEE Intern. Conf. on Acoustics, Speech and Signal Processing (ICASSP-2021)*, 2021, pp. 5904–5908. doi:10.1109/ICASSP39728.2021.9413535
48. **Barcowski A., Jain R., Corcoran P.** A comparative analysis between Conformer-Transducer, Whisper, and wav2vec2 for improving the child speech recognition. *Proc. of 2023 Intern. Conf. on Speech Technology and Human-Computer Dialogue (SpeD-2023)*, 2023, pp. 42–47. doi:10.1109/SpeD59241.2023.10314867
49. **Rekesh D., Koluguri N. R., Kriman S., Majumdar S., Noroozi V., Huang H., Hrinchuk O., Puvvada K., Kumar A., Balam J., Ginsburg B.** Fast Conformer with linearly scalable attention for efficient speech recognition. *Proc. of 2023 IEEE Automatic Speech Recognition and Understanding Workshop (ASRU-2023)*, pp. 1–8. doi:10.1109/ASRU57964.2023.10389701
50. **Nakatani T.** Improving transformer-based end-to-end speech recognition with connectionist temporal classification and language model integration. *Proc. of the 20th Annual Conf. of the International Speech Communication Association (Interspeech-2019)*, 2019, pp. 1408–1412. doi:10.21437/Interspeech.2019-1938
51. **Schneider S., Baevski A., Collobert R., Auli M.** Wav2vec: unsupervised pre-training for speech recognition. *Proc. of the 20th Annual Conf. of the International Speech Communication Association (Interspeech-2019)*, 2019, pp. 3465–3469. doi:10.21437/Interspeech.2019-1873
52. **Jovanović M., Campbell M.** Generative artificial intelligence: Trends and prospects. *Computer*, 2022, vol. 55, no. 10, pp. 107–112. doi:10.1109/MC.2022.3192720
53. **Baevski A., Zhou Y., Mohamed A., Auli M.** Wav2vec 2.0: A framework for self-supervised learning of speech representations. *Proc. of the 34th Intern. Conf. on Neural Information Processing Systems (NIPS-2020)*, 2020, pp. 12449–12460.
54. **Safonova A., Yudina T., Nadimanov E., Davenport C.** Automatic speech recognition of low-resource languages based on Chukchi. *arXiv preprint*, 2022. arXiv:2210.05726. doi:10.48550/arXiv.2210.05726

UDC 004.934.2

doi:10.31799/1684-8853-2024-5-2-15

EDN: MWTGXE

Analytical survey of transformer-based end-to-end speech recognition models and strategiesK. L. Kapusta^a, Programmer, orcid.org/0009-0008-8623-0101I. S. Kipyatkova^a, PhD, Tech., Associate Professor, Senior Researcher, orcid.org/0000-0002-1264-4458, kipyatkova@ias.spb.suI. A. Kagiroy^a, Research Fellow, orcid.org/0000-0003-1196-1117^aSt. Petersburg Federal Research Center of the RAS, 39, 14th Line, 199178, Saint-Petersburg, Russian Federation

Introduction: One of the trends in natural language processing is the shift from modular architectures to end-to-end models. These systems combine various processing stages such as acoustics, language, lexicon modeling, and decoding into a unified architecture. One of the most currently used state-of-the-art architectures for end-to-end speech recognition is transformer architecture and its modifications. **Purpose:** To make a detailed survey of decoding models and strategies in the context of end-to-end approaches in natural language processing. **Results:** The analysis of various decoding strategies has led to several conclusions. Connectionist Temporal Classification is effective when there is no alignment between the speech signal and text transcriptions, but its use is impractical when the length of the input data is shorter than that of the output. The main drawback of models using Connectionist Temporal Classification is the assumption of independence between output symbols. More promising are transducers, which take into account the preceding context for each output symbol, as well as Attention-based encoder-decoder models, which capture long-term dependencies and context. However, the latter strategy has the downside of being slow, limiting its use in real-time applications. Thus, each of the strategies reviewed in the paper has its own strengths, but they perform best when applied to specific types of tasks. **Practical relevance:** This survey paper contributes to the study of rapidly developing end-to-end speech recognition area, irrespective to particular languages. The results of the work can find application in the field of the development of automatic speech recognition systems in natural languages, including low-resource languages. **Discussion:** The current trend of increasing model sizes makes hybrid solutions the most promising ones, as they account for the need to use speech recognition systems in real-time while requiring fewer computational resources.

Keywords — end-to-end model, transformer, transducer, decoding, automatic speech recognition.

For citation: Kapusta K. L., Kipyatkova I. S., Kagiroy I. A. Analytical survey of transformer-based end-to-end speech recognition models and strategies. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2024, no. 5, pp. 2–15 (In Russian). doi:10.31799/1684-8853-2024-5-2-15, EDN: MWTGXE

Financial support

This survey was financially supported by budgetary theme No. FFZF-2022-0005.

References

1. Malik M., Malik M. K., Mehmood K., Makhdoom I. Automatic speech recognition: A survey. *Multimedia Tools and Applications*, 2021, no. 80, pp. 9411–9457. doi:10.1007/s11042-020-10073-7
2. Adolphi F., Bowers J. S., Poeppel D. Successes and critical failures of neural networks in capturing human-like speech recognition. *Neural Networks*, 2023, vol. 162, pp. 199–211. doi:10.1016/j.neunet.2023.02.032
3. Prabhavalkar R., Hori T., Sainath T. N., Schlüter R., Watanabe S. End-to-end speech recognition: A survey. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 2024, no. 32, pp. 325–351. doi:10.1109/TASLP.2023.3328283
4. Markovnikov N. M., Kipyatkova I. S. An analytic survey of end-to-end speech recognition systems. *SPIIRAS Proceedings*, 2018, vol. 58, no. 3, pp. 77–110 (In Russian). doi:10.15622/sp.58.4
5. Zhang Z., Geiger J., Pohjalainen J., Mousa A. E.-D., Jin W., Schuller B. Deep learning for environmentally robust speech recognition: An overview of recent developments. *ACM Transactions on Intelligent Systems and Technology*, 2018, vol. 9, no. 5, Article 49. doi:10.1145/31781
6. Najafian M., Russell M. Automatic accent identification as an analytical tool for accent robust automatic speech recognition. *Speech Communication*, 2020, no. 122, pp. 44–55. doi:10.1016/j.specom.2020.05.003
7. Radford A., Kim J. W., Xu T., Brockman G., Mcleavy C., Sutskever I. Robust speech recognition via large-scale weak supervision. *Proc. of the 40th Intern. Conf. on Machine Learning Intern. Conf. on Machine Learning (PMLR-2023)*, 2023, pp. 28492–28518.
8. Conneau A., Ma M., Khanuja S., Zhang Y., Axelrod V., Dalmia S., Riesa J., Rivera C., Bapna A. FLEURS: Few-shot learning evaluation of universal representations of speech. *Proc. of 2022 IEEE Spoken Language Technology Workshop (SLT-2022)*, 2022, pp. 798–805. doi:10.1109/SLT54892.2023.10023141
9. Jelinek F., Bahl L., Mercer R. Design of a linguistic statistical decoder for the recognition of continuous speech. *IEEE Transactions on Information Theory*, 1975, vol. 21, no. 3, pp. 250–256.
10. Wang S., Li G. Overview of end-to-end speech recognition. *Journal of Physics: Conference Series*, 2019, vol. 1187, Article 052068. doi:10.1088/1742-6596/1187/5/052068
11. Chuchupal V. Y. Neural language models for automatic speech recognition. *Speech Technology*, 2020, no. 1, pp. 27–47 (In Russian).
12. Seki H., Yamamoto K., Nakagawa S. A deep neural network integrated with filterbank learning for speech recognition. *Proc. of 2017 IEEE Intern. Conf. on Acoustics, Speech and Signal Processing (ICASSP-2017)*, 2017, pp. 5480–5484. doi:10.1109/ICASSP.2017.7953204
13. Ittichaichareon C., Suksri S., Yingthawornsuk T. Speech recognition using MFCC. *Proc. of Intern. Conf. on Computer Graphics, Simulation and Modeling*, 2012, pp. 135–138.
14. Hori T., Cho J., Watanabe S. End-to-end speech recognition with word-based RNN language models. *Proc. of 2018 IEEE Spoken Language Technology Workshop (SLT-2018)*, 2018, pp. 389–396. doi:10.1109/SLT.2018.8639693
15. Vaswani A., Shazeer N., Parmar N., Uszkoreit J., Jones L., Gomez A. N., Kaiser Ł., Polosukhin I. Attention is all you need. *Proc. of the 31st Intern. Conf. on Neural Information Processing Systems (NIPS-2017)*, 2017, pp. 6000–6010.
16. Hochreiter S. The vanishing gradient problem during learning recurrent neural nets and problem solutions. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 1998, vol. 06, no. 02, pp. 107–116.
17. Karita S., Chen N., Hayashi T., Hori T., Inaguma H., Jiang Z., Someki M., Soplin N. E. Y., Yamamoto R., Wang X., Watanabe S., Yoshimura T., Zhang W. A comparative study on transformer vs RNN in speech applications. *Proc. of 2019 IEEE Automatic Speech Recognition and Understanding Workshop (ASRU-2019)*, 2019, pp. 449–456. doi:10.1109/ASRU46091.2019.9003750
18. Latif S., Zaidi A., Cuayahuitl H., Shamshad F., Shoukat M., Qadir J. Transformers in speech processing: A survey. *arXiv preprint*, 2023. arXiv:2303.11607. doi:10.48550/arXiv.2303.11607
19. Zeyer A., Bahar P., Irie K., Schluter R., Ney H. A comparison of transformer and LSTM encoder decoder models for ASR. *Proc. of 2019 IEEE Automatic Speech Recognition and Understanding Workshop (ASRU-2019)*, 2019, pp. 8–15. doi:10.1109/ASRU46091.2019.9004025
20. Li J., Wu Y., Gaur Y., Wang C., Zhao R., Liu S. On the comparison of popular end-to-end models for large scale speech recognition. *Proc. of the 21st Annual Conf. of the International Speech Communication Association (Interspeech-2020)*, 2020, pp. 1–5. doi:10.21437/Interspeech.2020-2846

21. Wu C., Wang Y., Shi Y., Yeh C.-F., Zhang F. Streaming Transformer-based acoustic models using self-attention with augmented memory. *Proc. of the 21st Annual Conf. of the International Speech Communication Association (Interspeech-2020)*, 2020, pp. 2132–2136. doi:10.21437/Interspeech.2020-2079
22. Zhang Y., Chen G., Yu D., Yao K., Khudanpur S., Glass J. Highway long short-term memory RNNs for distant speech recognition. *Proc. of 2016 IEEE Intern. Conf. on Acoustics, Speech and Signal Processing (ICASSP-2016)*, 2016, pp. 5755–5759. doi:10.1109/ICASSP.2016.7472780
23. Gulati A., Qin J., Chiu C.-C., Parmar N., Zhang Y., Yu J., Han W., Wang S., Zhang Z., Wu Y., Pang R. Conformer: Convolution-augmented Transformer for speech recognition. *Proc. of the 21st Annual Conf. of the International Speech Communication Association (Interspeech-2020)*, 2020, pp. 5036–5040. doi:10.21437/Interspeech.2020-3015
24. Park D. S., Chan W., Zhang Y., Chiu C.-C., Zoph B., Cubuk E. D., Le Q. V. SpecAugment: A simple data augmentation method for automatic speech recognition. *Proc. of the 20th Annual Conf. of the International Speech Communication Association (Interspeech-2019)*, 2019, pp. 2613–2617. doi:10.21437/Interspeech.2019-2680
25. Radfar M., Lyskawa P., Trujillo B., Xie Y., Zhen K., Heymann J., Filimonov D., Strimel G., Susanj N., Mouchtaris T. Conmer: Streaming conformer without self-attention for interactive voice assistants. *Proc. of the 24th Annual Conf. of the International Speech Communication Association (Interspeech-2023)*, 2023, pp. 2198–2202. doi:10.21437/Interspeech.2023-2228
26. Mai F., Zuluaga-Gomez J., Parcollet T., Motlice P. HyperConformer: Multi-head HyperMixer for efficient speech recognition. *Proc. of the 24th Annual Conf. of the International Speech Communication Association (Interspeech-2023)*, 2023, pp. 2213–2217. doi:10.21437/Interspeech.2023-1611
27. Mai F., Pannatier A., Fehr F., Chen H., Marelli F., Fleuret F., Henderson J. Hypermixer: An MLP-based low cost alternative to transformers. *Proc. of the 61st Annual Meeting of the Association for Computational Linguistics*, 2023, vol. 1: Long Papers, pp. 15632–15654. doi:10.18653/v1/2023.acl-long.871
28. Carvalho C., Abad A. Memory-augmented conformer for improved end-to-end long-form ASR. *Proc. of the 24th Annual Conf. of the International Speech Communication Association (Interspeech-2023)*, 2023, pp. 2218–2222. doi:10.21437/Interspeech.2023-893
29. Graves A., Fernández S., Gomez F., Schmidhuber J. Connectionist temporal classification: labelling unsegmented sequence data with recurrent neural networks. *Proc. of the 23rd Intern. Conf. on Machine Learning (ICML-2006)*, 2006, pp. 369–376. doi:10.1145/1143844.1143891
30. Moritz N., Hori T., Le J. Streaming automatic speech recognition with the Transformer model. *Proc. of 2020 IEEE Intern. Conf. on Acoustics, Speech and Signal Processing (ICASSP-2020)*, 2020, pp. 6074–6078. doi:10.1109/ICASSP40776.2020.9054476
31. Burchi M., Vielzeuf V. Efficient Conformer: Progressive downsampling and grouped attention for automatic speech recognition. *Proc. of 2021 IEEE Automatic Speech Recognition and Understanding Workshop (ASRU-2021)*, 2021, pp. 8–15. doi:10.1109/ASRU51503.2021.9687874
32. Guo H., Chen Y., Xie X., Xu G., Guo W. Efficient Conformer-based CTC model for intelligent cockpit speech recognition. *Proc. of 13th Intern. Symp. on Chinese Spoken Language Processing (ISCSLP-2022)*, 2022, pp. 522–526. doi:10.1109/ISCSLP57327.2022.10037993
33. Chan W., Jaitly N., Le Q. V., Vinyals O. Listen, attend and spell: A neural network for large vocabulary conversational speech recognition. *Proc. of 2016 IEEE Intern. Conf. on Acoustics, Speech and Signal Processing (ICASSP-2016)*, 2016, pp. 4960–4964. doi:10.1109/ICASSP.2016.7472621
34. Chorowski J. K., Bahdanau D., Serdyuk D., Cho K., Bengio Y. Attention-based models for speech recognition. *Proc. of the 28th Intern. Conf. on Neural Information Processing Systems (NIPS-2015)*, 2015, vol. 1, pp. 577–585.
35. Chiu C.-C., Raffel C. Monotonic chunkwise attention. *arXiv preprint*, 2018. arXiv:1712.05382. doi:10.48550/arXiv.1712.05382
36. Inaguma H., Kawahara T. Alignment knowledge distillation for online streaming attention-based speech recognition. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 2021, vol. 31, pp. 1371–1385. doi:10.1109/TASLP.2021.313332
37. Lugosh L. *Sequence-to-sequence learning with Transducers*. Available at: <https://lorenlugosch.github.io/posts/2020/11/transducer> (accessed 08 August 2024).
38. Graves A. Sequence transduction with recurrent neural networks. *arXiv preprint*, 2012. arXiv:1211.3711. doi:10.48550/arXiv.1211.3711
39. Karmakar P., Teng S. W., Lu G. Thank you for attention: A survey on attention-based artificial neural networks for automatic speech recognition. *Intelligent Systems with Applications*, 2024, vol. 23, Article 200406. doi:10.1016/j.iswa.2024.200406
40. Chiu C.-C., Narayanan A., Han W., Prabhavalkar R., Zhang Y., Jaitly N., Pang R., Sainath T. N., Nguyen P., Cao L., Wu Y. RNN-T models fail to generalize to out-of-domain audio: Causes and solutions. *Proc. of 2021 IEEE Spoken Language Technology Workshop (SLT-2021)*, 2021, pp. 873–880. doi:10.1109/SLT48900.2021.9383518
41. Shi Y., Wang Y., Wu C., Yeh C.-F., Chan J., Zhang F., Le D., Seltzer M. Emformer: Efficient memory Transformer based acoustic model for low latency streaming speech recognition. *Proc. of 2021 IEEE Intern. Conf. on Acoustics, Speech and Signal Processing (ICASSP-2021)*, 2021, pp. 6783–6787. doi:10.1109/ICASSP39728.2021.9414560
42. Tsunoo E., Kashiwagi Y., Watanabe S. Streaming Transformer ASR with blockwise synchronous beam search. *Proc. of 2021 IEEE Spoken Language Technology Workshop (SLT-2021)*, 2021, pp. 22–29. doi:10.1109/SLT48900.2021.9383517
43. Zhang Q., Lu H., Sak H., Tripathi A., McDermott E., Koo S., Kumar S. Transformer Transducer: A streamable speech recognition model with Transformer Encoders and RNN-T loss. *Proc. of 2020 IEEE Intern. Conf. on Acoustics, Speech and Signal Processing (ICASSP-2020)*, 2020, pp. 7829–7833. doi:10.1109/ICASSP40776.2020.9053896
44. Liu C., Zhang F., Le D., Kim S., Saraf Y., Zweig G. Improving RNN Transducer based ASR with auxiliary tasks. *Proc. of 2021 IEEE Spoken Language Technology Workshop (SLT-2021)*, 2021, pp. 172–179. doi:10.1109/SLT48900.2021.9383548
45. Tripathi A., Kim J., Zhang Q., Lu H., Sak H. Transformer Transducer: One model unifying streaming and non-streaming speech recognition. *arXiv preprint*, 2020. arXiv:2010.03192. doi:10.48550/arXiv.2010.03192
46. Cui M., Kang J., Deng J., Yin X., Xie Y., Chen X., Liu X. Towards effective and compact contextual representation for Conformer Transducer speech recognition systems. *Proc. of the 24th Annual Conf. of the International Speech Communication Association (Interspeech-2023)*, 2023, pp. 2223–2227. doi:10.21437/Interspeech.2023-552
47. Chen X., Wu Y., Wang Z., Liu S., Li J. Developing real-time streaming Transformer Transducer for speech recognition on large-scale dataset. *Proc. of 2021 IEEE Intern. Conf. on Acoustics, Speech and Signal Processing (ICASSP-2021)*, 2021, pp. 5904–5908. doi:10.1109/ICASSP39728.2021.9413535
48. Barcovschi A., Jain R., Corcoran P. A comparative analysis between Conformer-Transducer, Whisper, and wav2vec2 for improving the child speech recognition. *Proc. of 2023 Intern. Conf. on Speech Technology and Human-Computer Dialogue (SpD-2023)*, 2023, pp. 42–47. doi:10.1109/SpD59241.2023.10314867
49. Reakesh D., Koluguri N. R., Krivan S., Majumdar S., Noroozi V., Huang H., Hrinchuk O., Puvvada K., Kumar A., Balam J., Ginsburg B. Fast Conformer with linearly scalable attention for efficient speech recognition. *Proc. of 2023 IEEE Automatic Speech Recognition and Understanding Workshop (ASRU-2023)*, pp. 1–8. doi:10.1109/ASRU57964.2023.10389701
50. Nakatani T. Improving transformer-based end-to-end speech recognition with connectionist temporal classification and language model integration. *Proc. of the 20th Annual Conf. of the International Speech Communication Association (Interspeech-2019)*, 2019, pp. 1408–1412. doi:10.21437/Interspeech.2019-1938
51. Schneider S., Baevski A., Collobert R., Auli M. Wav2vec: unsupervised pre-training for speech recognition. *Proc. of the 20th Annual Conf. of the International Speech Communication Association (Interspeech-2019)*, 2019, pp. 3465–3469. doi:10.21437/Interspeech.2019-1873
52. Jovanović M., Campbell M. Generative artificial intelligence: Trends and prospects. *Computer*, 2022, vol. 55, no. 10, pp. 107–112. doi:10.1109/MC.2022.3192720
53. Baevski A., Zhou Y., Mohamed A., Auli M. Wav2vec 2.0: A framework for self-supervised learning of speech representations. *Proc. of the 34th Intern. Conf. on Neural Information Processing Systems (NIPS-2020)*, 2020, pp. 12449–12460.
54. Safonova A., Yudina T., Nadimanov E., Davenport C. Automatic speech recognition of low-resource languages based on Chukchi. *arXiv preprint*, 2022. arXiv:2210.05726. doi:10.48550/arXiv.2210.05726



Модель формирования образа сцены в телевизионном модуле мобильного робота, движущегося по пересеченной местности

Т. А. Акименко^а, канд. техн. наук, доцент, orcid.org/0000-0003-1204-2657, tantan72@mail.ru

Е. В. Ларкин^а, доктор техн. наук, профессор, orcid.org/0000-0001-7471-1524

А. В. Богомолов^б, доктор техн. наук, профессор, orcid.org/0000-0002-7582-1802

А. Н. Привалов^в, доктор техн. наук, профессор, orcid.org/0000-0003-3311-0751

^аТульский государственный университет, Ленина пр., 92, Тула, 300012, РФ

^бФедеральный исследовательский центр «Информатика и управление» РАН, Вавилова ул., 44, кор. 2, Москва, 119333, РФ

^вТульский государственный педагогический университет им. Л. Н. Толстого, Ленина пр., 125, Тула, 300026, РФ

Введение: эффективность применения мобильных роботов для телевизионного мониторинга оперативной обстановки определяется качеством регистрируемого видеоряда, в связи с чем требуется сформировать математический аппарат, позволяющий прогнозировать качество изображения. **Цель:** разработать модель формирования образа сцены в телевизионном модуле мобильного робота, движущегося по пересеченной местности, адекватно учитывающую особенности механики движения мобильного робота и принципы функционирования его телевизионного модуля. **Результаты:** разработана математическая модель перемещения мобильного робота, осуществляющего телевизионный мониторинг обстановки, по пересеченной местности. Робот рассмотрен как платформа, закрепленная на двух рядах вязкоупругих опор, которые демпфируют колебания платформы при движении робота. Описан механизм возникновения продольных и поперечных колебаний платформы с установленной на ней стойкой с телевизионным модулем. Показано влияние колебаний платформы на изображение сцены, регистрируемой в фокальной плоскости объектива телевизионного модуля. Получена зависимость параметров перемещения изображения от механических параметров мобильного робота, угловых координат установки телевизионного модуля на платформе и автокорреляционной функции, описывающей траекторию робота. Описано динамическое размытие изображения, формируемого в фотоэлектронных преобразователях, в зависимости от скорости перемещения объекта в фокальной плоскости объектива телевизионного модуля и времени накопления заряда в ячейках фотоэлектронного преобразователя. **Практическая значимость:** построенная модель позволяет адекватно прогнозировать свойства изображения, получаемого бортовым телевизионным модулем мобильного робота, что обуславливает необходимость ее применения при проектировании и эксплуатации мобильных роботов телевизионного мониторинга обстановки.

Ключевые слова — мобильный робот, телевизионный модуль, продольные колебания платформы, поперечные колебания платформы, анализ сцен, обработка изображения, размытие изображения.

Для цитирования: Акименко Т. А., Ларкин Е. В., Богомолов А. В., Привалов А. Н. Модель формирования образа сцены в телевизионном модуле мобильного робота, движущегося по пересеченной местности. *Информационно-управляющие системы*, 2024, № 5, с. 16–24. doi:10.31799/1684-8853-2024-5-16-24, EDN: VWRHRF

For citation: Akimenko T. A., Larkin E. V., Bogomolov A. V., Privalov A. N. Model of scene image formation in the television module of a mobile robot moving over rough terrain. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2024, no. 5, pp. 16–24 (In Russian). doi:10.31799/1684-8853-2024-5-16-24, EDN: VWRHRF

Введение

В настоящее время мобильные роботы (МР) широко используются как средства телевизионного мониторинга оперативной обстановки [1–4]. Эффективность их применения определяется, в том числе, возможностями обеспечения качественного видеоряда при движении по пересеченной местности [5–7]. В наибольшей степени этим требованиям соответствуют МР на гусеничных шасси, состоящих из N вязкоупругих опор, на которых расположена платформа с закрепленным на ней телевизионным модулем (ТВ-модулем) [8, 9].

При движении МР изображение сцены, формируемое объективом ТВ-модуля в плоскости фотоэлектронных преобразователей, перемещается относительно центра преобразователя, поэтому динамика движения пространства этого оптического сигнала суммируется с динамикой фотоэлектронного процесса измерения интенсивности света. Результатом суперпозиции этих процессов является искажение изображения [10–12].

При наблюдении сцены человеком-оператором потери визуальной информации, связанные с искажениями, могут быть компенсированы способностью человека адаптироваться к усло-

виям наблюдения [13]. При автоматическом анализе сцен полезная информация, зашумленная искажениями, теряется гораздо чаще [14, 15]. Поэтому работу МР необходимо организовать так, чтобы минимизировать потери зрительной информации.

Для этого при разработке систем телевизионного мониторинга оперативной обстановки с применением МР необходима модель формирования образа сцены в ТВ-модуле МР, движущегося по пересеченной местности, адекватно учитывающая особенности механики движения МР и принципы работы ТВ-модуля [16]. Методы аналитического описания наблюдения сцены с МР не получили широкого распространения в инженерной практике, что обуславливает актуальность исследований в этой предметной области.

Математическое моделирование колебаний платформы мобильного робота

Динамическое размытие изображения, формируемое ТВ-модулем, вызвано поперечными колебаниями платформы МР, возникающими при движении вязкоупругих опор, на которых она установлена, по неровной подстилающей поверхности, причем каждая опора организована согласно кинематической схеме, представленной на рис. 1 [17–20].

На рисунке $xOyz$ обозначена геодезическая система координат, $x'O'y'z'$ – система координат, связанная с платформой (для простоты центры O и O' систем координат совпадают и помещены в центр масс платформы). Также принято допущение о том, что МР как транспортное средство симметричен относительно продольной вертикальной плоскости, проходящей через центр масс платформы, причем опоры левого (l) и правого (p) бортов расположены симметрично на расстоянии $\pm L_y$ от продольной плоскости, а опо-

ры с каждой стороны с одинаковым номером расположены по координате x' на расстояниях L_{x1}, \dots, L_{xN} . Центр масс платформы находится в плоскости точек крепления опор, а ТВ-модуль размещается в центре масс и может вращаться соответственно системе координат $x'O'y'z'$ на углы ψ' (азимут) и ϑ' (возвышение).

Опоры имеют единственную степень подвижности относительно платформы, а именно ортогональную плоскость узлов крепления. Платформа не имеет продольной и поперечной подвижности относительно опор шасси, поэтому центры колес могут перемещаться только по прямым, перпендикулярным плоскости подвески платформы, а координаты центров колес описываются уравнениями [20, 21]

$$\begin{cases} x'_{ln} = L_{xn} = \text{const}; \\ y'_{ln} = L_y = \text{const}; \\ z'_{ln} = -S_{ln}^0 + \delta_{hln}(t); \end{cases} \quad \begin{cases} x'_{rn} = L_{rn} = \text{const}; \\ y'_{rn} = -L_y = \text{const}; \\ z'_{rn} = -S_{rn}^0 + \delta_{hrn}(t), \end{cases}$$

где S_{ln}^0, S_{rn}^0 – начальные длины опор, измеренные при расположении МР на плоской горизонтальной плоскости; $\delta_{hln}(t), \delta_{hrn}(t)$ – значения шероховатости под n -м левым (правым) колесом, $1 \leq n \leq N$; t – время.

Пространственное положение ТВ-модуля определяется углами касания ϑ и качения γ , поэтому вектор (x'_v, y'_v, z'_v) имеет геодезические координаты:

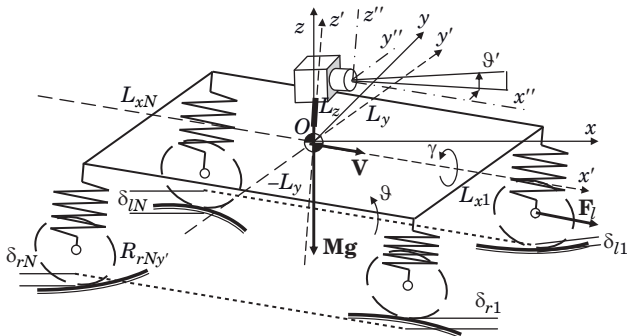
$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \mathbf{A}_\gamma \mathbf{A}_\vartheta \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}; \quad \begin{pmatrix} x'_v \\ y'_v \\ z'_v \end{pmatrix} = \mathbf{A}_\gamma \mathbf{A}_\vartheta \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix},$$

где $\mathbf{A}_\gamma, \mathbf{A}_\vartheta$ – матрицы поворота на углы ϑ, γ соответственно:

$$\mathbf{A}_\gamma = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \gamma & -\sin \gamma \\ 0 & \sin \gamma & \cos \gamma \end{pmatrix}; \quad \mathbf{A}_\vartheta = \begin{pmatrix} \cos \vartheta & 0 & \sin \vartheta \\ 0 & 1 & 0 \\ -\sin \vartheta & 0 & \cos \vartheta \end{pmatrix}.$$

Согласно кинематической схеме система дифференциальных уравнений для описания продольных и поперечных колебаний платформы МР имеет следующий вид [22–25]:

$$\begin{aligned} M\ddot{\delta}_z + 2\eta \sum_{n=1}^N (\dot{\delta}_z + \dot{\delta}_\vartheta L_{xn}) + 2c \sum_{n=1}^N (\delta_z + \delta_\vartheta L_{xn}) = \\ = \sum_{n=1}^N [\eta \dot{\delta}_{hl1}(t - \tau_n) + \eta \dot{\delta}_{hr1}(t - \tau_n) + \\ + c\delta_{hl1}(t - \tau_n) + c\delta_{hr1}(t - \tau_n)]; \end{aligned}$$



■ **Рис. 1.** Кинематическая схема мобильного робота
 ■ **Fig. 1.** Kinematic diagram of a mobile robot

$$J_y \ddot{\delta}_g + 2\eta \sum_{n=1}^N (\dot{\delta}_z + \dot{\delta}_g L_{xn}) L_{xn} + 2c \sum_{n=1}^N (\delta_z + \delta_g L_{xn}) L_{xn} = \sum_{n=1}^N [\eta \dot{\delta}_{hl1}(t - \tau_n) + \eta \dot{\delta}_{hr1}(t - \tau_n) + c \delta_{hl1}(t - \tau_n) + c \delta_{hr1}(t - \tau_n)] L_{xn};$$

$$J_x \ddot{\delta}_\gamma + 2\eta N \dot{\delta}_\gamma L_y^2 + 2c N \delta_\gamma L_y^2 = L_y \sum_{n=1}^N [\eta \dot{\delta}_{hl1}(t - \tau_n) + c \delta_{hl1}(t - \tau_n) - \eta \dot{\delta}_{hr1}(t - \tau_n) - c \delta_{hr1}(t - \tau_n)],$$

где M – масса платформы; J_y, J_x – моменты инерции платформы относительно осей y и x соответственно; $\ddot{\delta} \dots = \frac{d^2 \delta \dots}{dt^2}$; $\dot{\delta} \dots = \frac{d \delta \dots}{dt}$; N – количество рядов опор; $\delta_z, \delta_g, \delta_\gamma$ – отклонение линейных z и угловых ϑ, γ координат от их устойчивых значений; $\eta_{l1} = \eta_{r1} = \dots = \eta_{ln} = \eta_{rn} = \eta_{1N} \dots \eta_{rN} = \eta$ – коэффициент вязкого трения; $c_{l1} = c_{r1} = \dots = c_{ln} = c_{rn} = c_{1N} \dots c_{rN}$ – коэффициенты жесткости; $\tau_n = \frac{L_{xn} - L_{x1}}{V}$ – задержка прохождения некоторой точки траектории n -й парой колес по сравнению с ее прохождением первым рядом колес; V – скорость движения МР.

Преобразование Фурье записанной системы дифференциальных уравнений выглядит следующим образом:

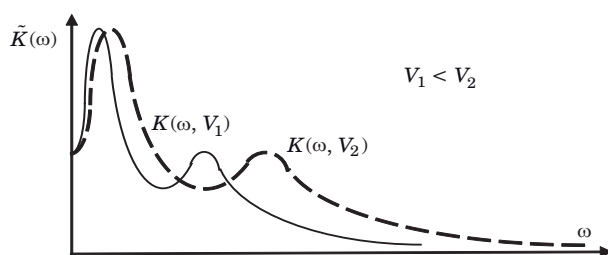
$$\begin{aligned} \tilde{\delta}_z(i\omega) \left[M(i\omega)^2 + 2N\eta i\omega + 2c \right] + \tilde{\delta}_g(i\omega) \left(2 \sum_{n=1}^N L_{xn} \right) [\eta i\omega + c] &= \\ = \sum_{n=1}^N [\tilde{\delta}_{hln}(i\omega) + \tilde{\delta}_{hrn}(i\omega)] [\eta i\omega + c] \exp(-\tau_n i\omega); \\ \tilde{\delta}_g(i\omega) \left[J_y(i\omega)^2 + 2\eta \sum_{n=1}^N L_{xn}^2 i\omega + 2c \sum_{n=1}^N L_{xn}^2 \right] + \tilde{\delta}_z(s) \left(2\eta \sum_{n=1}^N L_{xn} i\omega + 2c \sum_{n=1}^N L_{xn} \right) &= \\ = \sum_{n=1}^N L_{xn} [\tilde{\delta}_{hln}(i\omega) + \tilde{\delta}_{hrn}(i\omega)] (\eta i\omega + c) \exp(-\tau_n i\omega); \\ \tilde{\delta}_\gamma(i\omega) \left[J_x(i\omega)^2 + 2\eta N L_y^2 i\omega + 2c N L_y^2 \right] &= \\ = L_y \sum_{n=1}^N [\tilde{\delta}_{hln}(i\omega) - \tilde{\delta}_{hrn}(i\omega)] (\eta i\omega + c) \exp(-\tau_n i\omega), \end{aligned} \quad (1)$$

где ω – круговая частота; $i = \sqrt{-1}$; $\tilde{\delta} \dots(i\omega) = F[\delta \dots(t)]; F[\delta \dots(t)]$ – прямое преобразование Фурье [26, 27]. Система уравнений (1) может быть решена относительно $\delta_z(s), \delta_g(s), \delta_\gamma(s)$ следующим образом:

$$\tilde{\delta}_z(i\omega) = \frac{\Delta_z(i\omega)}{\Delta(i\omega)}; \tilde{\delta}_g(i\omega) = \frac{\Delta_g(i\omega)}{\Delta(i\omega)}; \tilde{\delta}_\gamma(i\omega) = \frac{\Delta_\gamma(i\omega)}{\Delta(i\omega)},$$

где $\Delta(i\omega)$ – главный определитель системы; $\Delta_z(i\omega), \Delta_g(i\omega), \Delta_\gamma(i\omega)$ – определители $\delta_z(i\omega), \delta_g(i\omega)$ и $\delta_\gamma(i\omega)$ соответственно, которые можно получить заменой первого, второго и третьего столбца на столбец $\Delta(i\omega)$;

$$C = \begin{pmatrix} \sum_{n=1}^N [\delta_{hln}(i\omega) + \delta_{hrn}(i\omega)] (\eta i\omega + c) \exp(-\tau_n i\omega) \\ \sum_{n=1}^N L_{xn} [\delta_{hln}(i\omega) + \delta_{hrn}(i\omega)] (\eta i\omega + c) \exp(-\tau_n i\omega) \\ L_y \sum_{n=1}^N [\delta_{hln}(i\omega) - \delta_{hrn}(i\omega)] (\eta i\omega + c) \exp(-\tau_n i\omega) \end{pmatrix};$$



■ **Рис. 2.** Спектральная плотность продольных колебаний
 ■ **Fig. 2.** Spectral density of longitudinal vibrations

$$\Delta(s) = \begin{vmatrix} M(i\omega)^2 + 2N\eta i\omega + 2c & \left(2 \sum_{n=1}^N L_{xn}\right)(\eta i\omega + c) & 0 \\ 2\eta \sum_{n=1}^N L_{xn} i\omega + 2c \sum_{n=1}^N L_{xn} & J_y(i\omega)^2 + 2\eta \sum_{n=1}^N L_{xn}^2 i\omega + 2c \sum_{n=1}^N L_{xn}^2 & 0 \\ 0 & 0 & J_x(i\omega)^2 + 2\eta N L_y^2 i\omega + 2c N L_y^2 \end{vmatrix}.$$

Возникновение колебаний вызвано продольным и поперечным перемещением МР по пересеченной местности, неровность которой может быть описана автокорреляционной функцией

$$K_{l(r)}(\Xi) = \int_{-\infty}^{\infty} \delta_{hl(r)1}(\xi) \delta_{hl(r)1}(\Xi + \xi) d\xi,$$

показывающей связь клиренса под первым левым (правым) колесом с высотой под этим же колесом в точках $\delta_{hl(r)1}(\xi)$, сдвинутых относительно начальной продольной координаты ξ на расстояние Ξ .

Автокорреляционная функция позволяет построить частотный спектр механического сигнала, вызывающего поперечные колебания платформы МР (рис. 2). Спектральная плотность напрямую зависит от скорости V продольного движения МР: увеличение скорости V приводит к увеличению ширины плотности (показано пунктирной линией).

Обратное преобразование Фурье характеризует продольные и поперечные колебания платформы МР с размещенным на ней ТВ-модулем:

$$\delta_z(t) = F^{-1}[\tilde{\delta}_z(i\omega)]; \delta_y(t) = F^{-1}[\tilde{\delta}_y(i\omega)]; \delta_\gamma(t) = F^{-1}[\tilde{\delta}_\gamma(i\omega)],$$

где $F^{-1}[\dots]$ – обратное преобразование Фурье [12, 13]. Скорость изменения $\delta_z(t)$, $\delta_y(t)$ и $\delta_\gamma(t)$ можно определить как

$$\dot{\delta}_z(t) = F^{-1}[i\omega \tilde{\delta}_z(i\omega)]; \dot{\delta}_y(t) = F^{-1}[i\omega \tilde{\delta}_y(i\omega)]; \dot{\delta}_\gamma(t) = F^{-1}[i\omega \tilde{\delta}_\gamma(i\omega)],$$

а для расчета оценки динамического размытия необходимо найти максимальные значения $\dot{\delta}_z(t)$, $\dot{\delta}_y(t)$ и $\dot{\delta}_\gamma(t)$.

Анализ динамического размытия изображения

Пусть МР наблюдает за неподвижным точечным источником света B , имеющим геодезические координаты x_B, y_B, z_B . В системе координат $x''O''y''z''$, центр O'' которой связан с центральным входным зрачком объектива ТВ-модуля, оси x'' и z'' направлены соответственно к источнику и вверх, а ось y'' дополняет систему до правой ортогональной системы координат, источник B имеет координаты

$$\begin{pmatrix} x''_B \\ y''_B \\ z''_B \end{pmatrix} = \mathbf{A}_{g'} \mathbf{A}_{\psi'} \left[\mathbf{A}_\gamma^{-1} \mathbf{A}_g^{-1} \begin{pmatrix} x_B \\ y_B \\ z_B \end{pmatrix} - \begin{pmatrix} 0 \\ 0 \\ L_z \end{pmatrix} \right] - \begin{pmatrix} 0 \\ 0 \\ z_c \end{pmatrix},$$

где \mathbf{A}_{ϑ} , \mathbf{A}_{ψ} – матрицы поворота ТВ-модуля относительно платформы по углу места ϑ' и углу азимута ψ' соответственно:

$$\mathbf{A}_{\vartheta'} = \begin{pmatrix} \cos \vartheta' & 0 & -\sin \vartheta' \\ 0 & 1 & 0 \\ \sin \vartheta' & 0 & \cos \vartheta' \end{pmatrix};$$

$$\mathbf{A}_{\psi'} = \begin{pmatrix} \cos \psi' & \sin \psi' & 0 \\ -\sin \psi' & \cos \psi' & 0 \\ 0 & 0 & 1 \end{pmatrix};$$

L_z – координата расположения ТВ-модуля, связанного с платформой; z_c – координата центра масс.

Дифференцирование по времени уравнения координат источника B дает составляющие скорости точечного источника света относительно ТВ-модуля, вызванные колебаниями платформы относительно опор:

$$\begin{pmatrix} \dot{x}_B'' \\ \dot{y}_B'' \\ \dot{z}_B'' \end{pmatrix} = \frac{d}{dt} \left\{ \mathbf{A}_{\vartheta'} \mathbf{A}_{\psi'} \left[\mathbf{A}_{\gamma}^{-1} \mathbf{A}_{\vartheta}^{-1} \begin{pmatrix} x_B \\ y_B \\ z_B \end{pmatrix} - \begin{pmatrix} 0 \\ 0 \\ L_z \end{pmatrix} \right] - \begin{pmatrix} 0 \\ 0 \\ z_c \end{pmatrix} \right\}. \quad (2)$$

В соответствии с уравнением (2) переменными, которые следует дифференцировать, являются γ , ϑ , ψ , входящие в \mathbf{A}_{γ} , \mathbf{A}_{ϑ} , \mathbf{A}_{ψ} и z_B . Если расстояние до точечного источника намного больше фокусного расстояния объектива, то скорость движения изображения точечного источника в фокальной плоскости равна

$$(\dot{Y}_B, \dot{Z}_B) = f \frac{d}{dt} \begin{pmatrix} y_B'' \\ x_B'' \end{pmatrix},$$

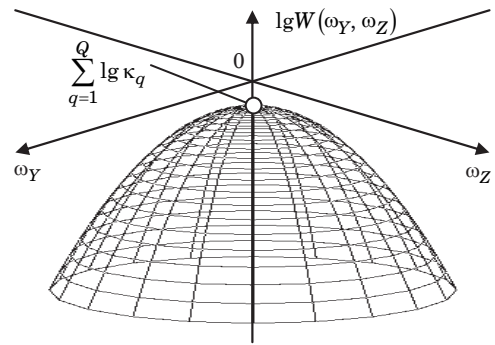
где f – фокусное расстояние линзы.

Известно, что если по пути оптического сигнала ТВ-модуля расположены Q элементов, каждый из которых способствует ухудшению разрешающей способности ТВ-модуля, то его логарифмическую пространственно-частотную характеристику можно определить как (рис. 3)

$$\lg W(\omega_Y, \omega_Z) = \sum_{q=1}^Q \lg \kappa_q - 2\pi^2 (\omega_Y^2 + \omega_Z^2) \sum_{q=1}^Q \sigma_q^2,$$

где ω_Y, ω_Z – пространственные частоты по координатам Y и Z соответственно; κ_q – статический логарифмический коэффициент передачи; σ_q – параметр, определяющий пропускную способность q -го элемента.

В фотоэлектронных преобразователях, построенных на приборах с зарядовой связью (ПЗС), за время накопления происходит объеди-



■ **Рис. 3.** Стандартная логарифмическая пространственно-частотная характеристика ТВ-модуля

■ **Fig. 3.** Standard logarithmic spatial frequency response of a TV module

нение зарядов, образовавшихся под действием света от точечного источника, поэтому помимо учтенных Q на пути оптического сигнала появляется $(Q + 1)$ -й элемент, который фильтрует только сигналы, динамически перемещающиеся в пространстве.

Предположим, что при накоплении заряда в ПЗС-фотоэлектронном преобразователе несколько изменяются как интенсивность точечного источника, так и составляющие \dot{Y}_B, \dot{Z}_B скорости. Это означает, что в момент времени T оптический сигнал преобразуется в электрическую энергию оптического сигнала, равномерно распределенную по координатам Y и Z на расстояниях $\dot{Y}_B T, \dot{Z}_B T$ соответственно. Дисперсии

таких распределений равны $\sigma_{Q+1,Y}^2 = \frac{(\dot{Y}_B T)^2}{12}$ и $\sigma_{Q+1,Z}^2 = \frac{(\dot{Z}_B T)^2}{12}$ [28–30]. Таким образом, полоса пропускания $(Q + 1)$ -го элемента универсально пропорциональна времени накопления заряда и скорости перемещения изображения сцены в фокальной плоскости объектива.

Примеры динамического размытия изображения

Изображения, полученные с МР, движущегося по пересеченной местности, показаны на рис. 4, а–г.

Видно, что продольные и поперечные колебания приводят к ухудшению изображения, в результате чего оно становится размытым при определенных параметрах. Колебания платформы с ТВ-модулем по углу места практически незаметны визуально (см. рис. 4, б), но существенно усложняют алгоритмы анализа сцен, поскольку в этом случае пропадают некоторые особенности сцены, содержащиеся в высоких частотах верти-



■ **Рис. 4.** Изображения сцены, полученные для разных условий наблюдения: *a* – исходное изображение, полученное с неподвижного МР; *б* – изображение при продольных вертикальных колебаниях платформы МР; *в* – поперечные горизонтальные колебания платформы; *г* – совмещение продольных и поперечных колебаний платформы

■ **Fig. 4.** Images of the scene obtained under different observation conditions: *a* – original image obtained from a stationary MR; *б* – image with longitudinal vertical oscillations of the MR platform; *в* – transverse horizontal oscillations of the platform; *г* – combination of longitudinal and transverse oscillations of the platform

кального пространства. Поперечные колебания ТВ-модуля приводят к ухудшению качества по краям изображения (см. рис. 4, в). Это означает, что элементы сцены, содержащиеся в этой области, могут быть пропущены при обработке.

При наличии в кадре движущегося объекта предложенное динамическое описание размытия изображения работает аналогичным образом: ТВ-модуль мобильного робота в совокупности с платформой, расположенной на подвижном основании, считается единым прибором, формирующим цифровой образ сцены безотносительно того, подвижный или неподвижный объект на ней расположен. Поскольку даже в неподвижном ТВ-модуле проекция динамического объекта на плоскость расположения фоточувствительных элементов может перемещаться относительно плоской системы координат, связанной с фоточувствительной матрицей, динамическое размытие изображений присутствует практически всегда. Суммарный эффект искажения изображения определяется суммой векторов скорости

точек проекции, создаваемых собственными движениями ТВ-модуля и объекта в земной системе координат. Вопросы получения численных оценок величин $\tilde{\delta}_z(t)$, $\tilde{\delta}_g(t)$ и $\tilde{\delta}_\gamma(t)$ требуют исследования, а их максимумы могут быть найдены как максимумы производных от обратного преобразования Фурье величин

$$\tilde{\delta}_z(i\omega) = \frac{\Delta_z(i\omega)}{\Delta(i\omega)}; \quad \tilde{\delta}_g(i\omega) = \frac{\Delta_g(i\omega)}{\Delta(i\omega)};$$

$$\tilde{\delta}_\gamma(i\omega) = \frac{\Delta_\gamma(i\omega)}{\Delta(i\omega)}.$$

Размытие изображений, вызываемое рассмотренными выше причинами, затрудняет задачу выделения заданных объектов в потоке формируемых видеоданных. Поэтому при разработке требований к алгоритмическому и программному обеспечению систем искусственного интеллекта, производящих анализ сцен, наблюдаемых

ТВ-модулем, размещенным на подрессоренной платформе МР, следует учитывать как конструктивные особенности робота, так и условия его эксплуатации, а именно характер пересеченной местности и скорость, на которой производится формирование видеосигнала [31].

Заключение

Разработанная модель формирования образа сцены в бортовом ТВ-модуле МР, движущегося по пересеченной местности, имеет существенное значение при проектировании системы наблюдения МР, поскольку она дает возможность про-

гнозировать качество изображения, получаемого ТВ-модулем. Это в свою очередь позволяет создать оптимальные алгоритмы управления МР при планировании его применения в заданных условиях движения и наблюдения.

Финансовая поддержка

Исследование выполнено при финансовой поддержке Российского научного фонда, грант № 24-29-20256 «Методы и алгоритмы синтеза иерархических систем цифрового управления многоконтурными объектами».

Литература

1. Mehmszow U. *Mobile Robotics: A Practical Introduction*. Springer-Verlag, London Ltd, 2003. 273 p.
2. Siciliano B., Khatib O. *Springer Handbook of Robotics*. Springer, 2016. 2155 p.
3. Hervas J. R., Reyhanoglu M., Tang H. Automatic landing control of unmanned aerial vehicles on moving platforms. *Industrial Electronics: IEEE 23rd Intern. Symp.*, IEEE, 2014, pp. 69–74.
4. Хрипунов С. П., Чиров Д. С., Благодарящев И. В. Военная робототехника: современные тренды и векторы развития. *Тренды и управление*, 2015, № 4, с. 410–422. doi:10.7256/2307-9118.2015.4.17117
5. Larkin E., Bogomolov A., Privalov A. Data buffering in information-measuring system. *Intern. Ural Conf. on Measurements (UralCon)*, Chelyabinsk, 16–19 October 2017, IEEE, 2017, pp. 118–123. doi:10.1109/URALCON.2017.8120697
6. Ronzhin A., Ngo T., Vu Q., Nguyen V. Theoretical Foundations to Control Technological and Robotic Operations with Physical Manipulations of Agricultural Products. *Ground and Air Robotic Manipulation Systems in Agriculture. Series: Intelligent Systems Reference Library*. Springer, Cham, 2022, vol. 214. https://doi.org/10.1007/978-3-030-86826-0_5
7. Чиров Д. С., Чертова О. Г., Потапчук Т. Н. Методика обоснования требований к системе технического зрения робототехнического комплекса. *Тр. СПИИРАН*, 2017, т. 2, № 51, с. 152–176. doi:10.15622/sp.51.7
8. Larkin E., Bogomolov A., Gorbachev D., Privalov A. About the approach of the transactions flow to control one in robot control systems. *Intern. Conf. on Interactive Collaborative Robotics. Series: Lecture Notes in Computer Science*, 2017, vol. 10459, pp. 113–122. doi:10.1007/978-3-319-66471-2_13
9. Larkin E., Bogomolov A., Privalov A., Antonov M. About one approach to robot control system simulation. *Third Intern. Conf. ICR 2018, Series: Lecture Notes in Computer Science*, Leipzig, Germany, 18–22 September 2018, vol. 11097, pp. 159–169. doi:10.1007/978-3-319-99582-3_17
10. Kopeika N. S. *A System Engineering Approach to Imaging*. SPIE Optical Engineering Press, Bellingham, WA, 1998. 704 p. doi:10.1117/3.2265069
11. Nguyen V., Saveliev A., Ronzhin A. Mathematical modeling of control and simultaneous stabilization of 3-dof aerial manipulation system. *Intern. Conf. on Interactive Collaborative Robotics. Series: Lecture Notes in Computer Science*, 2020, vol. 12336, pp. 253–264. doi:10.1007/978-3-030-60337-3_25
12. Варлашин В. В., Лопота А. В. Оптимизация параметров проецирования системы кругового обзора с использованием реперных меток. *Мехатроника, автоматизация, управление*, 2022, т. 23(2), с. 97–103. doi:10.17587/mau.23.97-103
13. Левин Г. Г., Минаев В. Л., Иванов А. Д., Голополов А. А., Горяинова И. В. Метрологическое обеспечение видеоизмерительных систем. *Измерительная техника*, 2023, № 1, с. 24–30. doi:10.32446/0368-1025it.2023-1-24-30
14. Мяо Ч., Рубцов В. И., Орлов Е. А. Повышение достоверности модели внешней среды с помощью телевизионного изображения. *Научный альманах Центрального Черноземья*, 2022, № 2-8, с. 74–84. EDN: GEXXWC
15. Жданов М. А., Сериков С. А. Аналитический обзор существующих платформ и конструктивных типов мобильных роботов-разведчиков. *Colloquium-Journal*, 2019, № 18-2 (42), с. 35–40. EDN: BGUCWW
16. Janesik J. R. *Scientific Charge-Coupled Devices*. SPIE Press, Bellingham, WA, 2001, 924 p. doi:10.1117/12.7974139
17. Dreizier R. M., Ludde C. S. *Theoretical Mechanics. Theoretical Physics 1*. Springer-Verlag, Berlin, Heidelberg, 2010. 402 p.
18. Kamaldar M., Mahjoob M., Yazdi M., Vahid-Alizadeh H., Ahmadizadeh Saeed. A control synthesis for reducing lateral oscillations of a spherical robot. *IEEE Intern. Conf. on Mechatronics*, 2011, pp. 546–551. doi:10.1109/ICMECH.2011.5971346

19. Larkin E., Bogomolov A., Antonov M. Modeling of increased rigidity of industrial manipulator. *5th Intern. Conf., ICR 2020*, Saint-Petersburg, Russia, October 7–9, 2020. *Lecture Notes in Computer Science*, 2020, vol. 12336 LNAI, pp. 170–178. doi:10.1007/978-3-030-60337-3_17
20. Плотников А. В. Кинестатическая модель захвата объектов сервисными роботами при активном перемещении мобильной платформы. М., ИПМ им. М. В. Келдыша, 2023. 21 с. (Препринт / ИПМ им. М. В. Келдыша; № 14). doi:10.20948/prepr-2023-14. https://keldysh.ru/papers/2023/prep2023_14.pdf (дата обращения: 15.02.2024).
21. Larkin E., Bogomolov A., Privalov A. Discrete model of mobile robot assemble fault-tolerance. *Lecture Notes in Computer Science*, 2019, vol. 11659 LNAI, pp. 204–215. doi:10.1007/978-3-030-26118-4_20
22. Janarthanan B., Padmanabhan C., Sujatha C. Longitudinal dynamics of a tracked vehicle: simulation and experiment. *Journal of Terramechanics*, 2012, vol. 49, no. 2, pp. 63–72.
23. Hu H., Li E., Zhao X., Liang Z., Yu W. Modeling and simulation of folding-boom aerial platform vehicle based on the flexible multi-body dynamics. *Intern. Conf. on Intelligent Control and Information Processing*, Dalian, 2010, pp. 798–802.
24. Nunez P., Jones R., Shoop S. High-fidelity ground platform and terrain mechanics modeling for military applications involving vehicle dynamics and mobility analysis. *Army tank automotive research development and Engineering Center Warren MI*, 2004, no. TAR-DEC-14339.
25. Степанов М. Ф., Степанов А. М., Степанова О. М., Петров Д. Ю. Разработка и исследование законов оптимального управления движением образовательного робота на колесной платформе. *Математические методы в технологиях и технике*, 2021, № 8, с. 67–74. doi:10.52348/2712-8873_ММТТ_2021_8_67
26. Schiff J. L. *The Laplace Transform: Theory and Applications*. USA, NY, Springer Verlag, 1991. 233 p.
27. Karnopp D. C., Margolis D. L., Rosenberg R. C. *System Dynamics: Modeling, Simulation and Control of Mechatronic Systems*. John Wiley & Sons, New Jersey, 2012. 636 p.
28. Kobayashi H., Marl B. L., Turin W. *Probability, Random Processes and Statistical Analysis*. Cambridge University Press, 2012. 812 p.
29. Larkin E. V., Akimenko T. A., Bogomolov A. V. Modeling the reliability of the onboard equipment of a mobile robot. *Izvestiya of Saratov University. Mathematics. Mechanics. Informatics*, 2021, vol. 21, no. 3, pp. 390–399. doi:10.18500/1816-9791-2021-21-3-390-399
30. Ronzhin A., Saveliev A., Basov O., Solyonyj S. Conceptual model of cyberphysical environment based on collaborative work of distributed means and mobile robots. *Lecture Notes in Computer Science*, 2016, vol. 9812, pp. 32–39. doi:10.1007/978-3-319-43955-6_5
31. Kahar S., Sulaiman R., Prabuwo A. S., Akma N., Ahmad S. A., Abu Hassan M. A. A Review of wireless technology usage for mobile robot controller. *Intern. Proc. of Computer Science and Information Technology IPCSIT*, 2012, vol. 34, pp. 7–12.

UDC 004.932

doi:10.31799/1684-8853-2024-5-16-24

EDN: VWRHRF

Model of scene image formation in the television module of a mobile robot moving over rough terrainT. A. Akimenko^a, PhD, Tech., Associate Professor, orcid.org/0000-0003-1204-2657, tantan72@mail.ruE. V. Larkin^a, Dr. Sc., Tech., Professor, orcid.org/0000-0001-7471-1524A. V. Bogomolov^b, Dr. Sc., Tech., Professor, orcid.org/0000-0002-7582-1802A. N. Privalov^c, Dr. Sc., Tech., Professor, orcid.org/0000-0003-3311-0751^aTula State University, 92, Lenina Av., 300012, Tula, Russian Federation^bFederal Research Center “Computer Science and Control” of RAS, 44, cor. 2, 119333, Vavilova St., Moscow, Russian Federation^cTula State Pedagogical University named after L. N. Tolstoy, 125, Lenina Av., 300026, Tula, Russian Federation

Introduction: The effectiveness of using mobile robots for television monitoring of the operational situation is determined by the quality of the recorded video, in connection with this, it is necessary to create a mathematical apparatus that makes it possible to predict image quality. **Purpose:** To develop a model for the formation of a scene image in the television module of a mobile robot moving over rough terrain, with this model adequately taking into account the peculiarities of the mechanics of the movement of a mobile robot and the principles of operation of its television module. **Results:** We develop a mathematical model for moving a mobile robot that carries out television monitoring of the situation over rough terrain. The robot is considered as a platform mounted on two rows of viscoelastic supports, which dampen platform vibrations when the robot moves. We describe the mechanism for the occurrence of longitudinal and transverse vibrations of the platform with a stand with a television module installed on it. We show the effect of platform vibrations on the image of a scene recorded in the focal plane of the television module lens. We have obtained the dependence of the image movement parameters on the mechanical parameters of the mobile robot, the angular coordinates of the installation of the television module on the platform and the autocorrelation function describing the robot's trajectory. We describe the dynamic blurring of the image formed in photoelectronic converters, depending on the speed of the movement of the object in the focal plane of the television module lens and the time of charge accumulation in the cells of the photoelectronic converter. **Practical relevance:** The constructed model allows us to

adequately predict the properties of the image received by the on-board television module of a mobile robot, which necessitates its use in the design and operation of mobile robots for television monitoring of the situation.

Keywords – mobile robot, television module, longitudinal vibrations of the platform, transverse vibrations of the platform, scene analysis, image processing, image blur.

For citation: Akimenko T. A., Larkin E. V., Bogomolov A. V., Privalov A. N. Model of scene image formation in the television module of a mobile robot moving over rough terrain. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2024, no. 5, pp. 16–24 (In Russian). doi:10.31799/1684-8853-2024-5-16-24, EDN: VWRHRF

Financial support

The research was financially supported by the Russian Science Foundation, grant No. 24-29-20256 “Methods and algorithms for the synthesis of hierarchical digital control systems for multi-circuit objects”.

References

- Mehmszow U. *Mobile Robotics: A Practical Introduction*. Springer-Verlag, London Ltd, 2003. 273 p.
- Siciliano B., Khatib O. *Springer Handbook of Robotics*. Springer, 2016. 2155 p.
- Hervas J. R., Reyhanoglu M., Tang H. Automatic landing control of unmanned aerial vehicles on moving platforms. *Industrial Electronics: IEEE 23rd Intern. Symp.*, IEEE, 2014, pp. 69–74.
- Khripunov S. P., Chirov D. S., Blagodaryashchev I. V. Military robotics: modern trends and development vectors. *Trendy i upravlenie* [Trends and Management], 2015, no. 4, pp. 410–422 (In Russian). doi:10.7256/2307-9118.2015.4.17117
- Larkin E., Bogomolov A., Privalov A. Data buffering in information-measuring system. *Intern. Ural Conf. on Measurements (UralCon)*, Chelyabinsk, IEEE, 2017, pp. 118–123. doi:10.1109/URALCON.2017.8120697
- Ronzhin A., Ngo T., Vu Q., Nguyen V. *Theoretical Foundations to Control Technological and Robotic Operations with Physical Manipulations of Agricultural Products*. In: *Ground and Air Robotic Manipulation Systems in Agriculture. Series: Intelligent Systems Reference Library*. Springer, Cham, 2022, vol. 214. https://doi.org/10.1007/978-3-030-86826-0_5
- Chirov D. S., Chertova O. G., Potapchuk T. N. Methods of study requirements for the complex robotic vision system. *SPIRAS Proceedings*, 2017, vol. 2, no. 51, pp. 152–176 (In Russian). doi:10.15622/sp.51.7
- Larkin E., Bogomolov A., Gorbachev D., Privalov A. About the approach of the transactions flow to control one in robot control systems. *Intern. Conf. on Interactive Collaborative Robotics. Series: Lecture Notes in Computer Science*, 2017, vol. 10459, pp. 113–122. doi:10.1007/978-3-319-66471-2_13
- Larkin E., Bogomolov A., Privalov A., Antonov M. About one approach to robot control system simulation. *Third Intern. Conf. ICR 2018, Series: Lecture Notes in Computer Science*, Leipzig, 2018, vol. 11097, pp. 159–169. doi:10.1007/978-3-319-99582-3_17
- Kopeika N. S. *A System Engineering Approach to Imaging*. SPIE Optical Engineering Press, Bellingham, WA, 1998. 704 p. doi:10.1117/3.2265069
- Nguyen V., Saveliev A., Ronzhin A. Mathematical modeling of control and simultaneous stabilization of 3-dof aerial manipulation system. *Intern. Conf. on Interactive Collaborative Robotics. Series: Lecture Notes in Computer Science*, 2020, vol. 12336, pp. 253–264. doi:10.1007/978-3-030-60337-3_25
- Varlashin V. V., Lopota A. V. Optimization of projection parameters of an all-round viewing system using reference marks. *Mechatronics, Automation, Control*, 2022, vol. 23, no. 2, pp. 97–103 (In Russian) doi:10.17587/mau.23.97-103
- Levin G. G., Minaev V. L., Ivanov A. D., Goloposov A. A., Goryainova I. V. Metrological support for video measuring systems. *Measuring Technology*, 2023, no. 1, pp. 24–30 (In Russian). doi:10.32446/0368-1025it.2023-1-24-30
- Miao Ch., Rubtsov V. I., Orlov E. A. Increasing the reliability of the model of the external environment with the help of a television image. *Nauchnyi al'manah Central'nogo Chernozem'ya* [Scientific Almanac of the Central Chernozem Region], 2022, no. 2-8, pp. 74–84 (In Russian). EDN: GEXXWC
- Zhdanov M. A., Serikov S. A. Analytical overview of existing platforms and design types of mobile robots-scouts. *Colloquium-Journal*, 2019, no. 18-2 (42), pp. 35–40 (In Russian). EDN: BGUCWW
- Janesik J. R. *Sensitive Charge-Coupled Devices*. SPIE Press, Bellingham, WA, 2001. 924 p. doi:10.1117/12.7974139
- Dreizler R. M., Ludde C. S. *Theoretical Mechanics. Theoretical Physics 1*. Springer-Verlag, Berlin, Heidelberg, 2010. 402 p.
- Kamalidar M., Mahjoob M., Yazdi M., Vahid-Alizadeh H., Ahmadizadeh Saeed. A control synthesis for reducing lateral oscillations of a spherical robot. *IEEE Intern. Conf. on Mechatronics*, 2011, pp. 546–551. doi:10.1109/ICMECH.2011.5971346
- Larkin E., Bogomolov A., Antonov M. Modeling of increased rigidity of industrial manipulator. *5th Intern. Conf. ICR 2020*, Saint-Petersburg, 2020. *Lecture Notes in Computer Science*, 2020, vol. 12336 LNAI, pp. 170–178. doi:10.1007/978-3-030-60337-3_17
- Plotnikov A. V. *Kinestatische model' zahvata ob'ektov servisyymi robotami pri aktivnom peremeshenii mobil'noi platform* [Kinetostatic model of object capture by service robots, with active movement of the mobile platform]. Preprints of Keldysh IPM RAS, 2023, no. 14. 21 p. (In Russian). doi:10.20948/prepr-2023-14. Available at: https://keldysh.ru/papers/2023/prep2023_14.pdf (accessed 15 February 2024).
- Larkin E., Bogomolov A., Privalov A. Discrete model of mobile robot assemble fault-tolerance. *Lecture Notes in Computer Science*, 2019, vol. 11659 LNAI, pp. 204–215. doi:10.1007/978-3-030-26118-4_20
- Janarthanan B., Padmanabhan C., Sujatha C. Longitudinal dynamics of a tracked vehicle: Simulation and experiment. *Journal of Terramechanics*, 2012, vol. 49, no. 2, pp. 63–72.
- Hu H., Li E., Zhao X., Liang Z., Yu W. Modeling and simulation of folding-boom aerial platform vehicle based on the flexible multi-body dynamics. *Intern. Conf. on Intelligent Control and Information Processing*, Dalian, 2010, pp. 798–802.
- Nunez P., Jones R., Shoop S. High-fidelity ground platform and terrain mechanics modeling for military applications involving vehicle dynamics and mobility analysis. *Army tank automotive research development and Engineering Center Warren MI*, 2004, no. TARDEC-14339.
- Stepanov M. F., Stepanov A. M., Stepanova O. M., Petrov D. Yu. Development and research of optimal control laws of the movement of the educational robot on the wheel platform. *Mathematical Methods in Technologies and Technics*, 2021, no. 8, pp. 67–74 (In Russian). doi:10.52348/2712-8873_MMTT_2021_8_67
- Schiff J. L. *The Laplace Transform: Theory and Applications*. USA, NY, Springer Verlag, 1991. 233 p.
- Karnopp D. C., Margolis D. L., Rosenberg R. C. *System Dynamics: Modeling, Simulation and Control of Mechatronic Systems*. John Wiley & Sons, New Jersey, 2012. 636 p.
- Kobayashi H., Marl B. L., Turin W. *Probability, Random Processes and Statistical Analysis*. Cambridge University Press, 2012. 812 p.
- Larkin E. V., Akimenko T. A., Bogomolov A. V. Modeling the reliability of the onboard equipment of a mobile robot. *Izvestiya of Saratov University. Mathematics. Mechanics. Informatics*, 2021, vol. 21, no. 3, pp. 390–399. doi:10.18500/1816-9791-2021-21-3-390-399
- Ronzhin A., Saveliev A., Basov O., Solyonyj S. Conceptual model of cyberphysical environment based on collaborative work of distributed means and mobile robots. *Lecture Notes in Computer Science*, 2016, vol. 9812, pp. 32–39. doi:10.1007/978-3-319-43955-6_5
- Kahar S., Sulaiman R., Prabuwoño A. S., Akma N., Ahmad S. A., Abu Hassan M. A. A review of wireless technology usage for mobile robot controller. *Intern. Proc. of Computer Science and Information Technology IPCSIT*, 2012, vol. 34, pp. 7–12.



Модель системы стабилизации стохастического биоинженерного объекта на множестве целевых состояний и синергетической теории управления

С. И. Колесникова^а, доктор техн. наук, профессор, orcid.org/0000-0001-7158-2747, skolesnikova@yandex.ru

А. А. Фоменкова^а, канд. техн. наук, доцент, orcid.org/0000-0002-5570-6743

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

Введение: разработка систем управления биотехническими объектами, характеризующимися плохой формализуемостью моделей и большим числом неконтролируемых параметров, является одной из актуальных задач современной биоинженерии. Недавно разработанные методы аналитического конструирования нелинейных регуляторов, апробированные на технических объектах, являются робастными по отношению к неопределенностям и возмущениям. Построение алгоритмов стабилизации биосистем на основе опорных синергетических регуляторов представляет вполне определенный практический интерес для повышения надежности и увеличения времени функционирования биотехнических объектов с требуемым качеством. **Цель:** разработать модель множественного управления и реализующую ее алгоритмы достижения и стабилизации заданных целевых свойств биотехнического объекта управления. **Методы:** используется новый метод стохастического синергетического управления. **Результаты:** предложены два алгоритма конструирования опорных регуляторов для реализации целевых макросостояний, обладающих свойством аттрактивности, и сформулированы их асимптотические свойства. Получены стохастический и на основе интегральной адаптации регуляторы для тестового объекта – анаэробного биореактора, представлены результаты численного моделирования, подтверждающие их робастность. Разработана функциональная модель для автоматизации процесса множественного регулирования и правила переключения регуляторов, апробированные на наборе из скалярного и векторного регуляторов для заданных целевых макросостояний биореактора: выхода биогаза до заданного уровня, обеспечения заданной степени очистки (не менее 90 %) и одновременного достижения указанных целевых требований. Энергосбережение множественного синергетического регулирования по сравнению с отдельным достигло 15–18 % (на тестовом объекте с нормальным стандартным шумом по каналу управления). **Практическая значимость:** физическая интерпретируемость и простота логики аналитического конструирования опорных регуляторов представляют инженерный интерес, а модель множественного управления есть потенциальный инструмент автоматизации функционирования различных биоинженерных систем с мультистабильностью.

Ключевые слова – нелинейный многомерный биоинженерный объект, оценка состояния, целевая макропеременная, макросостояние, инвариант системы, киберфизическая система, система синергетического стохастического управления, множественное управление, правило переключения регуляторов.

Для цитирования: Колесникова С. И., Фоменкова А. А. Модель системы стабилизации стохастического биоинженерного объекта на множестве целевых состояний и синергетической теории управления. *Информационно-управляющие системы*, 2024, № 5, с. 25–40. doi:10.31799/1684-8853-2024-5-25-40, EDN: NJUWOF

For citation: Kolesnikova S. I., Fomenkova A. A. Model of the stabilization system of a stochastic bioengineering object on a set of target states and synergetic control theory. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2024, no. 5, pp. 25–40 (In Russian). doi:10.31799/1684-8853-2024-5-25-40, EDN: NJUWOF

Введение

Основным целевым свойством киберфизических систем является их способность к самоорганизованной жизнедеятельности, направленной к устойчивому выполнению основных функций [1–6]. Как известно, свойство устойчивости целевых состояний в объектах управления является основой конструирования надежного и целенаправленного воздействия на них [7–14], в том числе и для систем детерминированного хаоса [11, 15, 16], а также для систем, функционирование которых отягощено возмущениями разного характера и неизбежной неточностью моделей их описания.

В данной статье концепция киберфизической системы используется как основа конструиро-

вания модели стабилизации биоинженерного объекта, суть которой сводится к управляемому «блужданию» между несколькими устойчивыми экспертно формализованными макросостояниями с аттрактивным свойством, технологические показатели которых удовлетворяют практическим требованиям и состоянию объекта управления, квалифицированному как работоспособное.

Термин «макросостояние», введенный для отличия от классически понимаемого динамического состояния, обязан синергетическому принципу, согласно которому управление, обеспечивающее достижение определенной окрестности аттрактивного многообразия, ищется из экстремали — закона изменения макроперемен-

ной, которая доставляет минимум определенному функционалу качества [9, 17].

В основу модели стабилизации функционирования биоинженерного объекта положена синергетическая теория управления (СТУ; synergetic control theory, SCT) [9] как инструмент обеспечения гомеостатического процесса за счет набора нескольких практически приемлемых аналитически сформулированных целевых макросостояний и сопоставленных им аналитически сконструированных опорных регуляторов, обеспечивающих стабилизацию в окрестности одного из них. При этом допускается наличие ограниченных неконтролируемых случайных и (или) систематических возмущений по управляющим переменным. Последнее предположение является естественным: неточность и (или) неопределенность модели отражается на неточности управления.

Развитие СТУ связано как с расширением области применения алгоритмов проектирования технических систем управления (например, энергетических), так и с моделированием биохимических, физических и экологических процессов, в частности, в технических и природных системах [18–25].

Предлагаемая в статье функциональная модель создана для решения основных задач анализа, мониторинга и управления в режиме реального времени макросостояниями биоинженерного объекта с неопределенностью в описании с целью предотвратить практически нежелательные макросостояния. Ее отличительной особенностью является множественность аналитически построенных на принципах СТУ стохастических регуляторов, реализующих соответствующую множественность целевых макросостояний. Основное предназначение такой

системы управления – повысить надежность синтезируемой системы управления по критерию времени безотказной работы с требуемым качеством выполнения основных функций объектом управления.

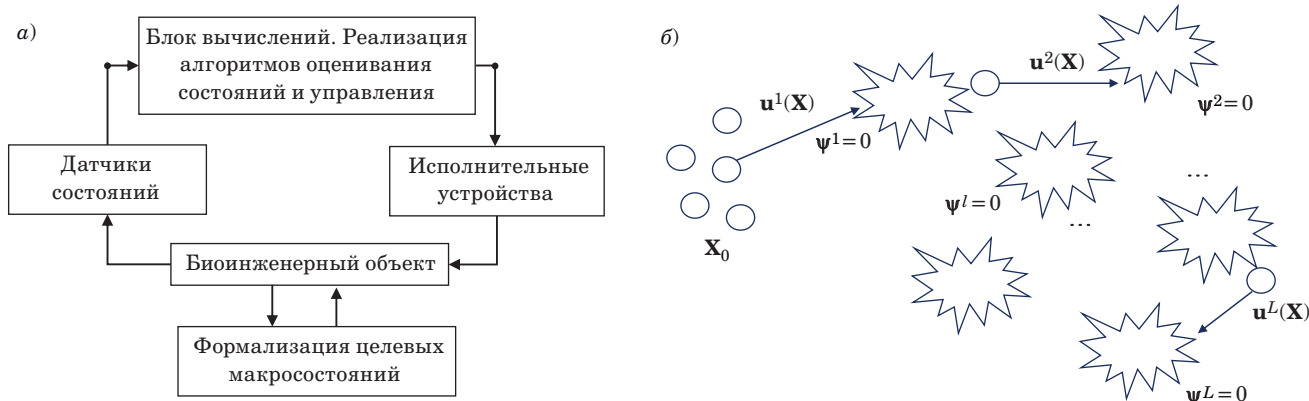
Используемый при этом подход как основа создания технологии достижения и поддержания стабильности сложного объекта (рис. 1, а) не ограничивается спецификой рассмотренных ниже примеров и может быть применен для любого объекта с известным набором пар <устойчивое равновесное макросостояние $\psi(\mathbf{X}(t)) \rightarrow 0, t \rightarrow \infty$; регулятор, обеспечивающий его достижение $\mathbf{u}(\mathbf{X}) >$ (рис. 1, б).

Отметим, что в работах школы А. А. Колесникова [9] решается задача конструирования управления в пространстве состояний (синтез регулятора) с учетом моделирования возмущения как решения дифференциального уравнения для придания системе управления замкнутости в предположении, что целевое состояние обладает аттрактивным свойством. Однако вопрос случайной неопределенности, актуальный для биофизических объектов, оставался открытым для аналитического синтеза системы управления в пространстве состояний.

В работе [17] представлен алгоритм (и доказаны его свойства) аналитического конструирования стохастического управления на принципах СТУ и выбора стратегий, минимизирующих дисперсию выходной макропеременной как заданной функции от динамического состояния и определяющей целевое макросостояние многомерной системы иммунологии с запаздыванием по нескольким переменным.

В настоящей статье:

1) приведена формализованная постановка задачи дискретного управления системой со слу-



■ **Рис. 1.** Иллюстрация алгоритмизации решения задачи стабилизации стохастического биоинженерного объекта в пространстве целевых макросостояний: а – концептуальная схема киберфизической системы; б – изменение ($\mathbf{u}^l, l = 1, 2, \dots, L$) целевых макросостояний ($\psi^l = 0$ при $t \rightarrow \infty, l = 1, 2, \dots, L$) для увеличения надежности функционирования биоинженерного объекта

■ **Fig. 1.** Illustration of the solution algorithmization to the problem of stabilization of a stochastic bioengineering plant in the space of target macrostates: а – conceptual diagram of a cyber-physical system; б – change ($\mathbf{u}^l, l = 1, 2, \dots, L$) of target macrostates ($\psi^l = 0$ with $t \rightarrow \infty, l = 1, 2, \dots, L$) to increase the reliability of the functioning of a plant controlled

чайным возмущением по каналу управления и дан алгоритм конструирования стохастического множественного СТУ-регулятора для биоинженерного объекта;

2) сделан краткий исторический обзор основных предпосылок, послуживших мотивацией к созданию информационной технологии стабилизации биотехнической системы, плохо поддающейся управлению [19–25];

3) предложена принципиальная функциональная схема модели множественного стабилизирующего управления на базе стратегий СТУ и представлено правило переключения между синергетическими регуляторами.

Особенности управления биоинженерными объектами и мотивация к постановке задачи множественного управления

Предмет и объект рассмотрения имеют отношение к так называемым «неудобным задачам»: в системах, в названии которых присутствует префикс «био», существует огромное число параметров, связей между ними, учесть которые не представляется возможным (и вряд ли целесообразным в силу их многосвязности и большого количества). Это обстоятельство приводит к явлению «немоделируемой динамики» и сопротивлению объекта управления к внешнему целенаправленному воздействию. Характерным является то, что математическое описание таких объектов зачастую включает уравнения вида «хищник-жертва» (например, [26]) со всеми вытекающими отсюда проблемами прогноза поведения объекта в условиях возмущений и синтеза регулятора, «удерживающего» объект управления в окрестности целевого состояния, отражающего технологические требования объекта управления.

Как известно, в динамических режимах хаотических объектов теоретически может возникнуть за конечное время бесконечное число значений переменных, поэтому своевременные прогнозы и оценки состояния для коррекции управляющего воздействия затруднительны, и алгоритмы обеспечения стабилизации неустойчивых биоинженерных объектов, функционирование которых неизбежно сопряжено с неопределенностью, должны учитывать слабую управляемость таких систем. Кроме того, «удержание» объекта в заданной окрестности целевого состояния с флуктуирующими параметрами если и удаётся, то может быть сопряжено с высокими энергозатратами.

Понятия гомеостаза и гомеостатических систем восходит к трудам С. Bernard (К. Бернара),

W. B. Cannon (В. Б. Кэннона), Н. А. Бернштейна (гипотеза «повторение без повторений» в биомеханике), к исследованиям П. К. Анохина, к классификации всех систем на детерминистские, стохастические и сложные живые (хаос) согласно W. Weaver (В. Уиверу) и далее к синергетике и ее принципу самоорганизации в гомеостатических системах (см., например, обзоры в [27–30]). Начиная с трудов И. Р. Пригожина в естествознании в целом устойчиво вводится понятие теории сложных эмерджентных биосистем, а проблема гомеостатического регулирования становится основой для науки в целом. Так, установлен важный результат (например, [9, 29]): стационарные режимы отсутствуют не только у природных систем, и особенно у устойчивого поведения являются флуктуации параметров состояния системы в притягивающей окрестности инвариантных макросостояний и траекторий. Сложные биосистемы демонстрируют непрерывное движение вектора состояния в этом пространстве, где хаотически изменяются все статистические характеристики выборки [27–29]. Относительно последнего также важно, что закон больших чисел в переходных процессах не работает [29].

С учетом неполной формализуемости объекта (наличия немоделируемой динамики) и особенностей цифрового регулирования (влияния округления числовых значений на нежелательное отклонение от намеченного поведения) весьма актуальна задача принципиальной стабилизации объекта в некоторой окрестности не одного, а нескольких практически приемлемых макросостояний, обладающих аттрактивным свойством. Естественно возникает задача конструирования системы множественного стабилизирующего управления динамическим объектом в пространстве заданных макросостояний, что дает надежду на повышение времени безотказной (безаварийной) работы («создание надежной схемы из ненадежных элементов» по К. Шеннону). Данный показатель — один из наиболее важных в силу повышенных экологических рисков в случае отказа объекта функционировать на планируемом уровне.

Предполагается, что математическое описание динамического объекта для конструирования СТУ-управления имеет вид

$$\dot{\mathbf{X}} = \mathbf{F}(\mathbf{X}) + \zeta(t) + \mathbf{u}(t), \mathbf{X}(t_0) = \mathbf{X}_0, t = t_0 + 1, \dots, (1)$$

где $\mathbf{F}(\mathbf{X})$ — непрерывная вектор-функция, описание которой определяет специфику динамической системы; $\zeta(t)$ — неконтролируемое ограниченное возмущение; $\mathbf{X}_0, \mathbf{X} \in \mathbf{X}^{st} \subset \mathbb{R}^n$ — множество начальных условий и допустимых состояний системы; $\mathbf{u} \in \mathbb{R}^m, m \leq n$ — управление.

Пример. Для анаэробного биореактора со взвешенно-седиментированной биомассой и скалярного управления описание (1) может быть следующим:

$$\begin{aligned} \mathbf{X} &= (S, B_1, P, B_2, G, \theta, Q)^T; \\ \zeta(t) &= (0, 0, 0, 0, 0, \zeta, 0)^T; \\ \mathbf{u}(t) &= (0, 0, 0, 0, 0, u, 0)^T; \\ F_1 &= Q(t)(S_{in}(t) - S(t)) - k_1(\theta)B_1(t) - k_2(\theta)\frac{S(t)B_1(t)}{k_3(\theta) + S(t)}; \\ F_2 &= -Q(t)B_1(t) + k_4(\theta)\frac{S(t)B_1(t)}{k_3(\theta) + S(t)}; \\ F_3 &= -Q(t)P(t) + k_5(\theta)B_1(t) + \\ &+ k_6(\theta)\frac{S(t)B_1(t)}{k_3(\theta) + S(t)} - k_7(\theta)B_2(t) - k_8(\theta)\frac{P(t)B_2(t)}{k_9(\theta) + P(t)}; \\ F_4 &= -Q(t)B_2(t) + k_{10}(\theta)\frac{P(t)B_2(t)}{k_9(\theta) + P(t)}; \\ F_5 &= -G(t) + k_{11}(\theta)\frac{P(t)B_2(t)}{k_9(\theta) + P(t)} - \frac{k_{12}(\theta)}{k_{12}(\theta) + P(t)}; \\ F_6 &= 0; F_7 = 0, \end{aligned} \quad (2)$$

где \mathbf{X} – вектор состояний (концентрации исходного субстрата S , кислотогенной биомассы B_1 , промежуточных продуктов брожения P , метаногенной биомассы B_2 , побочного продукта – метана – G , рабочая температура в биореакторе θ , скорость разбавления среды в биореакторе Q); $k_i(\theta)$, $i = 1, 12$ – кинетические параметры процесса с известным законом изменения $k_i(\theta) = k_{i35^\circ\text{C}} \exp(c_0(\theta - 35^\circ\text{C}))$. Тогда u , ζ – искомым закон регулирования по температуре и неизвестная функция времени соответственно.

Для объекта (2) можно указать несколько макросостояний $\psi(\mathbf{X}(t)) \rightarrow 0$, $\psi(\mathbf{X}(t)) \in \mathbb{R}^m$, стабилизация объекта управления в окрестности которых обеспечит требуемое технологическое состояние управляемого объекта [31]. Приведем формулировки некоторых целевых макропеременных для анаэробного биореактора с описанием (2), полагая единицы измерения условными:

- минимизация содержания летучих жирных кислот, выражаемая как $\psi_2 = P - P_{\min} \rightarrow 0$;
- обеспечение желательной производительности системы по биогазу (метану) [м³/сут] через соотношение $\psi_4 = G - G^* \rightarrow 0$;
- стабилизация нагрузки по органическому веществу (Organic Load Rate, OLR), характеризующая соотношением

$$\begin{aligned} \psi_3 &= OLR_{\max} - \frac{S_{in}}{HRT} \rightarrow 0, \\ OLR &= S_{in}/HRT = S_{in}Q/V, \end{aligned}$$

где HRT – гидравлическое время пребывания стока на очистке в сутки (Hydraulic Retention Time); V – объем биореактора.

Постановка задачи векторного локального СТУ-управления для объекта (2) может быть описана на следующем описании:

$$\begin{aligned} \mathbf{X} &= (S, B_1, P, B_2, G, \theta, Q)^T; \\ \zeta(t) &= (0, 0, 0, 0, 0, \zeta_1, \zeta_2)^T; \\ \mathbf{u}(t) &= (0, 0, 0, 0, 0, u_1, u_2)^T. \end{aligned}$$

В соответствии с векторным управлением (например, по температуре θ и скорости входного потока) возможен аналитический синтез системы управления с векторной макропеременной, например:

$$\begin{aligned} \psi_1(t) &= S(t) + P(t) - S^* \rightarrow 0; \\ \psi_2(t) &= G(t) - G^* \rightarrow 0, \end{aligned} \quad (3)$$

где знак «*» указывает на заданные экспертом значения. Первое условие в (3) выражает минимизацию на выходе биореактора концентрации загрязнений $S_{out}(t) = S(t) + P(t)$ потока очищенной воды, состоящей из остатков исходного субстрата с концентрацией $S(t)$ и промежуточных продуктов анаэробного брожения с концентрацией $P(t)$ (согласно модели анаэробного брожения ADM1 [32]).

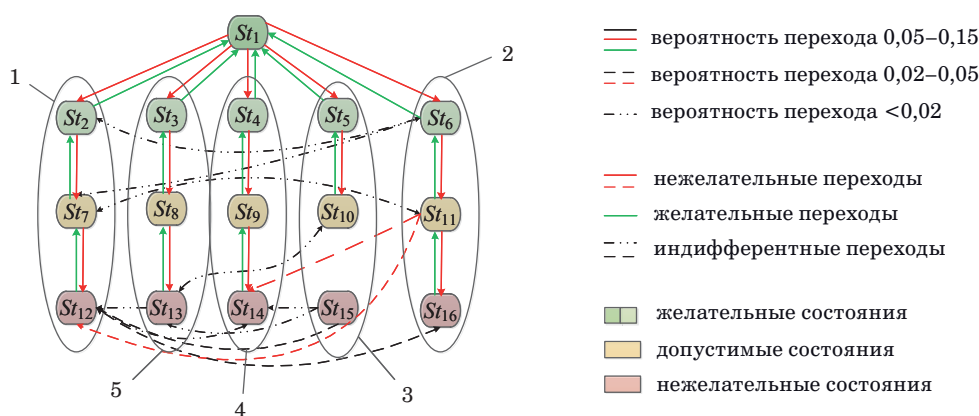
Требования к качеству опорного управления, конструирование которого основано на алгоритме интегральной адаптации как расширении СТУ [33] для объектов, описание которых содержит неконтролируемые возмущения, задаются функционалом качества при $m = 2$:

$$\Phi_C = \int_0^{\infty} \sum_{i=1}^m (\psi_i^2 + \omega_i^2 \dot{\psi}_i^2(t)) dt \rightarrow \min, \quad (4)$$

где индекс « C » указывает на непрерывность описания объекта управления; ω_i , $i = 1, m$ – параметры настройки скорости достижения предельного макросостояния $\psi(\mathbf{X}(t)) \rightarrow 0$, $\psi(\mathbf{X}(t)) \in \mathbb{R}^m$.

Актуальность задачи стабилизации в пространстве практически целесообразных макросостояний объекта вида (2) иллюстрирована рис. 2, где зеленым цветом обозначены целесообразные состояния, формализация и достижение которых могут быть обеспечены СТУ-алгоритмами (см. примеры синтеза регуляторов в приложении).

Из рис. 2 следует, что мотивация к рассмотрению здесь техники стабилизации биоинженерного объекта обусловлена множественностью практически целесообразных макросостояний, дости-



■ **Рис. 2.** Частоты наступления 16 состояний в анаэробном биореакторе и возможные переходы между ними
 ■ **Fig. 2.** Frequencies of occurrence of 16 states in an anaerobic bioreactor and possible transitions between them

жимость которых обеспечена СТУ-алгоритмами, реализующими принцип самоорганизации [9, 34].

Так, состояние St_1 является наиболее предпочтительным при длительной эксплуатации, поскольку характеризуется нормальным развитием анаэробной биомассы без признаков ее деградации; соответствующая ему стационарная точка \mathbf{X}^* со значениями переменных состояния (0,4125; 0,1709; 1,3217; 0,2195; 0,0033) является устойчивой (значения действительной части коэффициентов Ляпунова отрицательны: $\lambda = (-1; -0,0294; -0,0294; -0,0626; -0,0245)$);

Стационарная точка с координатами (0,5476; 0,033; 0,195; 1,3217; 0,094; 0,00064) относится к желательному состоянию St_3 , при этом это состояние является неустойчивым (значения действительной части коэффициентов Ляпунова $\lambda = (-1; -0,0148; -0,0148; 0,0001; -0,0205)$).

Среди 16 состояний выделены группы устойчивых переходов от удовлетворительного состояния к нежелательному (практически недопустимому) по причинам возможных процессов: перегрузки по органике – 1, повышенной нагрузки на биомассу со стороны изменения состава сточной воды – 2, выхода из температурного режима – 3, разрушения клеточных агрегатов биомассы – 4, недостатка питания биомассы – 5 и пр. Различают как желательные, так и нежелательные переходы, при этом те и другие имеют разную частоту наступления событий.

Несмотря на прозрачность логики аналитических алгоритмов конструирования систем управления с заданными свойствами на основе СТУ [9], гарантировать желаемый результат не всегда возможно в силу не только специфичности нелинейных систем, но и возможного приобретения неустойчивых свойств целевой системы за счет неизбежной флуктуации параметров и возмущения, которому подвергаются перемен-

ные биохимических объектов с неустойчивой динамикой.

В этой связи естественно рассмотреть задачу стабилизации в следующей постановке: для составляющих набора

$$\left\{ \left\{ \psi^l(\mathbf{X}) \right\}_{l=1, \overline{N_\psi}}, \left\{ \mathbf{u}^j(\mathbf{X}) \right\}_{j=1, \overline{N_u}}, \left\{ \mathbf{X}(0) \right\} \right\} -$$

практически приемлемых экспертно сформулированных целевых макропеременных, допустимых законов управлений, обеспечивающих достижение целевых макросостояний, и множества допустимых начальных состояний соответственно, сконструировать систему множественного управления для нелинейного динамического объекта с мультистабильностью, обеспечивающую стабилизацию объекта управления в пространстве равновесных состояний с аттрактивным свойством.

Решение задачи конструирования множественного управления на основе стохастического расширения СТУ

Представляемая ниже модель управления биоинженерной системой при наличии нескольких устойчивых состояний (рис. 3) предполагает выполнение условий, аналогичных требованиям при конструировании регуляторов на принципах СТУ [9], реализующей концепцию физической теории управления (по А. А. Красовскому):

1) описание динамики объекта управления в виде системы обыкновенных дифференциальных уравнений (или разностных уравнений);

2) существование устойчивых макросостояний, математическая модель которых известна в форме предельных равенств $\psi(\mathbf{X}(t)) = 0$, где $\psi(\mathbf{X}(t))$ – экспертно определяемая функция

(макропеременная) от состояния объекта управления;

3) существование и ограниченность решений исходной системы уравнений;

4) аттрактивность макросостояний;

5) ограниченность неконтролируемых возмущений по каналам управления.

Существенным расширением основного метода СТУ является его распространение на стохастические объекты [17], т. е. возмущение может носить систематический, случайный или смешанный характер.

Краткие комментарии к отдельным блокам структурной схемы (см. рис. 3), отражающей специфику представленной информационно-функциональной модели стабилизации многомерного, многосвязного объекта с неустойчивыми предельными режимами, заключаются в следующем:

– размерность множества формализованных экспертно задаваемых целевых макропеременных не является жесткой и по мере накопления истории об объекте динамически изменяется;

– априорно полученная на обучении траектория изображающей точки системы стабилизации, целенаправленно «блуждающей» под воздействием опорных регуляторов между различными целевыми макросостояниями, может подчиняться критерию энергосбережения, характерному для СТУ-регулирования, но в реальных условиях фактор стабильности всей системы является приоритетным в силу естественной глобальной цели – увеличения времени безотказной работы;

– система адаптивна и изменяет структуру регулятора в соответствии с глобальным критерием оценки качества управления и в ответ на

факторы характера возмущения и (или) на присутствующую биоинженерным объектам внутреннюю неустойчивость;

– система управляется и стабилизируется в соответствии с принципом направленной самоорганизующейся динамики сложного объекта [9, 17, 33].

В статье не преследуется цель дать полное алгоритмическое описание блоков 1–18 структурной схемы и их взаимосвязей, но важно отметить, что построение данной модели как основы автоматизации поддержания надежного функционирования сложных объектов стало возможным за счет двух СТУ-алгоритмов аналитического конструирования робастных регуляторов для объектов, динамика которых усложнена возмущениями как систематического, так и случайного характера.

Далее следует краткое описание двух основных блоков (см. рис. 3).

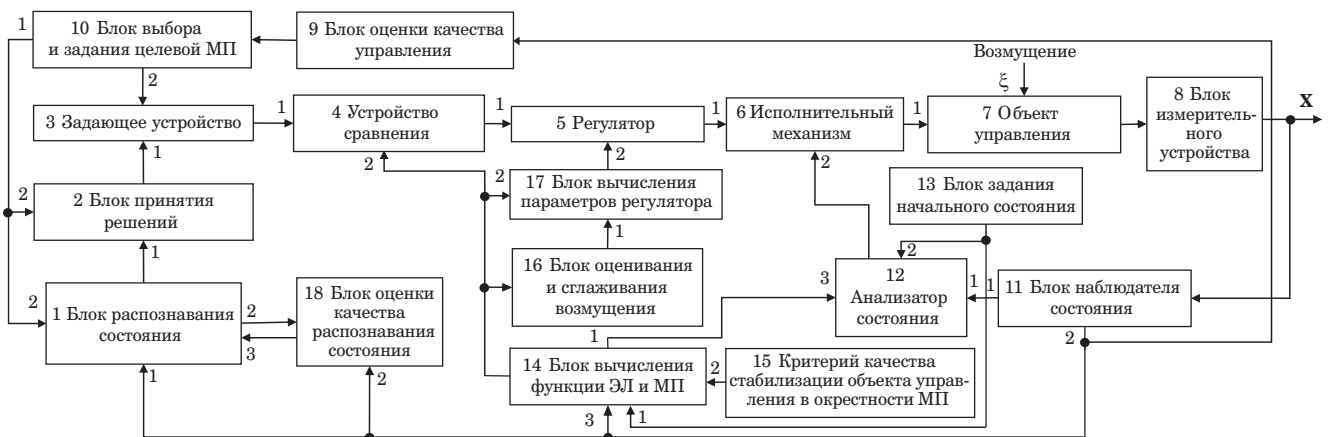
Блок 5 «Регулятор» реализован двумя алгоритмами конструирования СТУ-регуляторов: на основе методов интегральной адаптации [33] и стохастического расширения СТУ [17].

Алгоритм 1 конструирования непрерывного СТУ-управления с учетом неконтролируемого возмущения на основе метода интегральной адаптации

Этап 1. Расширение фазового пространства за счет моделирования возмущения, система (1) дополняется уравнениями вида

$$\dot{z}_l(t) = \eta_l \psi_l(t), \quad \eta_l > 0, \quad l = \overline{1, m},$$

где $z_l(t)$ – аддитивная модель возмущения $\zeta_l(t)$ по l -му каналу управления; $\psi_l(t)$ – l -я компонента



■ **Рис. 3.** Функциональная схема организации достижения и поддержания стабильного состояния биоинженерного объекта в условиях возмущений на множестве целевых макросостояний: *МП* – макропеременная; *ЭЛ* – (функция) Эйлера – Лагранжа

■ **Fig. 3.** Functional diagram of the organization of achieving and maintaining a stable state of a bioengineering object under conditions of disturbances in a set of target macrostates: *MP* – macro variable; *EL* – Euler – Lagrange function

целевой макропеременной $\psi(t)$; m – размерность вектора управления.

Этап 2. Осуществление детерминированного СТУ-синтеза [9] для полученной замкнутой системы с тем лишь различием, что на конечном этапе синтеза достигаемая l -я цель управления будет иметь описание $\psi_l^*(t) = \psi_l(t) + z_l(t) = 0, t \rightarrow \infty$ (см. подробно, например, в [33]).

Пример последовательного применения этапов алгоритма 1 – в приложении.

Утверждение 1. Управление, полученное по методу интегральной адаптации, если существует, обеспечивает асимптотическую устойчивость объекту управления в окрестности $\psi(t) = 0, t \rightarrow \infty$.

Замечание 1. Существование управления обусловлено выполнением требований СТУ-конструирования, характером нелинейностей в математической модели динамического объекта, аттрактивностью многообразия $\psi(t) = 0, t \rightarrow \infty$.

Алгоритм 2 конструирования стохастического дискретного управления на основе СТУ

Обозначим для удобства $\mathbf{Y}[k] := \mathbf{Y}[t_k], t_k = kh, h > 0, k = k_0, k_0+1, \dots$, где h – параметр дискретизации времени [например, параметр явной схемы Эйлера, если объект управления задан непрерывным описанием вида (1)].

В качестве СТУ-управления будем понимать последовательность $\{\mathbf{u}[1], \mathbf{u}[2], \dots\}$, где векторная функция $\mathbf{u}[k]: \mathbb{R}^n \rightarrow \mathbb{R}^m$ определяет закон управления в данный момент $t_k, k \geq k_0$ и приводит объект управления к реализации цели управления $\psi[k] \xrightarrow[k \rightarrow \infty]{} \mathbf{0}, \psi \in \mathbb{R}^m$ посредством решений последовательности вариационных задач $(\Phi_{D,j}, \psi_j^A), j = 1, j_{final}$, отвечающих СТУ-методологии и определяемых функционалами качества $\Phi_{D,j}$ и ограничениями [9] на промежуточные макропеременные ψ_j^A .

Основным методом СТУ [9] является метод аналитического конструирования агрегированных регуляторов, этим объясняется наличие верхнего индекса «А» у некоторых переменных здесь и ниже. Нижний индекс «D» в обозначении $\Phi_{D,j}$ указывает на дискретный синтез управления.

Основные положения алгоритма синтеза стохастического СТУ-регулятора следующие.

1. Представляем динамику объекта управления в дискретном виде:

$$\begin{aligned} \mathbf{Y}[k+1] &= \mathbf{H}[k] + \mathbf{u}[k] + \xi[k+1] + c\xi[k], \\ \mathbf{Y}[k_0] &= \mathbf{Y}_0, k = k_0, k_0+1, \dots, \end{aligned} \quad (5)$$

где $\mathbf{Y}[k] = (Y_1[k], \dots, Y_n[k])^T, \mathbf{H}[k] := \mathbf{H}(\mathbf{Y}[k]) \in \mathbb{R}^n, \mathbf{u} \in \mathbb{R}^m, m \leq n$ – векторы состояний, нелиней-

ной функции описания динамики изменения состояний объекта и управления соответственно; $\xi[k] \in \mathbb{R}^l, l \leq m$ – случайные некоррелированные функции;

$$\mathbf{E}\{\xi_i[k]\} = 0, \mathbf{D}\{\xi_i[k]\} = \sigma^2, i = \overline{1, l}, |c| < 1.$$

Формулируем задачу управления в форме СТУ. Для стохастического дискретного объекта критерий качества имеет вид

$$\begin{aligned} \mathbf{E}\{\psi[k]\} &= 0, \mathbf{D}\{\psi_j[k+1] + \omega_j \psi_j[k]\} \rightarrow \min, \\ |\omega_j| &< 1, j = \overline{1, m}, k \rightarrow \infty; \\ \mathbf{E}\{\Phi_D\} &= \\ = \mathbf{E}\left\{ \sum_{k=1}^{\infty} \sum_{j=1}^m \left(\alpha_j^2 (\psi_j[k])^2 + (\Delta \psi_j[k])^2 \right) \right\} &\rightarrow \min, \\ \Delta \psi_j[k] &= \psi_j[k] - \psi_j[k-1]. \end{aligned} \quad (6)$$

Здесь $\mathbf{E}\{\}, \mathbf{D}\{\}$ – операции математического ожидания и дисперсии соответственно. Постоянные ω_j, α_j есть параметры регулятора, связь между которыми вполне определенная (см., например, в [17]).

2. Ищем структуру детерминированного СТУ-регулятора $\hat{\mathbf{u}}^A[k], k = k_0, k_0+1, \dots$, при фиксированном возмущении $\xi[k] \in \mathbb{R}^l, l \leq m$ на основе детерминированного дискретного аналога алгоритма 1. С этой целью планируем поиск закона управления как функции предыдущих состояний и управлений:

$$\begin{aligned} \mathbf{u}[k] &= \mathbf{U}(\mathbf{Y}[k], \mathbf{Y}[k-1], \dots, \mathbf{Y}[k_0]); \\ \mathbf{u}[k-1], \mathbf{u}[k-2], \dots, \mathbf{u}[k_0] &), k = k_0, k_0+1, \dots \end{aligned}$$

3. Применяем операцию условного математического ожидания для управления $\hat{\mathbf{u}}^A[k], k = k_0, k_0+1, \dots$ при условии $\xi^k = (\xi[0], \xi[1], \dots, \xi[k])$, где

$$\begin{aligned} \hat{\mathbf{u}}[k] &= \mathbf{E}\{\hat{\mathbf{u}}^A[k] | \xi^k\}, \\ \xi[k] &= (\xi_1[k], \dots, \xi_m[k]), k = k_0, k_0+1, \dots \end{aligned}$$

4. Осуществляем декомпозицию исходного описания (5) с учетом полученных СТУ-управлений $\hat{\mathbf{u}}[k], k = k_0, k_0+1, \dots$

5. Находим оценку возмущения $\hat{\xi}[k]$ на основе шага 4.

6. Заменяем переменную $\xi[k]$ ее оценкой $\hat{\xi}[k]$ как функцией текущего макросостояния первого уровня иерархии СТУ-синтеза в выражении для полученного в п. 2 управления $\hat{\mathbf{u}}[k] = \mathbf{E}\{\hat{\mathbf{u}}^A[k] | \hat{\xi}^k\} \Rightarrow \mathbf{u}[k]$. Синтез системы стохастического СТУ-управления закончен.

Утверждение 2. Управление, полученное на основе стохастического расширения СТУ-конструирования, если существует, обеспечивает асимптотическую устойчивость в среднем объекту управления в окрестности $E\{\psi[k]\} = 0, k \rightarrow \infty$. При этом дисперсия компонент векторной макропеременной будет минимальна.

Замечание 2. Опорные СТУ-регуляторы, таким образом, могут быть сконструированы на основе представленных выше алгоритмов для объектов с непрерывным и дискретным описанием.

«Блок принятия решений» 2 реализует решение об изменении цели регулирования при одновременном выполнении требований: приемлемые значения радиуса окрестности текущего целевого макросостояния и ограничения на значение управляющего воздействия, формализация которых может быть следующей.

Алгоритм 3 конструирования правила переключений стохастических СТУ-регуляторов

Пусть на текущем этапе выбрана целевая макропеременная $\psi^j(\mathbf{X})$, и достижение соответствующего макросостояния $\psi^j(\mathbf{X}) \rightarrow 0$ реализует закон опорного управления $\mathbf{u}^j = \mathbf{u}^j(\mathbf{X}, \mathbf{X}_0, \alpha, \lambda)$.

1. Задаются значение δ_j радиус-окрестности целевого макросостояния $\psi^j(\mathbf{X})$, $\mathbf{X} = \mathbf{X}(t), t \geq 0$ и ограничения на управление в его окрестности $|\mathbf{u}^j(\psi^j(\mathbf{X}))| < \gamma_j$:

$$|\psi(\mathbf{X}) - \psi^j(\mathbf{X})| < \delta_j, |\mathbf{u}^j(\mathbf{X})| < \gamma_j. \quad (7)$$

2. Проверяется правило: если в момент наблюдения имеют место неравенства (7), то исходный объект стабилизирован и переключения регулятора не требуется; если хотя бы одно из

данных условий не выполнено, то производятся следующие действия:

а) полагается

$$\mathbf{X}_0 := \mathbf{X}(t), j_0 := \arg \min_{i=1, N_\psi, i \neq j} \left\{ \rho \left(\Omega_{\psi^i(\mathbf{X})}, \psi^j(\mathbf{X}) \right) < w \right\},$$

$$w = \text{const},$$

где $\Omega_{\psi^i(\mathbf{X})}$ – условное обозначение окрестности бассейна аттрактора $\psi^i(\mathbf{X}) \rightarrow 0$; $\rho(A, B)$ – одна из метрик сравнения величин A и B ; w – пороговая величина, определяемая (вместе с величинами δ_j, γ_j) на обучении и имитационном моделировании процессов в биоинженерном объекте управления;

б) осуществляется переключение на управление, отвечающее за реализацию цели $\psi^{j_0}(\mathbf{X}) \rightarrow 0$ с номером j_0 .

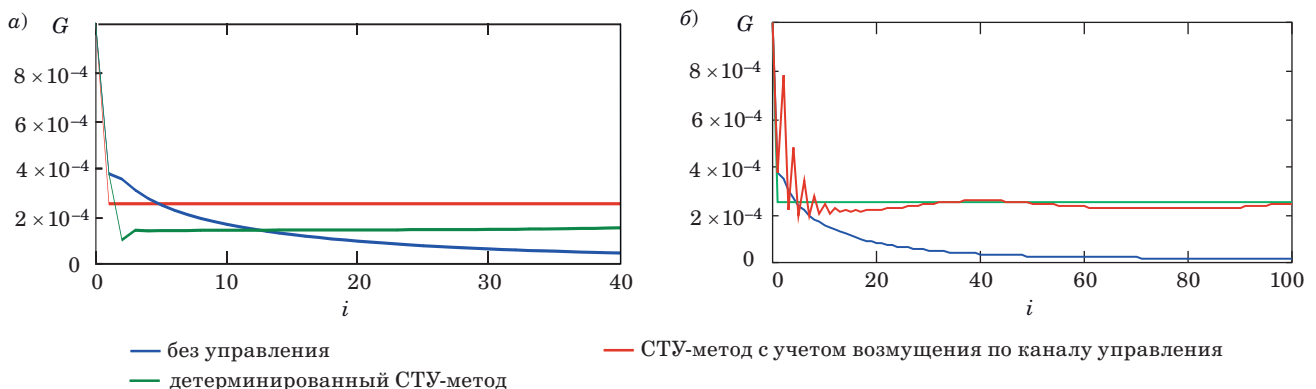
Для конкретной модели биотехнического объекта на обучении создаются правила запрета некоторых переходов и уточняются метрики сходства состояний.

Замечание 3. Для реализации п. а) алгоритма 3 представляется удобным применять подход из [35].

Численное моделирование алгоритма множественного управления и сравнение с одиночным управлением по критерию времени безотказной работы

Рассматривался набор из трех целевых макропеременных и соответствующих законов управления, полученных на основе алгоритмов 1, 2 (рис. 4–6):

– скалярный закон управления-1 по температуре $\theta(t)$, обеспечивающий выход биогаза $G(t)$ до заданного уровня G^* (G^* – расчетное количество



■ **Рис. 4.** Сравнение количества выхода биогаза без управления, с управлением по детерминированному СТУ-методу и по СТУ-методу с учетом постоянного возмущения по каналу управления: а – возмущение, равное 1 усл. ед.; б – возмущение, равное 10 усл. ед.

■ **Fig. 4.** Comparison of biogas output without control, with control using the deterministic SCT-method and the SCT-method for objects with disturbances along the control channel with a constant disturbance to a – disturbance equals 1 conventional units; б – disturbance equal 10 conventional units

ство биогаза, которое возможно получить из заданного сырья);

– векторный закон управления-2 по скорости входного потока $Q_{in}(t)$ и температуре $\theta(t)$, обеспечивающий достижение определенной степени очистки входного потока

$$\eta = \frac{S_{in} - (S + P)}{S_{in}} 100 = \eta^* (\eta^* \geq 80\%)$$

в условиях заданных входной концентрации загрязнений сточной воды S_{in} , объема реактора V и верхней границы времени пребывания сточной воды на очистке для реактора-смесителя $VQ_n^{-1} \leq 40$;

– векторный закон управления-3 по скорости входного потока $Q_{in}(t)$ и температуре $\theta(t)$, обеспечивающий значение концентрации органического загрязнения S_{in} до (или ниже) заданного нормативного значения S^* и заданный уровень биогаза на выходе:

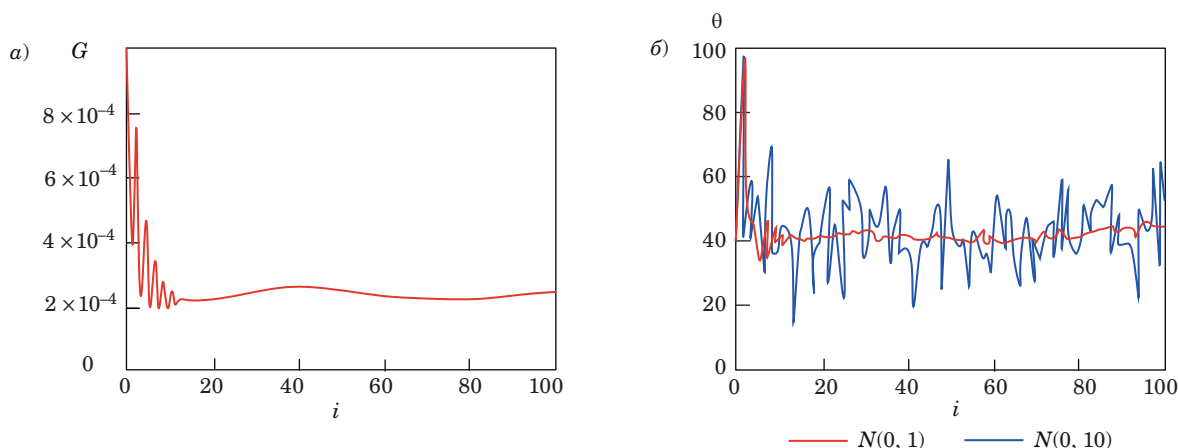
$$\psi_1(t) = S(t) - S^* \rightarrow 0; \psi_2(t) = G(t) - G^* \rightarrow 0.$$

При этом желательно максимально полно преобразовать загрязнения в биогаз.

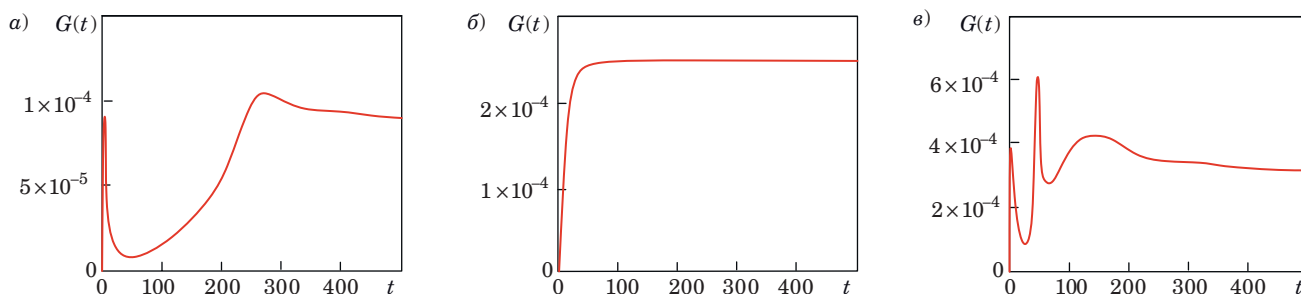
Однако случайное и гармоническое возмущения при неудачно подобранных параметрах регулятора могут привести к физически недопустимому диапазону значений управляющего воздействия (см. рис. 5 а, б), но целевая макропеременная при этом может находиться в среднем в некоторой окрестности целевого значения (здесь $2,394 \times 10^{-4}$).

Характер поведения переходных процессов при использовании определенных выше регуляторов показан на рис. 6: применение скалярного закона управления-1 приводит к снижению выработки биогаза со временем (см. рис. 6, а); векторный закон управления-2 стабилизирует выходную переменную за счет чрезмерно энергоемкого управления (см. рис. 6, б); векторный закон управления-3 приводит к максимальному выходу биогаза, но с большим перерегулированием (см. рис. 6, в).

Переключение полученных регуляторов в очередности 1 → 2 → 3 привело к итоговому вы-



■ **Рис. 5.** Изменение целевой переменной (а) и управления (б) в результате стохастического СТУ-регулирования
 ■ **Fig. 5.** Changing of the target variable (а) and control (б) as a result of stochastic SCT regulation



■ **Рис. 6.** Влияние отдельных регуляторов на усредненное количество выхода биогаза при практически приемлемом качестве очистки входного потока: а – скалярного управления-1; б – векторного управления-2; в – векторного управления-3
 ■ **Fig. 6.** The influence of individual regulators on the average amount of biogas output with practically acceptable quality of input flow cleaning: а – scalar control-1; б – vector control-2; в – vector control-3

ходу биогаза $2,58 \times 10^{-4}$ усл.ед. При этом степень очистки η в конечный момент отрезка моделирования достигала значения 90,9 %, что оказалось выше отдельно взятого опорного регулятора; время безотказной работы совпало со временем моделирования; энергоэффективность достигала 15–18 % по сравнению с одиночным регулированием при шуме вида $N(0, 1)$.

Замечание 4. На рис. 2 представлены возможные переходы между наиболее вероятными практическими стационарными состояниями, полученными на имитационной модели. Часть из них обладает свойством устойчивости, и рассматриваемая система (2) является мультистабильной. Тогда основное предназначение модели стабилизации (см. рис. 3) — достигнуть желаемого равновесного состояния, а затем «удержать» объект управления на некотором множестве целевых макросостояний. При этом используется определенное решающее правило переключения режимов регулирования, на которое может влиять и наиболее предпочтительная траектория целенаправленного «блуждания» [36], полученная, например, на обучении (реализована в [37]).

Заключение

Статья посвящена проблеме повышения качества управления объектами биоинженерной направленности.

Представлена синергетическая концепция решения сопутствующих этой проблеме локальных задач, ее алгоритмическое сопровождение с численной апробацией на тестовом примере — модели анаэробного биореактора (без ограничения общности в силу алгоритмов, изложенных для достаточно общего описания динамических объектов).

Основная решаемая задача — обеспечение стабилизации неустойчивого биохимического объекта в пространстве нескольких практически «благоприятных» целевых состояний, для достижения которых существуют СТУ-регуляторы, названные опорными. При этом стабилизация объекта управления в окрестности одного экспертно назначенного состояния будет частным случаем. Переходы между макросостояниями обеспечиваются последовательностью действий: оценкой и распознаванием текущего состояния; оценкой близости текущего образа состояния к другим макросостояниям, множество которых априорно задано; принятием решений об изменении цели в процессе функционирования объекта и управлением перехода от одного макросостояния к другому. Представлены алгоритмы, обеспечивающие функционирование си-

стемы множественного управления биоинженерным объектом.

Разработана структурная схема множественного динамически настраиваемого управления, отвечающая концепции И. Пригожина о применимости так называемых «универсальных законов» только к частным задачам, если рассматриваются объекты с неравновесной динамикой.

Благодарности

Авторы благодарят С. А. Цветницкую, кандидата технических наук, доцента Томского государственного университета, за проведение численных экспериментов и обсуждение результатов.

Финансовая поддержка

Научные исследования поддержаны грантом РФФИ (проект № 23-29-00336).

ПРИЛОЖЕНИЕ

Пример конструирования регулятора на основе алгоритма 1

Рассмотрим для краткости случай скалярного управления, поскольку векторный случай сводится к покомпонентному применению первого. Также для простоты положим относительно входного потока $Q(t) = Q = \text{const}$.

Пусть управление осуществляется по температуре θ и соответствующее уравнение в (1), (2) принимает вид

$$\frac{d\theta(t)}{dt} = \xi_1(t) + u_1(t). \quad (\text{П.1})$$

Целевая переменная $\psi_1^*(t)$, цель управления $\psi_1^*(t) \xrightarrow{t \rightarrow \infty} 0$. Пусть для определенности $\psi_1^*(t) = G(t) - G^*$, G^* — заданное значение.

Шаг 1. Осуществляем расширение фазового пространства с целью перевести модель возмущения в фазовую переменную согласно методу интегральной адаптации, исходная система (2) становится замкнутой:

$$\begin{aligned} \frac{dS(t)}{dt} &= \frac{Q}{V} (S_{in}(t) - S(t)) - k_1(\theta) \times \\ &\times B_1(t) - k_2(\theta) \frac{S(t)B_1(t)}{k_3(\theta) + S(t)}; \\ \frac{dB_1(t)}{dt} &= -\frac{Q}{V} B_1(t) + k_4(\theta) \frac{S(t)B_1(t)}{k_3(\theta) + S(t)}; \end{aligned}$$

$$\begin{aligned} \frac{dP}{dt} &= -\frac{Q}{V}P + k_5(\theta)B_1 + \\ &+ k_6(\theta)\frac{SB_1}{k_3(\theta)+S} - k_7(\theta)B_2 - k_8(\theta)\frac{PB_2}{k_9(\theta)+P}; \\ \frac{dB_2(t)}{dt} &= -\frac{Q}{V}B_2(t) + k_{10}(\theta)\frac{P(t)B_2(t)}{k_9(\theta)+P(t)}; \\ \frac{dG}{dt} &= -G(t) + k_{11}(\theta)\frac{P(t)B_2(t)}{k_9(\theta)+P(t)}\frac{k_{12}(\theta)}{k_{12}(\theta)+P(t)}; \\ \frac{d\theta(t)}{dt} &= z_1(t) + u_1(t); \\ \dot{z}_1(t) &= \eta_1\psi_1^*(t), \quad \eta_1 > 0. \end{aligned} \quad (\text{П.2})$$

Далее система (П.2) является исходной для алгоритма синтеза векторного регулятора. Здесь $z_1(t)$ – волновая модель [38] возмущения $\xi_1(t)$.

Шаг 2. Вводим вспомогательную макропеременную ψ_1 для формирования внутреннего управления φ_1 , где $\varphi_1 = \varphi_1(G, z_1)$:

$$\psi_1 = \theta - \varphi_1(G, z_1). \quad (\text{П.3})$$

Там, где контекст понятен, далее аргументы будем опускать.

Шаг 3. Выписываем уравнение для устойчивых экстремалей вариационной задачи $(\Phi_{C,1}, \psi_1)$ [5] с учетом вида функционала $\Phi_{C,1}$ из (4) при $m = 1$:

$$T_1\dot{\psi}_1(t) + \psi_1 = 0, \quad (\text{П.4})$$

T_1 – параметр регулятора, влияющий на скорость достижения целевого макросостояния.

Подставляем (П.3) в (П.4) и учитываем уравнения (П.2), получаем

$$\begin{aligned} T_1\dot{\psi}_1(t) + \psi_1 &= T_1\dot{\theta} - T_1\dot{\varphi}_1 + \psi_1 = 0, \\ \dot{\varphi}_1(G, z_1) &= \frac{\partial\varphi_1}{\partial G}\dot{G} + \frac{\partial\varphi_1}{\partial z_1}\dot{z}_1 = \frac{\partial\varphi_1}{\partial G}f_G + \frac{\partial\varphi_1}{\partial z_1}\eta_1\psi_1^*, \end{aligned} \quad (\text{П.5})$$

где f_G – правая часть описания в системе (П.2).

Из (П.5) следуют формулы для внешнего управления (согласно терминологии СТУ):

$$\begin{aligned} T_1\dot{\theta} - T_1\dot{\varphi}_1 + \psi_1 &= T_1(z_1 + u_1) - T_1\dot{\varphi}_1 + \psi_1 \Rightarrow \\ \Rightarrow u_1 &= \dot{\varphi}_1 - T_1^{-1}\psi_1 - z_1 = \\ &= \frac{\partial\varphi_1}{\partial G}f_G + \frac{\partial\varphi_1}{\partial z_1}\eta_1\psi_1^* - T_1^{-1}\psi_1 - z_1. \end{aligned} \quad (\text{П.6})$$

Формула (П.6) определяет структуру регулятора для (2) с точностью до неизвестной функции внутреннего управления $\varphi_1(G, z_1)$ и ее частных производных по переменным G, z_1 .

Шаг 4. Редуцируем модель (П.2) на многообразии $\psi_1 = 0$, получаем систему уравнений, ис-

ходную для дальнейшего синтеза внутреннего управления φ_1 . Заменяем в (П.2) $\theta = \varphi_1$:

$$\begin{aligned} \dot{S}(t) &= \frac{Q}{V}(S_{in}(t) - S(t)) - \\ &- k_1(\varphi_1)B_1(t) - k_2(\varphi_1)\frac{S(t)B_1(t)}{k_3(\varphi_1)+S(t)}; \\ \dot{B}_1(t) &= -\frac{Q}{V}B_1(t) + k_4(\varphi_1)\frac{S(t)B_1(t)}{k_3(\varphi_1)+S(t)}; \\ \dot{P}(t) &= -\frac{Q}{V}P + k_5(\theta)B_1 + \\ &+ k_6(\theta)\frac{SB_1}{k_3(\theta)+S} - k_7(\theta)B_2 - k_8(\theta)\frac{PB_2}{k_9(\theta)+P}; \\ \dot{B}_2(t) &= -\frac{Q}{V}B_2(t) + k_{10}(\varphi_1)\frac{P(t)B_2(t)}{k_9(\varphi_1)+P(t)}; \\ \dot{G}(t) &= -G(t) + k_{11}(\varphi_1)\frac{P(t)B_2(t)}{k_9(\varphi_1)+P(t)}\frac{k_{12}(\varphi_1)}{k_{12}(\varphi_1)+P(t)}; \\ \dot{z}_1(t) &= \eta_1\psi_1^*(t), \quad \eta_1 > 0. \end{aligned} \quad (\text{П.7})$$

Система (П.7) – исходная для второго этапа иерархического синтеза регулятора.

Шаг 5. Вводим вспомогательную макропеременную второго этапа синтеза регулятора $\psi_2 = \psi_1^* + \mu_1 z_1$ и выписываем уравнение для устойчивых экстремалей вариационной задачи

$$(\Phi_{C,2}, \psi_2), \quad \Phi_{C,2} = \int_0^\infty (\psi_2^2 + T_2^2\dot{\psi}_2^2(t))dt \rightarrow \min :$$

$$\begin{aligned} T_2\dot{\psi}_2 + \psi_2 &= T_2(\dot{\psi}_1^* + \mu_1\dot{z}_1) + \psi_2 = 0, \\ \psi_2 &= \psi_1^* + \mu_1 z_1. \end{aligned} \quad (\text{П.8})$$

Подставляя $\psi_1^*(t) = G(t) - G^*$ в (П.8), с учетом уравнений (П.7) получим выражение для внутреннего управления φ_1 :

$$\begin{aligned} \tilde{\varphi}_1 &= \frac{\tilde{C}P(h_9 + h_{12}) + \\ &+ P\sqrt{\tilde{C}^2(h_9 + h_{12})^2 + 4h_{12}\tilde{C}(h_{11}PB_2 - h_9\tilde{C})}}{2h_{12}(h_{11}PB_2 - h_9\tilde{C})}, \\ \tilde{C} &= -T_2^{-1}(G - G^* + \mu_1 z_1) + G = \\ &= (1 - T_2^{-1})G + T_2^{-1}(G^* - \mu_1 z_1), \\ \varphi_1 &= c_0^{-1} \ln(\tilde{\varphi}_1) + 35 \text{ }^\circ\text{C}, \quad c_0 = \text{const.} \end{aligned} \quad (\text{П.9})$$

Последнее равенство справедливо из имеющейся связи $\tilde{\varphi}_1 = \exp(c_0(\varphi_1 - 35 \text{ }^\circ\text{C}))$.

Из (П.9) имеем выражения для частных производных:

$$(\tilde{\varphi}_1)'_G = (1 - T_2^{-1})\Upsilon; (\tilde{\varphi}_1)'_{z_1} = -T_2^{-1}\mu_1\Upsilon;$$

$$\Upsilon = P(h_{11}PB_2 - h_9\tilde{C})^{-2} \times \left(\begin{array}{c} P(h_9 + h_{12})h_{11}B_2 + \\ + K(\tilde{C})(h_{11}PB_2 - h_9\tilde{C}) + h_9\Theta \end{array} \right),$$

$$K(\tilde{C}) = \tilde{C}(h_9 + h_{12})^2 + 2h_{12}(h_{11}PB_2 - 2h_9\tilde{C}),$$

$$\Theta = (\tilde{C}^2(h_9 + h_{12})^2 + 4h_{12}\tilde{C}(h_{11}PB_2 - h_9\tilde{C})). \quad (\text{П.10})$$

Система опорного управления для объекта с непрерывным описанием и компенсацией возмущения по каналу управления есть совокупность уравнений (1), (2), (П.6), (П.9), (П.10).

Пример конструирования стохастического регулятора на основе алгоритма 2

Алгоритм 2 изложим для той же целевой макропеременной (максимизация биогаза на выходе системы управления со случайным возмущением по температуре), для которой получен детальный вывод непрерывного регулятора (алгоритм 1), полагая:

$$\psi^*[k] = G[k] - G^*, \quad k = k_0, k_0 + 1, \dots$$

Цель и требования к управлению сформулированы при общем описании алгоритма 2: стабилизировать дискретный объект с описанием вида (5) в некоторой окрестности $\psi^*[k] = G[k] - G^* \rightarrow 0$ с учетом (6). Исходная система для получения закона управления согласно операции дискретизации по явному методу Эйлера с постоянным размером шага $h > 0$, примененной к (1) с учетом (2), примет вид

$$\mathbf{Y}[k] = (S[k], B_1[k], P[k], B_2[k], G[k], \theta[k])^T;$$

$$Y_j[k+1] = Y_j[k] + hF_j[k], \quad j = \overline{1, 5};$$

$$Y_6[k+1] = \theta[k+1] = \theta[k] + h(F_6[k] + u[k] + \xi[k+1] + c\xi[k]), \quad (\text{П.11})$$

где функции $F_j[k], j = \overline{1, 6}$ удовлетворяют описанию (2); $\xi[k] \in \mathbb{R}$ — случайные некоррелированные функции, $\mathbf{E}\{\xi[k]\} = 0, \mathbf{D}\{\xi[k]\} = \sigma^2, i = \overline{1, l}, |c| < 1$.

Шаг 1. Фиксируем в системе (П.11) случайные функции $\xi[k], k = k_0, k_0+1, \dots$ и осуществляем детерминированный СТУ-синтез дискретной си-

стемы управления на основе решения вариационной задачи вида

$$\Phi_D = \sum_{k=1}^{\infty} \left(\alpha^2 \psi^2[k] + (\Delta\psi[k])^2 \right) \rightarrow \min,$$

$$\Delta\psi[k] = \psi[k] - \psi[k-1], \quad k = k_0 + 1, \dots \quad (\text{П.12})$$

При этом можно показать (подробный вывод имеется, например, в [17]), что экстремали для функционала Φ_D из (П.12) удовлетворяют дискретному уравнению вида

$$\psi[k+1] + \omega\psi[k] = 0, \quad |\omega| < 1, \quad k \rightarrow \infty, \quad (\text{П.13})$$

а параметры α, ω связаны соотношением

$$\omega = 0,5 \left(2 + \alpha^2 - \left((2 + \alpha^2)^2 - 4 \right)^{1/2} \right).$$

Синтез СТУ-управления является иерархическим, на каждом уровне иерархии решается соответствующая вариационная задача. В данном примере мы имеем два уровня иерархии, поэтому функционал качества примет вид $\mathbf{E}\{\Phi\} = \mathbf{E}\{\Phi_{D,1} + \Phi_{D,2}\}$.

1.1. Вводим вспомогательную переменную иерархии 1-го уровня

$$\psi_1[k] = \theta[k] - \phi[k], \quad (\text{П.14})$$

где $\phi[k]$ — внутреннее управление (согласно терминологии СТУ), являющееся некоторой функцией состояния $\phi[k] = \Psi(S[k], B_1[k], P[k], B_2[k], G[k])$ конструируемой системы управления в дискретный момент k , подлежащей определению. Согласно вариационной задаче 1-го этапа иерархии синтеза синергетического регулятора

$$\Phi_{D,1} = \sum_{k=1}^{\infty} \left(\alpha_1^2 \psi_1^2[k] + (\Delta\psi_1[k])^2 \right) \rightarrow \min,$$

$$\Delta\psi_1[k] = \psi_1[k] - \psi_1[k-1], \quad k = k_0 + 1, \dots, \quad (\text{П.15})$$

на основе соотношений (П.11)–(П.15) получим

$$\psi_1[k+1] + \omega_1\psi_1[k] = 0, \quad |\omega_1| < 1 \Rightarrow$$

$$\Rightarrow \theta[k+1] - \phi[k+1] + \omega_1\psi_1[k] = \theta[k] + h(F_6[k] + u[k] + \xi[k+1] + c\xi[k]) - \phi[k+1] + \omega_1\psi_1[k] = 0.$$

Из последнего выражения определяем предварительную структуру внешнего регулятора. Для удобства указания поэтапного определения закона управления здесь и далее для u введены вспомогательные обозначения в виде дополнительных индексов:

$$u_{\xi}^A[k] := h^{-1}(\hat{\phi}[k] - \omega_1 \psi_1[k] - \theta[k]) - \xi[k+1] - c\xi[k], \quad k = k_0, k_0 + 1, \dots \quad (\text{П.16})$$

Здесь функция

$$\hat{\phi}[k] = \Psi(F_1[k], F_2[k], F_3[k], F_4[k], F_5[k]).$$

1.2. Формируем систему для 2-го уровня конструирования управления. Декомпозируем систему (П.11) на достигнутой области, выражаемой предельным уравнением вида $\psi_1[k] = \theta[k] - \phi[k] = 0 \Rightarrow \theta[k] = \phi[k], k \rightarrow \infty$, и введем соответствующий нижний индекс в обозначении переменных. Система (П.11) примет вид

$$\begin{aligned} \mathbf{Y}_{\psi_1}[k] &= (S_{\psi_1}[k], B_{1\psi_1}[k], P_{\psi_1}[k], \\ & B_{2\psi_1}[k], G_{\psi_1}[k])^T; \\ Y_{j\psi_1}[k+1] &= Y_{j\psi_1}[k] + hF_{j\psi_1}[k], \quad j = \overline{1, 5}, \quad (\text{П.17}) \end{aligned}$$

где, например, уравнение для биогаза

$$\begin{aligned} G_{\psi_1}[k+1] &= G_{\psi_1}[k] + \\ & + h \left(-G_{\psi_1}[k] + k_{11}(\theta) \frac{P_{\psi_1}[k] B_{2\psi_1}[k]}{k_9(\phi[k]) + P_{\psi_1}[k]} \times \right. \\ & \left. \times \frac{k_{12}(\phi[k])}{k_{12}(\phi[k]) + P_{\psi_1}[k]} \right), \\ k_i(\phi[k]) &= k_{i35^\circ\text{C}} \exp(c_0(\phi[k] - 35^\circ\text{C})). \end{aligned}$$

1.3. Вводим переменную иерархии 2-го уровня

$$\psi_2[k] = \psi_{\psi_1}^*[k] = G_{\psi_1}[k] - G^*, \quad k = k_0, k_0 + 1, \dots$$

Согласно вариационной задаче 2-го этапа иерархии СТУ-синтеза

$$\Phi_{D,2} = \sum_{k=1}^{\infty} \left(\alpha_2^2 \psi_2^2[k] + (\Delta \psi_2[k])^2 \right) \rightarrow \min,$$

Литература

1. Ронжин А. Л., Басов О. О., Соколов Б. В., Юсупов Р. М. Концептуальная и формальная модели синтеза киберфизических систем и интеллектуальных пространств. *Приборостроение*, 2016, № 11, с. 897–905.
2. Biffi S., Lder A., Gerhard D. *Multi-Disciplinary Engineering for Cyber-Physical Production Systems: Data Models and Software Solutions for Handling Complex Engineering Projects*. 1st. ed. Springer Publishing Company Incorporated, 2017. 472 p.

на основе функционального уравнения для экстремали функционала $\Phi_{D,2}$ получим

$$\begin{aligned} \psi_2[k+1] + \omega_2 \psi_2[k] &= 0, \quad |\omega_2| < 1 \Rightarrow \\ \Rightarrow G_{\psi_1}[k+1] - G^* + \omega_2 \psi_2[k] &= 0 \Rightarrow \\ \Rightarrow G_{\psi_1}[k] + hF_{5\psi_1} - G^* + \omega_2 \psi_{\psi_1}^*[k] &= 0. \quad (\text{П.18}) \end{aligned}$$

Последнее уравнение в (П.18) служит для определения закона внутреннего управления $\phi[k], k \rightarrow \infty$. Структура регулятора (П.16) при фиксированных возмущениях определена.

Шаг 2. Рассматриваем полученный закон управления (П.16) как случайную функцию и применяем к ней операцию условного математического ожидания:

$$\begin{aligned} u^A[k] &= \mathbf{E}\{u_{\xi}^A[k] | \xi^k\}, \quad \xi^k = (\xi[k_0], \dots, \xi[k]), \\ k &= k_0, k_0 + 1, \dots \end{aligned}$$

В силу независимости величин $\xi[k+1], \xi[k]$ и $\mathbf{E}\{\xi[k]\} = 0$ получаем

$$\begin{aligned} u^A[k] &:= h^{-1}(\hat{\phi}[k] - \omega_1 \psi_1[k] - \theta[k]) - c\xi[k], \\ k &= k_0, k_0 + 1, \dots \quad (\text{П.19}) \end{aligned}$$

Шаг 3. Подставляем полученный закон управления (П.19) в описание объекта управления (П.11), получим оценку возмущения как функцию от измеряемых макропеременных в виде

$$\begin{aligned} h^{-1}(\psi_1[k] + \omega_1 \psi_1[k-1]) &= \hat{\xi}[k], \\ k &= k_0 + 1, k_0 + 2, \dots \quad (\text{П.20}) \end{aligned}$$

Шаг 4. Подставляем (П.20) в (П.19), получим итоговое выражение стохастического регулятора

$$\begin{aligned} u[k] &:= h^{-1}(\hat{\phi}[k] - \omega_1 \psi_1[k] - \theta[k]) - c\hat{\xi}[k], \\ k &= k_0, k_0 + 1, \dots \quad (\text{П.21}) \end{aligned}$$

3. Верзилин Д. Н., Соколов Б. В., Юсупов Р. М. Неокибернетика: состояние исследований и перспективы развития. *Системный анализ в проектировании и управлении: сб. науч. тр. XXIII Междунар. науч.-практ. конф.*, 2019, т. XXIII, № 1, с. 81–98.
4. Sampayo M., Peças P. CPSD2: A new approach for cyber-physical systems design and development. *Journal of Industrial Information Integration*, 2022, vol. 28, 100348. <https://doi.org/10.1016/j.jii.2022.100348>
5. Estrada-Jimenez L. A., Pulikottil T., Peres R. S., Nighadam-Hojjati S., Barata J. Complexity theory

- and self-organization in Cyber-Physical Production Systems. *Procedia CIRP*, 2021, vol. 104, pp. 1831–1836. <https://doi.org/10.1016/j.procir.2021.11.309>
6. Sun Y.-P., Guang-Hong Yang, Dimirovski G. M. Dynamic event-triggered control for multi-channel cyber-physical systems under denial-of-service attacks. *Journal of the Franklin Institute*, 2024, vol. 361, iss. 11, 106933. <https://doi.org/10.1016/j.jfranklin.2024.106933>
 7. Андриевский Б. Р., Фрадков А. Л. Управление хаосом: методы и приложения. II. Приложения. *Автоматика и телемеханика*, 2004, № 4, с. 3–34.
 8. Isidori A. *Lectures in Feedback Design for Multivariable Systems. Series: Advanced Textbook in Control and Signal Processing*. London, Springer, 2016. 414 p.
 9. Колесников А. А. *Синергетические методы управления сложными системами: теория системного синтеза*. М., Либроком, 2019. 240 с.
 10. Lazarević M. P. Elements of mathematical phenomenology of self-organization nonlinear dynamical systems: Synergetics and fractional calculus approach. *International Journal of Non-Linear Mechanics*, 2015, vol. 73, pp. 31–42. <https://doi.org/10.1016/j.ijnonlinmec.2014.11.011>
 11. Abed E. H. Bifurcation control: new directions motivated by biology and complex engineered systems. *IFAC Proceedings Volumes*, 2007, vol. 40, iss. 12, pp. 757–763. <https://doi.org/10.3182/20070822-3-ZA-2920.00125>
 12. Diaz-Seoane S., Blas A. B., Villaverde A. F. Controlability and accessibility analysis of nonlinear biosystems. *Computer Methods and Programs in Biomedicine*, 2024, vol. 245, 108015.
 13. Xi R., Zhang H., Wan Y., Sun S. Overview of the recent research progress for stability and control on random nonlinear systems. *Annual Reviews in Control*, 2022, vol. 53, pp. 70–82. <https://doi.org/10.1016/j.arcontrol.2022.03.005>
 14. Li W., Krstic M. Prescribed-time control of stochastic nonlinear systems with reduced control effort. *Journal of Systems Science and Complexity*, 2021, vol. 34, pp. 1782–1800. <https://doi.org/10.1007/s11424-021-1217-7>
 15. Астахов В. В., Безручко Б. П., Гуляев Ю. В., Селезнев Е. П. Мультистабильные состояния диссипативно связанных фейгенбаумовских систем. *Письма в ЖТФ*, 1989, т. 15, вып. 3, с. 60–65.
 16. Иваницкий Г. Р. Самоорганизующаяся динамическая устойчивость биосистем, далеких от равновесия. *Успехи физических наук*, 2017, т. 187, № 7, с. 757–784.
 17. Kolesnikova S., Kustova E. Application of a stochastic extension of the analytical design of aggregated regulators to a multidimensional biomedical object. *Mathematics*, 2023, vol. 11, no. 21, 4484. <https://doi.org/10.3390/math11214484>
 18. Fomichev A. O., Vavilin V. A. The reduced model of self-oscillating dynamics in an anaerobic system with sulfate-reduction. *Ecol. Modelling*, 1997, vol. 95, pp. 133–144.
 19. Garcia-Gen S., Santos L. O., Wouwer A. V. Application of a nonlinear model predictive controller to the anaerobic digestion of readily biodegradable wastes. *IFAC-PapersOnLine*, 2022, vol. 55, iss. 7, pp. 909–914. <https://doi.org/10.1016/j.ifacol.2022.07.560>. Прототип.
 20. Грудяева Е. К., Душин С. Е. *Моделирование управляемых процессов биологической очистки сточных вод*. СПб., СПбГЭТУ «ЛЭТИ», 2017. 231 с.
 21. Tawai A., Sriariyanun M. Nonlinear optimization-based robust control approach for a two-stage anaerobic digestion process. *Hindawi Journal of Chemistry*, 2022, 8966350. <https://doi.org/10.1155/2022/8966350>
 22. He D., Wang H., Tian Y., Christov N., Simeonov I. Trajectory tracking of two-stage anaerobic digestion process: A predictive control with guaranteed performance and saturated input, based on ultra-local model. *Journal of Process Control*, 2023, vol. 129, 103039.
 23. Mendiola-Rodriguez T. A., Ricardez-Sandoval L. A. Integration of design and control for renewable energy systems with an application to anaerobic digestion: A deep deterministic policy gradient framework. *Energy*, 2023, vol. 274, 127212.
 24. Kurth A.-C., Sawodny O. Control of age-structured population dynamics with intraspecific competition in context of bioreactors. *Automatica*, 2023, vol. 152, 110944. <https://doi.org/10.1016/j.automatica.2023.110944>
 25. Mendiola-Rodriguez T. A., Ricardez-Sandoval L. A. Robust control for anaerobic digestion systems of tequila vinasses under uncertainty: A deep deterministic policy gradient algorithm. *Digital Chemical Engineering*, 2022, vol. 3, 100023.
 26. Diz-Pita É., Otero-Espinar M. V. Predator–Prey Models: A review of some recent advances. *Mathematics*, 2021, vol. 9(15), 1783. doi:10.3390/math9151783
 27. Haken G. *Advanced Synergetics. Instability Hierarchies of Self-Organizing Systems and Devices*. Springer Berlin, Heidelberg, 1983. 356 p.
 28. Еськов В. В. Термодинамика неравновесных систем I. R. Prigogine и энтропийный подход в физике живых систем. *Вестник новых медицинских технологий*, 2017, т. 24, № 2, с. 7–15.
 29. Моисеев Н. Н. *Избранные труды в 2 т. Т. 1. Гидродинамика и механика. Оптимизация, исследование операций и теория управления*. М., Тайдекс Ко, 2003. 376 с.
 30. Малинецкий Г. Г. Среда обитания: параметры порядка, самоорганизация. *Природа*, 2023, № 3 (1291), с. 26–35. <https://doi.org/10.7868/S0032874X23030031>
 31. Kolesnikova S. I., Fomenkova A. A. Dynamic strategies for monitoring quality control at a complex bioengineering facility. *Информационно-управляющие системы*, 2023, № 2, с. 51–60. doi:10.31799/1684-8853-2023-2-51-60, EDN: WUHAW
 32. Mo R., Guo W., Batstone D., Makinia J., Li Y. Modifications to the anaerobic digestion model no. 1

(ADM1) for enhanced understanding and application of the anaerobic treatment processes – A comprehensive review. *Water Research*, 2023, vol. 244, 120504. <https://doi.org/10.1016/j.watres.2023.120504>

33. Колесников А. А. *Новые нелинейные методы управления полетом*. М., Физматлит, 2013. 196 с.
34. Красовский А. А., Мисриханов М. Ш. Самоорганизующиеся контроллеры экстраполяции для энергосистем: история и перспективы. *Автоматика и телемеханика*, 2006, № 5, с. 48–53.
35. Datseris G. Effortless basins of attraction codebase. *Zenodo*, 2021. <https://doi.org/10.5281/zenodo.5806212> (дата обращения: 16.05.2024).

36. Дивеев А. И., Шмалько Е. Ю., Барабаш А. В. Полное решение задачи оптимального управления. *Надежность и качество: тр. Междунар. симп.*, 2023, т. 1, с. 96–102.
37. Свидетельство о государственной регистрации программы для ЭВМ № 2024664073. *Программа для реализации векторного управления нелинейным биотехническим объектом с неустойчивыми предельными состояниями*, Колесникова С. И. (RU), Фоменкова А. А. (RU); заявл. 17.06.24; опублик. 17.06.24.
38. *Фильтрация и стохастическое управление в динамических системах*/ под ред. К. Т. Леондеса. М., Мир, 1980. 406 с.

UDC 519.8, 519.7
doi:10.31799/1684-8853-2024-5-25-40
EDN: NJUWOF

Model of the stabilization system of a stochastic bioengineering object on a set of target states and synergetic control theory

S. I. Kolesnikova^a, Dr. Sc., Tech., Professor, orcid.org/0000-0001-7158-2747, skolesnikova@yandex.ru
A. A. Fomenkova^a, PhD, Tech., Associate Professor, orcid.org/0000-0002-5570-6743

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

Introduction: Due to the poor formalizability of models of bioengineering objects and a large number of uncontrolled parameters, the development of control systems is not a fully resolved problem. Recently obtained methods of controllers synergetic synthesis are robust with respect to uncertainties and disturbances in technical systems. Construct of biosystems stabilization algorithms on their basis is of practical interest for increasing reliability and the time of their functioning. **Purpose:** Construct a model of multiple control and algorithms implementing it for achieving and stabilizing the specified goal properties of the biotechnical system. **Methods:** We use the new stochastic synergetic control method to design local controllers, which are base for multiple control. **Results:** Two algorithms for construct local controller for implementing target macrostates are proposed, and their asymptotic properties are formulated. Stochastic controller and controller, based on integral adaptation method for an anaerobic bioreactor are obtained. The results of numerical simulation confirming their robustness are presented. The functional model for automating the process of multiple regulation and the rule for switching controllers have been developed, tested on a set of scalar and vector controllers that provide specified target macrostates of the bioreactor, namely: increasing the biogas yield to a specified level, a specified degree of purification (at least 90%), and simultaneously achieving these indicators. The energy efficiency of multiple synergetic control is estimated at 15–18% on a test plant with normal standard noise on the control channel relative to the use of a local controller. **Practical relevance:** Physical interpretability and simplicity of design algorithms for synergetic controllers are of engineering interest, and the multiple control model is a potential tool for automating various biotechnical systems with multistability.

Keywords – nonlinear multidimensional bioengineering object, state assessment, target macrovariable, macrostate, system invariant, cyber-physical system, synergetic stochastic control system, multiple control, controller switching rule.

For citation: Kolesnikova S. I., Fomenkova A. A. Model of the stabilization system of a stochastic bioengineering object on a set of target states and synergetic control theory. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2024, no. 5, pp. 25–40 (In Russian). doi:10.31799/1684-8853-2024-5-25-40, EDN: NJUWOF

Financial support

Scientific research was supported by a grant from the Russian Science Foundation (project No. 23-29-00336).

References

- Ronzhin A. L., Basov O. O., Sokolov B. V., Yusupov R. M. Conceptual and formal models of synthesis of cyber-physical systems and cyber-physical intellectual spaces. *Journal of Instrument Engineering*, 2016, no. 11, pp. 897–905 (In Russian).
- Biffi S., Lder A., Gerhard D. *Multi-Disciplinary Engineering for Cyber-Physical Production Systems: Data Models and Software Solutions for Handling Complex Engineering Projects*. 1st. ed. Springer Publishing Company Incorporated, 2017. 472 p.
- Verzilin D. N., Sokolov B. V., Yusupov R. M. Neocybernetics: State of research and development prospects. *Trudy XXIII Mezhdunarodnoj nauchno-prakticheskoy konferencii "Sistemnyj analiz v proektirovanii i upravlenii"* [Proc. XXIII Intern. Scient. and Pract. Conf. "System analysis in design and management"]. Saint-Petersburg, 2019, vol. XXIII, no. 1, pp. 81–98 (In Russian).
- Sampayo M., Peças P. CPSD2: A new approach for cyber-physical systems design and development. *Journal of Industrial Information Integration*, 2022, vol. 28, 100348. <https://doi.org/10.1016/j.jii.2022.100348>
- Estrada-Jimenez L. A., Pulikottil T., Peres R. S., Nikghadam-Hojjati S., Barata J. Complexity theory and self-organization in Cyber-Physical Production Systems. *Procedia CIRP*, 2021, vol. 104, pp. 1831–1836. <https://doi.org/10.1016/j.procir.2021.11.309>
- Sun Y.-P., Guang-Hong Yang, Dimirovski G. M. Dynamic event-triggered control for multi-channel cyber-physical systems under denial-of-service attacks. *Journal of the Franklin Institute*, 2024, vol. 361, iss. 11, 106933. <https://doi.org/10.1016/j.jfranklin.2024.106933>
- Andrievskii B. R., Fradkov A. L. Control of Chaos: Methods and Applications. II. Applications. *Automation and Remote Control*, 2004, vol. 65, no. 4, pp. 505–533.

8. Isidori A. *Lectures in Feedback Design for multivariable Systems. Advanced Textbook in Control and Signal Processing*. London, Springer, 2016. 414 p.
9. Kolesnikov A. A. *Sinergeticheskie metody upravleniya slozhnyimi sistemami: teoriya sistemnogo sinteza* [Synergistic methods of control complex systems: the theory of system synthesis]. Moscow, Librokombiz., 2019. 240 p. (In Russian).
10. Lazarević M. P. Elements of mathematical phenomenology of self-organization nonlinear dynamical systems: Synergetics and fractional calculus approach. *International Journal of Non-Linear Mechanics*, 2015, vol. 73, pp. 31–42. <https://doi.org/10.1016/j.ijnonlinmec.2014.11.011>
11. Abed E. H. Bifurcation control: new directions motivated by biology and complex engineered systems. *IFAC Proceedings Volumes*, 2007, vol. 40, iss. 12, pp. 757–763. <https://doi.org/10.3182/20070822-3-ZA-2920.00125>
12. Diaz-Seoane S., Blas A. B., Villaverde A. F. Controllability and accessibility analysis of nonlinear biosystems. *Computer Methods and Programs in Biomedicine*, 2024, vol. 245, 108015.
13. Xi R., Zhang H., Wan Y., Sun S. Overview of the recent research progress for stability and control on random nonlinear systems. *Annual Reviews in Control*, 2022, vol. 53, pp. 70–82. <https://doi.org/10.1016/j.arcontrol.2022.03.005>
14. Li W., Krstic M. Prescribed-time control of stochastic nonlinear systems with reduced control effort. *Journal of Systems Science and Complexity*, 2021, vol. 34, pp. 1782–1800. <https://doi.org/10.1007/s11424-021-1217-7>
15. Astakhov V. V., Bezruchko B. P., Gulyaev Yu. V., Seleznev E. P. Multistate states of dissipatively coupled Feigenbaum systems. *Pis'ma v ZHTF*, 1989, vol. 15, no. 3, pp. 60–65 (In Russian).
16. Ivanitsky G. R. Self-organizing dynamic stability of biosystems far from equilibrium. *Physics-Uspeski*, 2017, vol. 60, iss. 7, pp. 705–730.
17. Kolesnikova S., Kustova E. Application of a stochastic extension of the analytical design of aggregated regulators to a multidimensional biomedical object. *Mathematics*, 2023, vol. 11, no. 21, 4484. <https://doi.org/10.3390/math11214484>
18. Fomichev A. O., Vavilin V. A. The reduced model of self-oscillating dynamics in an anaerobic system with sulfate-reduction. *Ecol. Modelling*, 1997, vol. 95, pp. 133–144.
19. García-Gen S., Santos L. O., Wouwer A. V. Application of a nonlinear model predictive controller to the anaerobic digestion of readily biodegradable wastes. *IFAC-PapersOnLine*, 2022, vol. 55, iss. 7, pp. 909–914. <https://doi.org/10.1016/j.ifacol.2022.07.560>. Прототип.
20. Grudyaeva E. K., Dushin S. E. *Modelirovanie upravlyaemykh processov biologicheskoy ochistki stochnykh vod* [Modeling of controlled processes of biological wastewater treatment]. Saint-Petersburg, SPbGETU «LETI» Publ., 2017. 231 p. (In Russian).
21. Tawai A., Sriariyanun M. Nonlinear optimization-based robust control approach for a two-stage anaerobic digestion process. *Hindawi Journal of Chemistry*, 2022, 8966350. <https://doi.org/10.1155/2022/8966350>
22. He D., Wang H., Tian Y., Christov N., Simeonov I. Trajectory tracking of two-stage anaerobic digestion process: A predictive control with guaranteed performance and saturated input, based on ultra-local model. *Journal of Process Control*, 2023, vol. 129, 103039.
23. Mendiola-Rodriguez T. A., Ricardez-Sandoval L. A. Integration of design and control for renewable energy systems with an application to anaerobic digestion: A deep deterministic policy gradient framework. *Energy*, 2023, vol. 274, 127212.
24. Kurth A.-C., Sawodny O. Control of age-structured population dynamics with intraspecific competition in context of bioreactors. *Automatica*, 2023, vol. 152, 110944. <https://doi.org/10.1016/j.automatica.2023.110944>
25. Mendiola-Rodriguez T. A., Ricardez-Sandoval L. A. Robust control for anaerobic digestion systems of tequila vinasses under uncertainty: A deep deterministic policy gradient algorithm. *Digital Chemical Engineering*, 2022, vol. 3, 100023.
26. Diz-Pita E., Otero-Espinar M. V. Predator–prey models: A review of some recent advances. *Mathematics*, 2021, vol. 9(15), 1783. doi:10.3390/math9151783
27. Haken G. *Advanced Synergetics. Instability Hierarchies of Self-Organizing Systems and Devices*. Springer Berlin, Heidelberg, 1983. 356 p.
28. Eskov V. V. Thermodynamics of the I. R. Prigogine non-equilibrium systems and entropy approach in the physics of living systems. *Bulletin of New Medical Technologies*, 2017, vol. 24, no. 2, pp. 7–15 (In Russian).
29. Moiseev N. N. *Izbrannye trudy v 2 tomah. T. 1. Gidrodinamika i mekhanika. Optimizaciya, issledovanie operacij i teoriya upravleniya* [Selected works in 2 volumes. Vol. 1. Hydrodynamics and mechanics. Optimization, operations research and control theory]. Moscow, Taideks Co. Publ., 2003. 376 p. (In Russian).
30. Malinetsky G. G. Habitat: parameters of order, self-organization. *Priroda*, 2023, no. 3 (1291), pp. 26–35 (In Russian). doi:10.7868/S0032874X23030031
31. Kolesnikova S. I., Fomenkova A. A. Dynamic strategies for monitoring quality control at a complex bioengineering facility. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2023, no. 2, pp. 51–60. doi:10.31799/1684-8853-2023-2-51-60, EDN: WUHAW
32. Mo R., Guo W., Batstone D., Makinia J., Li Y. Modifications to the anaerobic digestion model no. 1 (ADM1) for enhanced understanding and application of the anaerobic treatment processes – A comprehensive review. *Water Research*, 2023, vol. 244, 120504. <https://doi.org/10.1016/j.watres.2023.120504>
33. Kolesnikov A. A. *Novye nelinejnye metody upravleniya poletom* [New nonlinear methods of flight control]. Moscow, Fizmatlit Publ., 2013. 196 p. (In Russian).
34. Krasovsky A. A., Misrikhanov M. Sh. Self-organizing extrapolation controllers for power systems: history and prospects. *Automation and Remote Control*, 2006, vol. 67, iss. 5, pp. 721–725. <https://doi.org/10.1134/S0005117906050043>
35. Datsers G. Effortless basins of attraction codebase. *Zenodo*, 2021. Available at: <https://doi.org/10.5281/zenodo.5806212> (accessed 16 May 2024).
36. Diveev A. I., Shmalko E. Yu., Barabash A. V. Complete solution of the optimal control problem. *Trudy Mezhdunarodnogo simpoziuma "Nadezhnost' i kachestvo"* [Proc. Int. Symp. "Reliability and quality"], 2023, vol. 1, pp. 96–102 (In Russian).
37. Kolesnikova S. I. (RU), Fomenkova A. A. (RU) *Programma dlya realizacii vektornogo upravleniya nelinejnym bioinzhenernym ob'ektom s neustojchivymi predel'nymi sostoyaniami* [Program for the implementation of vector control of a nonlinear bioengineering object with unstable limit states]. Certificate RU of state registration of computer program, no. 2024664073, 2024.
38. *Fil'traciya i stohasticheskoe upravlenie v dinamichestkih sistemah* [Filtering and stochastic control in dynamical systems]. K. T. Leonides Ed. Moscow, Mir Publ., 1980. 406 p. (In Russian).



Обнаружение сетевых атак ботнетов на основе технологий машинного обучения и переноса знаний

Н. М. Башмаков^а, аспирант, orcid.org/0000-0002-8647-6821

В. И. Васильев^а, доктор техн. наук, профессор, orcid.org/0000-0002-6105-5481

А. М. Вульфин^{а,б}, доктор техн. наук, профессор, orcid.org/0000-0001-5857-2413, vulfin.am@ugatu.su

В. М. Картак^а, доктор физ.-мат. наук, профессор, orcid.org/0000-0001-8167-8291

А. Д. Кириллова^а, канд. техн. наук, старший преподаватель, orcid.org/0009-0000-4164-2526

^аУфимский университет науки и технологий, Заки Валиди ул., 32, Уфа, 450076, РФ

^бОмский государственный технический университет, Мира пр., 11, Омск, 644050, РФ

Введение: совершенствование сетевых средств защиты информации неразрывно связано с развитием инструментов интеллектуального мониторинга состояния и сетевого взаимодействия, повышающих наблюдаемость корпоративных информационных систем. Актуальной проблемой является оценка применимости предварительно обученных моделей машинного обучения к новым наборам данных сетевого трафика (с применением переноса обучения) и возможности их эксплуатации в реальных инфраструктурах для обнаружения узкого класса сетевых атак на примере взаимодействия скомпрометированных хостов с серверами управления и контроля ботнетов. **Цель:** совершенствование моделей и алгоритмов обнаружения сетевого трафика инфраструктур управления и контроля ботнетов в корпоративных информационных системах на основе технологий машинного обучения (в том числе глубокого обучения). **Результаты:** разработан прототип интеллектуальной системы обнаружения сетевых атак, позволяющей решать задачи сбора и предобработки данных сетевых сессий, обеспечивать взаимодействие с центром оперативного управления и мониторинга информационной безопасности, готовить данные для обучения локальных моделей анализа и управлять их жизненным циклом. Предложен алгоритм подготовки, предобработки трафика и оптимизации гиперпараметров бинарных классификаторов. Результаты экспериментов ($F1\text{-мера}=0,71$) подтверждают, что предлагаемые модели, обученные на одном наборе данных, могут успешно применяться на другом наборе узкоспециализированного домена трафика управления ботнетами. Отличительной особенностью является применение переноса обучения для глубоких нейросетевых моделей, что позволяет повысить эффективность обнаружения (величину $F1\text{-меры}$) специализированных сетевых атак на 16–21 %. **Практическая значимость:** применение переноса обучения обеспечивает возможность аккумулировать знания о проводимых атаках на различные информационные инфраструктуры в рамках единой нейросетевой модели, что позволяет повысить оперативность и достоверность обнаружения трафика управления ботнетами, тем самым усилить защищенность клиентских корпоративных информационных систем. **Обсуждение:** дальнейший подъем эффективности обнаружения специализированных сетевых атак возможен за счет применения более сложных нейросетевых моделей при использовании технологий федеративного трансферного обучения.

Ключевые слова — обнаружение сетевых атак, ботнеты, трафик управления, машинное обучение, глубокое обучение, трансферное обучение.

Для цитирования: Башмаков Н. М., Васильев В. И., Вульфин А. М., Картак В. М., Кириллова А. Д. Обнаружение сетевых атак ботнетов на основе технологий машинного обучения и переноса знаний. *Информационно-управляющие системы*, 2024, № 5, с. 41–56. doi:10.31799/1684-8853-2024-5-41-56, EDN: SWCOYY

For citation: Bashmakov N. M., Vasilyev V. I., Vulfin A. M., Kartak V. M., Kirillova A. D. Detection of network botnet attacks based on machine learning and knowledge transfer technologies. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2024, no. 5, pp. 41–56 (In Russian). doi:10.31799/1684-8853-2024-5-41-56, EDN: SWCOYY

Введение

В последние годы наблюдается устойчивая динамика роста количества сетевых атак на информационную инфраструктуру предприятий и организаций, возрастает сложность и разнообразие сценариев их реализации (<https://ics-cert.kaspersky.ru/publications/reports/2024/03/19/threat-landscape-for-industrial-automation-systems-statistics-for-h2-2023/>, <https://www.ptsecurity.com/ru-ru/research/analytics/cybersecurity-threatscape-2023-q3/>). Согласно прогнозам (<https://www.crime-research.org/news/24.01.2024/4132>), объем ущерба от киберпреступлений в мире к 2025 г. достигнет 12 трлн долл. Увеличилось число мультивектор-

ных сетевых атак, направленных на выведение из строя сразу нескольких ключевых компонент информационной инфраструктуры корпоративных систем. Одним из основных инструментов реализации подобных атак являются ботнеты. Общее число устройств в ботнет-сетях, задействованных в атаках на ИТ-системы российских компаний в I квартале 2024 г., выросло в 1,6 раза и приближается к 5 млрд. Актуальной задачей является обнаружение сетевого трафика инфраструктур управления и контроля ботнетов (Command & Control, C&C) и пресечение их деятельности на ранних стадиях проникновения в корпоративные информационные инфраструктуры.

Наиболее современными средствами защиты корпоративных информационно-телекоммуникационных систем сегодня являются многоуровневые комплексы, включающие в себя средства обнаружения и предотвращения вторжений (Intrusion Detection System, IDS), расширенного обнаружения и реагирования (Endpoint Detection and Response, EDR/XDR) и анализа сетевого трафика (Network Traffic Analysis, NTA), интегрированные с системой управления событиями безопасности (Security Information and Event Management, SIEM), в составе центров оперативного управления и мониторинга информационной безопасности (ИБ). При этом классические сетевые IDS недостаточно хорошо справляются с обнаружением новых или модифицированных сценариев реализации специализированных типов сетевых атак (например, Advanced Persistent Threat, APT), они неспособны анализировать трафик защищенных соединений (TLS, VPN) и размещаются, как правило, на периметре сети. NTA-решения способны анализировать трафик внутри периметра сети, обладают сигнатурными и адаптивными моделями обнаружения сетевых атак, широко используются при расследовании инцидентов и активном поиске угроз (Threat Hunting).

Совершенствование сетевых средств защиты неразрывно связано с развитием инструментов интеллектуального мониторинга состояния и сетевого взаимодействия, повышающих наблюдаемость корпоративных информационных систем. Согласно [1], наибольшее число исследований применения технологий искусственного интеллекта в задачах обеспечения безопасности информационных систем посвящено именно вопросам создания средств обнаружения вторжений и анализа сетевого трафика.

Значительное количество исследователей использует для этих целей публичные наборы данных сетевого трафика с размеченными типами атак (например, KDD99, NSL-KDD, CIDDS, UNSW-NB15, Bot-IoT, ToN_IoT и т. п.). При этом многие популярные наборы данных сетевого трафика устарели [2] в связи с изменением ландшафта сетевых атак и эволюцией инфраструктуры корпоративных систем, при проведении исследований необходимо использовать актуальные данные. Еще одной особенностью является оценка эффективности предлагаемых решений: количественные и качественные характеристики моделей обнаружения атак получены, как правило, для данных из одного набора, примеры полноценного внедрения разработанных подходов для обнаружения сетевых атак в реальной инфраструктуре практически не освещаются. Таким образом, актуальной проблемой сегодня является оценка качества моделей обнаружения

специализированных сетевых атак на различных наборах данных и возможности анализа данных в инфраструктуре реальной корпоративной сети.

Целью работы является совершенствование моделей и алгоритмов обнаружения сетевого трафика инфраструктур управления и контроля ботнетов в корпоративных информационных системах на основе технологий машинного обучения (в том числе глубокого обучения) и переноса знаний.

Анализ проблемы обнаружения сетевых атак ботнетов на основе технологий машинного обучения и переноса знаний

Обзор существующих публичных наборов данных сетевого трафика

В работе [2] проведен анализ 52 наиболее широко известных на сегодня наборов данных, которые могут быть использованы для исследования возможностей применения методов интеллектуального анализа и машинного обучения в задачах обнаружения вредоносного сетевого трафика, в том числе трафика инфраструктур управления и контроля ботнетов. Отмечено, что наиболее часто используемыми наборами данных являются KDD99 и NSL-KDD, но эти наборы данных устарели, и следует отказаться от их использования в пользу более актуальных. Однако авторы не указывают форматы представления анализируемых наборов данных, способы выделения ключевых признаков и используемое для этой цели программное обеспечение, что существенно затрудняет анализ «дрейфа данных», «дрейфа концепции» и возможности переноса обучения для конструируемых моделей обнаружения сетевых атак.

Сделаем далее допущение, что актуальными можно назвать наборы данных, появившиеся в последнее десятилетие, и произведем анализ форматов представления данных, перечня признаков и инструментальных средств их выделения (табл. 1).

Анализ актуальных наборов данных сетевого трафика показал, что для этапов предобработки, выделения сетевых сессий между конечными системами и последующего извлечения признаков сетевого взаимодействия конечных систем в составе вычислительной сети используется несколько основных подходов:

– для выделения сетевых сессий из дампа трафика канального уровня применяются как разнообразные сетевые инструменты, так и специализированные модульные инструменты обнаружения вторжений (например, Zeek, Bro IDS и пр.);

■ **Таблица 1.** Наборы данных сетевого трафика за 2015–2024 гг. [3–6]

■ **Table 1.** Network traffic datasets for the period 2015–2024 [3–6]

Название (год)	Используемые форматы	Формат набора данных	Количество признаков	Инструментальные средства обработки и извлечения признаков	Примечание
UNSW-NB15 (2015)	pcap, bro, argus, CSV (Netflow), отчеты	CSV	49	IXIA Perfect-Storm, Tcpdump, Argus, Bro-IDS	–
NDSec (2016)	pcap, CSV (журналы работы ОС)	pcap, CSV	–	YAF	Дамп pcap без извлечения сетевых сессий
DDoS 2016 (2016)	arff	CSV	28	Не указано	Недоступен
NGIDS-DS (2016)	pcap, CSV (журналы работы ОС)	pcap, CSV	–	IXIA Perfect Storm	Дамп pcap без извлечения сетевых сессий
UGR'16 (2016)	pcap, CSV (Netflow)	CSV	12	nfdump	–
Witty Worm (2016)	pcap	pcap	–	Не указано	Дамп pcap без извлечения сетевых сессий; исходный трафик собран в 2004 г.
Unified Host and Network (2016)	Netflow, JSON (журналы работы ОС)	CSV	11	Не указано	Нет разметки
CIDDS-001 (2017)	CSV (Netflow)	CSV	16	OpenStack	–
CIDDS-002 (2017)	CSV (Netflow)	CSV	16	OpenStack	–
CICIDS 2017 (2017)	pcap, CSV (Netflow)	CSV	71	CICFlowMeter	–
SUEE 2017 (2017)	pcap	pcap	–	Не указано	Дамп pcap без извлечения сетевых сессий
ISOT HTTP Botnet (2017)	pcap	pcap	–	Не указано	Дамп pcap без извлечения сетевых сессий; недоступен, создан в 2010 г.
PUF (2018)	–	–	–	–	Недоступен
ISOT CID (2018)	–	–	–	OpenStack, Tcpdump	Недоступен; исходный трафик собран в 2004 г.
CICDDoS 2019 (2019)	pcap, CSV (Netflow)	CSV	87	CICFlowMeter-V3	–
BoT-IoT (2019)	pcap, argus, csv (Netflow), отчеты	CSV	47	Не указано	–
MTA-KDD19 (2019)	CSV	CSV	33	–	–
IoT-23 (2020)	pcap, capinfos, dnstop, passivedns, tcpdstat, weblogng, zeek, bro	Zeek conn. log file	23	Zeek	2018–2019 гг.
InSDN (2020)	pcap, CSV (Netflow)	CSV	84	CICFlowMeter	–
CIRA-CIC-DoHBrw 2020 (2020)	pcap, CSV (Netflow)	CSV	28	DoHMeter	–
OPCUA (2020)	CSV (Netflow)	CSV	32	Python Scapy, Pyshark (Tshark)	–
TON_IoT (2021)	pcap, CSV	Zeek conn. log file	44/46	Zeek	–
VHS22 (2022)	CSV	–	48	–	–

- Окончание табл. 1
- End of Table 1

Название (год)	Используемые форматы	Формат набора данных	Количество признаков	Инструментальные средства обработки и извлечения признаков	Примечание
BH-KSU23 (2023)	CSV	–	78	CICFlowMeter	–
Trojan Detection (2023)	CSV	–	86	CICFlowMeter	–

— для описания сетевых сессий не используется единый набор признаков, например, предлагаемый семейством протоколов Netflow (xFlow), следовательно, количество и состав признаков существенно варьируются;

— достаточно часто применяются дополнительные алгоритмы конструирования признаков описания сетевого взаимодействия конечных систем (например, извлекается информация о времени доставки пакетов (information about interarrival time, IAT), определяются статистические параметры IAT, рассчитывается общее количество установленных флагов TCP, на сетевом уровне определяется общее количество подключений к (от) заданного хоста или соотношение подключений к (от) заданного хоста по отношению к общему количеству активных потоков [4]).

Следовательно, прямое сравнение эффективности работы моделей обнаружения сетевых атак на различных наборах данных без предварительного этапа преобразования признаков затруднительно.

В ряде случаев авторы наборов данных также публикуют исходные pcap-файлы (дампы трафика канального уровня), которые можно преобразовать в необходимый формат с использованием, например, таких сетевых инструментов, как tshark или cicflowmeter. Исследователи из австралийского университета Квинсленда опубликовали в едином формате Netflow такие актуальные наборы данных сетевого трафика, как UNSW-NB15, ToN-IoT, BoT-IoT, CSE-CICIDS2018. Таким образом, становится возможным прямое сравнение наборов данных и обученных на них моделей после унификации формата представления и набора признаков. Оценка эффективности моделей на новых наборах данных позволяет имитировать как условия появления новых атак (zero-days), так и «дрейф данных», связанный с изменениями сетевой инфраструктуры и характера взаимодействия конечных систем.

Анализ публикаций подтверждает широкое использование методов и алгоритмов машинного обучения для построения моделей обнаружения вредоносного сетевого трафика с чрезвычайно

высокими показателями качества на исходных наборах данных [7, 8]. Среди рассматриваемых методов часто встречаются классические методы и модели машинного обучения: метод k-средних (K-means), метод опорных векторов (SVM), случайный лес (Random Forest), деревья решений C4.5 и т. п.

В последние несколько лет создано множество новых наборов данных, результаты анализа которых в публикациях представлены гораздо меньше: MTA-KDD19, Trojan Detection, CTU-IoT, VHS22 (является комбинацией наборов данных ISOT, CICIDS-17, CTU-13 и трафика с сайта Malware Traffic Analysis), BH-KSU23 и др.

Однако почти не встречаются публикации, в которых производится оценка классических моделей при их переносе на иные возможные наборы данных со сходными признаками или в условиях их применения для реальной инфраструктуры вычислительных сетей, а также не проводится сравнение с сигнатурными системами обнаружения сетевых атак.

Анализ возможностей машинного обучения и переноса знаний в задаче обнаружения сетевого трафика инфраструктур управления и контроля ботнетов

При использовании моделей на основе глубокого обучения, напротив, становится возможным применение трансферного обучения, или «переноса обучения» (Transfer Learning, TL) [9] — возможности дообучить модель на подмножестве новых данных, сохранив накопленные ранее знания и обобщающую способность модели [10–13] для обнаружения модификаций сетевых атак. Исследование эффективности трансферного обучения раскрывается в первую очередь на примере сверточных глубоких нейросетевых моделей, а также моделей глубоких автоэнкодеров.

На основе анализа публикаций [9, 11–13] могут быть выделены следующие сценарии трансферного обучения для глубоких нейросетевых моделей (табл. 2).

Другой значимой проблемой оценки возможности переноса моделей является разнообразный состав типов сетевых атак и способов их

■ **Таблица 2.** Сценарии применения трансферного обучения в задаче обнаружения сетевых атак
 ■ **Table 2.** Scenarios for applying transfer learning to network attack detection

№	Исследование	Модель для переноса обучения	Исходный набор данных	Набор данных, на котором протестирована модель	Доля правильных ответов, %	Прочие метрики
1	Wu P., Guo H., Buckland R. [9]	CNN-CNN	UNSW-NB15	NSL-KDD	81	FPR = 98,65 (сбалансирован)
2	Masum M., Shahriar H. [14]	DNN-DNN	VGG-16	NSL-KDD	70	Precision = 82,82 Recall = 82,15
3	Singla A., Bertino E., Verma D. [15]	DNN-DNN	UNSW-NB15 (175 341 запись)	UNSW-NB15 (девять выборок из исходного набора для каждого из девяти типов атак, сбалансированных по количеству примеров)	98	—
4	Fan Y., Li Y., Zhan M., Cui H., Zhang Y. [16]	CNN-CNN	CICIDS 2017	Собственный набор данных	91	FPR = 0,034 TPR = 0,99 TNR = 0,96
5	Idrissi I., Azizi M., Moussaoui O. [17]	CNN-CNN	BoT-IoT	TON-IoT	99	Precision = 0,99 Recall = 0,99
6	Rodríguez E., et al. [18]	CNN	BoT-IoT	UNSW-NB15	99	Precision = 0,996 Recall = 0,991 FPR = 0,05 F1 = 0,992
7	Gebresilassie S. K., Rafferty J., Chen L., Cui Z., Abu-Tair M. [19]	CNN	Собственный набор данных	Собственный набор данных	99	Precision = 0,981 Recall = 0,981 F1 = 0,981
8	Yehezkel A., Elyashiv E., Soffer O. [20]	DAE	Собственный набор данных	Собственный набор данных	98	—

реализации в различных наборах данных. Для проанализированных ранее наборов данных типы сетевых атак приведены в табл. 3.

Большинство типов сетевых атак (сетевое сканирование, брутфорс, DDoS) либо с высокой долей вероятности обнаруживаются и блокируются классическими средствами анализа сетевого трафика (IDS, NGFW), либо не представлены сразу в нескольких вышеупомянутых наборах данных в достаточном объеме, что затрудняет их сравнительный анализ. Следовательно, особое внимание необходимо обратить на анализ возможностей обнаружения специализированных сетевых атак, что является наиболее сложной задачей.

Анализ особенностей сетевого трафика инфраструктур управления и контроля ботнетов

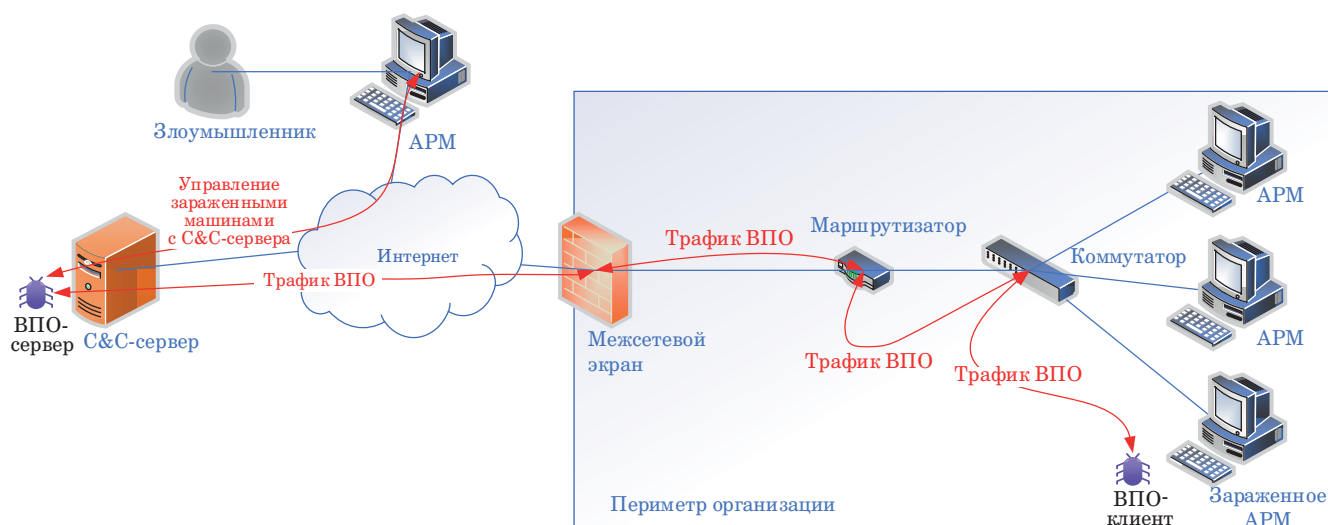
Среди специализированных типов атак следует выделить сетевой трафик инфраструктур управления и контроля ботнетов, который ха-

рактеризует взаимодействие серверов управления под контролем злоумышленника и скомпрометированных устройств. Взаимодействие со скомпрометированными устройствами заключается в отправке им команд и получении от них данных. Согласно базе знаний ATT & CK (Adversarial Tactics, Techniques & Common Knowledge) компании MITRE, сетевой трафик C&C является одной из тактик реализации сетевых атак. Злоумышленники могут использовать зараженные машины для проведения разведки внутри периметра, распространения вредоносного программного обеспечения или осуществления DDoS-атак.

Выделяют несколько топологий серверов C&C: звезда, звезда с несколькими серверами для обеспечения отказоустойчивости, иерархическая, случайная и P2P. Упрощенная (не приведены подробности внутренней сетевой инфраструктуры, C&C-сервер представлен в единственном экземпляре) схема взаимодействия C&C-сервера и зараженных устройств показана на рис. 1.

- **Таблица 3.** Типы сетевых атак в актуальных наборах данных сетевого трафика
- **Table 3.** Network attack types in current network traffic datasets

Набор данных	Количество типов сетевых атак	Типы сетевых атак в наборе данных
UNSW-NB15	9	Normal, Fuzzers, Analysis, Backdoors, DoS, Exploits, Generic, Reconnaissance, Shellcode, Worms
UGR'16	8	DoS, scan11, scan44, nerisbotnet, blacklist, anomaly-udpscan, anomaly-sshscan, anomaly-spam
Unified Host and Network	–	Нет разметки
CIDDS-001	4	Normal, portScan, DoS, pingScan, bruteForce
CIDDS-002	4	Normal, portScan, DoS, pingScan, bruteForce
CICIDS 2017	4	Benign, DDoS, PortScan, WebAttacks, Infiltration
CSE-CIC-IDS2018	6	Bruteforce attack, DoS attack, Web attack, Infiltration attack, Botnet attack, DDoS+PortScan
CICDDoS 2019	13	PortMap DDoS, NetBIOS DDoS, LDAP DDoS, MSSQL DDoS, UDP DDoS, UDP-Lag DDoS, SYN DDoS, NTP DDoS, DNS DDoS, SNMP DDoS, SSDP DDoS, WebDDoS, TFTP DDoS
Bot-IoT	4	Normal, DDoS, DoS, Reconnaissance, Theft
IoT-23 (CTU-IoT)	8	Benign, C&C, DDoS, FileDownload, HeartBeat, Mirai, Okiru, PartOfAHorizontalPortScan, Torii
InSDN	7	Normal, botnet, brute-force-attack, DoS, DDoS, Web_attack, Probe attack, Exploitation (R2L)
CIRA-CIC-DoHBrw	2	Non-DoH, Benign-DoH, Malicious-DoH
OPCUA	3	Normal, DoS, MITM, Impersonation
ToN-IoT	9	Normal, Backdoor, DDoS, DoS, Injection, Mitm, Password, Ransomware, Scanning, XSS
MTA-KDD19	1	Normal, Malware
Trojan Detection	1	Normal, Malware
VHS22	1	Normal, Malware
BH-KSU23	1	Normal, Malware



- **Рис. 1.** Упрощенная схема C&C-взаимодействия: АРМ – автоматизированное рабочее место; ВПО – вредоносное программное обеспечение
- **Fig. 1.** Simplified diagram of C&C interaction: АРМ – automated workstation; ВПО – malicious software

Особенности сетевого трафика C&C между зараженными устройствами и командным центром:

- сетевой трафик практически соответствует обычному использованию протоколов и аналогичен обычному трафику;
- объем трафика передачи управляющих команд небольшой;
- в анализируемой сети может быть очень мало зараженных устройств;
- использование шифрования сетевых C&C-сессий.

Злоумышленники часто используют различные техники для скрытия C&C-трафика, что усложняет его обнаружение классическим сигнатурным способом. Применение методов машинного обучения к обнаружению подобного рода сетевых атак, напротив, является перспективным ввиду их адаптивности и возможности обнаружения модифицированных сценариев их реализации за счет возможности дообучения на малых объемах данных.

Следовательно, актуальной является как оценка применимости предварительно обученных моделей машинного обучения к новым наборам данных сетевого трафика (с применением

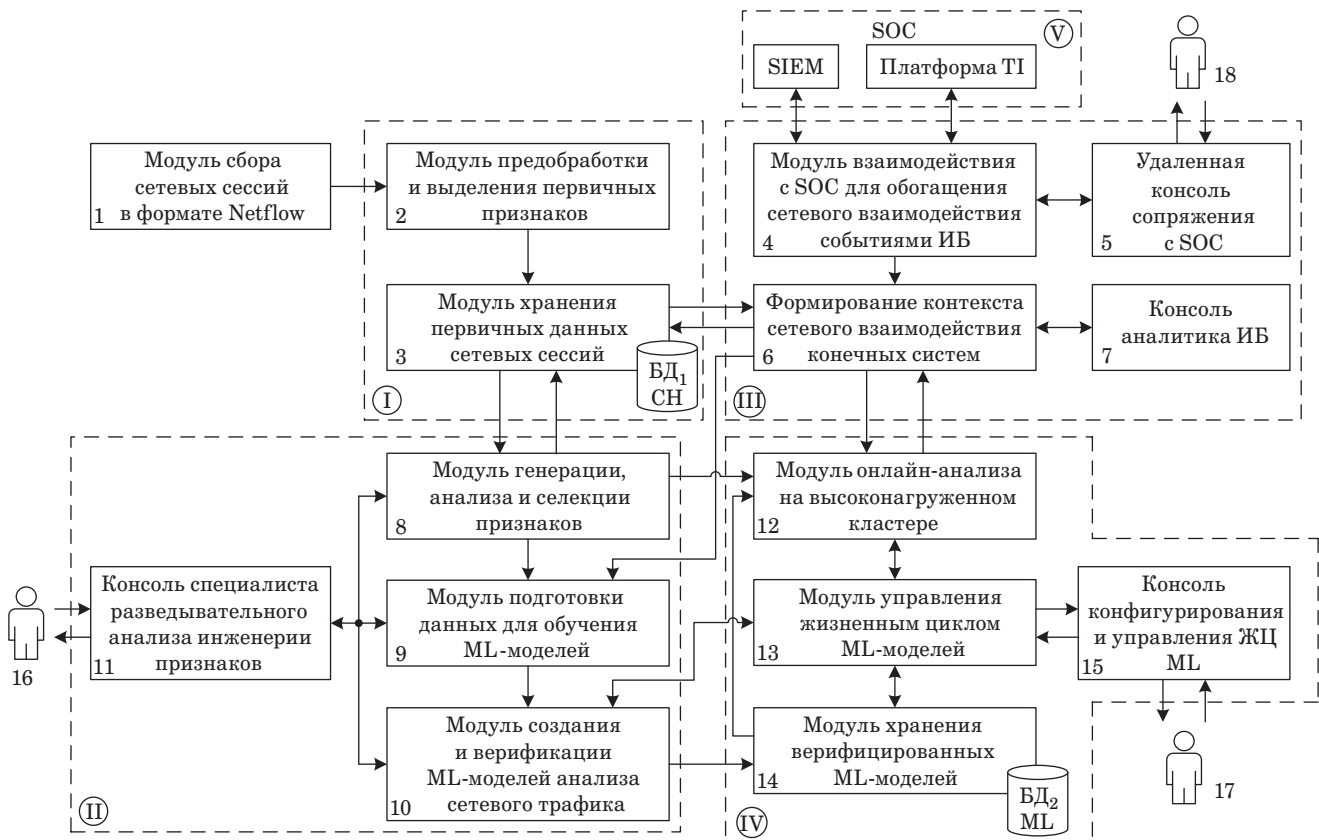
переноса обучения), так и возможность их эксплуатации в реальных инфраструктурах для обнаружения узкого класса сетевых атак на примере взаимодействия скомпрометированных хостов с C&C-серверами.

Разработка системы обнаружения сетевого трафика инфраструктур управления и контроля ботнетов

Структурная схема предлагаемой системы обнаружения сетевого трафика инфраструктур управления и контроля ботнетов представлена на рис. 2. Данная структурная схема состоит из четырех подсистем (I, II, III и IV, символом V на схеме обозначен центр мониторинга ИБ, который получает оповещения от системы обнаружения).

Основные подсистемы:

- подсистема I получает данные от внешнего модуля 1 сбора сетевых сессий в формате Netflow, далее в модуле 2 осуществляется предобработка поступивших данных, выделяются ключевые первичные признаки (например, сетевые адреса и порты источника и назначения, количество переданных и полученных пакетов и байт, про-



■ **Рис. 2.** Структурная схема системы обнаружения сетевого трафика инфраструктур управления и контроля ботнетов
 ■ **Fig. 2.** Structural diagram of network traffic detection system for botnet management and checking infrastructures

токолы сетевого и вышестоящих уровней, длительность и т. д.) и производится запись сетевых сессий в хранилище на основе колоночной базы данных Yandex ClickHouse (модуль 3). Специфика подсистемы в задаче обнаружения управляющего сетевого трафика C&C заключается в подборе узлов-источников Netflow (модуль 1), а также составлении экспертами перечня наблюдаемых узлов сети исходя из их роли;

– в подсистеме II осуществляется генерация, анализ и селекция признаков (модуль 8), выполняется подготовка данных для обучения моделей – создаются обучающие, проверочные и тестовые выборки по схеме k-fold перекрестной проверки (модуль 9), создаются и верифицируются модели машинного обучения (модуль 10), управление осуществляется специалистом разведывательного анализа данных и инженерии признаков (модуль 11);

– подсистема III предназначена для создания контекста анализа сетевых сессий (совокупность сведений о принадлежности взаимодействующих узлов к сегментам сети, типам запрашиваемых ресурсов, задействованном программном обеспечении, пользовательских учетных записях и т. д., характеризующая событие ИБ, с которым ассоциирована данная сетевая сессия) с учетом событий ИБ в ходе взаимодействия с центром оперативного управления и мониторинга ИБ (согласно базе знаний и набору эвристических правил проводится дополнительная оценка сетевых сессий на предмет отношения к C&C-трафику);

– подсистема IV предназначена для управления жизненным циклом созданных моделей машинного обучения (модуль 13) и обеспечением их работы для анализа поступающих событий в режиме, близком к реальному времени (модуль 12). Обновление моделей производится либо по заданному сценарию, либо по достижении ими срока устаревания. Реализована подсистема с использованием фреймворков Apache Airflow и Apache Spark.

Вычислительный эксперимент по обнаружению сетевого трафика инфраструктур управления и контроля ботнетов

Для проведения серии экспериментов были выбраны специализированные наборы данных BH-KSU23 и Trojan Detection, содержащие размеченный трафик C&C-сессий и сессии нормальной работы (табл. 4) виртуальной сетевой инфраструктуры модельных стендов.

Схема серии экспериментов (рис. 3) включает следующие основные шаги.

1. Наборы данных BH-KSU (B) и Trojan Detection (T) в бинарном колоночном формате parquet размещаются в хранилище, доступном для последующей обработки.

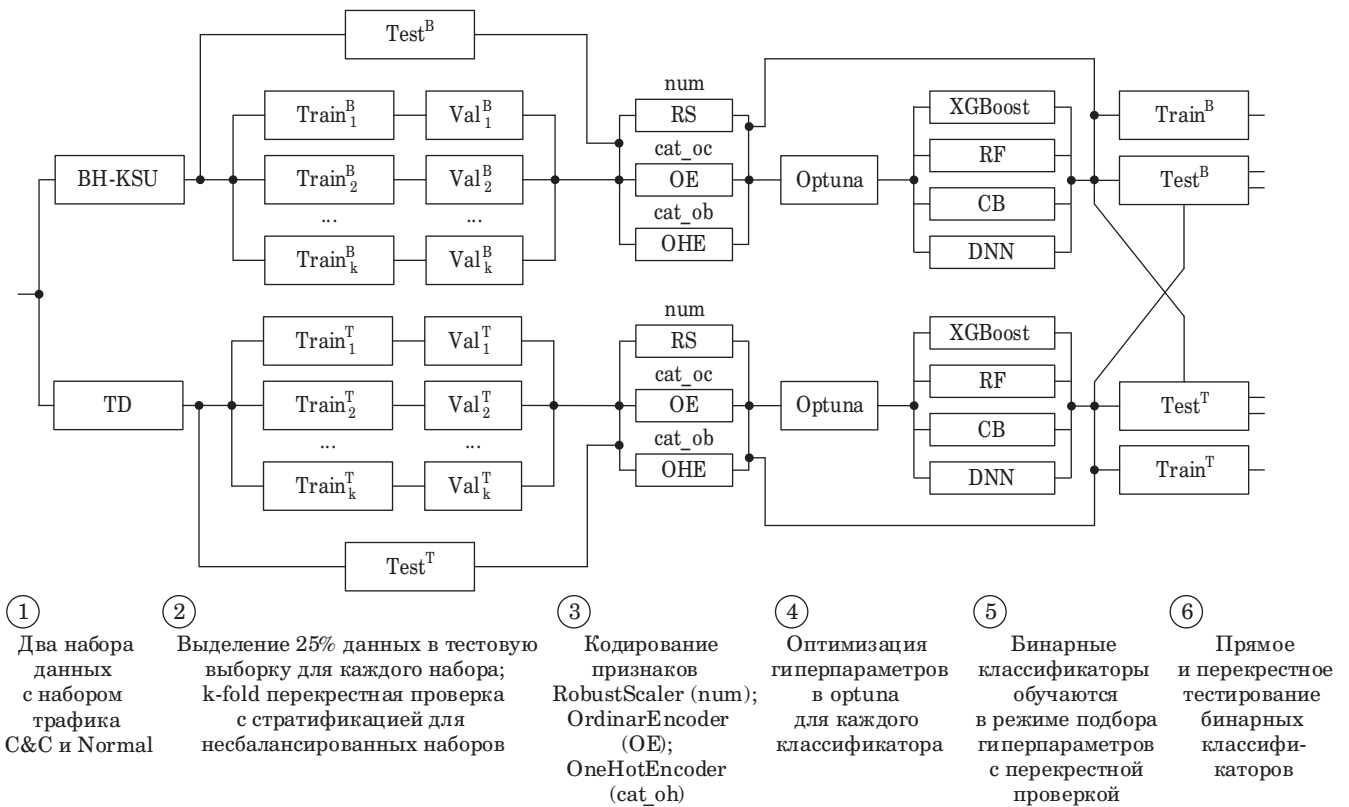
2. Из каждого набора данных выделяются 25 % примеров для тестовых выборок ($Test^B$ и $Test^T$) с сохранением пропорций между классами. На основе оставшихся данных готовятся выборки обучающих ($Train_i^B$ и $Train_j^T$) и проверочных (Val_i^B и Val_j^T) данных по схеме перекрестной проверки с пятью заходами с сохранением пропорций между классами (k-fold cross-validation в режиме стратификации).

3. Строится конвейер кодирования признаков. Все признаки разделяются на три группы: количественные признаки (используется RobustScaler, который при масштабировании вычитает из данных медиану и делит результат на интерквартильный размах), порядковые категориальные признаки (для кодирования используется OrdinarEncoder), номинальные признаки (для кодирования данных без внутренней иерархии используется OneHotEncoder) и категориальные порядковые признаки (для целевой переменной использован LabelEncoder). Параметры каждого типа кодировщиков настраиваются на обучающей выборке и применяются для преобразования проверочных и тестовых выборок во избежание «утечки данных».

■ Таблица 4. Характеристики наборов данных C&C

■ Table 4. Characteristics of C&C datasets

Набор данных	Источник	Количество признаков	Утилита извлечения признаков	Количество сетевых сессий	Особенности
BH-KSU23 (B)	Университет имени Короля Сауда (KSU), Саудовская Аравия	79	CICFlowmeter	400 000	Семь различных C&C-систем
Trojan Detection (T)	Университет Дрексела (Drexel), США	79	CICFlowmeter	180 000	1041 вариация серверной части C&C, 960 типов клиент-серверного взаимодействия C&C



■ **Рис. 3.** Схема проведения вычислительного эксперимента

■ **Fig. 3.** Scheme of conducting computational experiment

4. С помощью фреймворка Optuna выполняется оптимизация гиперпараметров четырех бинарных классификаторов на основе вероятностной оптимизации согласно алгоритму TPE.

5. Строятся бинарные классификаторы (табл. 5) с подобранными гиперпараметрами.

6. Выполняются прямая и перекрестная проверки обученных классификаторов на тестовых выборках из наборов данных T и B.

Бинарный классификатор на основе полносвязной глубокой нейронной сети (Deep Neural Network, DNN) используется в схеме трансферного обучения.

В схеме трансферного обучения модели DNN использованы 25 % данных из целевого набора (из подмножества Train) по следующим сценариям:

- полное дообучение модели (без «заморозки» весовых коэффициентов нейронов всех слоев);
- «заморозка» весовых коэффициентов первых двух слоев;
- «заморозка» весовых коэффициентов первых трех слоев.

Серия экспериментов показала, что наилучшие результаты достигаются в случае дообучения всей модели DNN, но применение «заморозки» первых двух слоев, выполняющих

функцию извлечения признаков, снижает качество классификации на 2–5 %, сокращая время на дообучение на 30–40 % (в зависимости от размера целевого набора). Таким образом, эффективность применения переноса обучения заключается:

- в сокращении времени на дообучение модели на 30–40 % в случае «заморозки» слоев извлечения признаков;
- в возможности использовать предварительно обученную на большом объеме данных модель в сценариях обнаружения модификаций специализированных сетевых атак (в том числе при наличии ограниченного набора размеченных данных для дообучения, собранных для конкретной сети).

При оценке качества моделей машинного обучения, используемых для классификации, применяются следующие метрики:

- True Positive (TP) – количество правильно предсказанных положительных случаев;
- True Negative (TN) – количество правильно предсказанных отрицательных случаев;
- False Positive (FP) – количество неправильно предсказанных положительных случаев;
- False Negative (FN) – количество неправильно предсказанных отрицательных случаев;

■ Таблица 5. Используемые модели машинного обучения

■ Table 5. Machine learning models used

Модель	Описание	Основные параметры модели, подобранные в результате оптимизации гиперпараметров			Особенности использования
Random Forest	В случайном лесе объединяется множество деревьев решений, каждое обучается на разных подвыборках данных. Окончательные предсказания делаются путем усреднения предсказаний каждого дерева	n_estimators	Количество деревьев	100	Алгоритм случайного леса обладает более высокой предсказательной точностью по сравнению с одним деревом решений; возможность работы с категориальными переменными
		max_features	Количество параметров, которые следует учитывать	\sqrt{n} , n – количество примеров	
		max_depth	Максимальная глубина дерева	50	
		min_samples_split	Минимальное количество выборок, необходимое для разделения внутреннего узла	6	
		min_samples_leaf	Минимальное количество выборок, которое должно находиться в листовом узле	1	
XGBoost	Ансамблевый метод, который объединяет несколько слабых моделей-классификаторов на основе деревьев решений для создания одной сильной модели	max_depth	Максимальная глубина дерева	16	Обладает высокой точностью предсказаний благодаря своей способности последовательно уменьшать ошибки, часто достигая более высокой точности по сравнению с другими алгоритмами; возможность работы с категориальными переменными
		learning_rate	Скорость обучения	0,029	
		n_estimators	Количество слабых классификаторов в ансамбле	316	
		min_child_weight	Минимальная сумма веса экземпляра, необходимая для дочернего элемента	2	
		lambda	Коэффициент регуляризации L2 по весам	2,840	
		alpha	Коэффициент регуляризации L1 по весам	1,391	
		eta	Уменьшение размера шага для предотвращения переобучения	0,051	
		gamma	Минимальное сокращение потерь, необходимое для создания дальнейшего разделения на конечном узле дерева	0,383	
		subsample	Соотношение подвыборок обучающих экземпляров	0,997	
CatBoost	Осуществление выбора случайных образцов из набора данных, создание дерева решений для каждого выбранного образца, получение предсказаний от каждого дерева, проведение голосования для каждого предсказанного результата и выбор наиболее часто встречающегося предсказания как окончательного результата	iterations	Максимальное количество деревьев, которое можно построить при решении задач машинного обучения	8550	Отличается высоким качеством предсказаний без необходимости тонкой настройки параметров, поддержкой категориальных признаков, быстрой и масштабируемой версией для GPU, улучшенной точностью за счет уменьшения переобучения, быстрыми предсказаниями и хорошей работой с малым
		learning_rate	Скорость обучения	0,025	
		l2_leaf_reg	Коэффициент при члене регуляризации L2 функции стоимости	2,681	
		random_strength	Степень случайности, используемая для оценки разделений при выборе древовидной структуры	0,0003	
		depth	Глубина деревьев	9	
		bagging_temperature	Определяет настройки байесовского бутстрапа	0,211	

- Окончание табл. 5
- End of Table 5

Модель	Описание	Основные параметры модели, подобранные в результате оптимизации гиперпараметров			Особенности использования
		od_type	Тип используемого детектора переобучения	IncToDec	объемом данных; специально разработан для работы с категориальными данными
		od_wait	Количество итераций для продолжения обучения после итерации с оптимальным значением метрики	45	
		min_data_in_leaf	Минимальное количество обучающих выборок в листе	21	
		leaf_estimation_iterations	Параметр регулирует количество шагов, выполняемых в каждом дереве при вычислении значений листьев	3	
		max_ctr_complexity	Максимальное количество функций, которые можно объединить	3	
DNN	Полносвязная глубокая нейронная сеть прямого распространения	learning_rate_init	Коэффициент скорости обучения	0,085	Проблемой является представление категориальных переменных в векторе входных признаков
		1 layer	Количество нейронов	98	
		2 layer	Количество нейронов	128	
		3 layer	3-й слой + дропаут	64	
		4 layer	4-й слой + дропаут	32	
		5 layer	5-й слой	4	
		output layer	Один нейрон + логистическая функция активации	—	
		—	Функция потерь – Binary Cross-Entropy With Logits	—	
		learning_rate	Скорость обучения	constant	
activation	Тип функции активации	relu			

– False Positive rate (FPR) – доля отрицательных объектов, неправильно предсказанных положительными:

$$FPR = FP / (FP + TN);$$

– Precision (точность) – показывает долю правильно предсказанных положительных случаев среди всех предсказанных положительных случаев:

$$Precision = \frac{TP}{TP + FP};$$

– Recall (полнота) – показывает долю правильно предсказанных положительных случаев среди всех реальных положительных случаев:

$$Recall = \frac{TP}{TP + FN};$$

– F1-мера – является гармоническим средним точности и полноты:

$$F1 = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall}.$$

Результаты оценки качества моделей приведены в табл. 6 и на рис. 4.

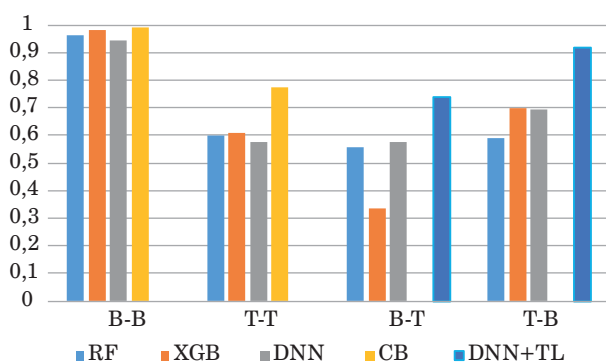
Высокие показатели F1-меры в экспериментах 1 и 2 ($F1 > 0,96$), а также ее относительно высокие показатели в экспериментах 6 и 7 ($F1 > 0,6$) указывают на то, что наборы данных BH-KSU23 и Trojan Detection имеют характеристики, позволяющие моделям машинного обучения успешно распознавать C&C-трафик.

Результаты экспериментов 9, 10 и 12, 13 (особенно эксперимент 13, где F1-мера достигла значения 0,70) подтверждают, что модели, обученные на одном наборе данных, могут успешно применяться на другом наборе в узкоспециализированном домене.

■ **Таблица 6.** Результаты серии экспериментов на наборах данных
 ■ **Table 6.** Results of a series of experiments on data sets

№	Модель	Исходный набор данных	Набор данных, на котором протестирована модель	Оценки качества модели на тестовом наборе			
				FPR	Precision	Recall	F1
1	Random Forest	BH-KSU23	BH-KSU23	0,042	0,949	0,978	0,964
2	Gradient Boosting	BH-KSU23	BH-KSU23	0,021	0,974	0,989	0,982
3	CatBoost	BH-KSU23	BH-KSU23	0,000	0,990	0,990	0,990
4	DNN	BH-KSU23	BH-KSU23	0,057	0,933	0,952	0,942
5	CatBoost	Trojan Detection	Trojan Detection	0,042	0,750	0,791	0,772
6	Random Forest	Trojan Detection	Trojan Detection	0,406	0,606	0,594	0,601
7	Gradient Boosting	Trojan Detection	Trojan Detection	0,322	0,652	0,574	0,611
8	DNN	Trojan Detection	Trojan Detection	0,342	0,670	0,656	0,663
9	Random Forest	BH-KSU23	Trojan Detection	0,647	0,504	0,629	0,559
10	Gradient Boosting	BH-KSU23	Trojan Detection	0,203	0,554	0,241	0,336
11	DNN	BH-KSU23	Trojan Detection	0,593	0,495	0,690	0,577
12	Random Forest	Trojan Detection	BH-KSU23	0,851	0,451	0,861	0,592
13	Gradient Boosting	Trojan Detection	BH-KSU23	0,517	0,581	0,883	0,701
14	DNN	Trojan Detection	BH-KSU23	0,523	0,572	0,879	0,693
15	DNN + TL	BH-KSU23	Trojan Detection	0,445	0,712	0,773	0,741
16	DNN + TL	Trojan Detection	BH-KSU23	0,221	0,876	0,966	0,912

Примечание: В моделях 1–14 TL не применяется, в моделях 15, 16 – применяется.



■ **Рис. 4.** Оценка F1-меры для серии экспериментов
 ■ **Fig. 4.** Estimation of F1-measure for a series of experiments

Результаты экспериментов с оптимизацией гиперпараметров с помощью фреймворка Optuna (эксперименты 3 и 5) демонстрируют, что модели CatBoost на наборах данных BH-KSU23 и Trojan Detection способны достичь высоких показателей точности, полноты и F1-меры после тщательной настройки гиперпараметров. Модель CatBoost на наборе данных Trojan Detection в эксперименте 5 показала высокие результаты, а на наборе данных BH-KSU23 – наилучшие результаты.

После завершения оптимизации гиперпараметров в экспериментах 3 и 5 были построены диаграммы оценки значимости признаков для классификаторов (результаты сведены в табл. 7).

Наиболее значимыми оказались признаки, характеризующие сетевой и транспортный уровни.

Результаты работы классификаторов с использованием глубокой нейронной сети сопоставимы с результатами моделей на основе комитетов деревьев решений. Однако при дообучении модели на 25 % данных из целевого набора результаты (значение F1-меры) на тестовой выборке значительно улучшаются (на 16,4 и 21,9 % соответственно) без существенного ухудшения результата на тестовом подмножестве исходного набора, что свидетельствует об эффективности трансферного обучения даже на сравнительно простых нейросетевых моделях.

Заключение

Для повышения эффективности систем обнаружения сетевого трафика инфраструктур управления и контроля ботнетов в корпоративных информационных системах предложено использовать модели и алгоритмы машинного

■ **Таблица 7.** Оценка значимости признаков для бинарной классификации
 ■ **Table 7.** Evaluation of the significance of features for binary classification

Эксперимент и целевой набор данных	Алгоритм оценки значимости признаков	Описание	Отобранные признаки	Описание отобранных признаков
3 ВНКСУ-23	Feature Importance (значимость параметра)	Алгоритм основан на оценке количества разбиений, которые использует признак в деревьях решений, и на уменьшении ошибки после использования признака	ECE_Flag_Cnt	Количество пакетов с флагом ECE
			Src_Port	Порт источника
			Pkt_Len_Var	Разница в длине пакета
5 Trojan Detection			Source_Port	Порт источника
			Bwd_URG_Flags	Количество раз, когда флаг URG устанавливался для пакетов, отправляемых в обратном направлении
			Fwd_Header_Length	Общее количество байт, используемых для заголовков в прямом направлении
3 ВНКСУ-23	Permutation Importance (важность перестановки)	Алгоритм оценивает, как перетасовка значений признака влияет на точность модели. Если перетасовка сильно снижает точность, значит, признак важен	Src_Port	Порт источника
			Dst_Port	Порт назначения
			Init_Bwd_Win_Byts	Общее количество байт, отправленных в исходном окне в обратном направлении
5 Trojan Detection			Init_Win_bytes_forward	Общее количество байт, отправленных в исходном окне в обратном направлении
			Source_Port	Порт источника
			Init_Win_bytes_backward	Общее количество байт, отправленных в исходном окне в обратном направлении

обучения, в том числе глубокого обучения и переноса знаний (трансферного обучения).

Разработан прототип интеллектуальной системы обнаружения сетевых атак, позволяющей решать задачи сбора и предобработки данных сетевых сессий, обеспечивать взаимодействие с центром оперативного управления и мониторинга ИБ, готовить данные для обучения локальных моделей анализа и управлять их жизненным циклом. Проведенные вычислительные эксперименты позволяют сделать вывод о высокой эффективности обнаружения C&C-трафика с помощью предлагаемого подхода.

В ходе эксперимента был рассмотрен один конкретный тип атак – трафик C&C, на котором производилось обучение бинарных классификаторов на наборах данных Trojan Detection и ВНКСУ23. В этом случае эффективность обнаружения вредоносной активности при перекрестном применении моделей оказалась достаточно высокой.

Это в первую очередь связано с тем, что сходные типы атак имеют похожее отражение в параметрах сетевого трафика, и, соответственно, в наборах данных, что позволяет с осторожным оптимизмом предположить, что модель, обу-

ченная для обнаружения специализированных атак, сможет обнаруживать сходные типы атак и в реальном трафике. Дальнейшее повышение эффективности обнаружения сетевых атак C&C возможно за счет:

- более тщательного выбора параметров для анализа (с применением для группы моделей алгоритма оценки значимости с перекрестной проверкой);
- использования еще одного набора данных с трафиком целевой атаки для перекрестной проверки эффективности моделей (например, МТА-KDD-19) и применения технологий трансферного обучения;
- применения глубоких сверточных нейросетевых моделей и моделей с долгой-краткосрочной памятью;
- применения технологий федеративного трансферного обучения.

Финансовая поддержка

Работа выполнена в ОмГТУ в рамках государственного задания Минобрнауки России на 2023–2025 годы № FSGF-2023-0004.

Литература

1. **Kaur R., Gabrijelčić D., Klobučar T.** Artificial intelligence for cybersecurity: Literature review and future research directions. *Information Fusion*, 2023, vol. 97, pp. 101804. doi:10.1016/j.inffus.2023.101804
2. **Yang Z., Liu X., Li T., Wu D., Wang J., Zhao Y., Han H.** A systematic literature review of methods and datasets for anomaly-based network intrusion detection. *Computers & Security*, 2022, vol. 116, pp. 102675. doi:10.1016/j.cose.2022.102675
3. **Moustafa N.** A new distributed architecture for evaluating AI-based security systems at the edge: Network TON_IoT datasets. *Sustainable Cities and Society*, 2021, vol. 72, pp. 102994. doi:10.1016/j.scs.2021.102994
4. **Szumelda P., Orzechowski N., Rawski M., Janicki A.** VHS-22 – a very heterogeneous set of network traffic data for threat detection. *2022 European Interdisciplinary Cybersecurity Conf.*, Barcelona, Spain, 15–16 June 2022, pp. 72–78. doi:10.1145/3528580.3532843
5. **Binsaeed K., Alaa-aldeen H.** *BH-KSU23: A Novel Dataset for Evaluating and Enhancing Intrusion Detection Systems Targeting Command-and-Control Traffic*. Mendeley Data, 2023, ver. 1. <https://data.mendeley.com/datasets/wjxc69xj3n/1> (дата обращения: 03.10.2023).
6. **Subhadeep Ch.** *Trojan Detection [Data set]*, 2021. doi:10.34740/KAGGLE/DSV/2625272. <https://www.kaggle.com/datasets/subhajournal/trojan-detection> (дата обращения: 03.10.2023).
7. **Vinayakumar R., Alazab M., Soman K. P., Poor-nachandran P., Al-Nemrat A., Venkatraman S.** Deep learning approach for intelligent intrusion detection system. *IEEE Access*, 2019, vol. 7, pp. 41525–41550. doi:10.1109/ACCESS.2019.2895334
8. **Jan S. U., Ahmed S., Shakhov V., Koo I.** Toward a lightweight intrusion detection system for the Internet of Things. *IEEE Access*, 2019, vol. 7, pp. 42450–42471. doi:10.1109/ACCESS.2019.2907965
9. **Wu P., Guo H., Buckland R.** A transfer learning approach for network intrusion detection. *2019 IEEE 4th Intern. Conf. on Big Data Analytics (ICBDA)*, IEEE, Suzhou, China, 15–18 March 2019, pp. 281–285. doi:10.1109/ICBDA.2019.8713213
10. **Беликов В. В.** Использование методов глубокого обучения с подкреплением для отбора признаков сетевого трафика при обнаружении компьютерных атак. *Программирование*, 2022, № 6, с. 3–13. doi:10.31857/S0132347422060024, EDN: KWOQRH
11. **Иогансон И.** Обзор методов федеративного обучения. *CEUR Workshop Proceeding*, 2023. <https://www.researchgate.net/publication/376083291> Обзор_методов_федеративного_обучения (дата обращения: 14.07.2024).
12. **Новикова Е. С., Котенко И. В., Мелешко А. В., Израйлов К. Е.** Обнаружение вторжений на основе федеративного обучения: архитектура системы и эксперименты. *Вопросы кибербезопасности*, 2023, № 6(58), с. 50–66. doi:10.21681/2311-3456-2023-6-50-66
13. **Новикова Е. С., Федорченко Е. В., Котенко И. В., Холод И. И.** Аналитический обзор подходов к обнаружению вторжений, основанных на федеративном обучении: преимущества использования и открытые задачи. *Информатика и автоматизация*, 2023, т. 22, № 5, с. 1034–1082. doi:10.15622/ia.22.5.4
14. **Masum M., Shahriar H.** TL-NID: Deep neural network with transfer learning for network intrusion detection. *2020 15th Intern. Conf. for Internet Technology and Secured Transactions (ICITST)*, IEEE, London, UK, 8 December 2020, pp. 1–7. doi:10.23919/ICITST51030.2020.9351317
15. **Singla A., Bertino E., Verma D.** Overcoming the lack of labeled data: Training intrusion detection models using transfer learning. *2019 IEEE Intern. Conf. on Smart Computing, IEEE*, Washington, USA, 12 June 2019, pp. 69–74. doi:10.1109/SMARTCOMP.2019.00031
16. **Fan Y., Li Y., Zhan M., Cui H., Zhang Y.** IoTDefender: A federated transfer learning intrusion detection framework for 5G IoT. *2020 IEEE 14th Intern. Conf. on Big Data Science and Engineering (BigDataSE)*, IEEE, Guangzhou, China, 31 December 2020, pp. 88–95. doi:10.1109/BigDataSE50710.2020.00020
17. **Idrissi I., Azizi M., Moussaoui O.** Accelerating the update of a DL-based IDS for IoT using deep transfer learning. *Indonesian Journal of Electrical Engineering and Computer Science*, 2021, vol. 23, no. 2, pp. 1059–1067. doi:10.11591/ijeecs.v23.i2.pp1059-1067
18. **Rodríguez E., Valls P., Otero B., Costa J. J., Verdú J., Pajuelo M. A., Canal R.** Transfer-learning-based intrusion detection framework in IoT networks. *Sensors*, 2022, vol. 22, no. 15, pp. 5621. doi:10.3390/s22155621
19. **Gebresilassie S. K., Rafferty J., Chen L., Cui Z., Abu-Tair M.** Transfer and CNN-based de-authentication (disassociation) DoS attack detection in IoT Wi-Fi networks. *Electronics*, 2023, vol. 12, no. 17, pp. 3731. doi:10.3390/electronics12173731
20. **Yehezkel A., Elyashiv E., Soffer O.** Network anomaly detection using transfer learning based on auto-encoders loss normalization. *Proceedings of the 14th ACM Workshop on Artificial Intelligence and Security*, Korea, 15 November 2021, pp. 61–71. doi:10.1145/3474369.3486869

UDC 004.056

doi:10.31799/1684-8853-2024-5-41-56

EDN: SWCOYY

Detection of network botnet attacks based on machine learning and knowledge transfer technologiesN. M. Bashmakov^a, Post-Graduate Student, orcid.org/0000-0002-8647-6821V. I. Vasilyev^a, Dr. Sc., Tech., Professor, orcid.org/0000-0002-6105-5481A. M. Vulfin^{a,b}, Dr. Sc., Tech., Professor, orcid.org/0000-0001-5857-2413, vulfin.am@ugatu.suV. M. Kartak^a, Dr. Sc., Phys.-Math., Professor, orcid.org/0000-0001-8167-8291A. D. Kirillova^a, PhD, Tech., Senior Lecturer, orcid.org/0009-0000-4164-2526^aUfa University of Science and Technology, 32, Z. Validi St., Ufa, 450076, Russian Federation^bOmsk State Technical University, 11, Mira Pr., 644050, Omsk, Russian Federation

Introduction: The improvement of network information protection tools is inextricably linked to the development of tools for intelligent monitoring of the state and network interaction, increasing the observability of corporate information systems. A pressing issue is to assess the applicability of pre-trained machine learning models to new network traffic datasets (using transfer learning) and the possibility of their exploitation in real infrastructures to detect a narrow class of network attacks using the example of interactions between compromised hosts and botnet control servers. **Purpose:** To improve models and algorithms for detecting network traffic of botnet management and control infrastructures in corporate information systems based on machine learning technologies (including deep learning). **Results:** We develop a prototype of an intelligent network attack detection system, which makes it possible to solve the problems of collecting and pre-processing network session data, ensuring interaction with the operational control and information security monitoring center, preparing data for training local analysis models and managing their life cycle. We propose an algorithm for preparation, preprocessing of traffic and optimization of hyperparameters of binary classifiers. The experimental results (F1-measure = 0.71) confirm that the proposed models trained on one dataset can be successfully applied to another dataset of a highly specialized botnet control traffic domain. A distinctive feature is the use of transfer learning for deep neural network models, which makes it possible to increase the efficiency of detecting specialized network attacks by 16–21%. **Practical relevance:** The use of transfer learning makes it possible to accumulate knowledge about attacks on various information infrastructures within a single neural network model, which allows one to increase efficiency and reliability of detecting botnet control traffic, as well as to increase the security of client corporate information systems. **Discussion:** Further improvement of the efficiency of detection of specialized network attacks is possible through the use of more complex neural network models involving federated transfer learning technologies.

Keywords — network attack detection, botnets, control traffic, machine learning, deep learning, transfer learning.

For citation: Bashmakov N. M., Vasilyev V. I., Vulfin A. M., Kartak V. M., Kirillova A. D. Detection of network botnet attacks based on machine learning and knowledge transfer technologies. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2024, no. 5, pp. 41–56 (In Russian). doi:10.31799/1684-8853-2024-5-41-56, EDN: SWCOYY

Financial support

The research was carried out in OmSU within the State assignment of the Ministry of Science and Higher Education of Russian Federation (theme No. FSGF-2023-0004).

References

- Kaur R., Gabrijelčić D., Klobučar T. Artificial intelligence for cybersecurity: Literature review and future research directions. *Information Fusion*, 2023, vol. 97, pp. 101804. doi:10.1016/j.inffus.2023.101804
- Yang Z., Liu X., Li T., Wu D., Wang J., Zhao Y., Han H. A systematic literature review of methods and datasets for anomaly-based network intrusion detection. *Computers & Security*, 2022, vol. 116, pp. 102675. doi:10.1016/j.cose.2022.102675
- Moustafa N. A new distributed architecture for evaluating AI-based security systems at the edge: Network TON IoT datasets. *Sustainable Cities and Society*, 2021, vol. 72, pp. 102994. doi:10.1016/j.scs.2021.102994
- Šzumelda P., Orzechowski N., Rawski M., Janicki A. VHS-22 – a very heterogeneous set of network traffic data for threat detection. *2022 European Interdisciplinary Cybersecurity Conf.*, Barcelona, 2022, pp. 72–78. doi:10.1145/3528580.3532843
- Binsaeed K., Alaa-aldeen H. *BH-KSU23: A Novel Dataset for Evaluating and Enhancing Intrusion Detection Systems Targeting Command-and-Control Traffic*. Mendeley Data, 2023, ver. 1. Available at: <https://data.mendeley.com/datasets/wjxc69xj3n/1> (accessed 3 October 2023).
- Subhadeep Ch. *Trojan Detection [Data set]*, 2021. doi:10.34740/KAGGLE/DSV/2625272. Available at: <https://www.kaggle.com/datasets/subhajournal/trojan-detection> (accessed 3 October 2023).
- Vinayakumar R., Alazab M., Soman K. P., Poornachandran P., Al-Nemrat A., Venkatraman S. Deep learning approach for intelligent intrusion detection system. *IEEE Access*, 2019, vol. 7, pp. 41525–41550. doi:10.1109/ACCESS.2019.2895334
- Jan S. U., Ahmed S., Shakhov V., Koo I. Toward a lightweight intrusion detection system for the Internet of Things. *IEEE Access*, 2019, vol. 7, pp. 42450–42471. doi:10.1109/ACCESS.2019.2907965
- Wu P., Guo H., Buckland R. A transfer learning approach for network intrusion detection. *2019 IEEE 4th Intern. Conf. on Big Data Analytics (ICBDA)*, Suzhou, IEEE, 2019, pp. 281–285. doi:10.1109/ICBDA.2019.8713213
- Belikov V. V. Using deep reinforcement learning for selecting network traffic features in intrusion detection systems. *Programming and Computer Software*, 2022, vol. 48, no. 6, pp. 359–368. doi:10.31857/S0132347422060024, EDN: KWO-QRH
- Ioganson I. Review of federated learning methods. *CEUR Workshop Proceeding*, 2023. Available at: <https://www.researchgate.net/publication/376083291> (Obzor metodov federativnogo obucenia (accessed 14 July 2024)) (In Russian).
- Novikova E. S., Kotenko I. V., Meleshko A. V., Izrailov K. E. Federated learning based intrusion detection: system architecture and experiments. *Cybersecurity Issues*, 2023, no. 6(58), pp. 50–66 (In Russian). doi:10.21681/2311-3456-2023-6-50-66
- Novikova E. S., Fedorchenko E. V., Kotenko I. V., Kholod I. I. Analytical review of intelligent intrusion detection systems based on federated learning: advantages and open challenges. *Informatics and Automation*, 2023, vol. 22, no. 5, pp. 1034–1082 (In Russian). doi:10.15622/ia.22.5.4
- Masum M., Shahriar H. TL-NID: Deep neural network with transfer learning for network intrusion detection. *2020 15th Intern. Conf. for Internet Technology and Secured Transactions (ICITST)*, London, IEEE, 2020, pp. 1–7. doi:10.23919/ICITST51030.2020.9351317
- Singla A., Bertino E., Verma D. Overcoming the lack of labeled data: Training intrusion detection models using transfer learning. *2019 IEEE Intern. Conf. on Smart Computing*,

- Washington, IEEE, 2019, pp. 69–74. doi:10.1109/SMART-COMP.2019.00031
16. Fan Y., Li Y., Zhan M., Cui H., Zhang Y. IoTDefender: A federated transfer learning intrusion detection framework for 5G IoT. *2020 IEEE 14th Intern. Conf. on Big Data Science and Engineering (BigDataSE)*, Guangzhou, IEEE, 2020, pp. 88–95. doi:10.1109/BigDataSE50710.2020.00020
 17. Idrissi I., Azizi M., Moussaoui O. Accelerating the update of a DL-based IDS for IoT using deep transfer learning. *Indonesian Journal of Electrical Engineering and Computer Science*, 2021, vol. 23, no. 2, pp. 1059–1067. doi:10.11591/ijeecs.v23.i2.pp1059-1067
 18. Rodriguez E., Valls P., Otero B., Costa J. J., Verdú J., Pajuelo M. A., Canal R. Transfer-learning-based intrusion detection framework in IoT networks. *Sensors*, 2022, vol. 22, no. 15, pp. 5621. doi:10.3390/s22155621
 19. Gebresilassie S. K., Rafferty J., Chen L., Cui Z., Abu-Tair M. Transfer and CNN-based de-authentication (disassociation) DoS attack detection in IoT Wi-Fi networks. *Electronics*, 2023, vol. 12, no. 17, pp. 3731. doi:10.3390/electronics12173731
 20. Yehezkel A., Elyashiv E., Soffer O. Network anomaly detection using transfer learning based on auto-encoders loss normalization. *Proc. of the 14th ACM Workshop on Artificial Intelligence and Security*, Korea, 2021, pp. 61–71. doi:10.1145/3474369.3486869
-

УВАЖАЕМЫЕ АВТОРЫ!

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

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

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



Выявление сетевых вторжений в промышленных киберфизических системах на основе сверточных нейронных сетей

Е. С. Новикова^а, канд. техн. наук, доцент, orcid.org/0000-0003-2923-4954, novikova@comsec.spb.ru

Е. О. Кузнецова^б, магистрант, orcid.org/0009-0008-2186-8630

С. А. Голубев^а, аспирант, orcid.org/0009-0000-4163-5326

^аСанкт-Петербургский Федеральный исследовательский центр РАН, 14-я линия В. О., 39, Санкт-Петербург, 191178, РФ

^бСанкт-Петербургский государственный электротехнический университет «ЛЭТИ», Профессора Попова ул., 5, Санкт-Петербург, 197022, РФ

Введение: одной из наиболее сложных проблем в области обнаружения вторжений является детектирование новых, ранее неизвестных атак. В последнее время для решения этой задачи активно исследуются и применяются методики на основе глубокого обучения, поскольку они способны эффективно извлекать пространственные и временные закономерности в данных. **Цель:** разработать методику выявления сетевых атак на основе сверточных нейронных сетей для повышения сетевой безопасности промышленных киберфизических систем. **Результаты:** исследованы и систематизированы подходы к выявлению сетевых атак, основанные на представлении сетевых данных в виде двумерной матрицы анализируемых атрибутов, т. е. в виде изображения. Предложена методика выявления сетевых атак на основе сверточной нейронной сети, отличительной особенностью которой является преобразование «сырых» сетевых потоков в двумерную матрицу с последующим формированием дополнительных атрибутов, представленных текстурными признаками Харалика. Разработана архитектура нейронной сети, выполняющей анализ матричного представления сетевого трафика и вектора признаков Харалика. Для демонстрации эффективности разработанной методики выполнена серия экспериментов с использованием набора данных SWaT, описывающего функционирование системы водоочистных сооружений. В ходе экспериментов исследовалось влияние каждого компонента методики на точность обнаружения сетевых атак. Кроме того, выполнен сравнительный анализ ее эффективности с эффективностью методики обнаружения вторжений, использующей алгоритм Random Forest и описательные статистики сетевых потоков в качестве анализируемых атрибутов. Полученные результаты показали, что предложенная методика имеет высокую точность обнаружения сетевых атак, связанных с извлечением (data exfiltration) и (или) подменой передаваемых данных (data manipulation), в частности, точность повысилась на 25 % по сравнению с методикой на основе Random Forest и составила 86,3 % на исследуемом наборе SWaT. **Практическая значимость:** разработанная методика может быть использована для выявления атак, связанных с подменой передаваемых данных и (или) их извлечением.

Ключевые слова — промышленные киберфизические системы, выявление сетевых атак, сетевые потоки, двумерные матрицы, изображения, признаки Харалика, сверточные нейронные сети.

Для цитирования: Новикова Е. С., Кузнецова Е. О., Голубев С. А. Выявление сетевых вторжений в промышленных киберфизических системах на основе сверточных нейронных сетей. *Информационно-управляющие системы*, 2024, № 5, с. 57–67. doi:10.31799/1684-8853-2024-5-57-67, EDN: NQLXNY

For citation: Novikova E. S., Kuznetsova E. O., Golubev S. A. Network intrusion detection based on convolutional neural networks in industrial cyber-physical systems. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2024, no. 5, pp. 57–67 (In Russian). doi:10.31799/1684-8853-2024-5-57-67, EDN: NQLXNY

Введение

В настоящее время киберфизические системы получили широкое применение в различных отраслях народного хозяйства: в транспорте, электроэнергетике, промышленности, медицине и т. д. Вместе с ростом уровня информатизации этих отраслей увеличивается и число различных информационных угроз. Их реализация может вызвать как нарушение основных функций системы, так и серьезные экономические и экологические последствия. Например, в 2019 г. кибератака на норвежскую компанию Norsk Hydro ASA коснулась 35 000 сотрудников

компании и привела к частичному переходу на ручное управление производством на период восстановления после атаки (<https://www.hydro.com/en/global/about-hydro/company-history/2018-present/2019-cyber-attack-on-hydro/>). В 2021 г. результатом кибератаки на водоочистные сооружения г. Олдсмар в США стало повышение уровня гидроксида натрия в воде (<https://pcsoweb.com/21-015-detectives-investigate-computer-software-intrusion-at-oldsmar%E2%80%99s-water-treatment-plant>). Кибератака в феврале 2024-го на немецкую компанию Varta по производству элементов питания привела к остановке пяти заводов, расположенных в разных частях

мира (<https://www.varta-ag.com/en/about-varta/news-press/details/varta-makes-good-progress-in-solving-the-cyberattack>). Таким образом, своевременное обнаружение вторжений в киберфизических системах является важной задачей. В последнее время для ее решения активно применяются модели глубокого машинного обучения, поскольку они отличаются способностью извлекать пространственные и временные закономерности в данных, что делает их привлекательными для разработки методик обнаружения новых, ранее неизвестных атак.

В статье предлагается методика обнаружения вторжений на основе сверточной нейронной сети, отличительной особенностью которой является преобразование «сырых» сетевых потоков в двумерную матрицу с последующим вычислением дополнительных атрибутов — текстурных признаков Харалика. Под сырым сетевым потоком в работе понимается сетевой поток на уровне протоколов TCP/IP, представленный в виде двоичного массива (дампов соответствующих сетевых пакетов). Такое решение позволяет выявлять скрытые зависимости между передаваемыми данными в потоке независимо от используемого сетевого протокола, поскольку не требует применения процедур по конструированию специальных признаков, для вычисления которых необходимы специальные экспертные знания. Примером таких признаков могут служить число передаваемых пакетов в потоке, число различных состояний сетевых соединений в потоке, типы используемых сетевых протоколов и т. д. Оценка эффективности разработанной методики выполнена с использованием набора данных SWaT, описывающего функционирование системы водоочистных сооружений [1].

Анализ релевантных работ

В настоящее время для обнаружения вторжений предложено большое число разнообразных подходов, в основе которых лежат методы статистического анализа данных, фрактального анализа временных рядов [2], машинного и глубокого обучения [2–6]. Например, в [4] представлена модель обнаружения вторжений в потоке данных от датчиков, которая представляет собой комбинацию сверточной нейронной сети и модулей долгой краткосрочной памяти. Х. Ван и В. Ли [5] разработали гибридную нейронную сеть DDosTC, объединяющую механизмы самовнимания и сверточные слои для обнаружения DDos-атак в программно определяемых сетях. Анрезини и др. [6] представили многоступенчатую систему обнаружения сетевых вторжений, в которой на первом этапе применяется

сверточная нейронная сеть, обученная в режиме с учителем, а на втором — два автокодировщика, обученные в режиме без учителя. Два автокодировщика используются для реконструкции нормальных и аномальных потоков, а также для формирования расширенной обучающей выборки. В [7] рассмотрен подход GraphDDoS для обнаружения низкоскоростных и высокоскоростных DDos-атак, в его основе лежит графовая нейронная сеть, анализирующая сетевые подключения и связи между ними.

Чаще всего в таких подходах в качестве входных данных используется одномерный вектор, элементы которого представлены статистическими параметрами сетевого потока. Однако в последнее время активно развиваются методы обнаружения вторжений, в которых одномерный входной вектор данных преобразуется в двумерную матрицу, что позволяет применять двумерные свертки, которые эффективнее извлекают пространственные зависимости между атрибутами. В [8] представлена таксономия методов обнаружения вторжений, использующих преобразования сетевых данных в двумерную матрицу. В научной литературе такие подходы часто обозначаются как методы на основе анализа изображений (image-based methods), поскольку получаемая матрица может быть интерпретирована как цветное изображение или изображение в серых оттенках.

В большинстве работ [9–14] двумерная матрица строится на основе одномерного вектора статистик, вычисленных для сетевого потока. Например, в [12] векторы с сетевыми статистиками объединяются в группы по 24 объекта, для которых строится цветное изображение, полученная матрица подается на вход нейронной сети-трансформеру ViT, архитектура которой была предложена специально для анализа изображений [13]. Очевидно, что ключевым моментом в данном подходе является предположение, что сгруппированные сетевые потоки, формирующие изображение, упорядочены во времени. Другим существенным недостатком является высокая требовательность модели к вычислительным ресурсам.

В [14] предложен оригинальный способ построения цветного изображения, он состоит из трех последовательных этапов. Сначала выбираются три различных нелинейных метода снижения размерности, выполняющих проекцию многомерных данных в двумерное пространство, например t-SNE, метод главных компонент с косинусной ядерной функцией и UMAP. Далее, для каждой полученной проекции исходной обучающей выборки определяется минимальный ограничивающий прямоугольник, который затем вращается для получения финаль-

ных координат проекций каждой точки данных. Полученные проекции приводятся к единому размеру 120×120 и объединяются в цветное изображение, в котором каждый цветовой канал — красный, зеленый и синий — задается одной из трех проекций. Полученное изображение анализируется нейронной сетью, состоящей из двух четырехслойных сверточных сетей, объединяемых через слой конкатенации, классификация изображений осуществляется с помощью полносвязной нейронной сети.

Другой подход к преобразованию сетевых данных в изображение представлен методами, в которых двумерная матрица строится непосредственно на основе исходных сетевых данных, причем изображения могут быть построены как для пакетов [15–17], так и для потоков [18, 19]. Например, в [17] цветное изображение строится для последовательности из k пакетов. Для каждого пакета вычисляется множество вспомогательных признаков, таких как IP-адреса, порты получателя и отправителя, тип сетевого протокола, и формируется вектор, состоящий из вспомогательных признаков и первых 1458 байт пакета. Данный вектор определяет один ряд пикселей в изображении. Цвет пикселя задается направлением пакета: значения байт входящих пакетов кодируются зеленым цветом, а исходящих — красным. Значение параметра k определяется опытным путем, например, для тестового набора данных CIC-IDS2017 оптимальное количество пакетов составило девять. В качестве модели выявления сетевых аномалий используется четырехслойная сверточная нейронная сеть. Достоинством данного подхода является возможность выявлять аномальную сетевую активность в режиме, близком к реальному времени.

Показано [8], что преобразование сетевого трафика на уровне пакетов является ресурсоемкой задачей, в том числе и на этапе формирования обучающей выборки, и предложено генерировать изображения на уровне сетевых потоков. Также были исследованы различные способы компоновки пикселей в изображении, в частности с помощью кривых, заполняющих пространство. Выполненные авторами эксперименты показали, что точность сверточной нейронной сети, используемой в качестве детектора сетевых аномалий, практически не зависит от способа формирования изображения, использование прямого последовательного преобразования байт в пиксель вычислительно эффективнее, что делает его применение на практике более предпочтительным.

Следует также отметить, что предложены гибридные подходы, объединяющие представления сетевых данных на уровне пакетов и потоков [19]. В их основе лежит предположение,

что анализ данных на уровне потоков позволяет выявить пространственные закономерности в сетевом трафике, а анализ данных на уровне пакетов — временные закономерности.

Таким образом, в настоящее время предложены различные методики преобразования сетевого трафика в двумерные матрицы. Их отличительной особенностью является возможность анализировать данные на уровне содержания сетевых пакетов и (или) потоков. Однако методики, выполняющие генерацию изображений на основе пакетов, вычислительно ресурсоемки в силу большого объема данных, который необходимо обработать, что делает их применение на практике нецелесообразным. Эффективность методик, в основе которых лежит преобразование сетевых потоков в изображения, исследована на наборах данных CIC-IDS2017 [12, 14, 17–19], UNSW-NB2015 [10–12, 18], в которых в основном представлены такие атаки, как отказ в обслуживании (DoS- и DDoS-атаки), сканирование портов, заражение бот-сетью. Между тем в [20] показано, что обнаружение подобных атак эффективно выполняется путем анализа статистических характеристик сетевых потоков. Следовательно, необходимо выполнить оценку эффективности методики, в основе которой лежит преобразование сетевых потоков в изображения, в задаче выявления атак другого типа, в частности атак вида извлечение данных (data exfiltration) и подмена передаваемых данных (data manipulation).

Стоит добавить, что в настоящей работе впервые предложено использовать дополнительные признаки, вычисляемые на основе сгенерированных изображений, для повышения эффективности выявления вредоносной активности в сети.

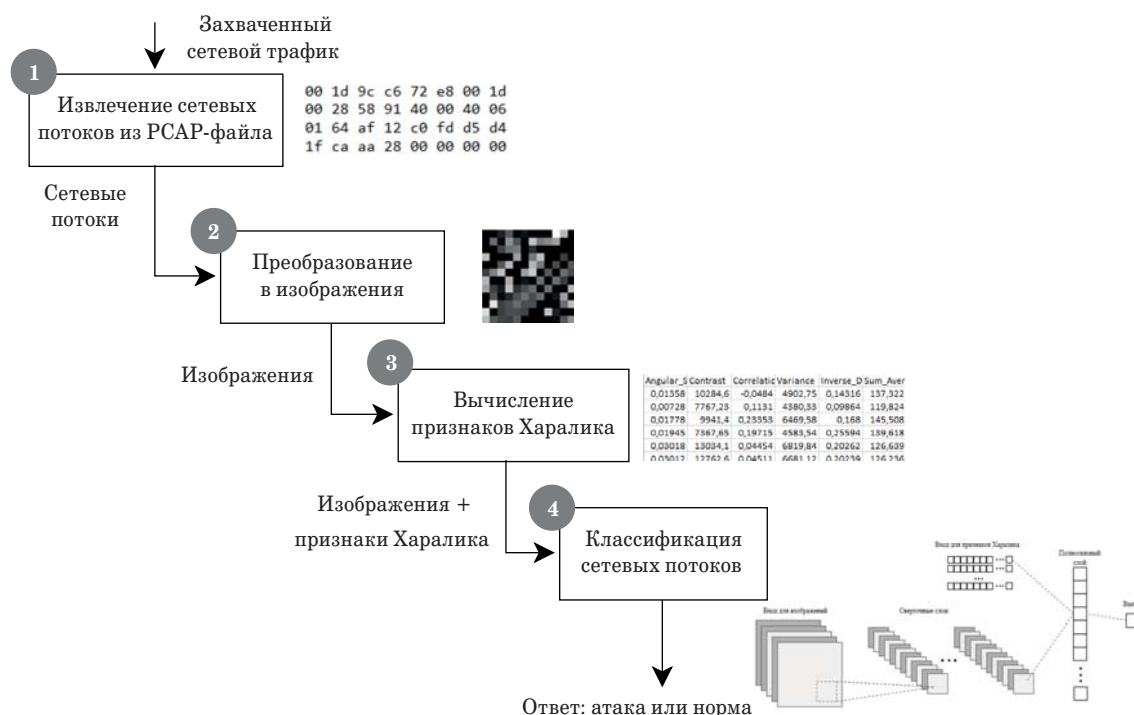
Методика обнаружения сетевых атак на основе сверточной нейронной сети

В основе разработанной методики лежат следующие предположения:

1) двумерная свертка лучше извлекает пространственные зависимости между атрибутами, чем одномерная [19];

2) непосредственный анализ содержимого сетевого потока позволяет выявлять атаки, которые не влияют на статистические характеристики потока, такие как число пакетов, средняя длина пакетов и т. д.;

3) преобразование данных в двумерную матрицу позволяет интерпретировать получаемую матрицу как изображение в оттенках серого, что дает возможность формировать новые признаки для выявления сетевых атак, которые вычисляются на основе анализа изображения.



■ **Рис. 1.** Схема методики к выявлению сетевых вторжений
 ■ **Fig. 1.** Scheme of the approach to network intrusion detection

Рассмотрим подробно каждый шаг представленной на рис. 1 схемы методики.

Извлечение сетевых потоков из дампа сетевого трафика

В работе используется следующее определение сетевого потока. Сетевой поток – это однонаправленная последовательность пакетов, объединенных семью общими свойствами: входным интерфейсом, IP-адресом источника, IP-адресом назначения, номером протокола IP, портом источника, портом назначения, типом IP-сервиса. При извлечении потоков также выполняется анонимизация пакетов, которая включает в себя удаление адресов. В [21] было доказано, что любые данные об адресах, такие как MAC-адрес и (или) IP-адрес источника и назначения, могут оказывать существенное влияние на эффективность обнаружения вторжений, поскольку эти признаки могут быть использованы моделью для определения класса атаки. В предлагаемом подходе анонимизация выполняется путем замены MAC- и IP-адресов на нули.

Преобразование сетевого потока в изображение

Перед определением функции преобразования сетевого потока в изображение необходимо задать три основных параметра: 1) размер изображения; 2) цветовой режим (оттенки серого

или RGB); 3) способ компоновки атрибутов (пикселей) в изображении.

Анализ соответствующих исследований показал, что не существует единого подхода к определению размера изображения, хотя этот параметр оказывает решающее влияние на точность модели. Часто для визуализации данных выбирается фиксированное число байт сетевого пакета, а размер изображения определяется с учетом формата входных данных для нейронной сети [16, 21, 22]. В настоящей работе предлагается для определения размера изображения $n \times n$ использовать формулу $n = \text{ceil}(\sqrt{P_{stat}})$, где ceil – функция округления до ближайшего целого, а P_{stat} – статистический показатель, определяемый на основе статистического анализа распределения длин потоков в обучающей выборке. В качестве P_{stat} может быть выбран средний размер сетевого потока, максимальный или минимальный размер потока, медиана и т. д. В настоящей работе параметр P_{stat} задается средним значением размера потока.

В разработанной методике изображения строятся в оттенках серого. Такие изображения имеют только один канал, который передает информацию об интенсивности света и принимает значения в диапазоне от 0 до 255. В этом случае алгоритм преобразования сетевых пакетов является довольно простым: каждый сетевой поток представляется в виде двоичной последователь-

ности, которая делится на байты. Затем каждый байт преобразуется в уровень серого в соответствии со следующим правилом (рис. 2):

$$0 \times 00 \rightarrow 0 \text{ (черный)}, \dots, 0 \times FF \rightarrow 255 \text{ (белый)}.$$

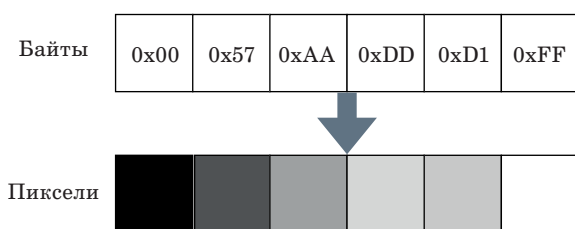
В этом случае не требуется извлекать такие атрибуты, как тип протокола, тип сервиса и др.

Существует несколько подходов к заполнению пространства изображения [8]. В нашей работе используется линейное заполнение, при котором каждый байт сетевого потока последовательно преобразуется в пиксель, начиная с верхнего левого угла изображения и заканчивая нижним правым углом. Когда строка заполняется, обрабатывается следующая строка. Это действие повторяется до тех пор, пока изображение не будет завершено. Если пакеты имеют меньшую длину, чем число элементов матрицы, то оставшиеся пиксели заполняются 0×00 .

Вычисление признаков Харалика

В качестве дополнительных анализируемых параметров используются признаки Харалика [23]. Они применяются для описания текстуры изображения и позволяют количественно оценить и описать визуальные и тактильные свойства поверхностей.

Признаки Харалика вычисляются с помощью специальной матрицы совместной встречаемости на уровне серого (Gray-Level Co-Occurrence Matrix, GLCM), которая показывает, как часто пиксель со значением интенсивности (уровня серого) i встречается в определенном пространственном соотношении с пикселем со значением j . Чтобы составить матрицу GLCM, в соответствующие элементы записывается число раз, когда пиксели определенной интенсивности находились рядом друг с другом. На рис. 3 представлена схема расчета элементов матрицы GLCM: пиксель интенсивности 3 находится справа от пикселя 2 один раз, поэтому элемент матрицы [3, 2] получает значение 1, а комбинация пикселей 3–2 встречается дважды, поэтому элемент матрицы [2, 3] выставляется в значение 2.



■ **Рис. 2.** Кодирование байт в оттенки серого
 ■ **Fig. 2.** Byte encoding in grayscale color

Очевидно, что оценить комбинации соседних пикселей можно не только слева направо, но и в других направлениях: справа налево, сверху вниз и по двум диагоналям (рис. 4). Таким образом, получается четыре матрицы GLCM, которые можно использовать для расчета признаков Харалика.

Признаки Харалика являются вторичными текстурными признаками, их перечень приведен в табл. 1.

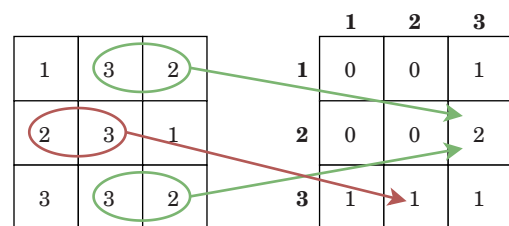
Классификация сетевых вторжений

Задача выявления сетевых вторжений нами рассматривается как задача бинарной классификации, в которой вторжения представлены потоками, соответствующими различным типам сетевых атак и обозначенными одной меткой – 1 (атака). Для ее решения была использована модель сверточной нейронной сети, которая является наиболее распространенной и эффективной архитектурой для анализа изображений.

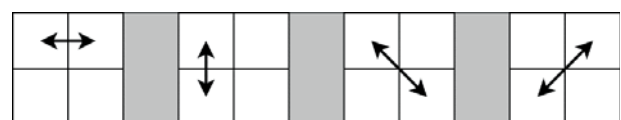
Поскольку для классификации сетевых потоков предлагается анализировать пиксели изображения и признаки Харалика, было реализовано два входных слоя для модели. Изображение обрабатывается тремя сверточными слоями, после каждого из которых следует подвыборочный слой с функцией max pool. Выход третьего подвыборочного слоя объединяется с входом для признаков Харалика, после чего подается для дальнейшей обработки на полносвязный слой (рис. 5).

Особенности программной реализации подхода

Исходные сетевые данные в формате PCAP-файла разбиваются на сетевые потоки с помощью специального программного инструмента



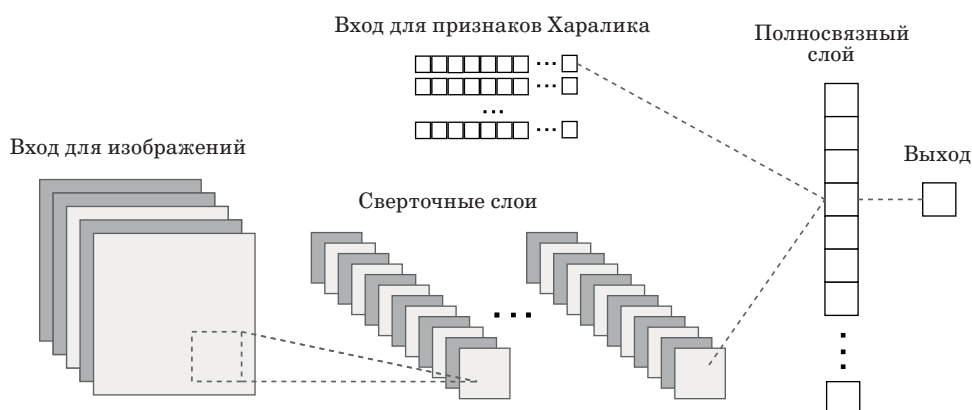
■ **Рис. 3.** Вычисление матрицы GLCM
 ■ **Fig. 3.** Calculation of GLCM matrix



■ **Рис. 4.** Способы вычисления матрицы GLCM
 ■ **Fig. 4.** Ways to calculate GLCM matrix

- **Таблица 1.** Описание текстурных признаков Харалика [23]
- **Table 1.** Description of the Haralick texture features [23]

Признак	Описание
Второй угловой момент (Angular Second Moment)	Измеряет локальную однородность уровней серого. Если пиксели очень похожи, значение будет большим
Контраст (Contrast)	Изменение интенсивности или уровня серого между опорным пикселем и соседним с ним. Большой контраст отражает большие различия в интенсивности в GLCM
Корреляция (Correlation)	Показывает линейную зависимость значений уровня серого в матрице GLCM
Однородность (Variance)	Измеряет неоднородность текстуры
Обратный момент разностей (Inverse Difference Moment)	Измеряет однородность текстуры
Сумма средних (Sum Average)	Измеряет сумму средних значений всех пикселей
Суммарная неоднородность (Sum Variance)	Признак неоднородности, который сильно коррелирует со статистической переменной первого порядка, такой как стандартное отклонение. Дисперсия увеличивается, когда значения уровня серого отличаются от их среднего значения
Суммарная энтропия (Sum Entropy)	Характеризует неоднородность изображения или сложность текстуры
Энтропия (Entropy)	Измеряет случайность интенсивности пикселей
Неоднородность разностей (Difference Variance)	Показывает неоднородность изображения
Энтропия разностей (Difference Entropy)	Отражает уровень случайности, отсутствие структуры или порядка в контрастности изображения
Информационные показатели корреляции 1, 2	Измеряют корреляцию параметров матрицы с использованием дополнительных методов



- **Рис. 5.** Архитектура разработанной нейронной сети
- **Fig. 5.** Architecture of the proposed neural network

NetFlow2Image (<https://github.com/EveNovikova/FedIDSExplorer>), разработанного на языке программирования Python с использованием библиотеки Scapy (<https://scapy.net/>). Инструмент позволяет сразу разметить дампы на аномальные сетевые потоки и норму, используя файл с разметкой в формате JSON, в котором указываются IP-адреса атакующих и атакуемых хостов и на-

чальное и конечное время атаки. Полученные сетевые потоки дальше преобразуются в черно-белые изображения с помощью библиотек NumPy и Pillow (<https://pypi.org/project/pillow/>). Пользователь имеет возможность задать различные настройки формирования изображения. Полученные изображения сетевых потоков сохраняются в формате PNG, при этом формиру-

ется иерархия директорий (аномальные сетевые потоки или норма) для дальнейшего проведения обучения нейронных сетей. Данная утилита также использует библиотеку CICFlowMeter (<https://github.com/ahlashkari/CICFlowMeter>), которая вычисляет различные статистики для сетевого потока, наиболее часто применяемые для выявления аномальной сетевой активности.

Расчет признаков Харалика осуществляется с помощью библиотеки Mahotas (<https://mahotas.readthedocs.io/en/latest/>), а для разработки нейронной сети была использована программная библиотека TensorFlow (<https://www.tensorflow.org/?hl=ru>).

Экспериментальная оценка

Целью эксперимента являлось определение эффективности предложенного подхода к обнаружению сетевых атак, для чего был разработан следующий сценарий.

На первом этапе выполнялась оценка влияния направления расчета признаков Харалика на точность обнаружения сетевых вторжений.

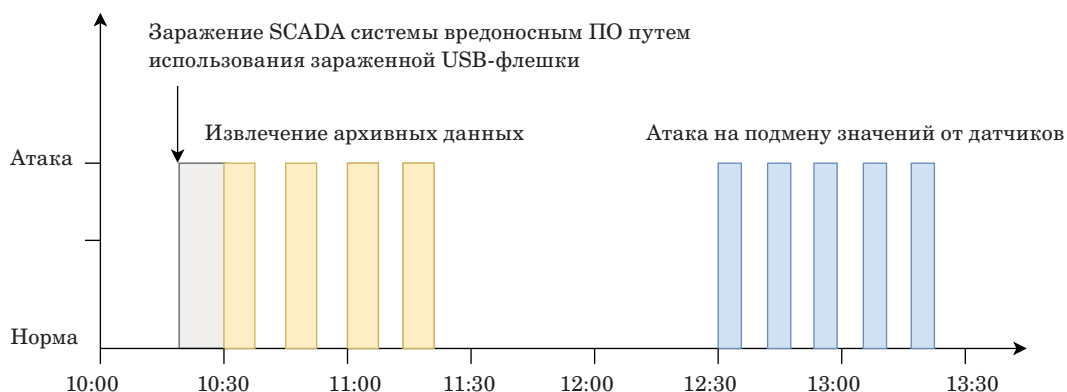
На втором этапе оценивалось влияние использования признаков Харалика в качестве дополнительных атрибутов. Для этого была проведена серия экспериментов, в которой сравнивалась точность обнаружения вторжений с использованием только сверточной нейронной сети и с использованием сверточной сети и признаков Харалика. Кроме того, для оценки целесообразности применения преобразования сетевого трафика в изображения было выполнено сравнение с моделью обнаружения вторжений на основе алгоритма Random Forest, обученного на классических признаках, описывающих статистические параметры сетевых потоков. Выбор этой модели объясняется тем, что она наиболее часто используется в задачах обнаружения сетевых атак.

Описание набора данных и типов атак

В качестве анализируемого набора данных был выбран набор SWaT версии 2019 г. [1], который получен с использованием водоочистительного полигона, созданного Центром исследований кибербезопасности iTrust Сингапурского университета технологий и дизайна в 2019 г. Этот полигон является уменьшенной копией водоочистных сооружений и моделирует современный процесс очистки воды. В состав полигона SWaT входят водоочистительное оборудование, многочисленные ПЛК, SCADA, автоматизированное рабочее место оператора и хранилище данных от технологического процесса.

Версия набора данных SWaT.A6_Dec 2019 состоит из PCAP-файлов с дампом сетевого трафика, файла в формате csv с показаниями датчиков и файла с описанием выполненных атак. PCAP-файлы содержат сетевой трафик, характеризующий как нормальный, так и аномальный режим работы полигона. Вредоносные сценарии представлены двумя типами атак: извлечением исторических данных (exfiltrate historian data) и подменой показателей сенсоров (disrupt sensor readings and process).

Атака на извлечение данных представляет собой преднамеренную несанкционированную передачу данных технологического процесса. Атака на подмену показателей датчиков относится к категории атак, целью которых является искажение или подмена данных, получаемых от оборудования. Последнее может быть выполнено путем физического вмешательства в работу датчика или через программное обеспечение, зараженное вредоносным кодом. Подобные атаки могут привести к неправильной работе систем управления и мониторинга, что особенно критично в промышленных и инфраструктурных объектах, где точность данных от физических устройств имеет первостепенное значение. Временная диаграмма атакующих воздействий,



■ **Рис. 6.** Схема атакующих сценариев в наборе данных SWaT.A6_Dec 2019

■ **Fig. 6.** Schema of attack scenarios in the SWaT.A6_Dec 2019 dataset

представленных в наборе SWaT.A6_Dec 2019, приведена на рис. 6 [1].

Параметры экспериментов

При обучении сверточной нейронной сети были использованы следующие настройки:

- 1) оптимизатор: Adam;
- 2) функция потерь: binary_crossentropy;
- 3) количество эпох обучения: 10.

Для обучения модели Random Forest настройка параметров осуществлялась с помощью функции GridSearchCV. В результате были использованы следующие параметры: n_estimators = 5 (число деревьев), max_depth = 9 (максимальная глубина деревьев), max_leaf_node = 9 (максимальное число листьев в дереве).

При выполнении эксперимента данные были анонимизированы: из анализа были исключены данные о IP-адресе сетевого потока. Для обеспечения баланса классов был использован механизм оверсэмплинга. Во всех экспериментах разделение на тренировочную и тестовую выборки производилось случайным образом в соотношении 80:20.

Оценка эффективности обнаружения вторжений осуществлялась с помощью метрик точность (precision), полнота (recall) и F1-мера.

Точность отражает долю объектов, которые действительно принадлежат данному классу относительно всех объектов, которым модель сопоставила этот класс. В рамках задачи обнаружения атак высокая точность означает, что система способна в большинстве случаев корректно детектировать атаки с сохранением низкого уровня ложных срабатываний. Полнота — доля выявленных моделью объектов, принадлежащих классу, относительно всех объектов этого класса. Высокое значение метрики полноты означает, что система способна в большинстве случаев корректно детектировать атаки с сохранением низкого уровня пропуска атак. F1-мера представляет собой среднее гармоническое между вышеуказанными метриками и выражается формулой

$$F1 = 2 \cdot (\text{Precision} \cdot \text{Recall}) / (\text{Precision} + \text{Recall}).$$

Анализ полученных результатов

Результаты точности выявления атак для разных способов формирования признаков Харалика представлены в табл. 2.

Лучшие результаты по всем метрикам были получены для угла поворота 45°. При данном способе извлечения признаков Харалика полнота обнаружения вторжений составила 98 %, а точность — 77 %.

■ **Табл. 2.** Результаты точности выявления атак при разных углах расчета признаков Харалика

■ **Table 2.** Experimental results for different angles of Haralick features

Угол расчета признаков Харалика	Precision	Recall	F1-мера
0	0,67	0,93	0,78
45	0,77	0,98	0,86
90	0,72	0,92	0,81
135	0,74	0,91	0,82
Усредненный	0,68	0,78	0,73

■ **Таблица 3.** Результаты экспериментов для различных моделей обнаружения вторжений

■ **Table 3.** Experimental results for different intrusion detection models

Тип модели обнаружения вторжений	Precision	Recall	F1-мера
Сверточная нейронная сеть (только изображения)	0,69	0,93	0,79
Сверточная нейронная сеть (только изображения + признаки Харалика)	0,77	0,98	0,86
Random Forest (статистики)	0,65	0,77	0,61

Результаты второго этапа эксперимента представлены в табл. 3.

Из нее следует, что сверточная нейронная сеть, обученная на сырых данных сетевых потоков, дает лучшие результаты обнаружения таких атак, как извлечение данных и модификация данных сенсора. Можно предположить, что это связано с тем, что такие атаки затрагивают в первую очередь содержание передаваемых данных и в меньшей степени влияют на статистические параметры сетевых потоков, как это происходит в случае атак типа сканирование портов и отказ в обслуживании.

Исследование влияния признаков Харалика на точность обнаружения сетевых атак показало, что их использование дает повышение точности решения задачи на 7 % по сравнению со сверточной нейронной сетью, обученной только на изображениях, и почти на 25 % по сравнению с моделью Random Forest. Следует отметить, что их применение позволяет в большей мере повысить точность, т. е. снизить число ложноположительных срабатываний, что важно при практическом применении разработанных моделей.

Однако вычисление признаков Харалика является достаточно ресурсоемкой задачей, вычислительная сложность которой прямо пропорциональна размеру анализируемого изображения. Эта проблема дает дальнейшее направление исследованиям, связанным как с определением временных показателей рассмотренного подхода, так и с исследованием других текстурных признаков, обладающих более высокой вычислительной эффективностью.

Заключение

В последнее время для обнаружения атак и аномалий в киберфизических системах было предложено большое число методов, в основе которых лежат модели машинного обучения, включая глубокие нейронные сети. В настоящей статье предложена методика выявления сетевых атак, отличающаяся способом преобразования сетевых потоков в двумерную матрицу с последующим формированием признаков Харалика. Детально представлены ее основные шаги: генерация изображения в оттенках серого на основе сетевого потока, вычисление признаков Харалика, классификация объекта с помощью сверточной нейронной сети.

Хотя некоторые исследователи считают, что достаточно использовать статистические параметры сетевых потоков [20, 22], а в преобразовании исходных сетевых данных в изображения нет необходимости, проведенные эксперименты с данными от тестового полигона водоочистных сооружений показали, что этот способ подготовки входных данных позволяет эффективно обнаруживать такие атаки, как подмена переда-

ваемых данных и (или) их извлечение даже при использовании простых сверточных нейронных сетей. Применение дополнительных признаков, которые оценивают текстуру формируемых изображений, позволяет снизить число ложноположительных срабатываний и тем самым повысить точность обнаружения сетевых атак, в частности на промышленные киберфизические системы. Кроме того, благодаря тому, что исследуемые входные данные формируются путем преобразования бинарного вектора в числовую матрицу, предлагаемая методика может считаться независимой от используемого сетевого протокола и применяться для анализа сетевого трафика, передаваемого по любому сетевому протоколу, основанному на TCP/IP, например по промышленному протоколу Modbus TCP.

Дальнейшее направление исследований связано с оценкой вычислительной эффективности разработанного подхода, апробацией на других наборах данных, сформированных для других систем и других сетевых протоколов, и анализом других текстурных признаков в качестве дополнительных анализируемых атрибутов. Также в задачи будущих исследований включен поиск и анализ других архитектур нейронных сетей, в частности одномерных сверточных сетей.

Финансовая поддержка

Работа выполнена при поддержке гранта Российского научного фонда № 23-11-20024 (<https://rscf.ru/project/23-11-20024/>) и Санкт-Петербургского научного фонда в СПб ФИЦ РАН.

Литература

1. Goh J., Adepu S., Junejo K. N., Mathur A. A dataset to support research in the design of secure water treatment systems. *Critical Information Infrastructures Security*, 2017, vol. 10242, pp. 88–99. doi:10.1007/978-3-319-71368-7_8
2. Kotenko I., Saenko I., Kribel A., Lauta A. A technique for early detection of cyberattacks using the traffic self-similarity property and a statistical approach. *Proc. of 29th Euromicro Intern. Conf. on Parallel, Distributed and Network-Based Processing, PDP 2021*, Virtual, Valladolid, 10–12 March 2021, pp. 281–284. doi:10.1109/PDP52278.2021.00052
3. Branitskiy A., Kotenko I., Saenko I. Applying machine learning and parallel data processing for attack detection in IoT. *IEEE Transactions on Emerging Topics in Computing*, 2021, vol. 9, pp. 1642–1653. doi:10.1109/TETC.2020.3006351
4. Gaber T., Awotunde J., Torky M., Ajagbe S., Hamoudeh M., Li W. Metaverse-IDS: Deep learning-based intrusion detection system for Metaverse-IoT networks. *Internet of Things*, 2023, vol. 24, no. 100977. doi:10.1016/j.iot.2023.100977
5. Wang H., Li W. DDosTC: A transformer-based network attack detection hybrid mechanism in SDN. *Sensors*, 2021, vol. 21, iss. 15, pp. 5047. doi:10.3390/s21155047
6. Andresini G., Appice A., Mauro N. D., Loglisci C., Malerba D. Multi-channel deep feature learning for intrusion detection. *IEEE Access*, 2020, vol. 8, pp. 53346–53359. doi:10.1109/ACCESS.2020.2980937
7. Li Y., Li R., Zhou Z., Guo J., Yang W., Du M., Liu Q. GraphDDoS: Effective DDoS attack detection using graph neural networks. *2022 IEEE 25th Intern. Conf. on Computer Supported Cooperative Work in Design, CSCWD*, 2022, Hangzhou, China, pp. 1275–1280. doi:10.1109/CSCWD54268.2022.9776097

8. Golubev S., Novikova E. Transformation of network flow data into images for intrusion detection using convolutional neural networks. *2023 Intern. Russian Automation Conf. (RusAutoCon)*, Sochi, Russian Federation, 2023, pp. 948–952. doi:10.1109/RusAutoCon58002.2023.10272890
9. Wang Q., Zhao W., Ren J. Intrusion detection algorithm based on image enhanced convolutional neural network. *J. Intell. Fuzzy Syst.*, 2021, vol. 41, no. 1, pp. 2183–2194. doi:10.3233/JIFS-210863
10. Masum M., Shahriar H., Haddad H. M. A transfer learning with deep neural network approach for network intrusion detection. *International Journal of Intelligent Computing Research (IJICR)*, 2021, vol. 12, pp. 087–1095. doi:10.20533/ijicr.2024.4655.2021.0132
11. Noever D. A., Noever S. E. M. Image classifiers for network intrusions. *CoRR*, 2021, arXiv:2103.07765. <https://arxiv.org/abs/2103.07765> (дата обращения: 05.04.2024).
12. Ho C. M. K., Yow K.-C., Zhu Z., Aravamuthan S. Network intrusion detection via flow-to-image conversion and vision transformer classification. *IEEE Access*, 2022, vol. 10, pp. 97780–97793. doi:10.1109/ACCESS.2022.3200034
13. Dosovitskiy A., Beyer L., Kolesnikov A., Weissenborn D., Zhai X., Unterthiner T., Dehghani M., Minderer M., Heigold G., Gelly S., Uszkoreit J., Houlsby N. An image is worth 16x16 words: Transformers for image recognition at scale. 2020, arXiv:2010.11929. <https://doi.org/10.48550/arXiv.2010.11929> (дата обращения: 05.04.2024).
14. Kim T., Pak W. Deep learning-based network intrusion detection using multiple image transformers. *Appl. Sci.*, 2023, vol. 13, no. 2754. doi:10.3390/app13052754
15. Hosler R., Sundar A., Zou X., Li F., Gao T. Unsupervised deep learning for an image based network intrusion detection system. *Proc. of 2023 IEEE Global Communications Conf.*, Kuala Lumpur, Malaysia, 2023, pp. 6825–6831. doi: 10.1109/GLOBECOM54140.2023.10437636
16. Golubev S., Novikova E., Fedorchenko E. Image-based approach to intrusion detection in cyber-physical objects. *Information*, 2022, vol. 13, iss. 12, pp. 553. doi:10.3390/info13120553
17. Ghadermazi J., Shah A., Bastian N. Towards real-time network intrusion detection with image-based sequential packets representation. *IEEE Transactions on Big Data*, 2024, no. 01, pp. 1–17. doi:10.1109/TBDATA.2024.3403394
18. Zhang X., Chen J., Zhou Y., Han L., Lin J. A multiple-layer representation learning model for network-based attack detection. *IEEE Access*, 2019, vol. 7, pp. 91992–92008. doi:10.1109/ACCESS.2019.2927465
19. Yu L., Dong J., Chen L., Li M., Xu B., Li Z., Qiao L., Liu L., Zhao B., Zhang B. Pbcnn: Packet bytes-based convolutional neural network for network intrusion detection. *Computer Networks*, 2021, vol. 194, no. 108117. doi:10.1016/j.comnet.2021.108117
20. Sharafaldin I., Lashkari A. H., Ghorbani A. A. Toward generating a new intrusion detection dataset and intrusion traffic characterization. *Proc of Intern. Conf. on Information Systems Security and Privacy*, Funchal, Madeira – Portugal, January 22–24, 2018, pp. 108–116.
21. Sun Y., Esaki H., Ochiai H. Adaptive intrusion detection in the networking of large-scale LANs with segmented federated learning. *IEEE Open Journal of the Communications Society*, 2021, vol. 2, pp. 102–112. doi:10.1109/OJCOMS.2020.3044323
22. Wu P., Guo H., Buckland R. A transfer learning approach for network intrusion detection. *Proc. of 2019 IEEE 4th Intern. Conf. on Big Data Analytics (ICBDA)*, Suzhou, China, 2019, pp. 281–285. doi:10.1109/ICBDA.2019.8713213
23. Haralick R., Shanmugam K., Dinstein I. Textural features for image classification. *IEEE TSMC*, 1973, vol. 3, iss. 6, pp. 610–621.

UDC 004.056

doi:10.31799/1684-8853-2024-5-57-67

EDN: NQLXNY

Network intrusion detection based on convolutional neural networks in industrial cyber-physical systemsE. S. Novikova^a, PhD, Tech., Associate Professor, orcid.org/0000-0003-2923-4954, novikova@comsec.spb.ruE. O. Kuznetsova^b, Master Student, orcid.org/0009-0008-2186-8630S. A. Golubev^a, Post-Graduate Student, orcid.org/0009-0000-4163-5326^aSt. Petersburg Federal Research Center of the RAS, 39, 14th Line, 199178, Saint-Petersburg, Russian Federation^bSaint-Petersburg Electrotechnical University «LETI», 5, Prof. Popov St., 197376, Saint-Petersburg, Russian Federation

Introduction: One of the most challenging problems in intrusion detection is the detection of new, previously unknown attacks. Recently, deep learning techniques have been extensively researched and applied to this problem because of their ability to efficiently extract spatial and temporal patterns in data. **Purpose:** To develop a methodology for detecting network attacks based on convolutional neural networks to improve network security of industrial cyber-physical systems. **Results:** We investigate and systematize approaches to detecting network attacks based on the representation of network data in the form of a two-dimensional matrix of analyzed attributes, i.e. in the form of an image. We propose a new approach to the detection of network attacks based on convolutional neural network, the distinctive feature of this is the transformation of “raw” network flows into a two-dimensional matrix with the subsequent formation of

additional attributes represented by Haralick texture features. We develop the architecture of a neural network that analyzes the matrix representation of network traffic and Haralick feature vector. To demonstrate the effectiveness of the developed approach, we perform a series of experiments using the SWaT dataset describing the operation of a water treatment plant system. During the experiments, we have investigated the impact of each component of the approach on the detection accuracy of network attacks. In addition, we perform a comparative performance analysis with an intrusion detection method using Random Forest algorithm and descriptive statistics of network flows as analyzed attributes. The results show that the proposed technique has a high accuracy in detecting network attacks related to data exfiltration and/or data manipulation, in particular, it has improved by 25% as compared to the Random Forest-based method and equals 86.3% on the SWaT set. **Practical relevance:** The developed methodology can be used to detect attacks related to spoofing of transmitted data and/or their extraction.

Keywords – industrial cyber-physical systems, network attack detection, network flows, two-dimensional matrices, images, Haralick features, convolutional neural networks.

For citation: Novikova E. S., Kuznetsova E. O., Golubev S. A. Network intrusion detection based on convolutional neural networks in industrial cyber-physical systems. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2024, no. 5, pp. 57–67 (In Russian). doi:10.31799/1684-8853-2024-5-57-67, EDN: NQLXNY

Financial support

The research is supported by the grant of Russian Science Foundation No. 23-11-20024 (<https://rscf.ru/project/23-11-20024/>), and St. Petersburg Science Foundation.

References

- Goh J., Adepu S., Junejo K. N., Mathur A. A dataset to support research in the design of secure water treatment systems. *Critical Information Infrastructures Security*, 2017, vol. 10242, pp. 88–99. doi:10.1007/978-3-319-71368-7_8
- Kotenko I., Saenko I., Kribel A., Lauta A. A technique for early detection of cyberattacks using the traffic self-similarity property and a statistical approach. *Proc. of 29th Euro-micro Intern. Conf. on Parallel, Distributed and Network-Based Processing, PDP 2021, Virtual, Valladolid*, 10–12 March 2021, pp. 281–284. doi:10.1109/PDP52278.2021.00052
- Branitskiy A., Kotenko I., Saenko I. Applying machine learning and parallel data processing for attack detection in IoT. *IEEE Transactions on Emerging Topics in Computing*, 2021, vol. 9, pp. 1642–1653. doi:10.1109/TETC.2020.3006351
- Gaber T., Awotunde J., Torky M., Ajagbe S., Hammoudeh M., Li W. Metaverse-IDS: Deep learning-based intrusion detection system for Metaverse-IoT networks. *Internet of Things*, 2023, vol. 24, no. 100977. doi:10.1016/j.iot.2023.100977
- Wang H., Li W. DDoS: A transformer-based network attack detection hybrid mechanism in SDN. *Sensors*, 2021, vol. 21, iss. 15, pp. 5047. doi:10.3390/s21155047
- Andresini G., Appice A., Mauro N. D., Loglisci C., Malerba D. Multi-channel deep feature learning for intrusion detection. *IEEE Access*, 2020, vol. 8, pp. 53346–53359. doi:10.1109/ACCESS.2020.2980937
- Li Y., Li R., Zhou Z., Guo J., Yang W., Du M., Liu Q. Graph-DDoS: Effective DDoS attack detection using graph neural networks. *2022 IEEE 25th Intern. Conf. on Computer Supported Cooperative Work in Design, CSCWD, 2022, Hangzhou, China*, pp. 1275–1280. doi:10.1109/CSCWD54268.2022.9776097
- Golubev S., Novikova E. Transformation of network flow data into images for intrusion detection using convolutional neural networks. *2023 Intern. Russian Automation Conf. (RusAutoCon)*, Sochi, Russian Federation, 2023, pp. 948–952. doi:10.1109/RusAutoCon58002.2023.10272890
- Wang Q., Zhao W., Ren J. Intrusion detection algorithm based on image enhanced convolutional neural network. *J. Intell. Fuzzy Syst.*, 2021, vol. 41, no. 1, pp. 2183–2194. doi:10.3233/JIFS-210863
- Masum M., Shahriar H., Haddad H. M. A transfer learning with deep neural network approach for network intrusion detection. *International Journal of Intelligent Computing Research (IJICR)*, 2021, vol. 12, pp. 087–1095. doi:10.20533/ijicr.2042.4655.2021.0132
- Noever D. A., Noever S. E. M. Image classifiers for network intrusions. *CoRR*, 2021, arXiv:2103.07765. Available at: <https://arxiv.org/abs/2103.07765> (accessed 5 April 2024).
- Ho C. M. K., Yow K.-C., Zhu Z., Aravamathan S. Network intrusion detection via flow-to-image conversion and vision transformer classification. *IEEE Access*, 2022, vol. 10, pp. 97780–97793. doi:10.1109/ACCESS.2022.3200034
- Dosovitskiy A., Beyer L., Kolesnikov A., Weissenborn D., Zhai X., Unterthiner T., Dehghani M., Minderer M., Heigold G., Gelly S., Uszkoreit J., Houlsby N. An image is worth 16x16 words: Transformers for image recognition at scale. 2020, arXiv:2010.11929. Available at: <https://arxiv.org/abs/2010.11929> (accessed 5 April 2024).
- Kim T., Pak W. Deep learning-based network intrusion detection using multiple image transformers. *Appl. Sci.*, 2023, vol. 13, no. 2754. doi:10.3390/app13052754
- Hosler R., Sundar A., Zou X., Li F., Gao T. Unsupervised deep learning for an image based network intrusion detection system. *Proc. of 2023 IEEE Global Communications Conf., Kuala Lumpur, Malaysia, 2023*, pp. 6825–6831. doi:10.1109/GLOBECOM54140.2023.10437636
- Golubev S., Novikova E., Fedorchenko E. Image-based approach to intrusion detection in cyber-physical objects. *Information*, 2022, vol. 13, iss. 12, pp. 553. doi:10.3390/info13120553
- Ghadermazi J., Shah A., Bastian N. Towards real-time network intrusion detection with image-based sequential packets representation. *IEEE Transactions on Big Data*, 2024, no. 01, pp. 1–17. doi:10.1109/TBDDATA.2024.3403394
- Zhang X., Chen J., Zhou Y., Han L., Lin J. A multiple-layer representation learning model for network-based attack detection. *IEEE Access*, 2019, vol. 7, pp. 91992–92008. doi:10.1109/ACCESS.2019.2927465
- Yu L., Dong J., Chen L., Li M., Xu B., Li Z., Qiao L., Liu L., Zhao B., Zhang B. Pbcnn: Packet bytes-based convolutional neural network for network intrusion detection. *Computer Networks*, 2021, vol. 194, no. 108117. doi:10.1016/j.comnet.2021.108117
- Sharafaldin I., Lashkari A. H., Ghorbani A. A. Toward generating a new intrusion detection dataset and intrusion traffic characterization. *Proc. of Intern. Conf. on Information Systems Security and Privacy, Funchal, Madeira – Portugal, January 22–24, 2018*, pp. 108–116.
- Sun Y., Esaki H., Ochiai H. Adaptive intrusion detection in the networking of large-scale LANs with segmented federated learning. *IEEE Open Journal of the Communications Society*, 2021, vol. 2, pp. 102–112. doi:10.1109/OJCOMS.2020.3044323
- Wu P., Guo H., Buckland R. A transfer learning approach for network intrusion detection. *Proc. of 2019 IEEE 4th Intern. Conf. on Big Data Analytics (ICBDA)*, Suzhou, China, 2019, pp. 281–285. doi:10.1109/ICBDA.2019.8713213
- Haralick R., Shanmugam K., Dinstein I. Textural features for image classification. *IEEE TSMC*, 1973, vol. 3, iss. 6, pp. 610–621.



Коды-произведения с полуслепым декодированием для одночастотной передачи в многолучевом канале с замираниями

Ф. А. Таубин^а, доктор техн. наук, профессор, orcid.org/0000-0002-8781-9531, ftaubin@yahoo.com

А. Н. Трофимов^а, канд. техн. наук, доцент, orcid.org/0000-0003-1233-5222

^аСанкт-Петербургский государственный университет аэрокосмического приборостроения,

Б. Морская ул., 67, Санкт-Петербург, 190000, РФ

Введение: построение кодовых конструкций для одночастотной передачи на основе кодов-произведений со сравнительно простыми внутренними кодами, обеспечивающими приемлемый уровень вероятности ошибки при отсутствии в приемнике оценок параметров канала и наличии в канале многолучевости, представляет значительный интерес в плане разработки спектрально-эффективных методов надежной передачи в каналах со случайными быстро меняющимися параметрами. **Цель:** разработать коды-произведения с полуслепым декодированием для одночастотной передачи в многолучевом канале с замираниями и выполнить анализ их помехоустойчивости. **Результаты:** приведено формализованное описание одночастотной передачи с фазовой модуляцией в многолучевом канале с замираниями; предложена кодовая конструкция в виде кода-произведения, ориентированная на использование полуслепой процедуры декодирования внутреннего кода в многолучевом канале с замираниями. Приведено описание полуслепое декодирования по максимуму правдоподобия внутреннего кода с использованием мягких решений; выполнен анализ помехоустойчивости предложенной конструкции; представлены примеры конкретных конструкций кодов-произведений, ориентированных на однолучевой и двухлучевой каналы, с последующим сравнительным анализом, включающим учет как требуемой скорости передачи, так и ряда параметров канала. Показано, что предложенные коды при полуслепое декодировании обеспечивают реализацию внутреннего разнесения, при существующем многолучевому каналу, что выражается появлением существенного энергетического выигрыша. **Практическая значимость:** предложенная конструкция кода-произведения с полуслепым декодированием позволяет эффективно интегрировать помехоустойчивое кодирование с оцениванием текущих параметров многолучевого канала с замираниями, что с учетом умеренной сложности представляет практический интерес при передаче в каналах со случайными быстро меняющимися параметрами. Гибкость конструкции обеспечивает возможность адаптации схемы кодирования к требуемой скорости передачи, количеству лучей, а также величине максимального доплеровского рассеяния. Выполненный анализ помехоустойчивости позволяет получить широкий диапазон обменных соотношений между требуемой скоростью передачи, вероятностью ошибки, сложностью реализации и основными параметрами канала, что способствует рациональному выбору конкретных параметров схемы передачи.

Ключевые слова – канал с замираниями и многолучевым распространением, доплеровское рассеяние, одночастотная передача, коды-произведения, полуслепое декодирование, аддитивная граница вероятности ошибки, граница Чернова, состояние Бхаттачария.

Для цитирования: Таубин Ф. А., Трофимов А. Н. Коды-произведения с полуслепым декодированием для одночастотной передачи в многолучевом канале с замираниями. *Информационно-управляющие системы*, 2024, № 5, с. 68–80. doi:10.31799/1684-8853-2024-5-68-80, EDN: ONWDIC

For citation: Taubin F. A., Trofimov A. N. Product codes with semi-blind decoding for single carrier transmission over multipath fading channel. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2024, no. 5, pp. 68–80 (In Russian). doi:10.31799/1684-8853-2024-5-68-80, EDN: ONWDIC

Введение

Технология беспроводных коммуникаций, поддерживающих обмен информацией между множеством пользователей и (или) устройствами, является одной из доминирующих в настоящее время и в обозримой перспективе. Растущий спрос на высокую скорость передачи данных и высокое качество обслуживания для широкого спектра мобильных приложений требуют разработки и использования гораздо более надежных и спектрально эффективных методов организации связи, включая исследование перспективных методов модуляции и кодирования [1–3].

Как известно, в беспроводных каналах передаваемые сигналы подвергаются замираниям, вызываемым наличием множества рассеивающих объектов между передатчиком и приемником, многолучевым распространением, а также относительным, и, как правило, случайным движением между передатчиком, приемником и рассеивающими объектами, порождающим доплеровское рассеяние [4, 5].

Одним из стандартных решений для компенсации влияния временного рассеяния является использование широкополосных сигналов на основе мультиплексирования с ортогональным частотным разделением (orthogonal frequency-

division multiplexing, OFDM) в качестве основного инструмента для схемы передачи на физическом уровне [4–6]. Как правило, при передаче в диапазоне до 6 ГГц длительность OFDM-сигналов оказывается много меньше величины интервала когерентности канала (которая в первом приближении обратна величине доплеровского рассеяния), поэтому полагают, что OFDM-сигнал передается по каналу со случайными, но неизменяющимися параметрами. Например, в стандарте 802.11a длительность OFDM-сигнала составляет 4 мкс, тогда как типичные значения интервала когерентности канала не опускаются ниже 0,5 мс. Технология OFDM позволяет обеспечить достаточно высокую спектральную эффективность при умеренной сложности реализации частотного компенсатора канальных искажений.

Вместе с тем технологии OFDM присущ ряд существенных недостатков, которые могут ограничить область ее возможного использования. Во-первых, это высокий уровень пик-фактора, что порождает высокую чувствительность OFDM-сигнала к нелинейным искажениям, вызванным усилителем мощности передатчика [7]. Во-вторых, помехоустойчивость OFDM-сигналов резко снижается в высокоомобильных беспроводных системах, при передаче в терагерцевом диапазоне и в нисходящих каналах спутниковых систем связи, где весьма существенным фактором оказывается влияние доплеровского рассеяния (величина которого может достигать нескольких кГц [8, 9]). При сильном доплеровском рассеянии интервал когерентности канала может оказаться сопоставимым с длительностью OFDM-сигнала (или даже меньше его), что приводит к появлению интерференции между поднесущими OFDM-сигнала и заметному возрастанию искажений, связанных как с оценкой текущих параметров канала, так и с компенсацией канальных искажений [10–13].

В ситуациях, когда недостатки, сопряженные с использованием технологии OFDM, оказываются решающими, альтернативным вариантом является одночастотная (single-carrier, SC) передача, при условии, конечно, если основные сопутствующие одночастотной передаче издержки, такие как меньшая по сравнению с OFDM спектральная эффективность и существенно более высокая сложность компенсации временного рассеяния, являются приемлемыми. Следует отметить также, что одночастотная передача рассматривается как наиболее подходящая в таких интенсивно развивающихся технологиях, как узкополосный интернет вещей (NB-IoT), усовершенствованная межмашинная связь (eMTC) и разработанный консорциумом 3GPP стандарт LTE-M для маломощной сотовой связи [14, 15]. Эти технологии ориентированы на передачу

в сравнительно узкой полосе частот — от 100 кГц (NB-IoT и eMTC) до 1 МГц (LTE-M).

Вне зависимости от используемого подхода — OFDM или SC, передаваемые по каналу с замираниями и частотно-временным рассеянием дискретные символы структурируются на передающей стороне посредством объединения в сравнительно короткие пакеты с встроенными в каждый пакет заранее известными в приемнике символами. С учетом указанного структурирования передаваемых символов возможны по меньшей мере три подхода к организации процедуры приема в канале со случайно изменяющимися параметрами при отсутствии информации о текущих параметрах канала.

Первый подход — использование префикса для оценивания текущих параметров канала на интервале, соответствующем длительности пакета, и последующее вынесение решений с применением полученных оценок параметров канала [16, 17]. Основным недостатком этого подхода связан с довольно большой вносимой избыточностью и сложностью учета погрешности оценивания изменяющихся (на интервале длительности пакета) параметров канала. Второй подход, в известном смысле противоположный первому, — слепое вынесение решений, при котором передаваемые пакеты не содержат каких-либо известных символов и решения выносятся только на основе вероятностной модели принятого сигнала [18, 19]. Основным и существенным недостатком слепого вынесения решений связан с тем, что для многих важных сценариев имеет место внутренне присущая этому подходу неоднозначность, проявляющаяся в том, что для нескольких различных (несовпадающих) передаваемых пакетов значения решающей функции могут оказаться идентичными. Кроме того, эффективные слепые процедуры требуют использования достаточно длинных пакетов, что сопровождается значительным ростом сложности приемника.

Промежуточное положение занимает третий подход — полуслепое вынесение решений (semi-blind detection), представляющий собой модификацию слепой процедуры в целях исключения присущей ей неоднозначности [20–22]. Для этого в передаваемый пакет, как и при первом подходе, включается префикс, который учитывается в вероятностной модели принятого сигнала, используемой для вынесения решения. Важным достоинством полуслепой процедуры вынесения решений является возможность использования более коротких префиксов. Кроме того, при анализе помехоустойчивости отсутствует необходимость учитывать погрешность оценивания параметров канала и их изменения на интервале длительности пакета. В целом полуслепая процедура вынесения решений при сравнительно

небольшой длине пакета представляется наиболее предпочтительной.

Далее в настоящей работе рассматривается сценарий одночастотной передачи пакетов с использованием фазовой модуляции (т. е. с минимальным значением пик-фактора) и с полуслепой процедурой вынесения решений. Наряду с рациональным выбором сигнального формата адекватный выбор схемы помехоустойчивого кодирования является ключевым фактором, определяющим требуемую помехоустойчивость передачи в каналах с замираниями и частотно-временным рассеянием. Подход к организации помехоустойчивого кодирования для таких каналов, основанный на использовании одного фиксированного кода, в настоящее время не рассматривается как приоритетный. Наибольшей эффективностью как по помехоустойчивости, так и по гибкости конструкции обладают кодовые схемы, основанные на использовании нескольких кодов [23–29]. К таким схемам относятся каскадные коды и коды-произведения (product codes). В обеих схемах кодирование осуществляется с использованием двух ступеней – внешней и внутренней. Основная идея использования двух ступеней кодирования применительно к каналу с замираниями и частотно-временным рассеянием состоит в следующем. Во-первых, с помощью подходящего сравнительно простого внутреннего кода, встроенного в пакет, обеспечить некоторый, достаточно приемлемый уровень вероятности ошибки при отсутствии в приемнике оценок параметров канала и наличия в канале многолучевости. Во-вторых, учитывая, что благодаря перемежению расширенный канал (включающий кодер и декодер внутреннего кода) оказывается практически каналом с независимыми ошибками, использовать в качестве внешнего достаточно мощный двоичный код для исправления независимых ошибок с умеренной сложностью реализации декодера.

Модель передачи

Пусть двоичная информационная последовательность кодируется с использованием кода-произведения и передается по каналу при помощи q -ичной фазовой модуляции (q -PSK). Рассматриваемый код-произведение G состоит из двух кодов: внешнего двоичного (n, k) -кода C и внутреннего кода V над кольцом Z_q целых чисел по модулю $q, q = 2^m, m$ – целое, имеющего длину N и объем (мощность) 2^K . Входные блоки, состоящие из kK двоичных символов, разбиваются на K подблоков, состоящих из k символов. Каждый подблок кодируется внешним двоичным (n, k) -кодом C со скоростью $R_{out} = k/n$. Полученные

K кодовых слов располагаются как строки прямоугольной матрицы размера $K \times n$. Каждый столбец полученной матрицы кодируется внутренним кодом V . Кодовое слово \mathbf{g} кода-произведения G представляет собой прямоугольную $(N \times n)$ -матрицу, состоящую из n столбцов $\mathbf{g} = (\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n), \mathbf{v}_i \in V, \mathbf{v}_i = (v_{i1}, \dots, v_{iN})'$, где символ «'» означает транспонирование. Скорость кода-произведения G составляет kK/nN [бит/символ].

Пусть (a_1, \dots, a_s) и (z_1, \dots, z_t) – две последовательности символов над кольцом Z_q , которые будут добавляться к каждому слову внутреннего кода V при передаче по каналу, образуя единый пакет. А именно последовательность (a_1, \dots, a_s) представляет собой префикс, добавляемый перед началом каждого кодового слова $\mathbf{v}_i \in V$, тогда как последовательность (z_1, \dots, z_t) – постфикс, добавляется в конце каждого кодового слова $\mathbf{v}_i \in V$:

$$\mathbf{v}_i' = (v_{i1}, \dots, v_{iN}) \mapsto \mathbf{b}_i = (a_1, \dots, a_s, v_{i1}, \dots, v_{iN}, z_1, \dots, z_t), \quad (1)$$

где $a_l, z_j \in Z_q, l = 1, \dots, s, j = 1, \dots, t, i = 1, \dots, n$. Префикс и постфикс могут рассматриваться как пилот-символы, встраиваемые во внутренний код V . Отображение (1) иллюстрирует преобразование внутреннего кода V в удлинненный код V_E с длиной $J = N + s + t$ и скоростью $R_{in} = K/J$ [бит/символ]. Результирующая скорость рассматриваемого кода-произведения с встроенными пилот-символами составляет $R = kK/(nJ)$ [бит/ q -ичный символ].

Отметим, что организация передачи с добавлением префикса к каждому кодовому слову $\mathbf{v}_i \in V$ не является единственно возможным вариантом. Альтернативный вариант – использование внутреннего кода с дополнительной избыточностью, позволяющей исключить неоднозначность, которая характерна, как указывалось выше, для слепого декодирования. В этом случае $s = 0$ и соотношение (1) приобретает вид $\mathbf{v}_i' = (v_{i1}, \dots, v_{iN}) \mapsto \mathbf{b}_i = (v_{i1}, \dots, v_{iN}, z_1, \dots, z_t)$.

Таким образом, передаваемый блок, соответствующий кодовому слову \mathbf{g} кода G , состоит из n подблоков $\mathbf{b}_1, \dots, \mathbf{b}_n$. Будем полагать, что эти подблоки в передаваемой последовательности блоков подвергаются перемежению перед передачей по каналу. Перемежение предназначено для исключения пакетирования ошибок на входе декодера внешнего кода C . Глубина перемежения определяется статистическими характеристиками канала и выбирается таким образом, чтобы ошибки в расширенном канале (включающем внутренний кодер/декодер, модулятор/демодулятор и собственно физический канал с замираниями) можно рассматривать как независимые.

Подблок $\mathbf{b} = (b_1, b_2, \dots, b_j)$, где $b_i \in Z_q, J = N + s + t$ и каждый символ подблока определен как

$$b_i = \begin{cases} a_i, & i = 1, 2, \dots, ; \\ v_{i-s}, & i = s+1, s+2, \dots, s+N; \\ z_{i-s-N}, & i = s+N+1, s+N+2, \dots, J, \end{cases} \quad (2)$$

передается по каналу с тактовой частотой $1/T$ с использованием q -ичной фазовой модуляции, T – длительность тактового интервала. Тактовая частота $1/T$ определяется скоростью w генерации источником двоичных символов и скоростью R кода-произведения G : $1/T = w/R$.

Ограничимся далее рассмотрением формализованной модели канала с дискретным временем. Для рассматриваемой модели принятый на i -м такте сигнал r_i имеет вид

$$r_i = \sum_{l=0}^{L-1} \mu_{il} u_{i-l} + n_i, \quad (3)$$

где $u_i = \sqrt{R\gamma_b} \exp(j2\pi b_i/q)$, γ_b – отношение сигнал/шум, приходящееся на передачу одного бита; n_i – комплексная гауссова случайная величина, $E[n_i] = 0$, $E[n_i n_i^*] = 1$, $E[\cdot]$ – оператор усреднения, $i = 1, 2, \dots, J$. В выражении (3) совокупность комплексных коэффициентов $(\mu_{i0}, \dots, \mu_{iL-1})$ представляет собой дискретную импульсную характеристику L -лучевого канала, соответствующую i -му такту передачи.

Будем полагать, что последовательность векторов $\{\mu_i\}$, $\mu_i = (\mu_{i0}, \dots, \mu_{iL-1})'$ есть L -мерный стационарный комплексный гауссов случайный процесс с нулевым средним и известной матричной корреляционной функцией $\Phi(\cdot)$. Матричная корреляционная функция $\Phi(\cdot)$ последовательности векторов $\{\mu_i\}$ определяется как $\Phi(k) = \Phi_k$, $k = \dots, -2, 1, 0, 1, 2, \dots$, где $L \times L$ корреляционная матрица $\Phi_k = E[\mu_i \mu_{i-k}^*]$, символ «*» означает эрмитово сопряжение. Принятый вектор $\mathbf{r} = (r_1, r_2, \dots, r_J)'$ является J -мерным комплексным гауссовым вектором с плотностью распределения

$$p_{\mathbf{r}|\mathbf{v}}(\mathbf{r}|\mathbf{v}) = \frac{1}{\pi^J \det \mathbf{K}_{\mathbf{v}}} \exp(-\mathbf{r}^* \mathbf{K}_{\mathbf{v}}^{-1} \mathbf{r}), \quad (4)$$

где $\mathbf{K}_{\mathbf{v}}$ – $J \times J$ корреляционная матрица вектора \mathbf{r} , соответствующая кодовому слову \mathbf{v} , т. е. $\mathbf{K}_{\mathbf{v}} = E[\mathbf{r} \mathbf{r}^* | \mathbf{v}]$. Будем полагать, что длина постфикса $t \geq L-1$; тогда элементы k_{lh} , $l, h = 1, \dots, J$ матрицы $\mathbf{K}_{\mathbf{v}}$ имеют следующий вид:

$$k_{lh} = R\gamma_b \sum_{d=0}^{L-1} \sum_{i=0}^{L-1} \varphi_{l-h}(d, i) \times \exp(j2\pi(b_{l-d} - b_{h-i})/q) + \delta_{lh}, \quad (5)$$

где $\varphi_{l-h}(d, i)$ есть (d, i) -й элемент матрицы Φ_{l-h} ; символы b_p при $1 \leq p \leq J$ определены правилом (2), а при $-(L-2) \leq p \leq 0$ являются символа-

ми постфикса: $b_{-p} = b_{J-p} = z_{t-p}$; δ_{lh} – дельта Кронекера.

Одной из наиболее распространенных моделей матричной корреляционной функции $\Phi(\cdot)$, соответствующей модели многолучевого распространения, предложенной Джейксом [1, 30, 31], является модель на основе функций Бесселя первого рода нулевого порядка. Вместе с тем во многих случаях (например, при малых значениях аргумента корреляционной функции $\Phi(\cdot)$) случайные процессы $\{\mu_{il}\}$, $l = 0, 1, \dots, L-1$ в модели Джейкса достаточно точно аппроксимируются авторегрессионной моделью небольшого порядка [32]. При использовании авторегрессионной модели первого порядка с независимыми лучами дискретная импульсная характеристика канала μ_i на i -м такте определяется как

$$\mu_i = \rho \mu_{i-1} + \sqrt{1-\rho^2} \chi_i A^{1/2}, \quad (6)$$

где параметр $\rho \in (0, 1)$; $\{\chi_i\}$ – L -мерный комплексный гауссов случайный процесс с нулевым средним и матричной корреляционной функцией $E[\chi_i \chi_j^*] = \delta_{ij} \mathbf{I}$, $\forall i, j$ (символ \mathbf{I} означает единичную матрицу); $A^{1/2} = \text{diag}(\alpha_0^{1/2}, \dots, \alpha_{L-1}^{1/2})$, $\sum_{l=0}^{L-1} \alpha_l = 1$, $\alpha_l \geq 0$, $l = 0, 1, \dots, L-1$. Из (6) следует, что $E[\mu_i \mu_j^*] = \delta_{ij} \mathbf{A}$, $\forall i, j$, т. е. главная диагональ $(\alpha_0, \dots, \alpha_{L-1})$ матрицы \mathbf{A} определяет распределение дисперсий коэффициентов дискретной импульсной характеристики канала. Корреляционная матрица Φ_k для этой модели имеет вид

$$\Phi_k = \rho^{|k|} \mathbf{A} = \rho^{|k|} \text{diag}(\alpha_0, \dots, \alpha_{L-1}). \quad (7)$$

Параметр ρ в (7) характеризует скорость изменения дискретной импульсной характеристики канала во времени. Как правило, этот параметр в модели (4) определяется нормализованным относительно тактовой частоты $1/T$ максимальным доплеровским рассеянием в канале f_d ; в частности, в модели Джейкса параметр $\rho = J_0(2\pi f_d T)$, где $J_0(\cdot)$ – функция Бесселя первого рода нулевого порядка. При $f_d T \ll 1$ $\rho \approx 1 - (\pi f_d T)^2$. Выражение (5) для элементов k_{lh} , $l, h = 1, 2, \dots, J$ матрицы $\mathbf{K}_{\mathbf{v}}$ с учетом равенства (6) упрощается и приобретает вид

$$k_{lh} = R\gamma_b \sum_{i=0}^{L-1} \alpha_i \rho^{|l-h|} \exp(j2\pi(b_{l-i} - b_{h-i})/q) + \delta_{lh}. \quad (8)$$

Для исключения интерференции между кодовыми словами внутреннего кода V в соседних (после перемежения) передаваемых блоках будем полагать, что длина постфикса $t \geq L-1$. Отметим также, что при введении перемежения/деперемежения подблоков (пакетов) статистическая структура последовательности векторов $\{\mu_i\}$, ассоциированных с каждым передаваемым

подблоком передаваемой последовательности $\mathbf{b}_1, \dots, \mathbf{b}_n$, остается неизменной. Кроме того, принимая во внимание, что величина нормализованного (относительно тактовой частоты $1/T$) максимального доплеровского рассеяния в канале примерно обратна нормализованному интервалу когерентности канала, можно заключить, что при глубине перемерения порядка $1/(f_d T)$ практически исключается пакетирование ошибок на входе декодера внешнего кода C .

Декодирование кода-произведения

Классическая процедура декодирования кода-произведения включает два этапа. На первом этапе декодируются принятые слова внутреннего кода V ; положим, что декодирование внутреннего кода V выполняется посредством процедуры, которую будем называть полуслепое декодирование по максимуму правдоподобия с использованием мягких решений. На втором этапе декодируются принятые слова внешнего двоичного (n, k) -кода C с использованием алгебраического декодера; будем полагать, что алгебраический декодер кода C позволяет исправить ошибки кратностью до половины минимального расстояния. Для рассматриваемого сценария передачи полуслепое декодирование по максимуму правдоподобия определяется следующим образом. Во-первых, в связи с отсутствием в декодере оценок текущих параметров канала при декодировании внутреннего кода V используются дополнительные сигналы, полученные при передаче префикса и постфикса; тем самым реализуется принцип полуслепое вынесения решений. Во-вторых, в качестве мягких решений используются компоненты вектора $\mathbf{r} = (r_1, r_2, \dots, r_j)'$, определяемые выражением (3). В-третьих, декодер по принятому вектору \mathbf{r} выносит решение $\hat{\mathbf{v}}$, определяемое как $\hat{\mathbf{v}} = \max_{\mathbf{v} \in V} p_{\mathbf{r}|\mathbf{v}}(\mathbf{r}|\mathbf{v})$, где условная плотность $p_{\mathbf{r}|\mathbf{v}}(\mathbf{r}|\mathbf{v})$ задается выражением (4). Таким образом, декодер внутреннего кода V в соответствии с выражением (4) для плотности распределения принятого вектора \mathbf{r} должен минимизировать метрику $m(\cdot)$ вида

$$m(\mathbf{v}) = \mathbf{r}^* \mathbf{K}_{\mathbf{v}}^{-1} \mathbf{r} + \ln \det \mathbf{K}_{\mathbf{v}} \quad (9)$$

по всем словам внутреннего кода V . Отыскание минимального значения метрики $m(\cdot)$ в (9) требует предварительного вычисления и хранения 2^K обратных матриц $\mathbf{K}_{\mathbf{v}}^{-1}, \mathbf{v} \in V$. Временные затраты, возникающие при непосредственной реализации алгоритма декодирования, основанного на полном переборе, определяются числом кодовых слов внутреннего кода. Это число равно 2^K , и оно, как следует из примеров, приведенных

в следующем разделе, оказывается очень небольшим, что приводит в итоге к практически реализуемым процедурам декодирования. Кроме того, отметим, что сложность вычисления значений эрмитовой квадратичной формы $\mathbf{r}^* \mathbf{K}_{\mathbf{v}}^{-1} \mathbf{r}$ в (9) может быть существенно уменьшена при использовании разложения Холецкого матрицы $\mathbf{K}_{\mathbf{v}}^{-1}$.

Анализ помехоустойчивости

Для рассматриваемой схемы передачи с использованием кода-произведения G анализ помехоустойчивости включает два этапа. Вначале оценивается вероятность ошибки на бит для внутренней ступени кодирования с кодом V . Затем полученная оценка вероятности ошибки на бит для кода V используется для вычисления оценки вероятности ошибки на бит для внешнего кода C . Верхняя граница вероятности ошибки на бит $P_b^{(in)}$ для внутреннего кода V имеет вид

$$P_b^{(in)} \leq \frac{1}{K} \sum_{\mathbf{v}, \mathbf{v}' \in V} p(\mathbf{v}) t(\mathbf{v}, \mathbf{v}') P(\mathbf{v} \rightarrow \mathbf{v}'), \quad (10)$$

где $p(\mathbf{v})$ — вероятность передачи кодового слова \mathbf{v} ; $t(\mathbf{v}, \mathbf{v}')$ — число бит, в которых отличаются сообщения, соответствующие кодовым словам \mathbf{v} и \mathbf{v}' внутреннего кода, и $P(\mathbf{v} \rightarrow \mathbf{v}')$ — вероятность ошибки декодирования для кода из двух слов \mathbf{v} и \mathbf{v}' внутреннего кода при условии, что было передано слово \mathbf{v} (попарная вероятность ошибки). Попарная вероятность ошибки в верхней границе (10) с учетом выражения для метрики (9) определяется следующим образом:

$$P(\mathbf{v} \rightarrow \mathbf{v}') = \Pr[m(\mathbf{v}) - m(\mathbf{v}') > 0] = \Pr[\mathbf{r}^* (\mathbf{K}_{\mathbf{v}}^{-1} - \mathbf{K}_{\mathbf{v}'}^{-1}) \mathbf{r} - \ln(\det \mathbf{K}_{\mathbf{v}'} / \det \mathbf{K}_{\mathbf{v}}) > 0], \quad (11)$$

где $\mathbf{K}_{\mathbf{v}}$ и $\mathbf{K}_{\mathbf{v}'}$ — корреляционные матрицы вектора \mathbf{r} , соответствующие кодовым словам \mathbf{v} и \mathbf{v}' . Точные формулы для попарной вероятности $P(\mathbf{v} \rightarrow \mathbf{v}')$ требуют вычисления собственных чисел матрицы $\mathbf{I} - \mathbf{K}_{\mathbf{v}} \mathbf{K}_{\mathbf{v}'}^{-1}$ и оказываются довольно сложными (с точки зрения вычислительных затрат) даже в случае простых собственных чисел этой матрицы и умеренных значений длины внутреннего кода. Кроме того, точное значение попарной вероятности весьма сложным образом зависит от конкретного вида кодовых слов, что существенно ограничивает возможности эвристического поиска подходящих внутренних кодов в совокупности с префиксом и постфиксом. В этом плане при анализе помехоустойчивости конкретных конструкций более приемлемым может оказаться использование подходящей верхней границы для вероятности $P(\mathbf{v} \rightarrow \mathbf{v}')$, например границы Чер-

нова. Применяя границу Чернова для правой части формулы (11), получаем следующую оценку сверху для попарной вероятности:

$$\begin{aligned} & P(\mathbf{v} \rightarrow \mathbf{v}') \leq \\ & \leq \mathbb{E} \left[\exp \left(\lambda \left(\mathbf{r}^* (\mathbf{K}_{\mathbf{v}}^{-1} - \mathbf{K}_{\mathbf{v}'}^{-1}) \mathbf{r} - \ln(\det \mathbf{K}_{\mathbf{v}'}/\det \mathbf{K}_{\mathbf{v}}) \right) \right) \right] = \\ & = \int_{\mathbf{r}} \frac{\exp(-\mathbf{r}^* \mathbf{K}_{\mathbf{v}}^{-1} \mathbf{r})}{\pi^J \det(\mathbf{K}_{\mathbf{v}})} \times \\ & \times \exp \left(\lambda \left(\mathbf{r}^* (\mathbf{K}_{\mathbf{v}}^{-1} - \mathbf{K}_{\mathbf{v}'}^{-1}) \mathbf{r} - \ln(\det \mathbf{K}_{\mathbf{v}'}/\det \mathbf{K}_{\mathbf{v}}) \right) \right) d\mathbf{r} = \\ & = e^{-\Delta(\mathbf{v}, \mathbf{v}', \lambda)}, \end{aligned} \quad (12)$$

где усреднение $\mathbb{E}[\cdot]$ производится по случайному вектору \mathbf{r} при передаваемом кодовом слове \mathbf{v} ; λ — параметр Чернова, $0 < \lambda < 1$; величина $\Delta(\mathbf{v}, \mathbf{v}', \lambda)$ в показателе экспоненты, определяемая как

$$\Delta(\mathbf{v}, \mathbf{v}', \lambda) = \ln \frac{\det((1-\lambda)\mathbf{K}_{\mathbf{v}} + \lambda\mathbf{K}_{\mathbf{v}'})}{(\det \mathbf{K}_{\mathbf{v}})^{1-\lambda} (\det \mathbf{K}_{\mathbf{v}'})^{\lambda}}, \quad 0 < \lambda < 1, \quad (13)$$

представляет собой λ -скошенное расстояние Бхаттачария между условными плотностями вектора \mathbf{r} , соответствующими кодовым словам \mathbf{v} и \mathbf{v}' :

$$P(\mathbf{v} \rightarrow \mathbf{v}') < \frac{(\det \mathbf{K}_{\mathbf{v}})^{1-\lambda} (\det \mathbf{K}_{\mathbf{v}'})^{\lambda}}{\det((1-\lambda)\mathbf{K}_{\mathbf{v}} + \lambda\mathbf{K}_{\mathbf{v}'})}, \quad 0 < \lambda < 1. \quad (14)$$

Принимая во внимание, что корреляционная матрица $\mathbf{K}_{\mathbf{v}}$ вектора \mathbf{r} при передаваемом кодовом слове \mathbf{v} может быть представлена как $\mathbf{K}_{\mathbf{v}} = \mathbf{H}_{\mathbf{v}} + \mathbf{I}$, где матрица $\mathbf{H}_{\mathbf{v}} = \mathbb{E}\{\mathbf{s}\mathbf{s}^* | \mathbf{v}\}$ есть корреляционная матрица вектора

$$\mathbf{s} = (s_1, s_2, \dots, s_J)', \quad s_i = \sqrt{R\gamma_b} \sum_{l=0}^{L-1} \mu_{il} \exp(j2\pi b_{i-l}/q),$$

λ -скошенное расстояние Бхаттачария (13) можно представить в виде

$$\begin{aligned} & \Delta(\mathbf{v}, \mathbf{v}', \lambda) = \\ & = \ln \frac{\det(\lambda\mathbf{H}_{\mathbf{v}} + (1-\lambda)\mathbf{H}_{\mathbf{v}'} + \mathbf{I})}{(\det(\mathbf{H}_{\mathbf{v}} + \mathbf{I}))^{\lambda} (\det(\mathbf{H}_{\mathbf{v}'} + \mathbf{I}))^{1-\lambda}}, \quad 0 < \lambda < 1. \end{aligned} \quad (15)$$

Введем величину $d_B(\mathbf{v}, \mathbf{v}')$, определяемую как $d_B(\mathbf{v}, \mathbf{v}') = \max_{0 < \lambda < 1} \lim_{\gamma_b \rightarrow \infty} \Delta(\mathbf{v}, \mathbf{v}', \lambda)$. Из (15) следует, что

$$d_B(\mathbf{v}, \mathbf{v}') = \max_{0 < \lambda < 1} \ln \frac{\det(\lambda\mathbf{H}_{\mathbf{v}} + (1-\lambda)\mathbf{H}_{\mathbf{v}'})}{(\det \mathbf{H}_{\mathbf{v}})^{\lambda} (\det \mathbf{H}_{\mathbf{v}'})^{1-\lambda}}.$$

Очевидно, что $d_B(\mathbf{v}, \mathbf{v}') < \infty$ для любой пары кодовых слов $(\mathbf{v}, \mathbf{v}')$. Можно полагать, что при

слабом шуме в первом приближении в качестве подходящего критерия качества внутреннего кода V (в сочетании, конечно, с префиксом и постфиксом) можно рассматривать минимальное расстояние Бхаттачария d_B , определяемое как $d_B = \min_{\mathbf{v}, \mathbf{v}' \in V} d_B(\mathbf{v}, \mathbf{v}')$.

Декодирование внешнего кода C , как отмечалось выше, выполняется алгебраическим декодером, позволяющим исправить ошибки кратностью до половины минимального расстояния. Будем полагать, что глубина перемежения при передаче последовательности подблоков $\mathbf{b}_1, \dots, \mathbf{b}_n$ имеет порядок, примерно равный величине $1/(f_d T)$. При такой глубине перемежения практически исключается пакетирование ошибок на входе декодера внешнего кода C , поэтому вероятность ошибки декодирования на бит на выходе внешнего декодера оценивается стандартным выражением

$$P_b^{(out)} \leq \sum_{i=t+1}^n \frac{i+t}{n} C_n^i (P_b^{(in)})^i (1 - P_b^{(in)})^{n-i}, \quad (16)$$

где n есть длина внешнего кода C ; t — максимальное число ошибок, исправляемых кодом C .

Примеры конструкций и численные результаты

Выбор внутренних кодов для рассматриваемого сценария передачи в существенной степени определяется следующими двумя факторами. Во-первых, величиной нормализованного (относительно тактовой частоты $1/T$) максимального доплеровского рассеяния в канале f_d , определяющей значение коэффициента корреляции ρ и, следовательно, глубину памяти в модели канала (3). В типичных беспроводных каналах сантиметрового и миллиметрового диапазонов максимальное доплеровское рассеяние f_d может принимать значения от нескольких герц до нескольких килогерц в зависимости от несущей частоты и скорости движения пользователей, а занимаемая полоса частот в узкополосных системах с одностотной передачей варьируется от сотни килогерц до нескольких мегагерц. Поэтому величина нормализованного максимального доплеровского рассеяния в канале $f_d T$ может находиться в пределах от $\sim 10^{-5}$ до $\sim 10^{-1}$. Это означает наличие весьма сильной памяти в канале, т. е. параметр ρ в (6), характеризующий силу памяти, может принимать значения в интервале от $\sim 0,9$ до величины, очень близкой к 1.

Второй фактор — наличие многолучевости в канале. Влияние этого фактора на выбор кода характеризуется тем, что для любых слов \mathbf{v} и \mathbf{v}' внутреннего кода матрицы $\mathbf{K}_{\mathbf{v}}$ и $\mathbf{K}_{\mathbf{v}'}$, определяющие

щие попарную вероятность $P(\mathbf{v} \rightarrow \mathbf{v}')$, представляют собой линейные комбинации, состоящие из L матричных слагаемых. Поэтому для обеспечения «хорошего» различия матриц \mathbf{K}_V и $\mathbf{K}_{V'}$, следует обеспечить «хорошее» попарное различие всех соответствующих слагаемых в этих линейных комбинациях (например, посредством увеличения длины префикса s).

Очевидно, что с учетом указанных факторов: 1) подходящие внутренние коды должны быть весьма короткими; 2) традиционные подходы к выбору кодов над кольцом Z_q целых чисел по

модулю q , основанные на рассмотрении метрик Хэмминга и Ли либо евклидовой метрики, для рассматриваемого сценария передачи не являются рациональными. Одним из возможных подходов в такой ситуации (с учетом того, что декодирование внутреннего кода выполняется перебором по всем 2^K словам) является эвристический поиск подходящих, в смысле минимального расстояния Бхаттачария d_B , коротких кодов посредством направленного перебора. В табл. 1 и 2 приведены примеры коротких внутренних кодов для однолучевого и двухлучевого

■ **Таблица 1.** Параметры внутреннего кода для однолучевого канала

■ **Table 1.** Inner code parameters for one-path channel

Код	q	K	Префикс a_1	Код V		Постфикс z_1	Скорость R_{in}	Минимальное расстояние d_B при	
				Порождающая матрица кода $V_0 \subseteq V$	Образующие смежных классов			$\rho = 0,9$	$\rho = 0,9999$
#1.1	2	1	0	(1)	–	–	1/2	1,7	8,5
#1.2	2	2	–	$\begin{pmatrix} 011 \\ 110 \end{pmatrix}$	–	–	2/3	1,7	8,5
#1.3	2	3	–	$\begin{pmatrix} 0111 \\ 0011 \\ 1011 \end{pmatrix}$	–	–	3/4	1,7	8,5
#1.4	4	2	0	(1)	–	–	1,0	1,1	7,8
#1.5	4	4	–	(012)	(101) (203) (300)	–	4/3	1,1	7,8
#1.6	8	3	–	(1)	–	–	1,5	0,5	6,6

■ **Таблица 2.** Параметры внутреннего кода для двухлучевого канала с равными интенсивностями лучей

■ **Table 2.** Inner code parameters for two-path channel with the equal path intensities

Код	q	K	Префикс $a_1 a_2$	Код V		Постфикс z_1	Скорость R_{in}	Минимальное расстояние d_B при	
				Порождающая матрица кода $V_0 \subseteq V$	Образующие смежных классов			$\rho = 0,9$	$\rho = 0,9999$
#2.1	2	1	0 1	(1)	–	0	1/4	2,2	15,3
#2.2	2	2	0 1	$\begin{pmatrix} 01 \\ 10 \end{pmatrix}$	–	0	2/5	2,3	15,3
#2.3	2	3	0 1	$\begin{pmatrix} 001 \\ 010 \\ 100 \end{pmatrix}$	–	0	1/2	2,7	16,0
#2.4	4	2	0 2	(1)	–	0	1/2	2,1	14,6
#2.5	4	4	0 2	(01)	(1 0) (2 0) (3 0)	0	4/5	2,1	14,6
#2.6	8	3	0 4	(1)	–	0	3/4	2,0	13,4

каналов. Каждый внутренний код V в таблицах является либо линейным кодом над кольцом Z_q , либо объединением нескольких смежных классов линейного подкода $V_0 \subset V$, включая сам подкод V_0 . Значения минимального расстояния Бхаттачария d_B приведены в таблицах для двух значений параметра ρ (0,9 и 0,9999) для иллюстрации существенного влияния значения этого параметра на величину d_B .

В качестве внешних кодов при рассмотрении примеров конкретных конструкций были выбраны коды Боуза – Чоудхури – Хоквингема (коды БЧХ), которые являются одним из основных кандидатов для внешней ступени кодирования (при умеренной длине внешнего кода) благодаря приемлемой сложности реализации и хорошим обменным соотношениям между скоростью и числом исправляемых ошибок [33, 34]. Во многих случаях при умеренных длинах коды БЧХ оказываются лучшими линейными кодами; кроме того, для кодов БЧХ несложным образом могут быть получены весьма точные оценки помехоустойчивости. Для четырех значений требуемой скорости в однолучевом канале (0,25, 0,5, 0,75 и 1,0 бит/символ) и для двух значений требуемой скорости в двухлучевом канале (0,25 и 0,5 бит/символ) были сформированы соответствующие наборы вариантов конструкций с внутренними кодами, приведенными в табл. 1 и 2, и внешними кодами БЧХ. Для примера в табл. 3 приведены рассмотренные наборы вариантов при требуемой скорости $R = 0,5$ бит/символ.

Анализ помехоустойчивости рассмотренных вариантов кодов-произведений показал следующее. Для однолучевого канала при требуемых скоростях $R = 0,25$ бит/символ и $R = 0,5$ бит/символ наименьшая вероятность ошибки достигается при использовании внутреннего кода #1.4; при требуемых скоростях $R = 0,75$ бит/символ и $R = 1,0$ бит/символ предпочтительными оказываются варианты с внутренними кодами #1.5 и

#1.6 соответственно. Для двухлучевого канала с равными интенсивностями лучей при требуемой скорости $R = 0,25$ бит/символ наименьшая вероятность ошибки достигается при использовании внутреннего кода #2.2; при требуемой скорости $R = 0,5$ бит/символ предпочтительным оказывается вариант с внутренним кодом #2.5. Важно отметить, что с увеличением требуемой скорости передачи и (или) величины L более предпочтительными оказываются коды с большим значением модуля q кольца Z_q .

Графики зависимости вероятности ошибки $P_b^{(out)}$, вычисленной по формулам (16), (10) и (14), от отношения сигнал/шум на бит γ_b для лучших вариантов при различных значениях параметра ρ приведены на рис. 1 для однолучевого канала и на рис. 2 для двухлучевого канала. Кроме того, показаны оценки вероятности ошибки декодирования внутреннего кода, полученные путем имитационного моделирования.

Графики позволяют сделать следующие выводы.

1. С возрастанием величины параметра ρ (т. е. уменьшением величины нормализованного максимального доплеровского рассеяния в канале $f_d T$) вероятность ошибочного декодирования $P_b^{(out)}$ снижается; при этом значение $\rho = 0,9999$ (что соответствует $f_d T \approx 3 \cdot 10^{-3}$) можно рассматривать как пороговое, так как при возрастании ρ сверх 0,9999 вероятность ошибки снижается незначительно, тогда как при снижении ρ до величины 0,999 вероятность ошибки резко возрастает.

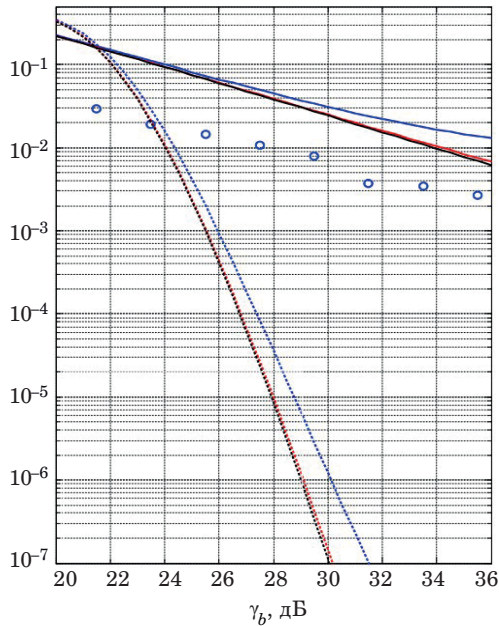
2. При фиксированной вероятности ошибки порядка $10^{-5} \dots 10^{-6}$ увеличение скорости передачи с 0,25 до 0,5 бит/символ в однолучевом канале сопровождается дополнительными энергетическими затратами около 1,5 дБ; дальнейшее увеличение скорости вдвое, до 1,0 бит/символ, требует около 3 дБ. При передаче в двухлучевом канале увеличение скорости передачи с 0,25 до

■ **Таблица 3.** Параметры вариантов конструкций при требуемой скорости $R = 0,5$ бит/символ

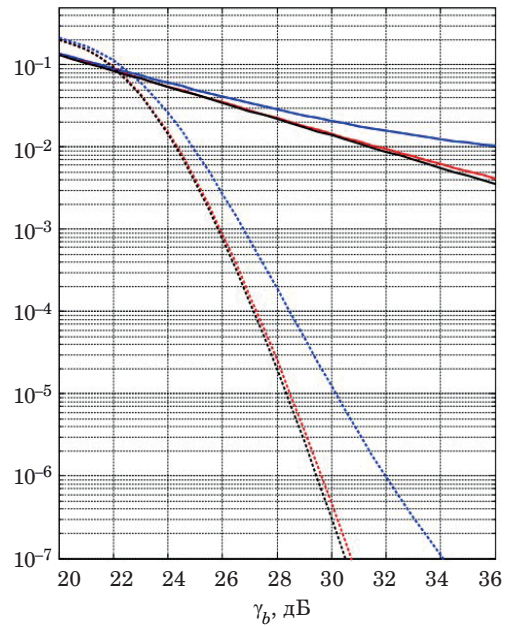
■ **Table 3.** Parameters of the code construction for rate $R = 0.5$ bit/symbol

L	q	Внутренний код		Внешний код		Фактическая скорость R , бит/символ
		Номер	R_{in}	БЧХ	R_{out}	
1	2	1.2	2/3	(127,92)	0,724	0,483
		1.2	2/3	(127,99)	0,780	0,520
		1.3	3/4	(127,85)	0,669	0,502
	4	1.4	1,0	(127,64)	0,504	0,504
		1.5	4/3	(127,50)	0,394	0,525
		1.6	1,5	(127,43)	0,338	0,507
2	4	2.5	4/5	(127,78)	0,614	0,491
	8	2.6	3/4	(127,85)	0,6693	0,502

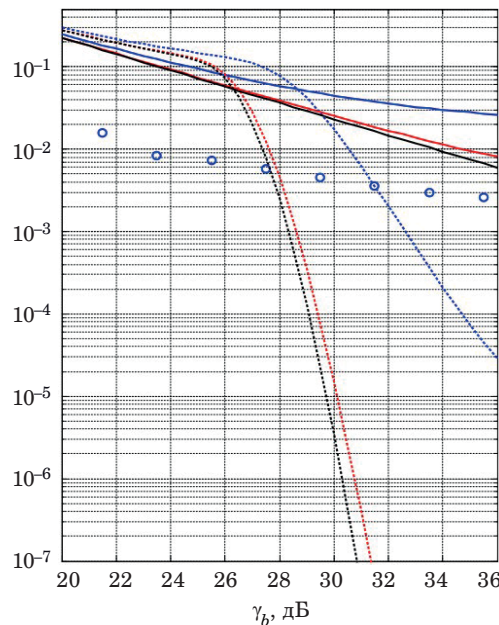
$P_b^{(out)}$, $q = 4$; внутр. код: #1.4 (2,2);
внешний код: (63,18); $R = 0,286$ бит/символ; $L = 1$



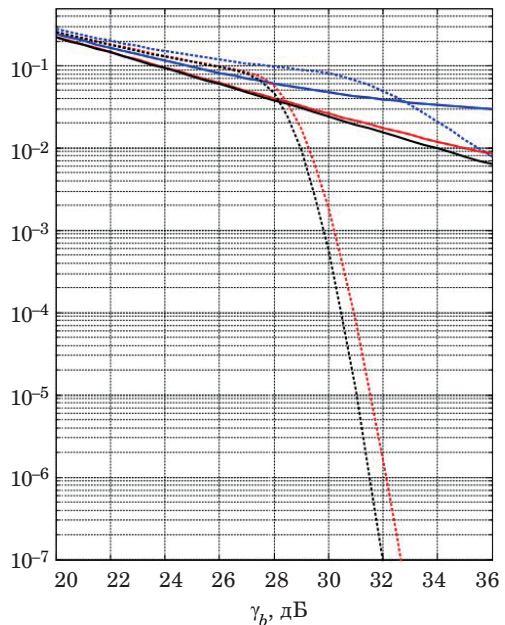
$P_b^{(out)}$, $q = 4$; внутр. код: #1.4 (2,2);
внешний код: (127,64); $R = 0,504$ бит/символ; $L = 1$



$P_b^{(out)}$, $q = 4$; внутр. код: #1.5 (3,4);
внешний код: (511,286); $R = 0,746$ бит/символ; $L = 1$



$P_b^{(out)}$, $q = 8$; внутр. код: #1.6 (2,3);
внешний код: (1023,678); $R = 0,994$ бит/символ; $L = 1$



○ $\rho = 0,999, P_b^{(in)}$, модель
 --- $\rho = 0,999, P_b^{(out)}$
 --- $\rho = 0,9999, P_b^{(out)}$
 --- $\rho = 0,99999, P_b^{(out)}$
— $\rho = 0,999, P_b^{(in)}$
 — $\rho = 0,9999, P_b^{(in)}$
 — $\rho = 0,99999, P_b^{(in)}$

■ **Рис. 1.** Зависимости вероятности ошибки $P_b^{(in)}$ и $P_b^{(out)}$ от отношения сигнал/шум на бит γ_b для однолучевого канала

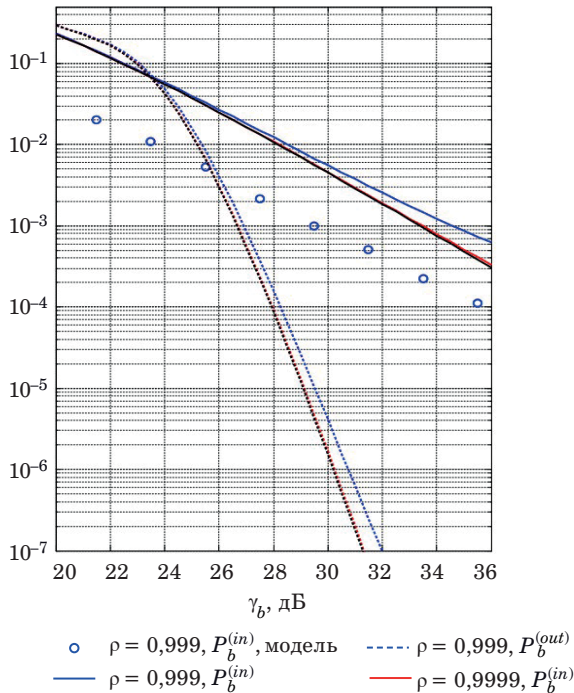
■ **Fig. 1.** Error probabilities $P_b^{(in)}$ and $P_b^{(out)}$ as functions of signal-to-noise ratio per bit γ_b for the one-path channel

0,5 бит/символ требует меньших энергетических затрат – около 0,5 дБ.

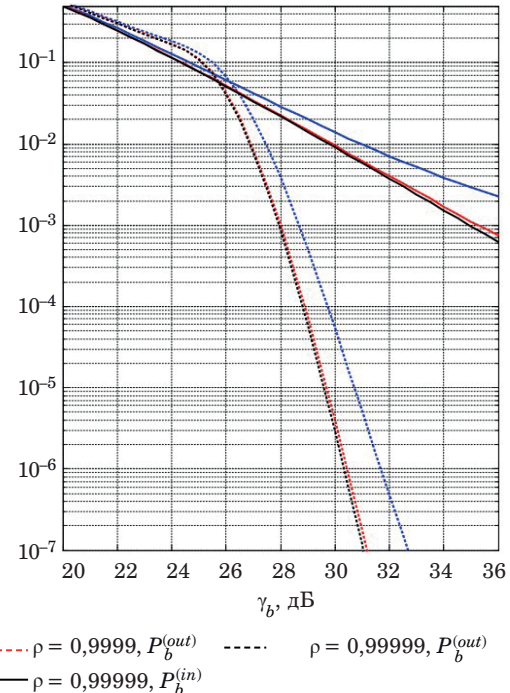
3. Представленные коды для двухлучевого канала со скоростями 0,25 и 0,5 бит/символ

обеспечивают при декодировании внутреннего кода по максимуму правдоподобия реализацию внутреннего разнесения в канале (inner diversity of the channel), что выражается появ-

$P_b^{(out)}$, $q = 2$; внутр. код: #2.2 (4,2); $t = 1$;
внешний код: (63,39); $R = 0,248$ бит/символ; $L = 2$



$P_b^{(out)}$, $q = 4$; внутр. код: #2.5 (4,4); $t = 1$;
внешний код: (127,78); $R = 0,491$ бит/символ; $L = 2$



■ **Рис. 2.** Зависимости вероятностей ошибки $P_b^{(in)}$ и $P_b^{(out)}$ от отношения сигнал/шум на бит γ_b для двухлучевого канала

■ **Fig. 2.** Error probabilities $P_b^{(in)}$ and $P_b^{(out)}$ as functions of signal-to-noise ratio per bit γ_b for the two-path channel

лением существенного энергетического выигрыша — порядка 5 дБ при вероятности ошибки $10^{-5} \dots 10^{-6}$.

Заключение

В настоящей работе представлена и проанализирована кодовая конструкция в виде кода-произведения, предназначенного для одночастотной передачи с использованием фазовой модуляции по многолучевому каналу с коррелированными замираниями при отсутствии оценок текущих параметров канала. Внешний код кода-произведения является линейным двоичным кодом, декодирование которого выполняется алгебраическим декодером, позволяющим исправить ошибки кратностью до половины минимального расстояния. Внутренний код строится над кольцом Z_q целых чисел по модулю q и может снабжаться двумя дополнительными фиксированными последовательностями — префиксом и постфиксом. В силу отсутствия оценок текущих параметров канала для декодирования внутреннего кода применяется полуслепой вариант декодирования по максимуму правдоподобия с использованием мягких решений.

Приведено описание процедуры полуслепого декодирования по максимуму правдоподобия внутреннего кода и выполнен анализ помехоустойчивости предложенной конструкции кода-произведения.

Проведенный анализ показал, что вероятность ошибки на выходе декодера внутреннего кода в существенной степени определяется двумя факторами: характерной для рассматриваемого сценария передачи весьма сильной памятью в канале, поэтому подходящие внутренние коды должны быть весьма короткими, и наличием многолучевости. Основной характеристикой внутреннего кода, в первом приближении определяющей его помехоустойчивость, является минимальное расстояние Бхаттачария. Сравнительный анализ предложенных конкретных конструкций с короткими внутренними кодами показал, что: а) с увеличением требуемой скорости передачи и (или) числа лучей более предпочтительными оказываются коды с большим значением модуля q кольца Z_q ; б) можно указать пороговые величины нормализованного максимального доплеровского рассеяния в канале $f_d T$, при превышении которых вероятность ошибки начинает резко возрастать; в) представленные коды для двухлучевого канала

позволяют реализовать внутреннее разнесение, присущее многолучевому каналу. Отметим, что стандартный подход к реализации внутреннего разнесения базируется на использовании широкополосных сигналов для разделения лучей.

Одним из основных инструментов при выполнении анализа помехоустойчивости предложенной конструкции кода-произведения было использование границы Чернова, что позволило получить верхнюю границу вероятности ошиб-

ки в замкнутой форме. Вместе с тем, учитывая, что точность границы Чернова и (или) ее модификаций во многих ситуациях оказывается невысокой, представляется важным получение для рассматриваемого сценария передачи точных значений попарной вероятности ошибки, вычисляемых с применением подхода, развитого в работах [35, 36], на основе аппарата характеристических функций логарифма отношения правдоподобия.

Литература

1. **Molisch A. F.** *Wireless Communications: From Fundamentals to Beyond 5G*. 3rd Edition. Wiley-IEEE Press, 2022. 1008 p.
2. **Durisi G., Koch T., Popovski P.** Towards massive, ultra-reliable, and low-latency wireless communications with short packets. *Proc. IEEE*, 2016, vol. 104, no. 9, pp. 1711–1726.
3. **Dahlman E., Parkvall S., Skold J.** *5G NR: The Next Generation Wireless Access Technology*. London, UK, Academic, 2020. 466 p.
4. **Goldsmith A. J.** *Wireless Communications*. Cambridge University Press, 2012. 644 p.
5. **Rappaport T. S.** *Wireless Communications: Principles and Practice*. 2nd Edition. Cambridge, Cambridge University Press, 2024. 732 p.
6. **Gopalam S., Pillai S. B., Whiting P., Inaltekin H., Collings I. B., Hanly S. V.** A new micro-subcarrier OFDM-based waveform for delay Doppler domain communication. *IEEE Access*, 2024, vol. 12, pp. 57879–57894. doi:10.1109/ACCESS.2024.3390682
7. **Kumari N., Kaur S.** A survey on various PAPR reduction techniques in OFDM communication systems. *Int. Res. J. Eng. Technol.*, 2015, no. 2, pp. 24–28.
8. **Rappaport T. S., MacCartney G. R., Samimi M. K., Sun S.** Wideband millimeter-wave propagation measurements and channel models for future wireless communication system design. *IEEE Trans. Commun.*, 2015, vol. 63, no. 9, pp. 3029–3056.
9. **Zeng Y., Zhang R., Lim T. J.** Wireless communications with unmanned aerial vehicles: opportunities and challenges. *IEEE Commun. Mag.*, 2016, vol. 54, no. 5, pp. 36–42.
10. **Wang T., Proakis J. G., Masry E., Zeidler J. R.** Performance degradation of OFDM systems due to Doppler spreading. *IEEE Trans. on Wireless Commun.*, 2006, vol. 5, no. 6, pp. 1422–1432.
11. **Eliseev S. N., Filimonova L. N.** On the estimation of the ICI in OFDM in conditions of multipath Doppler spreading and carrier frequencies offset. *2022 Systems of Signals Generating and Processing in the Field of On-Board Communications*, Moscow, 2022, pp. 1–4. doi:10.1109/IEEECONF53456.2022.9744325
12. **Arslan H.** *Wireless Communication Signals: A Laboratory-based Approach*. Wiley, 2021. 464 p.
13. **Wang X., Brink S. T.** SINR analysis of different OFDM-inspired waveforms over doubly dispersive channels. *IEEE Transactions on Vehicular Technology*, 2020, vol. 69, no. 9, pp. 9459–9468. doi:10.1109/TVT.2020.3004570
14. **Medina-Acosta G. A., Zhang L., Chen J., Uesaka K., Wang Y., Lundqvist O., Bergman J.** 3GPP Release-17 physical layer enhancements for LTE-M and NB-IoT. *IEEE Communications Standards Magazine*, 2022, vol. 6, iss. 4, pp. 80–86. doi:10.1109/MCOM-STD.0001.2100099
15. **Mahjoubi A. E., Mazri T., Hmina N.** NB-IoT and eMTC: Engineering results towards 5G/IoT mobile technologies. *2018 Intern. Symp. on Advanced Electrical and Communication Technologies (ISAECT)*, Rabat, Morocco, 2018, pp. 1–7. doi:10.1109/ISAECT.2018.8618818
16. **Vikalo H., Hassibi B., Kailath T.** Optimal training for frequency-selective fading channels. *IEEE Intern. Conf. on Acoustics, Speech, and Signal Processing. Proc. (Cat. No. 01CH37221)*, Salt Lake City, UT, USA, 2001, vol. 4, pp. 2105–2108. doi:10.1109/ICASSP.2001.940408
17. **Chiu L.-K., Wu S.-H.** An effective approach to evaluate the training and modeling efficacy in MIMO time-varying fading channels. *IEEE Trans. Commun.*, 2015, vol. 63, no. 1, pp. 140–155.
18. **Nabavi P., Yuksel M.** Blind detection of digital signals in MIMO communication. *2022 IEEE Wireless Communications and Networking Conf. (WCNC)*, Austin, TX, USA, 2022, pp. 2196–2201. doi:10.1109/WCNC51071.2022.9771848
19. **Walk P., Jung P., Hassibi B.** MOCZ for blind short-packet communication: Basic principles. *IEEE Transactions on Wireless*, 2019, vol. 18, no. 11, pp. 5080–5097. doi:10.1109/TWC.2019.2932668
20. **Giannakis G. B., Hua Y., Stoica P., Tong L.** Eds. *Signal Processing Advances in Wireless & Mobile Communications: Trends in Single and Multi-User Systems*. Prentice-Hall, New York, NY, USA, 2001. 448 p.
21. **Naraghi-Pour M., Rashid M., Vargas-Rosales C.** Semi-blind channel estimation and data detection for multi-cell massive MIMO systems on time-varying channels. *IEEE Access*, 2021, vol. 9, pp. 161709–161722. doi:10.1109/ACCESS.2021.3132263

22. Wenjing Y., Yuan X. Semi-blind channel-and-signal estimation for uplink massive MIMO with channel sparsity. *IEEE Access*, 2021, vol. 7, pp. 95008–95020.
23. Yavasoglu O., Akcam N., Okan T. Performance analysis of concatenated BCH and convolutional coded OFDM system. *International Journal of Electronics*, 2020, vol. 107, no. 10, pp. 1574–1587. doi:10.1080/00207217.2020.1727029
24. Sadkhan S. B. Performance evaluation of concatenated codes applied in wireless channels. *2019 1st AL-Noor Intern. Conf. for Science and Technology (NICST)*, Sulimanyiah, 2019, pp. 89–93. doi:10.1109/NICST49484.2019.9043836
25. Chen J., Chen H., Li Z. A double serial concatenated code using CRC-aided error correction for highly reliable communication. *Computer Networks*, 2022, vol. 216, iss. C, Article 109260. https://doi.org/10.1016/j.comnet.2022.109260
26. Liu B., Li Y., Rong B., Gui L., Wu Y. LDPC-RS product codes for digital terrestrial broadcasting transmission system. *IEEE Trans. Broadcasting*, 2014, vol. 60, no. 1, pp. 38–49. doi:10.1109/TBC.2013.2291359.60
27. Dong P., Xiang X., Liang Y., Wang P. A block-based concatenated LDPC-RS code for UAV-to-ground SC-FDE communication systems. *Electronics*, 2023, vol. 12, no. 14, 3143. https://doi.org/10.3390/electronics12143143
28. Mukhtar H., Al-Dweik A., Shami A. Turbo product codes: Applications, challenges, and future directions. *IEEE Communications Surveys & Tutorials*, 2016, vol. 18, no. 4, pp. 3052–3069. doi:10.1109/COMST.2016.2587863
29. Li Z., Miao M., Wang Z. Parallel coding scheme with turbo product code for mobile multimedia transmission in MIMO-FBMC system. *IEEE Access*, 2019, vol. 8, pp. 3772–3780. https://doi.org/10.1109/ACCESS.2019.2958482
30. Ferrand P., Amara M., Valentin S., Guillaud M. Trends and challenges in wireless channel modeling for evolving radio access. *IEEE Communications Magazine*, 2016, vol. 54, no. 7, pp. 93–99.
31. Liu H., Chen H. Simulation model of wireless channel. *Applied Mathematics*, 2015, vol. 6, no. 8, pp. 1318–1331. doi:10.4236/am.2015.68125
32. Baddour K., Beaulieu N. Autoregressive modeling for fading channel simulation. *IEEE Transactions on Wireless Communications*, 2005, vol. 4, no. 4, pp. 1650–1662.
33. Hussain G. A., Audah L. Downlink LTE system performance improvement by using BCH codes over LTE-MIMO channel. *International Journal of Integrated Engineering*, 2018, vol. 10, no. 4, pp. 95–101.
34. Freudenberger J., Nicolas Bailon D., Safieh M. Reduced complexity hard- and soft-input BCH decoding with applications in concatenated codes. *IET Circ. Devices Syst.*, 2021, vol. 15, no. 3, pp. 284–296.
35. Трофимов А. Н., Таубин Ф. А. Вычисление аддитивной границы вероятности ошибки декодирования с использованием характеристических функций. *Информационно-управляющие системы*, 2021, № 4, с. 71–85. doi:10.31799/1684-8853-2021-4-71-85
36. Трофимов А. Н., Таубин Ф. А. Улучшенная граница вероятности ошибки при оптимальном приеме в канале с межсимвольной интерференцией. *Информационно-управляющие системы*, 2023, № 5, с. 33–42. doi:10.31799/1684-8853-2023-5-33-42, EDN: MDHOXU

UDC 621.391

doi:10.31799/1684-8853-2024-5-68-80

EDN: ONWDIC

Product codes with semi-blind decoding for single carrier transmission over multipath fading channel

F. A. Taubin^a, Dr. Sc., Tech., Professor, orcid.org/0000-0002-8781-9531, ftaubin@yahoo.comA. N. Trofimov^a, PhD, Tech., Associate Professor, orcid.org/0000-0003-1233-5222^aSaint-Petersburg State University of Aerospace Instrumentation, 67, B. Morskaya St., 190000, Saint-Petersburg, Russian Federation

Introduction: Design of codes for single-carrier transmission based on product codes with relatively simple internal codes that provide an acceptable error probability in a multipath fading channel with no channel parameter estimates in the receiver is of considerable interest for developing spectrally efficient methods for reliable transmission in channels with random, rapidly changing parameters. **Purpose:** To develop product codes with semi-blind decoding for single-carrier transmission in a multipath fading channel and analyze their performance. **Results:** We propose a formal description of single-carrier transmission with PSK modulation in a multipath fading channel, as well as a product code exploiting a semi-blind procedure for decoding the inner code. We describe semi-blind maximum likelihood soft-decision decoding of the inner code, carry out the performance analysis of the proposed design and present examples of specific designs of product codes for one- and two-path channels, various transmission rates and the values for the channel parameters. We demonstrate that the proposed codes with semi-blind decoding ensure the implementation of the internal diversity inherent in a multipath channel, which is expressed in significant energy gain. **Practical relevance:** The proposed design of a product code with semi-blind decoding efficiently integrates the error-correcting coding with the channel parameter estimation and has moderate complexity, which is of practical interest for the transmission in channels with random, rapidly changing parameters. The flexibility of the design makes it possible to adapt it to the transmission rate, number of paths, and maximum Doppler spread. The performance analysis provides a wide range of possible trade-offs between transmission rate, error probability, implementation complexity, and channel parameter values, which can help to rationally select key system parameters.

Keywords – fading multipath channel, Doppler spread scattering, single carrier transmission, product codes, semi-blind detection, union bound on error probability, Chernoff bound, Bhattacharyya distance.

For citation: Taubin F. A., Trofimov A. N. Product codes with semi-blind decoding for single carrier transmission over multipath fading channel. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2024, no. 5, pp. 68–80 (In Russian). doi:10.31799/1684-8853-2024-5-68-80, EDN: ONWDIC

References

- Molisch A. F. *Wireless Communications: From Fundamentals to Beyond 5G*. 3rd Edition. Wiley–IEEE Press, 2022. 1008 p.
- Durisi G., Koch T., Popovski P. Towards massive, ultra-reliable, and low-latency wireless communications with short packets. *Proc. IEEE*, 2016, vol. 104, no. 9, pp. 1711–1726.
- Dahlman E., Parkvall S., Skold J. *5G NR: The Next Generation Wireless Access Technology*. London, UK, Academic, 2020. 466 p.
- Goldsmith A. J. *Wireless Communications*. Cambridge University Press, 2012. 644 p.
- Rappaport T. S. *Wireless Communications: Principles and Practice*. 2nd Edition. Cambridge, Cambridge University Press, 2024. 732 p.
- Gopalam S., Pillai S. B., Whiting P., Inaltekin H., Collings I. B., Hanly S. V. A new micro-subcarrier OFDM-based waveform for delay Doppler domain communication. *IEEE Access*, 2024, vol. 12, pp. 57879–57894. doi:10.1109/ACCESS.2024.3390682
- Kumari N., Kaur S. A survey on various PAPR reduction techniques in OFDM communication systems. *Int. Res. J. Eng. Technol.*, 2015, no. 2, pp. 24–28.
- Rappaport T. S., MacCartney G. R., Samimi M. K., Sun S. Wideband millimeter-wave propagation measurements and channel models for future wireless communication system design. *IEEE Trans. Commun.*, 2015, vol. 63, no. 9, pp. 3029–3056.
- Zeng Y., Zhang R., Lim T. J. Wireless communications with unmanned aerial vehicles: opportunities and challenges. *IEEE Commun. Mag.*, 2016, vol. 54, no. 5, pp. 36–42.
- Wang T., Proakis J. G., Masry E., Zeidler J. R. Performance degradation of OFDM systems due to Doppler spreading. *IEEE Trans. on Wireless Commun.*, 2006, vol. 5, no. 6, pp. 1422–1432.
- Eliseev S. N., Filimonova L. N. On the estimation of the ICI in OFDM in conditions of multipath Doppler spreading and carrier frequencies offset. *2022 Systems of Signals Generating and Processing in the Field of On-Board Communications*, Moscow, 2022, pp. 1–4. doi:10.1109/IEEECONF53456.2022.9744325
- Arslian H. *Wireless Communication Signals: A Laboratory-based Approach*. Wiley, 2021. 464 p.
- Wang X., Brink S. T. SINR analysis of different OFDM-inspired waveforms over doubly dispersive channels. *IEEE Transactions on Vehicular Technology*, 2020, vol. 69, no. 9, pp. 9459–9468. doi:10.1109/TVT.2020.3004570
- Medina-Acosta G. A., Zhang L., Chen J., Uesaka K., Wang Y., Lundqvist O., Bergman J. 3GPP Release-17 physical layer enhancements for LTE-M and NB-IoT. *IEEE Communications Standards Magazine*, 2022, vol. 6, iss. 4, pp. 80–86. doi:10.1109/MCOMSTD.0001.2100099
- Mahjoubi A. E., Mazri T., Hmina N. NB-IoT and eMTC: Engineering results towards 5G/IoT mobile technologies. *2018 Intern. Symp. on Advanced Electrical and Communication Technologies (ISAECT)*, Rabat, Morocco, 2018, pp. 1–7. doi:10.1109/ISAECT.2018.8618818
- Vikalo H., Hassibi B., Kailath T. Optimal training for frequency-selective fading channels. *IEEE Intern. Conf. on Acoustics, Speech, and Signal Processing. Proc. (Cat. No. 01CH37221)*, Salt Lake City, UT, USA, 2001, vol. 4, pp. 2105–2108. doi:10.1109/ICASSP.2001.940408
- Chiu L.-K., Wu S.-H. An effective approach to evaluate the training and modeling efficacy in MIMO time-varying fading channels. *IEEE Trans. Commun.*, 2015, vol. 63, no. 1, pp. 140–155.
- Nabavi P., Yuksel M. Blind detection of digital signals in MIMO communication. *2022 IEEE Wireless Communications and Networking Conf. (WCNC)*, Austin, TX, USA, 2022, pp. 2196–2201. doi:10.1109/WCNC51071.2022.9771848
- Walk P., Jung P., Hassibi B. MOCZ for blind short-packet communication: Basic principles. *IEEE Transactions on Wireless*, 2019, vol. 18, no. 11, pp. 5080–5097. doi:10.1109/TWC.2019.2932668
- Giannakis G. B., Hua Y., Stoica P., Tong L. Eds. *Signal Processing Advances in Wireless & Mobile Communications: Trends in Single and Multi-User Systems*. Prentice-Hall, New York, NY, USA, 2001. 448 p.
- Naraghi-Pour M., Rashid M., Vargas-Rosales C. Semi-blind channel estimation and data detection for multi-cell massive MIMO systems on time-varying channels. *IEEE Access*, 2021, vol. 9, pp. 161709–161722. doi:10.1109/ACCESS.2021.3132263
- Wenjing Y., Yuan X. Semi-blind channel-and-signal estimation for uplink massive MIMO with channel sparsity. *IEEE Access*, 2021, vol. 7, pp. 95008–95020.
- Yavasoglu O., Akcam N., Okan T. Performance analysis of concatenated BCH and convolutional coded OFDM system. *International Journal of Electronics*, 2020, vol. 107, no. 10, pp. 1574–1587. doi:10.1080/00207217.2020.1727029
- Sadkhan S. B. Performance evaluation of concatenated codes applied in wireless channels. *2019 1st AL-Noor Intern. Conf. for Science and Technology (NICST)*, Sulimanyiah, 2019, pp. 89–93. doi:10.1109/NICST49484.2019.9043836
- Chen J., Chen H., Li Z. A double serial concatenated code using CRC-aided error correction for highly reliable communication. *Computer Networks*, 2022, vol. 216, iss. C, Article 109260. https://doi.org/10.1016/j.comnet.2022.109260
- Liu B., Li Y., Rong B., Gui L., Wu Y. LDPC-RS product codes for digital terrestrial broadcasting transmission system. *IEEE Trans. Broadcasting*, 2014, vol. 60, no. 1, pp. 38–49. doi:10.1109/TBC.2013.2291359.60
- Dong P., Xiang X., Liang Y., Wang P. A block-based concatenated LDPC-RS code for UAV-to-ground SC-FDE communication systems. *Electronics*, 2023, vol. 12, no. 14, 3143. https://doi.org/10.3390/electronics12143143
- Mukhtar H., Al-Dweik A., Shami A. Turbo product codes: Applications, challenges, and future directions. *IEEE Communications Surveys & Tutorials*, 2016, vol. 18, no. 4, pp. 3052–3069. doi:10.1109/COMST.2016.2587863
- Li Z., Miao M., Wang Z. Parallel coding scheme with turbo product code for mobile multimedia transmission in MIMO-FBMC system. *IEEE Access*, 2019, vol. 8, pp. 3772–3780. https://doi.org/10.1109/ACCESS.2019.2958482
- Ferrand P., Amara M., Valentin S., Guillaud M. Trends and challenges in wireless channel modeling for evolving radio access. *IEEE Communications Magazine*, 2016, vol. 54, no. 7, pp. 93–99.
- Liu H., Chen H. Simulation model of wireless channel. *Applied Mathematics*, 2015, vol. 6, no. 8, pp. 1318–1331. doi:10.4236/am.2015.68125
- Baddour K., Beaulieu N. Autoregressive modeling for fading channel simulation. *IEEE Transactions on Wireless Communications*, 2005, vol. 4, no. 4, pp. 1650–1662.
- Hussain G. A., Audah L. Downlink LTE system performance improvement by using BCH codes over LTE-MIMO channel. *International Journal of Integrated Engineering*, 2018, vol. 10, no. 4, pp. 95–101.
- Freudenberger J., Nicolas Bailon D., Safieh M. Reduced complexity hard- and soft-input BCH decoding with applications in concatenated codes. *IET Circ. Devices Syst.*, 2021, vol. 15, no. 3, pp. 284–296.
- Trofimov A. N., Taubin F. A. Evaluation of the union bound for the decoding error probability using characteristic functions. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2021, no. 4, pp. 71–85 (In Russian). doi:10.31799/1684-8853-2021-4-71-85
- Trofimov A. N., Taubin F. A. Improved bound on optimal reception error probability for an intersymbol interference channel. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2023, no. 5, pp. 33–42 (In Russian). doi:10.31799/1684-8853-2023-5-33-42, EDN: MDH-OXUA

**АКИМЕНКО
Татьяна
Алексеевна**



Доцент кафедры промышленной автоматизации и робототехники Тульского государственного университета.

В 1995 году окончила Тульский государственный технический университет по специальности «Ракетостроение».

В 2003 году защитила диссертацию на соискание ученой степени кандидата технических наук. Является автором более 150 научных публикаций, четырех патентов на полезную модель и трех свидетельств о регистрации программы для ЭВМ.

Область научных интересов — теория устойчивости иерархических многоконтурных систем управления сложными технологическими объектами, информационно-измерительные и управляющие системы, теплотехника.
Эл. адрес: tantan72@mail.ru

**БАШМАКОВ
Наиль
Маратович**



Аспирант, ассистент кафедры вычислительной техники и защиты информации Уфимского университета науки и технологий.

В 2019 году окончил бакалавриат по специальности «Информационная безопасность», в 2021 году — магистратуру по специальности «Информатика и вычислительная техника» Уфимского государственного авиационного технического университета.

Является автором девяти научных публикаций.

Область научных интересов — информационная безопасность, интеллектуальные системы в информационной безопасности.

Эл. адрес: nail.bashmakov@gmail.com

**БОГОМОЛОВ
Алексей
Валерьевич**



Профессор, ведущий научный сотрудник Федерального исследовательского центра «Информатика и управление» РАН, Москва. Заслуженный деятель науки РФ, лауреат премий Правительства РФ в области науки и техники.

В 1996 году окончил Военно-воздушную инженерную академию им. профессора Н. Е. Жуковского по специальности «Автоматизированные системы управления». В 2003 году защитил диссертацию на соискание ученой степени доктора технических наук.

Является автором более 200 научных публикаций, 80 патентов на изобретения.

Область научных интересов — цифровые технологии сбора и обработки информации.

Эл. адрес: a.v.bogomolov@gmail.com

**ВАСИЛЬЕВ
Владимир
Иванович**



Профессор кафедры вычислительной техники и защиты информации Уфимского университета науки и технологий.

В 1980 году окончил Ленинградский электротехнический институт им. В. И. Ульянова (Ленина) по специальности «Электронные вычислительные машины».

В 1990 году защитил диссертацию на соискание ученой степени доктора технических наук.

Является автором 400 научных публикаций и 14 свидетельств о регистрации программ для ЭВМ. Область научных интересов — интеллектуальные системы управления, информационная безопасность.

Эл. адрес: vas0015@yandex.ru

**ВУЛЬФИН
Алексей
Михайлович**



Профессор кафедры вычислительной техники и защиты информации Уфимского университета науки и технологий.

В 2008 году окончил Уфимский государственный нефтяной технический университет по специальности «Программное обеспечение вычислительной техники и автоматизированных систем». В 2022 году защитил диссертацию на соискание ученой степени доктора технических наук.

Является автором 160 научных публикаций, 23 свидетельств о регистрации программ для ЭВМ и одного патента на изобретение. Область научных интересов — интеллектуальный анализ данных, моделирование сложных технических систем.

Эл. адрес: vulfin.am@ugatu.su

**ГОЛУБЕВ
Сергей
Александрович**



Аспирант лаборатории проблем компьютерной безопасности Санкт-Петербургского Федерального исследовательского центра РАН, ассистент кафедры информационных систем Санкт-Петербургского государственного электротехнического университета «ЛЭТИ», программист .Net компании «Терралинк девелопмент».

В 2021 году окончил Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» по специальности «Информационные системы и технологии».

Область научных интересов — информационная безопасность, информационные системы и технологии.

Эл. адрес: ser9800@mail.ru

КАГИРОВ
Ильдар
Амирович



Научный сотрудник лаборатории речевых и мультимодальных интерфейсов Санкт-Петербургского Федерального исследовательского центра РАН. В 2008 году окончил филологический факультет Санкт-Петербургского государственного университета по специальности «Лингвистика». Является автором свыше 40 научных публикаций. Область научных интересов – малоресурсные языки, человеко-машинное взаимодействие, синтаксис и грамматическая семантика естественных языков, корпусная лингвистика. Эл. адрес: kagirov@iias.spb.su

КАПУСТА
Константин
Леонидович



Программист лаборатории речевых и мультимодальных интерфейсов Санкт-Петербургского Федерального исследовательского центра РАН. В 2024 году окончил Университет ИТМО по специальности «Информационные системы и технологии». Является автором двух научных публикаций. Область научных интересов – речевые технологии, распознавание речи, обработка естественного языка. Эл. адрес: khome1518@gmail.com

КАРТАК
Вадим
Михайлович



Профессор, заведующий кафедрой вычислительной техники и защиты информации Уфимского университета науки и технологий. В 1995 году окончил Уфимский государственный авиационный технический университет по специальности «Программное обеспечение вычислительной техники и АС». В 2011 году защитил диссертацию на соискание ученой степени доктора физико-математических наук. Является автором 88 научных публикаций. Область научных интересов – большие данные, методы оптимизации, информационная безопасность. Эл. адрес: kartak.vm@ugatu.su

КИПЯТКОВА
Ирина
Сергеевна



Доцент, старший научный сотрудник лаборатории речевых и мультимодальных интерфейсов Санкт-Петербургского Федерального исследовательского центра РАН. В 2008 году окончила Санкт-Петербургский государственный университет аэрокосмического приборостроения по специальности «Роботы и робототехнические системы». В 2011 году защитила диссертацию на соискание ученой степени кандидата технических наук. Является автором 100 научных публикаций и одного патента на изобретение. Область научных интересов – автоматическое распознавание речи, нейронные сети, малоресурсное распознавание языков. Эл. адрес: kipyatkova@iias.spb.su

КИРИЛЛОВА
Анастасия
Дмитриевна



Старший преподаватель кафедры вычислительной техники и защиты информации Уфимского университета науки и технологий. В 2015 году окончила бакалавриат по специальности «Информационная безопасность», в 2017 году – магистратуру по специальности «Информатика и вычислительная техника» Уфимского государственного авиационного технического университета. В 2023 году защитила диссертацию на соискание ученой степени кандидата технических наук. Является автором 50 научных публикаций и девяти свидетельств о регистрации программ для ЭВМ. Область научных интересов – информационная безопасность, автоматизированные системы управления, управление рисками, когнитивное моделирование. Эл. адрес: kirillova.andm@gmail.com

КОЛЕСНИКОВА
Светлана
Ивановна



Профессор кафедры компьютерных технологий и программной инженерии Санкт-Петербургского государственного университета аэрокосмического приборостроения. В 1980 году окончила Томский госуниверситет по специальности «Прикладная математика». В 2011 году защитила диссертацию на соискание ученой степени доктора технических наук. Является автором 160 научных публикаций и двух патентов на изобретения. Область научных интересов – интеллектуальный анализ данных, управление нелинейными объектами с неопределенностью в описании и их применение в технических и биоинженерных исследованиях. Эл. адрес: skolesnikova@yandex.ru

**КУЗНЕЦОВА
Елизавета
Олеговна**



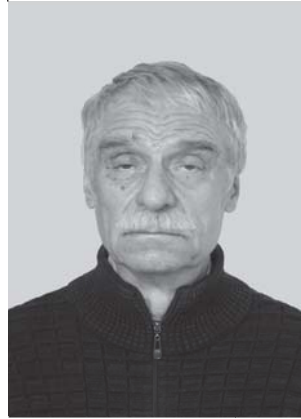
Магистрант Санкт-Петербургского государственного электротехнического университета «ЛЭТИ».

В 2024 году окончила бакалавриат Санкт-Петербургского государственного электротехнического университета «ЛЭТИ» по специальности «Информационные системы и технологии».

Область научных интересов — машинное обучение, искусственный интеллект.

Эл. адрес:
eokuznetsova@stud.etu.ru

**ЛАРКИН
Евгений
Васильевич**



Профессор-консультант кафедры промышленной автоматизации и робототехники Тульского государственного университета. Почетный работник высшего профессионального образования РФ.

В 1975 году окончил Тульский политехнический институт по специальности «Автоматизированные системы управления».

В 1996 году защитил диссертацию на соискание ученой степени доктора технических наук. Является автором более 200 научных публикаций, 70 патентов на изобретения.

Область научных интересов — теория устойчивости иерархических многоконтурных систем управления сложными технологическими объектами.

Эл. адрес: elarkin@mail.ru

**НОВИКОВА
Евгения
Сергеевна**



Старший научный сотрудник лаборатории проблем компьютерной безопасности Санкт-Петербургского Федерального исследовательского центра РАН, доцент Санкт-Петербургского государственного электротехнического университета «ЛЭТИ».

В 2007 году окончила Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» по специальности «Компьютерная безопасность».

В 2010 году защитила диссертацию на соискание ученой степени кандидата технических наук. Является автором 160 научных публикаций и трех патентов на изобретения.

Область научных интересов — информационная безопасность, обнаружение аномалий и вторжений, конфиденциальность машинное обучение.

Эл. адрес: novikova@comsec.spb.ru

**ПРИВАЛОВ
Александр
Николаевич**



Профессор, директор Института передовых информационных технологий Тульского государственного педагогического университета им. Л. Н. Толстого.

В 1983 году окончил Тульское высшее артиллерийское училище имени Тульского пролетариата по специальности «Электронные вычислительные машины».

В 2010 году защитил диссертацию на соискание ученой степени доктора технических наук.

Является автором 390 научных публикаций и 12 свидетельств о регистрации программ для ЭВМ. Область научных интересов — методы проектирования программного обеспечения информационных систем, тренажерных систем, компьютерное моделирование, распределенные вычисления.

Эл. адрес: privaloov.61@mail.ru

**ТАУБИН
Феликс
Александрович**



Профессор кафедры информационных систем Санкт-Петербургского государственного университета аэрокосмического приборостроения.

В 1968 году окончил Ленинградский институт авиационного приборостроения по специальности «Радиоэлектронные устройства систем управления».

В 1992 году защитил диссертацию на соискание ученой степени доктора технических наук.

Является автором более 80 научных публикаций, девяти патентов и двух изобретений.

Область научных интересов — цифровые системы связи, методы помехоустойчивого кодирования, широкополосные системы, беспроводные сети.

Эл. адрес: ftaubin@yahoo.com

**ТРОФИМОВ
Андрей
Николаевич**



Доцент кафедры инфокоммуникационных систем Санкт-Петербургского государственного университета аэрокосмического приборостроения.

В 1976 году окончил Ленинградский институт авиационного приборостроения по специальности «Автоматизированные системы управления».

В 1982 году защитил диссертацию на соискание ученой степени кандидата технических наук.

Является автором 41 научной публикации и шести патентов на изобретения.

Область научных интересов — теория передачи дискретных сообщений, теория информации, теория кодирования.

Эл. адрес:
andrei.trofimov@vu.spb.ru

ФОМЕНКОВА
Анастасия
Алексеевна



Доцент кафедры компьютерных технологий и программной инженерии Санкт-Петербургского государственного университета аэрокосмического приборостроения.

В 2012 году окончила Киевский политехнический университет по специальности «Оборудование фармацевтической и микробиологической промышленности».

В 2022 году защитила диссертацию на соискание ученой степени кандидата технических наук. Является автором 19 научных публикаций и одного патента на изобретения.

Область научных интересов – математическое и имитационное моделирование биотехнических систем.

Эл. адрес: a.a.fomenkova@mail.ru

ПАМЯТКА ДЛЯ АВТОРОВ

Поступающие в редакцию статьи проходят обязательное рецензирование.

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

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

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

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

Уважаемые авторы!

При подготовке рукописей статей необходимо руководствоваться следующими рекомендациями.

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

Текст рукописи должен быть оригинальным, а цитирование и самоцитирование корректно оформлено.

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

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

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

Простые формулы набирайте в Word, сложные с помощью редактора Mathtype или Equation. Для набора одной формулы не используйте два редактора; при наборе формул в формульном редакторе знаки препинания, ограничивающие формулу, набирайте вместе с формулой; для установки размера шрифта в Mathtype никогда не пользуйтесь вкладкой Other, Smaller, Larger, используйте заводские установки редактора, не подгоняйте размер символов в формулах под размер шрифта в тексте статьи, не растягивайте и не сжимайте мышью формулы, вставленные в текст; пробелы в формуле ставьте только после запятой при перечислении с помощью Ctrl+Shift+Space (пробел); не отделяйте пробелами знаки: + = - ×, а также пространство внутри скобок; для выделения греческих символов в Mathtype полужирным начертанием используйте Style → Other → bold.

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

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

Подробнее см. <http://i-us.ru/index.php/ius/author-guide>

Иллюстрации:

— рисунки, графики, диаграммы, блок-схемы предоставляйте в виде отдельных исходных файлов, подающихся редактированию, используя векторные программы: Visio (*.vsd, *.vsdx); Adobe Illustrator (*.ai); Coreldraw (*.cdr, версия не выше 15); Excel (*.xls); Word (*.docx); AutoCad, Matlab (экспорт в PDF, EPS, SVG, WMF, EMF); Компас (экспорт в PDF); веб-портал DRAW.IO (экспорт в PDF); Inkscape (экспорт в PDF);

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

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

В редакцию предоставляются:

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

— экспертное заключение;

— экспортное заключение.

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

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

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

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

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

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

Более подробно правила подготовки текста с образцами изложены на нашем сайте в разделе «Руководство для авторов» — <http://i-us.ru/index.php/ius/author-guide>.

Контакты

Куда: 190000, г. Санкт-Петербург, ул. Большая Морская, д. 67, лит. А, ГУАП, РИЦ

Кому: Редакция журнала «Информационно-управляющие системы»

Тел.: (812) 494-70-02

Эл. почта: ius.spb@gmail.com

Сайт: www.i-us.ru