

Издательский дом

МЕДИА ПАБЛИШЕР

Научный журнал "Т-Сomm: Телекоммуникации и транспорт"

Журнал включен в перечень периодических научных изданий, рекомендуемый ВАК Минобрнауки России для публикации научных работ, отражающих основное научное содержание кандидатских и докторских диссертаций

Журнал зарегистрирован Федеральной службой по надзору за соблюдением законодательства в сфере массовых коммуникаций и охране культурного наследия

Свидетельство о регистрации СМИ:

ПИ № ФС77-55956.

Дата выдачи: 07 ноября 2013 г.

Язык публикации: русский, английский.

Территория распространения:

Российская Федерация, зарубежные страны

Тираж: 1000 экз.

Периодичность выхода: 12 номеров в год

Стоимость одного экземпляра: 1000 руб.

Плата с аспирантов за публикацию рукописи не взимается

Предпечатная подготовка:

ООО "ИД Медиа Паблишер"

Мнения авторов не всегда совпадают с точкой зрения редакции.
За содержание рекламных материалов редакция ответственности не несет

Материалы, опубликованные в журнале — собственность ООО "ИД Медиа Паблишер". Перепечатка, цитирование, дублирование на сайтах допускаются только с разрешения издателя

© ООО "ИД Медиа Паблишер", 2026

Адрес редакции и издателя

111024, Россия, Москва, ул. Авиамоторная, д. 8, стр. 1, офис 323

e-mail: t-comm@media-publisher.ru

Тел.: +7 (495) 957-77-43

Адрес типографии

Москва, ул. Складочная, д. 3, корп. 6

Индексация журнала:

Ulrich's Periodicals Directory; RSCI; EBSCO; elibrary.ru (ПИНЦ)
Google Scholar; CyberLeninka (Open Science);
Bielefeld Academic Search Engine (BASE); OCLC WorldCat;
Registry of Open Access Repositories (ROAR)

Journal is registered by Federal Service for monitoring compliance with cultural heritage protection law

ISSN 2072-8743 (Online) ISSN 2072-8735 (Print)

Media Registration Certificate

PI No. FS77-55956. Date of issue: November 7, 2013

Publication language: Russian, English.

Distribution Territory: Russian Federation, foreign countries

All articles and illustrations are copyright. All rights reserved.

No reproduction is permitted in whole or part without the express consent of Media Publisher Joint-Stock Company

© "Media Publisher", 2026

Editorial and Publisher Address

111024, Russia, Moscow, Aviamotornaya str. 8, bloc 1, office 323

e-mail: t-comm@media-publisher.ru.

Tel.: +7 (495) 957-77-43

ИЗДАТЕЛЬСКИЙ ДОМ МЕДИА ПАБЛИШЕР



ПОЛНЫЙ ЦИКЛ ПОДГОТОВКИ КНИГ, ПЕРИОДИЧЕСКИХ ИЗДАНИЙ И РЕКЛАМНОЙ ПРОДУКЦИИ

ПРОФЕССИОНАЛЬНОЕ ЛИТЕРАТУРНОЕ И ТЕХНИЧЕСКОЕ РЕДАКТИРОВАНИЕ

ВЫСОКОКАЧЕСТВЕННАЯ ОФСЕТНАЯ И ЦИФРОВАЯ ПЕЧАТЬ

www.media-publisher.ru

Журнал включен в перечень периодических научных изданий, рекомендуемый ВАК Минобрнауки России для публикации научных работ, отражающих основное научное содержание кандидатских и докторских диссертаций

Учредитель

ООО "Издательский дом Медиа Паблишер"

Главный редактор

Тихвинский Валерий Олегович

Издатель

Дымкова Светлана Сергеевна

ds@media-publisher.ru

Редакционная коллегия

Аджемов Артём Сергеевич

(д.т.н., профессор МТУСИ), *Россия*

Анютин Александр Павлович

(д.ф.-м.н., профессор, член программного и оргкомитетов WSEAS), *Россия, Мексика*

Бестугин Александр Рюальдович

(д.т.н., профессор ГУАП), *Россия*

Вааль Альберт

(д.т.н., старший научный сотрудник Ганноверского университета им. Лейбница на кафедре коммуникационной техники), *Германия*

Варламов Олег Витальевич

(д.т.н., в.н.с. МТУСИ), *Россия*

Головачев Юлиус

(управляющий консультант Detecon International GmbH), *Германия*

Гребенников Андрей Викторович

(Sumitomo Electric Europe), *Великобритания*

Данилов Владимир Григорьевич

(д.ф.-м.н., профессор МИЭМ, НИУ ВШЭ), *Россия*

Дулкейтс Эрик

(д.т.н., старший исполнительный директор корпорации Detecon), *Силиконовая долина, США*

Елизаров Андрей Альбертович

(д.т.н., профессор МИЭМ, НИУ ВШЭ), *Россия*

Ибрагимов Байрам

(д.т.н., профессор Азербайджанского технического университета, АзТУ), *Азербайджан*

Корбетт Ровэлл

(д.т.н., директор по исследованиям в научно-исследовательском центре China Mobile Research Institute, профессор университета Назарбаева), *Гон-Конг (Китай), США*

Кузовкова Татьяна Алексеевна

(д.э.н., декан экономического факультета МТУСИ), *Россия*

Лазарева Галина Геннадьевна

(член-корр. РАН, д.ф.-м.н., профессор РАН, РУДН), *Россия*

Лернер Илья Михайлович

(д.т.н., КНИТУ-КАИ), *Россия*

Ныркв Анатолий Павлович

(д.т.н., профессор, ГУМРФ им. адмирала С.О. Макарова), *Россия*

Омельянов Георгий Александрович

(д.ф.-м.н., Университет де Сонора, факультет математики, Эрмосильо), *Мексика*

Самойлов Александр Георгиевич

(д.т.н., профессор Владимирского государственного университета им. А.Г. и Н.Г. Столетовых), *Россия*

Сысоев Николай Николаевич

(д.ф.-м.н., декан физического факультета МГУ им. М.В. Ломоносова), *Россия*

Чиров Денис Сергеевич

(д.т.н., профессор МТУСИ), *Россия*

Шаврин Сергей Сергеевич

(д.т.н., профессор МТУСИ), *Россия*

Шарп Майкл

(д.э.н., Европейский институт стандартизации – ETSI), *Великобритания*

Яшина Марина Викторовна

(д.т.н., профессор, МТУСИ), *Россия*

СОДЕРЖАНИЕ

ЭЛЕКТРОНИКА. РАДИОТЕХНИКА

Джиган В.И.

Сравнительный анализ эффективности компенсатора помех и адаптивной антенной решётки с линейными ограничениями в наземных навигационных приёмниках

4

**Дворников С.В., Бестугин А.Р., Киршина И.А.,
Васильева Д.В., Васильев А.П.**

Повышение контрастности признаков распознавания изображений на основе кратномасштабной обработки гистограмм распределения их яркости

18

Мосин С.Г.

Обеспечение безопасности беспроводных устройств IoT-инфраструктуры с помощью полупроводниковых физически неклонируемых функций

27

СВЯЗЬ

Бурков А.А., Тюрликов А.М.

Взаимосвязь мощности и энергии при передаче в режиме случайного множественного доступа с минимальной средней задержкой

36

ИНФОРМАТИКА

Петров А.Ю., Оленев В.Л., Тюрликов А.М.

Методика применения АРГМОП для программного разделения сетевых ресурсов в распределённой вычислительной сети

47

ПУБЛИКАЦИИ НА АНГЛИЙСКОМ ЯЗЫКЕ

ЭЛЕКТРОНИКА. РАДИОТЕХНИКА

Кудряшова А.Ю.

Разработка и отладка программного обеспечения для реализации резервного радиоканала управления БПЛА

59

ИНФОРМАТИКА

Карелина М.Ю., Акулов А.А., Талдыкин Д.С.

Исследование порога когнитивного насыщения модели YOLOv8 в условиях ограниченной обучающей выборки

68

CONTENT

ELECTRONICS. RADIO ENGINEERING

Djigan V.I.

Comparative analysis of interference canceller and linearly-constrained adaptive antenna array efficiency in ground-based navigation receivers

4

Dvornikov S.V., Bestugin A.R., Kirshina I.A., Vasilyeva D.V., Vasilyev A.P.

Contrast enhancement of image recognition features based on multiscale processing of histograms of their brightness Distribution

18

Mosin S.G.

The security ensuring of IoT-infrastructure's wireless devices using semiconductor physically unclonable functions

27

COMMUNICATIONS

Burkov A.A., Turlikov A.M.

Relationship between power and energy in random multiple access transmission with minimum average delay

36

COMPUTER SCIENCE

Petrov A.Yu., Olenov V.L., Tyurlikov A.M.

Methodology for Applying ARG MOR to Software-Based Partitioning of Network Resources in a Distributed Computing Network

47

PUBLICATIONS IN ENGLISH

ELECTRONICS. RADIO ENGINEERING

Kudryashova A.Yu.

Development and debugging of software for the implementation of a backup radio channel for UAV control

59

COMPUTER SCIENCE

Karelina M.Yu., Akulov A.A., Taldykin D.S.

Investigation of the cognitive saturation threshold of the YOLOv8 model under limited training data

68

T - C o m m

Telecommunications and transport

Volume 20. No. 9-2026

Release date: 20.09.2026

The journal is included in the list of scientific publications, recommended Higher Attestation Commission Russian Ministry of Education for the publication of scientific works, which reflect the basic scientific content of candidate and doctoral theses.

Founder: "Media Publisher", Ltd.

Publisher: Svetlana S. Dymkova
ds@media-publisher.ru

Editor in Chief: Dr. Valery O. Tikhvinskiy

Editorial board

Artem S. Adzhemov
Doctor of sciences, Professor MTUCI, *Russia*

Alexander P. Anyutin
Doctor of sciences, Professor, member of the program and organizing committee WSEAS, *Russia, Mexico*

Aleksandr R. Bestugin
Doctor of sciences, Professor SUAI, *Russia*

Corbett Rowell
Full Professor: Electronic & Electrical Engineering
Nazarbayev University, *Hong Kong (China), USA*

Denis S. Chirov
Doctor of sciences, MTUCI, *Russia*

Vladimir G. Danilov
Doctor of sciences, Professor MIEM, HSE, *Russia*

Eric Dulkeyts
Ph.D., chief executive officer of the corporation Detecon, *USA*

Julius Golovachyov
Managing Consultant Detecon International GmbH, *Germany*

Andrey Grebennikov
Ph.D., Sumitomo Electric Europe, *United Kingdom*

Bayram Ibrahimov
Ph.D., Professor of Azerbaijan Technical University (AzTU),
Azerbaijan

Tatyana A. Kuzovkova
Doctor of sciences, MTUCI, *Russia*

Galina G. Lazareva
Corresponding Member, RAS, Doctor of sciences, Professor RAS,
RUDN, *Russia*

Ilya M. Lerner
Doctor of sciences, KNRTU-KAI, *Russia*

Anatoliy P. Nyrkov
Doctor of sciences, Professor of Admiral Makarov State University of Maritime and Inland Shipping, *Russia*

Georgii A. Omel'yanov
Doctor of sciences, Universidad de Sonora,
Department of Mathematics, Hermosillo, *Mexico*

Alexander G. Samoilov
Doctor of sciences, VLSU, *Russia*

Michael Sharpe
PhD, European Standards Institute – ETSI, *United Kingdom*

Sergey S. Shavrin
Doctor of sciences, MTUCI, *Russia*

Nikolai N. Sysoev
Doctor of sciences, Dean of the Faculty of Physics
of Moscow State University Lomonosov, *Russia*

Oleg V. Varlamov
Doctor of sciences, MTUCI, *Russia*

Albert Waal
Ph.D., Senior Research Fellow University of Hanover. Leibniz
at the Department of Communications Technology, *Germany*

Marina V. Yashina
Doctor of sciences, Professor MTUCI, *Russia*

Andrey A. Yelizarov
Doctor of sciences, Professor MIEM, HSE, *Russia*

www.media-publisher.ru

СРАВНИТЕЛЬНЫЙ АНАЛИЗ ЭФФЕКТИВНОСТИ КОМПЕНСАТОРА ПОМЕХ И АДАПТИВНОЙ АНТЕННОЙ РЕШЕТКИ С ЛИНЕЙНЫМИ ОГРАНИЧЕНИЯМИ В НАЗЕМНЫХ НАВИГАЦИОННЫХ ПРИЕМНИКАХ

DOI: 10.36724/2072-8735-2026-20-9-4-17

Manuscript received 07 June 2026;

Accepted 12 August 2026

Работа проведена в рамках выполнения государственного задания НИЦ "Курчатовский институт"

Джиган Виктор Иванович,
Национальный исследовательский центр "Курчатовский институт",
Национальный исследовательский университет "Московский
институт электронной техники", Москва, Россия,
dzhigan_vi@nrcki.ru

Ключевые слова: Адаптивная антенная решётка, компенсатор помех, линейные ограничения, рекурсивный алгоритм по критерию наименьших квадратов, глобальные навигационные спутниковые системы, коэффициент усиления шума, коэффициент подавления помех, карты отношения сигнал-шум

Современные радиосистемы различного назначения часто функционируют в условиях помех, обусловленных другими радиосистемами, работающими в том же диапазоне радиочастот. Выделение полезного сигнала из помех путём частотной фильтрации в этом случае неэффективно, а кодовое разделение сигналов возможно, лишь если полезный сигнал обладает специальной структурой. Более универсальным способом борьбы с помехами является их пространственная фильтрация, если в качестве антенны радиосистемы используется адаптивная антенная решётка (ААР). Подавление помех с помощью ААР осуществляется путем образования провалов в её диаграмме направленности в направлениях на источники помех. Эти провалы определяются значениями весовых коэффициентов ААР, вычисляемых с помощью адаптивных алгоритмов в реальном масштабе времени. В работе рассмотрены особенности архитектур и вычислительных процедур двух разновидностей ААР: в виде компенсатора помех и с линейными ограничениями, – при их использовании в качестве антенн наземных радиоприёмников глобальных навигационных спутниковых систем (ГНСС). Представленные вычислительные процедуры используют рекурсивные алгоритмы адаптивной фильтрации по критерию наименьших квадратов на базе QR-разложения без операций извлечения квадратного корня. Моделированием ААР в одинаковых условиях работы установлено, что компенсатор помех демонстрирует в среднем лучшие показатели качества в терминах меньшего коэффициента усиления шума на выходе ААР, большего коэффициента подавления помех, а также большей доли направлений, в которых возможен приём спутниковых сигналов в установившемся режиме относительно всех направлений, окружающих ААР. Эти характеристики заметно различаются в рассматриваемых ААР, если число источников подавляемых помех меньше максимального числа помех, которое может подавить ААР. Рассмотренные ААР могут быть использованы не только в радиоприёмниках ГНСС, но и в других радиосистемах, где принимаемые радиосигналы скрыты шумами.

Информация об авторе:

Джиган Виктор Иванович, доктор технических наук, профессор, главный научный сотрудник Отдела методологии проектирования интегральных схем Отделения проблем проектирования в микроэлектронике Центра перспективной микроэлектроники Национального исследовательского центра "Курчатовский институт", профессор Института микровприборов и систем управления им. Л.Н. Преснухина Национального исследовательского университета "Московский институт электронной техники", Москва, Россия. ORCID 0000-0001-7485-1623

Для цитирования:

Джиган В.И. Сравнительный анализ эффективности компенсатора помех и адаптивной антенной решётки с линейными ограничениями в наземных навигационных приёмниках // Т-Comm: Телекоммуникации и транспорт. 2026. Том 20. №9. С. 4-17.

For citation:

V.I. Djigan, "Comparative analysis of interference canceller and linearly-constrained adaptive antenna array efficiency in ground-based navigation receivers," T-Comm, 2026, vol. 20, no. 9, pp. 4-17. (in Russian)

Введение

Радиосистемы различного назначения сегодня широко используются в повседневной деятельности современного человека. Ярким примером является массовый переход от проводной телефонной связи и проводного интернета к их беспроводным аналогам. Все чаще используются беспроводные датчики потребления электроэнергии, воды и тепла. Часто применяется и беспилотный транспорт. Также часто применяется непрерывное дистанционное наблюдение за текущим состоянием здоровья пациентов, страдающих специфическими заболеваниями. Все это лишь некоторые примеры использования беспроводных технологий передачи и приёма данных в бытовых приложениях, наряду с которыми, по-прежнему, существуют и активно эксплуатируются радиосистемы специального назначения, такие как радио- и телевидение, радиоуправление, радиолокация, радионавигация и радиосвязь. Сегодня полосы частот сигналов разных радиосистем часто пересекаются, так как частотный ресурс ограничен. Поэтому такие системы являются источниками помех друг-другу.

Существует несколько способов выделения полезного сигнала из его смеси с помехами, базирующихся на методах цифровой обработки сигналов (ЦОС). Простейшим из них является фильтрация сигналов по частоте. Однако, если полосы частот полезного сигнала и помех пересекаются, то такая фильтрация в большинстве случаев становится неэффективной. Также возможно кодовое разделение полезных сигналов и помех. Однако не все полезные сигналы обладают структурными особенностями, позволяющими осуществлять такое разделение. И, наконец, разделение сигналов можно осуществлять путем их пространственной фильтрации с помощью антенных решеток (АР) [1, 2] в случае их использования в качестве антенн радиоприёмных устройств.

Пространственная фильтрация сигналов с помощью приёмной АР обеспечивается за счет того, что её амплитудно-угловая характеристика, называемая диаграммой направленности (ДН), является неравномерной. Её форма может быть фиксированной подобно амплитудно-частотной характеристике (АЧХ) цифровых фильтров с фиксированными весовыми коэффициентами (ВК) [3, 4] или изменяться подобно АЧХ цифровых адаптивных фильтров с изменяемыми во времени ВК [5, 6].

Форма ДН АР определяется значениями её ВК. Если ВК изменяются в реальном масштабе времени, подстраивая ДН АР так, чтобы обеспечивать работу радиоприёмника в помеховой обстановке, то такая АР называется адаптивной (ААР) [7, 8]. ВК ААР рассчитываются с помощью различных адаптивных алгоритмов [5, 6] также в реальном масштабе времени.

Существует два основных класса адаптивных алгоритмов. Это алгоритмы, в основе которых находится градиентный поиск минимума некоторой целевой функции, например, функции ошибки, и алгоритмы, базирующиеся на решении системы линейных уравнений. Алгоритмы первого класса являются простыми с вычислительной точки зрения, но медленными с точки зрения скорости достижения установившегося режима. Наиболее эффективными адаптивными алгоритмами сегодня являются рекурсивные алгоритмы по критерию наименьших квадратов (Recursive Least Squares, RLS), принадлежащие ко второму классу. RLS-алгоритмы характеризуются сравнительно высокой вычислительной сложностью, которая,

однако, уже не является препятствием для их реализации на современной цифровой элементной базе [9, 10]. Одновременно эти алгоритмы характеризуются высокой скоростью сходимости, что позволяет им эффективно функционировать в условиях изменяющейся сигнальной обстановки. Такие алгоритмы уже используются в ряде приложений, например, в [11-13]. Под вычислительной сложностью алгоритмов ЦОС понимается число арифметических операций, требуемое для выполнения одной итерации алгоритма, длительность которой, как правило, совпадает с длительностью периода дискретизации обрабатываемых сигналов.

Для борьбы с сигналами пространственно-разнесённых источников помех в основном используются три разновидности ААР. Это компенсаторы боковых лепестков (Side-Lobe Canceller, SLC) [14], ААР с ограничением значений ДН в интересующих направлениях (Linearly Constrained, LC) [15], а также ААР, минимизирующие различные функции ошибок между выходным и так называемым требуемым сигналами [16].

Одним из востребованных приложений ААР сегодня является их использование в качестве антенн наземных радиоприёмников сигналов глобальных навигационных спутниковых систем (ГНСС) [17]. ААР обеспечивает возможность определения координат такого радиоприёмника в помеховой обстановке, в том числе при наличии преднамеренных помех. Особенностью функционирования ААР в этом приложении является то обстоятельство, что отношение сигнал-шум (ОСШ) на выходе приёмника сигналов ГНСС при расположении навигационного спутника в зените оценивается на уровне -20 дБ, а в других направлениях оно еще меньше.

Это означает, что сигнал со спутника скрыт шумами, а потому в ААР адаптивный алгоритм с требуемым сигналом использовать нельзя несмотря на то, что этот сигнал известен – это псевдослучайная последовательность (ПСП) [18], например, код Голда в случае ГНСС Global Positioning System, GPS [17]), посредством которой данные со спутника передаются на Землю. Кроме того, для решения навигационной задачи требуется принимать сигналы от нескольких спутников одновременно, а потому использовать лишь один требуемый сигнал не представляется возможным.

Решение [16], позволяющее использовать требуемый сигнал в ААР приёмников сигналов ГНСС за счет предварительной корреляционной обработки этих сигналов, является эффективным, но одновременно и ресурсозатратным, так как требует построения многолучевой ААР, лучи которой ориентируются на каждый из видимых спутников, в части вычисления ВК и выходного сигнала для каждого луча, а также в части корреляторов, число которых равно произведению числа лучей на число антенн ААР.

Поэтому в качестве ААР наземных приёмников сигналов ГНСС обычно используются либо ААР в виде компенсатора помех на базе SLC, либо LC ААР. В [19] было показано, что обе разновидности ААР демонстрируют примерно одинаковые характеристики. Одной из таких характеристик является относительная величина, определяемая как доля направлений, в которых в установившемся режиме, т.е. по окончании переходного процесса, на выходе ААР обеспечивается отношение сигнал-шум плюс подавленные помехи (ОСПП) не хуже -20 дБ относительно всех возможных направлений, образующих сферу, окружающую ААР. Здесь под шумом понимаются тепловые и другие шумы на выходах радиоприёмников каналов

ААР. Заключение об эффективности указанных двух разновидностей ААР было сделано по результатам их моделирования в условиях, когда число источников помех, сигналы которых можно подавить с помощью ААР, равно теоретически максимальному значению $M-1$ [7], где M – это число антенн ААР. Однако дальнейшие исследования показали, что данный показатель эффективности, а также другие показатели, такие как коэффициент усиления шума и коэффициент подавления помех, этих двух разновидностей ААР все же различаются в ситуациях, когда число источников помех меньше значения $M-1$, что часто встречается на практике.

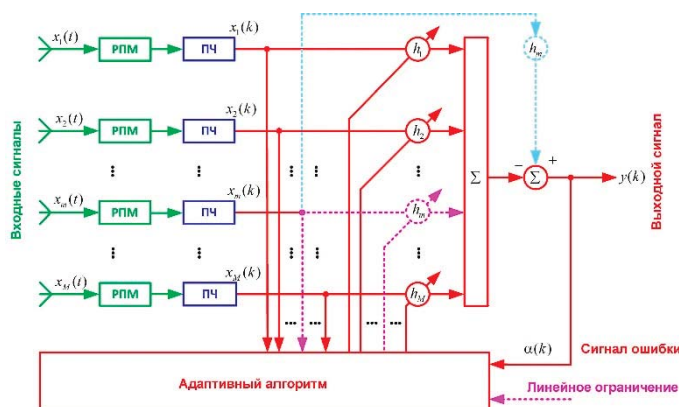


Рис. 1. Обобщенная архитектура компенсатора помех и LC ААР

Цель настоящей работы – путем моделирования сравнить характеристики компенсатора помех и LC ААР в различных помеховых обстановках: при разном числе помех и разных угловых расположениях их источников.

В работе приводятся архитектуры рассматриваемых ААР, их вычислительные процедуры на базе адаптивных RLS-алгоритмов, использующих QR-разложение без операций извлечения квадратного корня [5]. Также приводятся результаты численного моделирования, демонстрирующие характеристики этих ААР.

Архитектуры и вычислительные процедуры компенсатора помех и LC ААР

Обобщенная архитектура рассматриваемых в настоящей работе ААР приведена на рисунке 1. Здесь зеленым цветом обозначена аналоговая часть ААР, т.е. антенны и радиоприёмники (РПМ), осуществляющие приём, полосовую фильтрацию и усиление радиосигналов; а синим цветом – преобразователи частоты (ПЧ), которые преобразуют непрерывные во времени комплексные огибающие $x_m(t)$ входных сигналов ААР в дискретные отсчеты $x_m(k)$. Здесь и далее, $m = 1, 2, \dots, M$, t – это время, а k – номер отсчета. ПЧ могут быть аналоговыми или цифровыми. В последнем случае ПЧ строятся на принципах, используемых в программно-определяемом радио [20]. Остальная часть архитектуры, рис. 1, реализуется в цифровом виде и является цифровой АР (ЦАР) [21]. В ЦАР формирование отсчетов выходного сигнала $y(k)$ осуществляется цифровым способом путём умножения комплексных отсчетов $x_m(k)$ на комплексные ВК $h_m(k)$ как

$$y(k) = \mathbf{h}_M^H(k-1) \mathbf{x}_M(k),$$

где

$$\mathbf{x}_M(k) = [x_1(k), x_2(k), \dots, x_m(k), \dots, x_M(k)]^T$$

– это вектор отсчетов $x_m(k)$, а

$$\mathbf{h}_M(k) = [h_1(k), h_2(k), \dots, h_m(k), \dots, h_M(k)]^T$$

– вектор ВК ААР. Вычисление выходного сигнала и ВК ААР осуществляется в арифметике комплексных чисел, так как используемые в этих вычислениях сигналы $x_m(t)$ и их дискретные отсчеты $x_m(k)$ являются комплексными, поскольку характеризуются амплитудами и фазами [1, 2]. По этой же причине адаптивные алгоритмы, с помощью которых вычисляются ВК ААР, также реализуются в арифметике комплексных чисел.

Здесь и далее используются общепринятые векторные и матричные обозначения. Векторы обозначаются прямыми жирными строчными символами, а матрицы – прямыми жирными заглавными символами. Нижний индекс M в векторе обозначает число его элементов, а нижний индекс M в квадратной матрице – число её элементов, равное $M \times M$. Верхний индекс T обозначает операцию транспонирования вектора, а верхний индекс H – операцию эрмитово сопряжения вектора. Скалярные переменные, в том числе элементы векторов и матриц, обозначаются наклонными символами (соответственно строчными и заглавными), а операция их комплексного сопряжения обозначается верхним индексом $*$.

Если на рисунке 1 удалить элементы, обозначенные розовым цветом, то оставшаяся часть будет соответствовать компенсатору помех. Если же удалить элементы, обозначенные голубым цветом, то оставшаяся часть будет соответствовать LC ААР. Особенности этих двух архитектур заключаются в следующем.

SLC представляет собой устройство, в котором с помощью дополнительных антенн компенсируются сигналы помех, источники которых расположены в направлениях боковых лепестков отдельной направленной антенны. В компенсаторе же помех качестве такой антенны может быть использована одна из слабонаправленных антенн ААР. Остальные такие же антенны ААР используются как компенсационные. Значение ВК $h_m(k)$ канала ААР с выделенной антенной не изменяется во времени, так как этот ВК не вычисляется с помощью адаптивного алгоритма, то есть всегда $h_m(k) = h_m = 1$.

Так как принимаемые со спутников сигналы, как уже отмечалось, скрыты шумами радиоприёмников ААР, то их наличие и число не влияют на работу компенсатора помех. Это позволяет в компенсаторе помех использовать сигнал $x_m(k)$ канала с выделенной антенной в качестве требуемого сигнала $d(k)$ адаптивного алгоритма. В качестве алгоритма, вычисляющего ВК компенсатора помех, можно использовать любой адаптивный алгоритм, минимизирующий квадратичную функцию ошибки

$$\alpha(k) = d(k) - \mathbf{h}_{M-1}^H(k-1) \mathbf{x}_{M-1}(k) = x_m(k) - \mathbf{h}_{M-1}^H(k-1) \mathbf{x}_{M-1}(k).$$

Сигнал $\alpha(k)$ совпадает с выходным сигналом компенсатора помех $y(k)$, то есть в процессе вычисления ВК минимизируются все сигналы, принимаемые ААР, за исключением сигналов со спутников, так как они скрыты шумами, а потому на адаптивный процесс не влияют. Вектор ВК компенсатора помех $\mathbf{h}_{M-1}(k)$ содержит $M-1$ элементов. При его вычислении используется вектор $\mathbf{x}_{M-1}(k)$, содержащий все элементы вектора $\mathbf{x}_M(k)$, за исключением элемента $x_M(k)$.

Если число источников помех не превышает значения $M-1$, то, как и любая ААР [7], компенсатор помех может подавить эти помехи за счет формирования провалов в ДН в направлениях на их источники. Однако компенсатор помех не гарантирует никакого наперед заданного значения ДН в направлении на источник полезного сигнала (спутник), так как его ДН контролируется только в направлениях на источники помех, то есть в такой ДН не формируется основной луч в каком-либо наперед заданном направлении. В то же время, для решения навигационной задачи требуется одновременно принимать сигналы, как минимум, от четырех спутников, то есть из четырех направлений. Такой приём возможен и с помощью компенсатора помех, если уровни его ДН с учетом ДН антенн в направлениях на спутники таковы, что в установившемся режиме на его выходе для каждого из спутниковых сигналов обеспечивается ОСШП ≥ -20 дБ [17].

Это ОСШП, в первую очередь, обеспечивается для спутников, которые расположены в окрестности зенита. При прочих равных условиях принимаемые сигналы от этих спутников являются самыми мощными в силу кратчайшего расстояния от таких спутников до Земли. Обеспечить принудительную ориентацию луча ААР (основного лепестка ДН) можно за счет линейного ограничения на значение ДН в интересующем направлении, в данном случае в направлении зенита. Такое ограничение задается аналитически в алгоритме вычисления ВК LC ААР. Оно обеспечивается на каждой итерации этого алгоритма независимо от того, какие её ВК вычисляются в процессе подавления помех.

В LC ААР требуемый сигнал $d(k) = 0$, а потому сигнал ошибки $\alpha(k)$, минимизируемой в процессе адаптации, совпадает с выходным сигналом ААР

$$\alpha(k) = d(k) - y(k) = -\mathbf{h}_M^H(k-1)\mathbf{x}_M(k).$$

Линейное ограничение задается как

$$\mathbf{c}_M^H \mathbf{h}_M(k) = |F_{\text{ант}}(\theta_c, \phi_c)| M = f,$$

где

$$\mathbf{c}_M = \left[F_{\text{ант}}^{(1)}(\theta_c, \phi_c) e^{j\psi_1}, F_{\text{ант}}^{(2)}(\theta_c, \phi_c) e^{j\psi_2}, \dots, F_{\text{ант}}^{(M)}(\theta_c, \phi_c) e^{j\psi_M}, \dots, F_{\text{ант}}^{(M)}(\theta_c, \phi_c) e^{j\psi_M} \right]^T,$$

$F_{\text{ант}}^{(m)}(\theta_c, \phi_c)$ – это комплексные значения ДН антенн ААР в направлении ограничения, определяемого сферическими углами θ_c и ϕ_c , а ψ_m – пространственные набеги фаз относительно фазового центра ААР от источника сигнала, расположенного в дальней зоне в этом же направлении, до каждой из её антенн [1, 2]. В LC ААР вектор отсчетов входных сигналов $\mathbf{x}_M(k)$ и вектор ВК $\mathbf{h}_M(k)$ содержат по M элементов.

Ниже представлены вычислительные процедуры компенсатора помех и LC ААР для K итераций.

Вычислительная процедура цифрового компенсатора помех

Инициализация: $\mathbf{x}_M(k) = \mathbf{0}_M$, $\mathbf{h}_{M-1}(0) = \mathbf{c}_{M-1}$, $h_m(0) = 1$,

$$\bar{\mathbf{R}}_{M-1}^{-H}(0) = \delta^{-1} \mathbf{I}_{M-1}, \mathbf{K}_{M-1}^R(0) = \mathbf{I}_{M-1}$$

For $k = 1, 2, \dots, K$

$$\mathbf{x}_M(k) = [x_1(k), x_2(k), \dots, x_m(k), \dots, x_M(k)]^T$$

$$y(k) = \alpha(k) = x_m(k) - \mathbf{h}_{M-1}^H(k-1)\mathbf{x}_{M-1}(k)$$

$$\bar{u}_{M-1,j}^{(j-1)*}(k) = 0, \quad 1 \leq j \leq M-1, \quad K_{M-1}^{B(0)}(k) = 1$$

For $i = 1, 2, \dots, M-1$

$$\bar{a}_{M-1,i}(k) = \bar{\mathbf{R}}_{M-1}^{-H}(k-1) \left| \begin{matrix} \mathbf{x}_{M-1}(k) \\ i, 1:i \end{matrix} \right|_{1:i}$$

$$K_{M-1}^{B(i)}(k) = K_{M-1}^{B(i-1)}(k) + \lambda^{-1} K_{M-1,i}^R(k-1) \bar{a}_{M-1,i}^*(k) \bar{a}_{M-1,i}(k)$$

$$\bar{s}_{M-1,i}(k) = \lambda^{-1} K_{M-1,i}^R(k-1) \bar{a}_{M-1,i}^*(k) / K_{M-1}^{B(i)}(k)$$

$$\bar{c}_{M-1,i}(k) = K_{M-1}^{B(i-1)}(k) / K_{M-1}^{B(i)}(k)$$

For $j = 1, 2, \dots, i$

$$\bar{R}_{M-1,ij}^{-H}(k) = \bar{R}_{M-1,ij}^{-H}(k-1) - \bar{a}_{M-1,i}(k) \bar{u}_{M-1,j}^{(i-1)*}(k)$$

$$\bar{u}_{M-1,j}^{(i)*}(k) = \bar{s}_{M-1,i}(k) \bar{R}_{M-1,ij}^{-H}(k-1) + \bar{c}_{M-1,i}(k) \bar{u}_{M-1,j}^{(i-1)*}(k)$$

End for j

$$K_{M-1,i}^R(k) = \lambda^{-1} K_{M-1,i}^R(k-1) \bar{c}_{M-1,i}(k)$$

End for i

$$\mathbf{g}_{M-1}(k) = \bar{\mathbf{u}}_{M-1}^{(M-1)}(k)$$

$$\mathbf{h}_{M-1}(k) = \mathbf{h}_{M-1}(k-1) + \mathbf{g}_{M-1}(k) \alpha^*(k)$$

End

Вычислительная процедура цифровой LC ААР

Инициализация: $\mathbf{0}_M$, $\mathbf{q}_M(0) = \mathbf{c}_M [\mathbf{c}_M^H \mathbf{c}_M]^{-1}$,

$$\mathbf{h}_M(0) = \mathbf{q}_M(0) f, \bar{\mathbf{R}}_M^{-H}(0) = \delta^{-1} \mathbf{I}_M, \mathbf{K}_M^R(0) = \mathbf{I}_M$$

For $k = 1, 2, \dots, K$

$$\mathbf{x}_M(k) = [x_1(k), x_2(k), \dots, x_m(k), \dots, x_M(k)]^T$$

$$y(k) = \mathbf{h}_M^H(k-1)\mathbf{x}_M(k)$$

$$\bar{u}_{M,j}^{(j-1)*}(k) = 0, \quad 1 \leq j \leq M, \quad K_M^{B(0)}(k) = 1$$

For $i = 1, 2, \dots, M$

$$\bar{a}_{M,i}(k) = \bar{\mathbf{R}}_M^{-H}(k-1) \left| \begin{matrix} \mathbf{x}_M(k) \\ i, 1:i \end{matrix} \right|_{1:i}$$

$$K_M^{B(i)}(k) = K_M^{B(i-1)}(k) + \lambda^{-1} K_{M,i}^R(k-1) \bar{a}_{M,i}^*(k) \bar{a}_{M,i}(k)$$

$$\bar{s}_{M,i}(k) = \lambda^{-1} K_{M,i}^R(k-1) \bar{a}_{M,i}^*(k) / K_M^{B(i)}(k)$$

$$\bar{c}_{M,i}(k) = K_M^{B(i-1)}(k) / K_M^{B(i)}(k)$$

For $j = 1, 2, \dots, i$

$$\bar{R}_{M,ij}^{-H}(k) = \bar{R}_{M,ij}^{-H}(k-1) - \bar{a}_{M,i}(k) \bar{u}_{M,j}^{(i-1)*}(k)$$

$$\bar{u}_{M,j}^{(i)*}(k) = \bar{s}_{M,i}(k) \bar{R}_{M,j}^{-H}(k-1) + \bar{c}_{M,i}(k) \bar{u}_{M,j}^{(i-1)*}(k)$$

End for j

$$K_{M,i}^R(k) = \lambda^{-1} K_{M,i}^R(k-1) \bar{c}_{M,i}(k)$$

End for i

$$\mathbf{g}_M(k) = \bar{\mathbf{u}}_M^{(M)}(k)$$

$$\mathbf{v}(k) = \mathbf{c}_M^H \mathbf{g}_M(k)$$

$$\mathbf{v}^*(k) = \mathbf{x}_M^H(k) \mathbf{q}_M(k-1)$$

$$\mathbf{q}'_M(k) = \left[\mathbf{q}_M(k-1) - \mathbf{g}_M(k) \mathbf{v}^*(k) \right] \left[1 + \frac{\mathbf{v}(k) \mathbf{v}^*(k)}{1 - \mathbf{v}^* \mathbf{v}(k)} \right]$$

$$\mathbf{q}'_M(k) = \mathbf{q}'_M(k) + \mathbf{q}_M(0) \left[1 - \mathbf{c}_M^H \mathbf{q}'_M(k) \right]$$

$$\alpha(k) = -\gamma(k)$$

$$\mathbf{h}'_M(k) = \mathbf{h}_M(k-1) + \mathbf{g}_M(k) \alpha^*(k)$$

$$\mathbf{h}_M(k) = \mathbf{h}'_M(k) + \mathbf{q}_M(k) \left[f - \mathbf{c}_M^H \mathbf{h}'_M(k) \right]$$

End

В данных ААР используются вычисления векторов коэффициентов Калмана $\mathbf{g}_{M-1}(k)$ и $\mathbf{g}_M(k)$ с помощью RLS-алгоритма на базе QR-разложения без операций извлечения квадратного корня. В LC ААР используется LC версия этого алгоритма. Детали таких алгоритмов приведены в [5].

Используя вычислительные процедуры рассматриваемых ААР, было проведено их сравнительное моделирование, результаты которого представлены в следующем разделе.

Моделирование SLC и LC ААР

В моделировании рассматривалась прямоугольная АР с числом антенн (каналов и ВК), равным $M = 2 \times 4 = 8$. Антенны ААР были одинаковые, но их ДН $F_{\text{ант}}^{(m)}(\theta, \phi)$ слегка различались, что обусловлено взаимной связью между антеннами. В моделировании использовались таблицы значений ДН антенн, полученные с помощью средств электродинамического проектирования. Каждая из таблиц – это набор комплексных чисел, представляющих собою значения ненормированной ДН, рассчитанной в диапазоне сферических углов $\theta = 0^\circ \dots 180^\circ$ и $\phi = 0^\circ \dots 360^\circ$ с одинаковым дискретом 1° . Ширина ДН антенн по уровню -3 дБ составляла примерно 80° , а их максимальное значение составляло $|F_{\text{ант}}(\theta, \phi)|_{\text{max}} \approx 8.2$ дБ. ААР располагалась горизонтально, её луч в исходном состоянии ориентировался в направлении зенита $\theta_0 = 0^\circ$, $\phi_0 = 0^\circ$. Это направление совпадало с направлением на источник полезного сигнала и с направлением ограничения ДН θ_c , ϕ_c в LC ААР. Антенны располагались так, что их ДН также были ориентированы в направлении зенита. Использовалась ААР с эквидистантным размещением антенн на плоскости, с расстоянием между соседними антеннами, равным $\lambda_0 / 2$, где λ_0 – это длина волны несущего колебания ГНСС GPS.

Полезный сигнал моделировался одной из ПСП Голда, используемых в ГНСС GPS [17]. Сигналы семи источников помех моделировались независимыми случайными последовательностями белого гауссового шума. Расположение источников этих помех в каждом из экспериментов было случайным в

диапазоне углов $\theta = 90^\circ \dots 180^\circ$ и $\phi = 0^\circ \dots 360^\circ$, то есть при их расположении в нижней полусфере (под ААР). Такое расположение источников помех соответствует наиболее часто встречающимся на практике случаям. Мощности всех источников помех предполагались одинаковыми, с отношением сигнал/помеха, равным -70 дБ.

Все сигналы моделировались аналогично [22]. Дискретный сигнал комплексной огибающей каждого из каналов ААР определялся как

$$x_m(k) = \sum_{l=1}^L \tilde{x}_m^{(l)}(k) = \sum_{l=1}^L F_{\text{ант}}^{(m)}(\theta^{(l)}, \phi^{(l)}) A^{(l)}(k) e^{j\psi_m^{(l)}},$$

где $F_{\text{ант}}^{(m)}(\theta^{(l)}, \phi^{(l)})$ – это комплексное значение ДН m -й антенны в направлении l -го источника сигнала, $A^{(l)}(k)$ – комплексная огибающая этого сигнала, $\psi_m^{(l)}$ – пространственный набег фазы от l -го источника сигнала до m -й антенны относительно фазового центра ААР. Сигнал $\tilde{x}_m^{(l)}(k)$ в этой сумме – это полезный сигнал, а сигналы $\tilde{x}_m^{(l)}(k)$, $l = 2, \dots, L$ – помехи, где $L \leq M$. Шумы радиоприёмников в каналах ААР моделировались независимыми последовательностями отсчетов белого гауссового шума с

$$\text{ОСШ} = \sigma_{\tilde{x}^{(1)}}^2 / \sigma_{\text{ш}}^2 = -20 \text{ дБ},$$

где $\sigma_{\tilde{x}^{(1)}}^2$ – это дисперсия полезного сигнала, а $\sigma_{\text{ш}}^2$ – дисперсия шумов в радиоприёмных каналах ААР. Результаты моделирования сравниваемых ААР приведены на рисунках ниже.

На рисунке 2 показана исходная 3D ДН ААР плоскости. Эта ДН одинаковая для обеих ААР. Такое отображение отдельно в верхней и отдельно в нижней полусферах, окружающих ААР, является удобным при анализе ДН антенн наземной аппаратуры ГНСС, так как источники принимаемых сигналов (спутники) всегда располагаются в верхней полусфере, а источники помех – в основном в нижней полусфере. Здесь угол $\phi = 0^\circ \dots 360^\circ$ отсчитывается вдоль окружности, угол $\theta = 0^\circ \dots 90^\circ$ – от центра окружности до её края в ДН для верхней полусферы, и угол $\theta = 90^\circ \dots 180^\circ$ – от края окружности до её центра в ДН для нижней полусферы. Значения ДН обозначены цветом в соответствии с цветовой палитрой. Угловое расположение и номера источников принимаемых сигналов на рис. 2 обозначены цветными крестиками и цифрами: черным цветом для источников полезных сигналов и фиолетовым – для источников помех. Здесь приведен пример одного источника полезного сигнала, расположенного в направлении зенита (в верхней полусфере) и семи случайно расположенных источников помех в нижней полусфере. ААР с числом антенн $M = 8$, может подавить до $M - 1 = 7$ помех [7].

В исследованиях, результаты которых приведены в настоящей работе, ААР моделировались для десяти случаев случайного расположения источников помех в пределах, ранее обозначенного диапазона углов θ и ϕ . Один из таких примеров показан на рис. 2б.

В каждом из десяти экспериментов, число источников помех изменялось от одного до семи, причем их угловое положение соответствовало положению источников помех при максимальном их числе. Например, 1; 1, 2; 1, 2, 3; ... и 1, 2, 3, 4, 5, 6, 7.

Примеры 3D ДН в установившемся режиме приведены на рисунках 3 и 4.

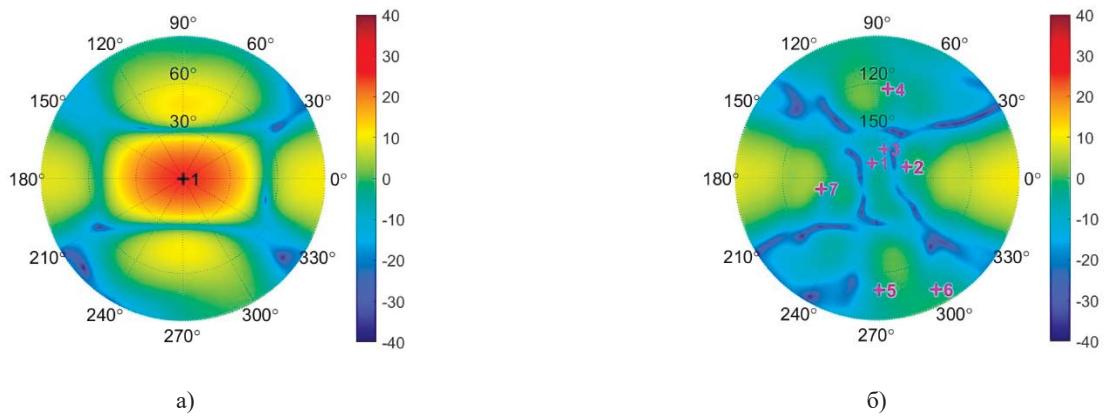


Рис. 2. ДН ААР в исходном состоянии: а) – в верхней полусфере; б) – в нижней полусфере

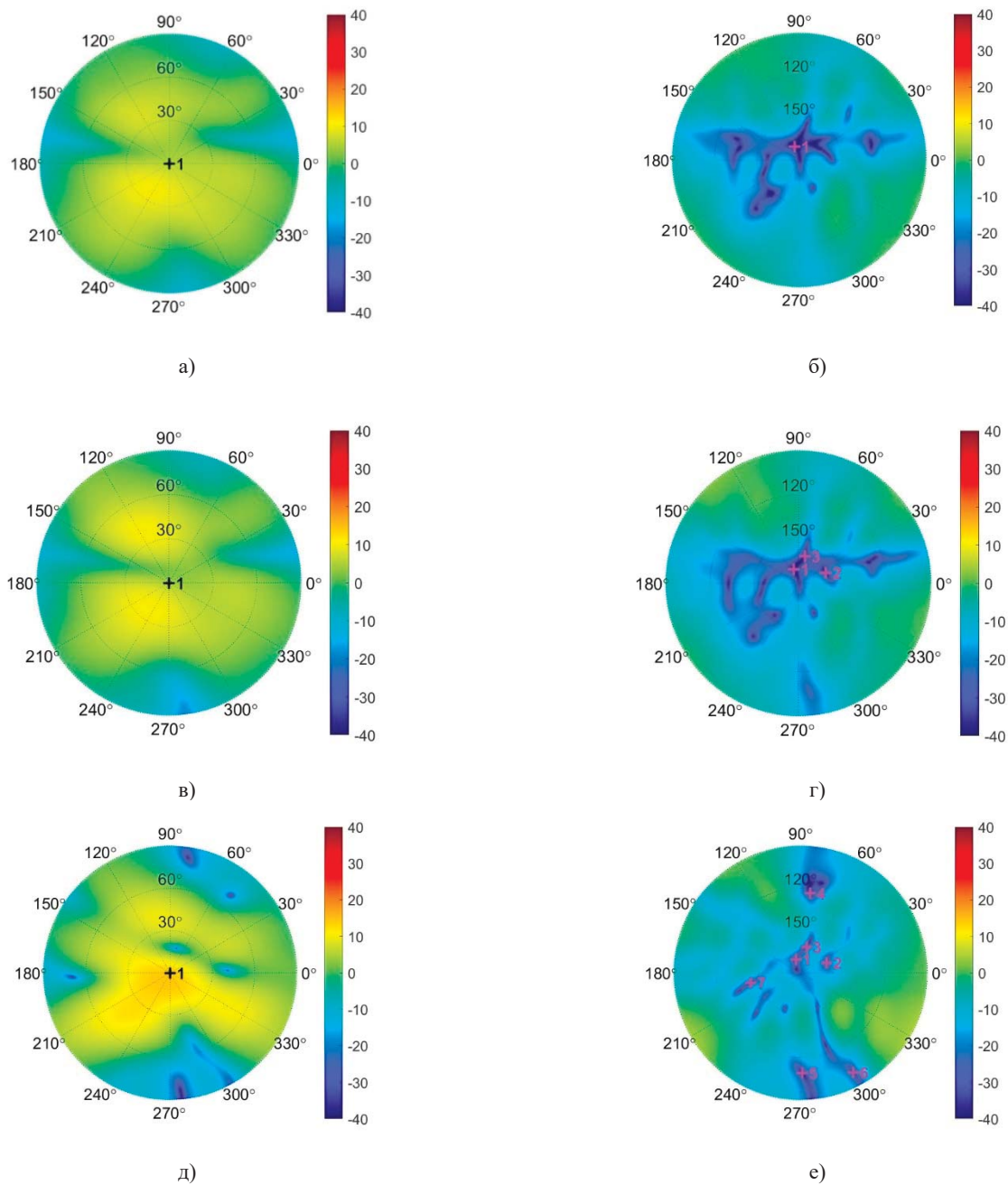


Рис. 3. ДН компенсатора помех в исходном состоянии: слева – в верхней полусфере; справа – в нижней полусфере: а), б) – одна помеха; в), г) – три помехи; д), е) – семь помех

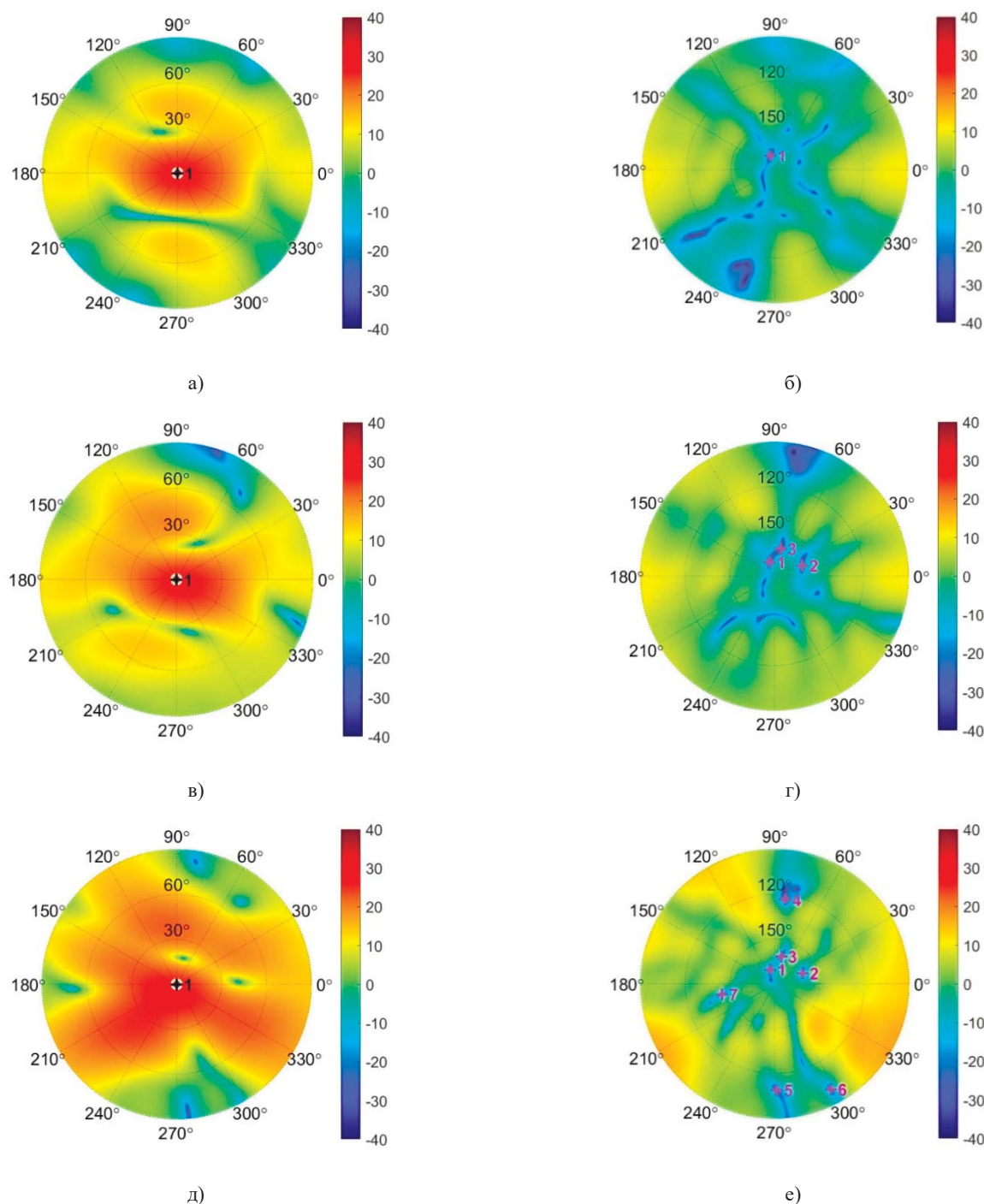


Рис. 4. ДН LC ААР в исходном состоянии: слева – в верхней полусфере; справа – в нижней полусфере; а), б) – одна помеха; в), г) – три помехи; д), е) – семь помех

Белый кружок в направлении источника полезного сигнала на рис. 4 указывает на направление ограничения ДН ААР. Из этих рисунков видно, что в установившемся режиме обе рассматриваемые ААР подавляют все помехи, но их ДН различаются. Они разные в зависимости от вида ААР и от числа подавляемых помех. Видно, что в среднем значения ДН LC ААР во всех направлениях, кроме направлений на источники помех, больше, чем в компенсаторе помех. Это объясняется следующим образом. В компенсаторе помех, ДН контролируется

только в направлениях на источники помех, но не контролируется в каком-либо другом направлении, например, в направлении первоначального луча. Напротив, ДН LC ААР сохраняет свое значение в направлении ограничения независимо от значений ВК, вычисляемых в процессе подавления помех. Из рисунка 4 следует, что чем больше помех, тем значения ДН становятся «большими во всех направлениях» (это видно из Белый кружок в направлении источника полезного сигнала на рис. 4 указывает на направление ограничения ДН ААР.

Из этих рисунков видно, что в установившемся режиме обе рассматриваемые ААР подавляют все помехи, но их ДН различаются. Они разные в зависимости от вида ААР и от числа подавляемых помех. Видно, что в среднем значения ДН LC ААР во всех направлениях, кроме направлений на источники помех, больше, чем в компенсаторе помех. Это объясняется следующим образом. В компенсаторе помех, ДН контролируется только в направлениях на источники помех, но не контролируется в каком-либо другом направлении, например, в направлении первоначального луча. Напротив, ДН LC ААР

сохраняет свое значение в направлении ограничения независимо от значений ВК, вычисляемых в процессе подавления помех. Из рисунка 4 следует, что чем больше помех, тем значения ДН становятся «большими во всех направлениях» (это видно из цветного обозначения значений ДН), что обусловлено увеличением значений вычисляемых ВК ААР.

Изменения ДН ААР в направлениях каждого из источников принимаемых сигналов в процессе адаптации показаны на рисунке 5.

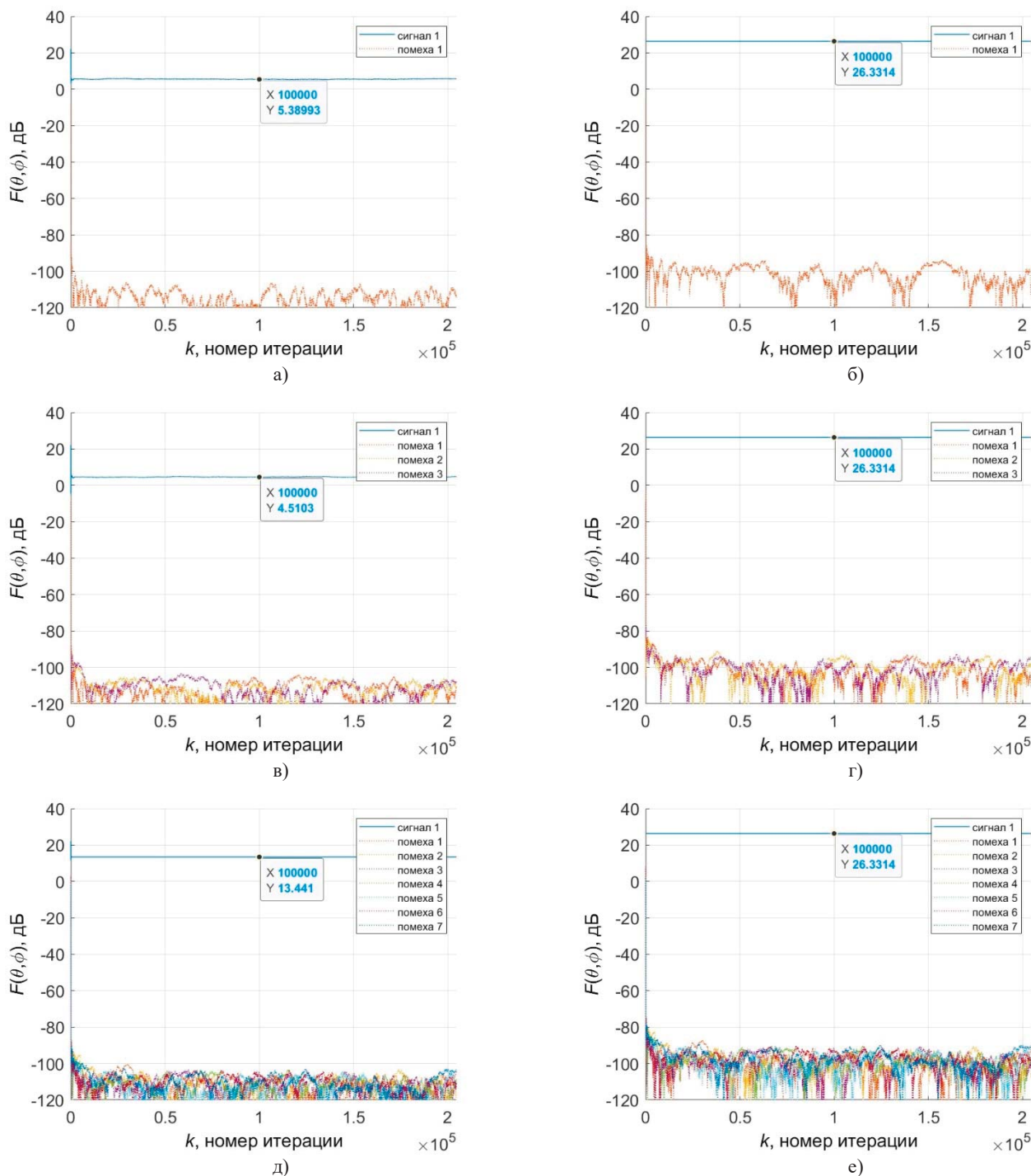


Рис. 5. Переходные процессы в терминах значений ДН ААР в направлениях на все источники принимаемых сигналов: слева – компенсатор помех; справа – LC ААР; а), б) – одна помеха; в), г) – три помехи; д), е) – семь помех

Видно, что уровень ДН компенсатора помех в направлении на источник полезного сигнала зависит от числа источников подавляемых помех, в то время как в LC ААР он в этом же направлении постоянный и равен значению ограничения

$$20 \log_{10} M + 20 \log_{10} |F(\theta, \phi)|_{\max} = 18.1 + 8.2 \approx 26.3 \text{ дБ.}$$

Оценить возможность приёма информационного сигнала после подавления помех можно с помощью графиков взаимокорреляционной функций (ВКФ) ПСП этого сигнала и выходного сигнала ААР. Графики этих функций показаны рис. 6 в увеличенном масштабе на начальном участке адаптации.

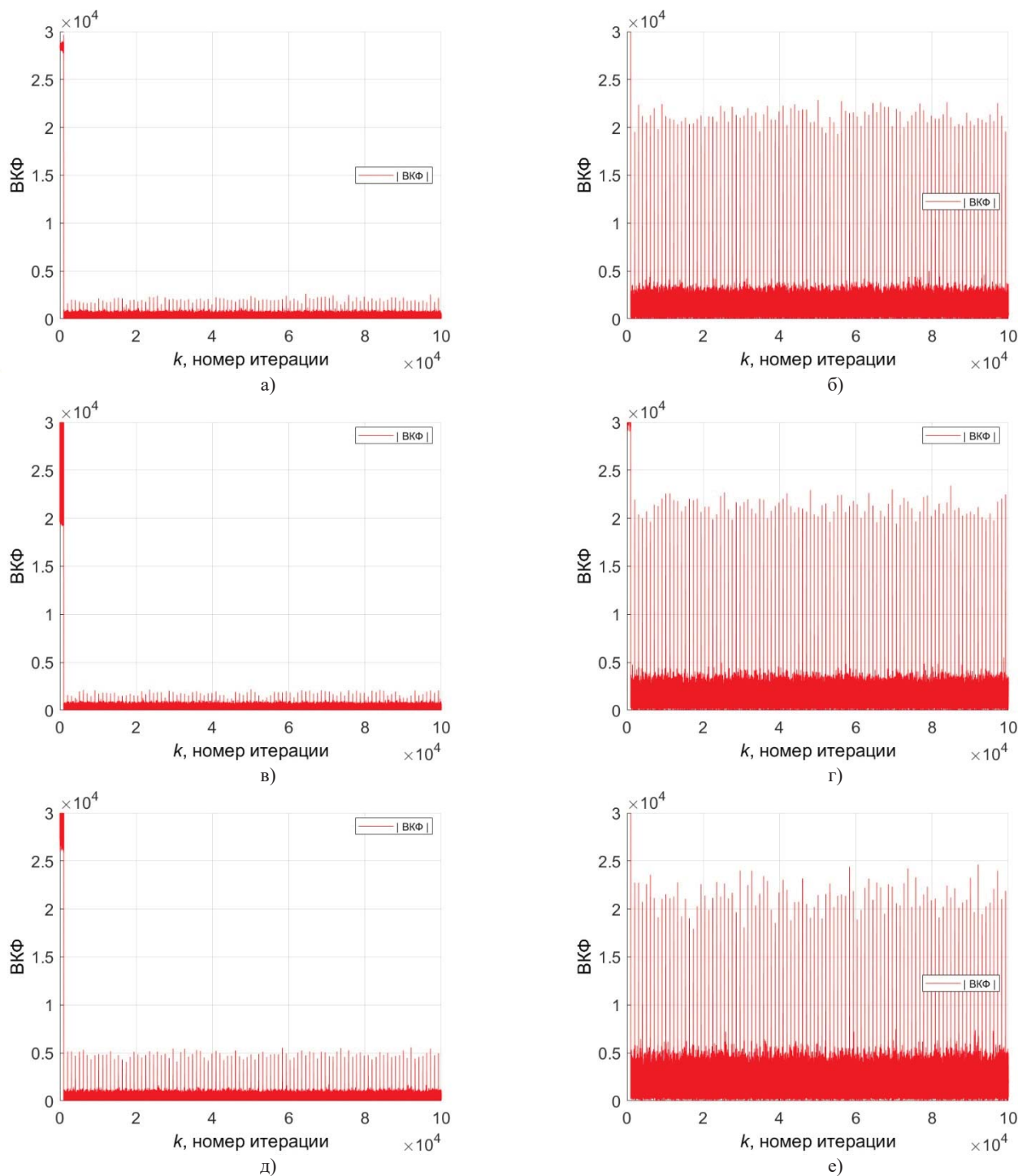


Рис. 6. ВКФ сигнала на выходе ААР и ПСП сигнала со спутника: слева – компенсатор помех: справа – LC ААР; а), б) – одна помеха; в), г) – три помехи; д), е) – семь помех

Можно заметить, что пики ВКФ хорошо видны, а значит, обе ААР принимают сигналы спутника в рассматриваемых помеховых обстановках при расположении спутника в направлении зенита. Одновременно видно, что в случае LC ААР происходит не только усиление полезного сигнала за счет увеличения значений ВК, но и усиление шума на выходе

ААР, что может приводить к ухудшению ОСШ спутниковых сигналов, принимаемых из других направлений.

Оценить возможность такого приёма можно с помощью карт ОСШП [23], примеры которых приведены на рис. 7 для рассматриваемых случаев расположения помех.

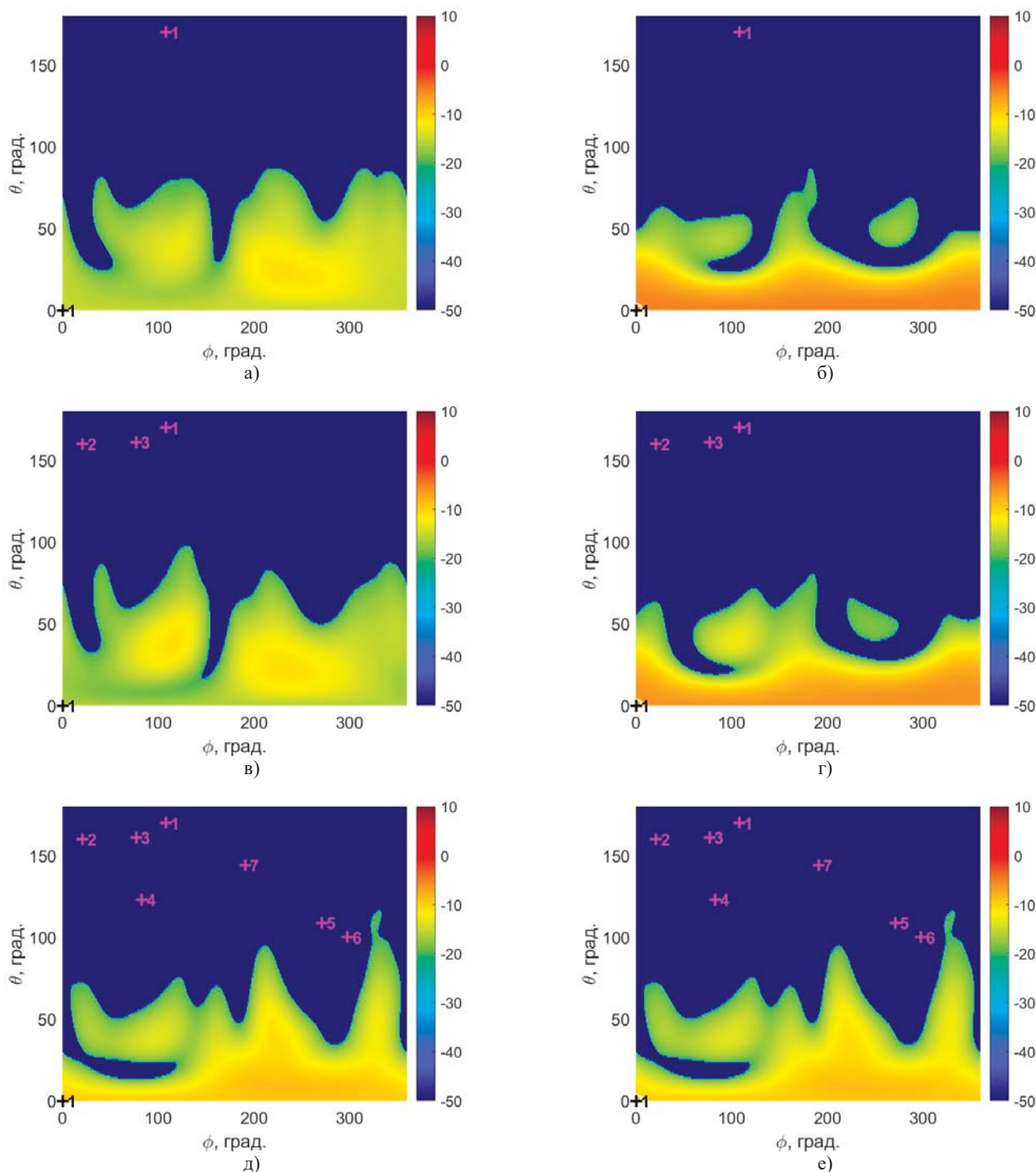


Рис. 7. Карты ОСШП: слева – компенсатор помех; справа – LC ААР:
а), б) – одна помеха; в), г) – три помехи; д), е) – семь помех

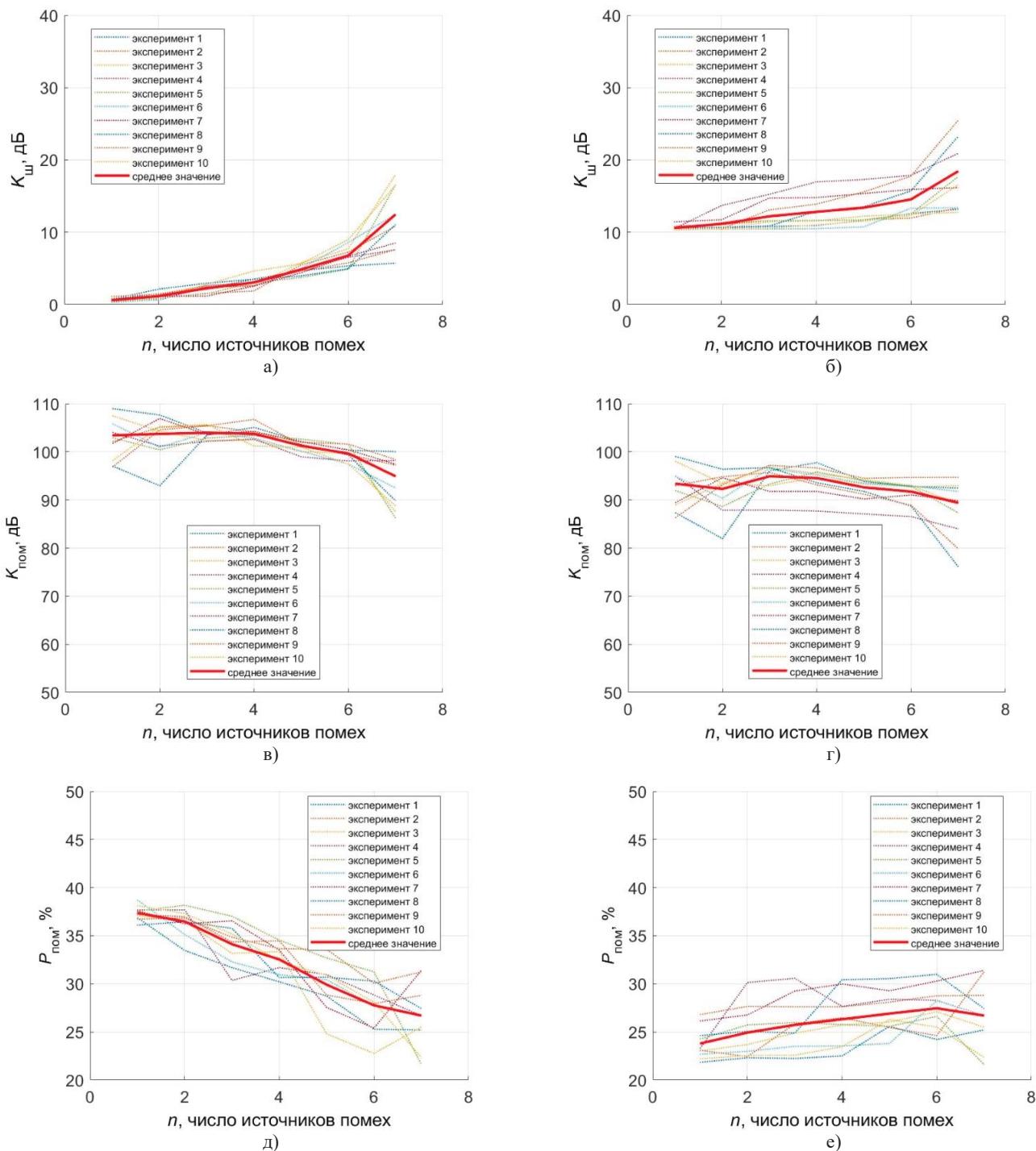


Рис. 8. Результаты статистического моделирования: слева – компенсатор помех; справа – LC AAR:
а), б) – коэффициент шума $K_{ш}$; в), г) – коэффициент подавления помех $K_{п}$; д), е) – параметр P

Эти карты строятся аналогично 3D ДН в диапазоне углов $\theta = 0^\circ \dots 180^\circ$ и $\phi = 0^\circ \dots 360^\circ$. Они отображают области приёма полезного сигнала с направлений, где в установившемся режиме на выходе ААР обеспечивается $\text{ОСШП} \geq -20$ дБ. Такие области обозначены разными цветами в соответствии с цветовой палитрой. Сплошным одинаковым темно-синим цве-

том обозначена область, где приём спутниковых сигналов невозможен, то есть, где на выходе ААР $\text{ОСШП} < -20$ дБ независимо от его конкретного значения.

На каждом из графиков рисунка 7, отношение площади графика $S_{\text{ОСШП} \geq -20 \text{ дБ}}$, где $\text{ОСШП} \geq -20$ дБ, ко всей его площади S , выраженное в процентах, то есть $P = \frac{S_{\text{ОСШП} \geq -20 \text{ дБ}}}{S} 100\%$,

является оценкой возможности приёма спутниковых сигналов в рассматриваемой помеховой обстановке в установившемся режиме ААР (при полученных значениях её ВК) для критерия приёма $ОСШП \geq -20$ дБ [23].

Из рисунка 7 видно, что приём полезного информационного сигнала обеспечивается в большей части направлений, образующих верхнюю полусферу. Области приёма различаются при числе помех $M-1 < 7$, и практически не различаются при числе помех $M-1 = 7$, для рассматриваемых ААР. Также видно, что при числе помех $M-1 < 7$, компенсатор помех обеспечивает приём спутниковых сигналов из большего числа направлений чем обеспечивает LC ААР. Объяснение этих результатов и выводов получено с помощью статистического моделирования процессов адаптации обеих ААР, которое было описано в начале настоящего раздела.

Результаты этого моделирования приведены на рисунке 8. Здесь тонкими кривыми показаны результаты каждого из 10 экспериментов, а их средние значения – жирными красными кривыми. Численные значения усредненных кривых на рисунке 8 приведены в таблице 1.

Таблица 1

Численные значения средних значений рисунка 8

Параметр	Сценарий: число помех						
	1	2	3	4	5	6	7
$K_{ш}$, дБ, SCL	0.6	1.2	2.2	3.1	4.8	6.8	12.5
$K_{ш}$, дБ, LC ААР	10.6	11.2	12.2	12.8	13.4	14.5	18.4
$K_{ш}$, дБ, SCL	103.4	103.7	104.0	103.8	101.3	99.7	94.9
$K_{ш}$, дБ, LC ААР	93.4	92.3	94.9	94.6	92.6	91.7	89.4
P , %, SCL	37.4	36.5	34.1	32.5	29.9	27.8	26.7
P , %, LC ААР	23.8	24.9	25.7	26.3	26.9	27.5	26.7

Здесь,

$$K_{ш} = 10 \log_{10} \left(\sum_{m=1}^M |h_m|^2 \sigma_{ш}^2 / \sigma_{ш}^2 \right) = 10 \log_{10} \left(\sum_{m=1}^M |h_m|^2 \right)$$

– это коэффициент усиления шума ААР, h_m – ВК ААР в установившемся режиме.

Коэффициент подавления помех определяется как,

$$K_{п} = 10 \log_{10} (p_{вх} / p_{вых}),$$

где $p_{вх}$ – это суммарная мощность всех помех на входе ААР (на входе антенн ААР), а $p_{вых}$ – суммарная мощность всех помех на выходе ААР.

Рисунки 8а и 8б показывают, что ранее отмеченное усиление шума имеет место быть в обеих рассматриваемых ААР, оно растёт с увеличением числа подавляемых помех, но в LC ААР оно заметно больше, чем в компенсаторе помех.

Так как аддитивный шум, добавляемый к сигналам $x_m(k)$, участвующим в вычислении ВК ААР (на рис. 1 и в вычислительных процедурах ААР он не показан) смещает рассчитываемые ВК от их оптимальных значений, то и коэффициент подавления помех $K_{п}$ LC ААР в среднем на 5...10 дБ меньше, чем в компенсаторе помех, согласно рисункам 8в, 8г и таблице 1. Аналогичный результат сказывается и на значениях показателя

P . Лишь только при $M-1 = 7$ значения $K_{п}$ и P примерно одинаковые для обеих ААР, что следует рис. 7 и 8, а также было отмечено в [19] для показателя P .

Следовательно, согласно приведенным результатам, компенсатор помех демонстрирует в среднем лучшие характеристики по сравнению с LC ААР в одинаковых условиях функционирования при использовании этих ААР в качестве антенн наземных приёмников ГНСС.

Заключение

Таким образом, в работе рассмотрены особенности архитектур и вычислительные процедуры двух разновидностей ААР: в виде компенсатора помех и с линейным ограничением ДН, – для использования в качестве антенн наземных радиоприёмников ГНСС. Путем моделирования установлено, что в среднем компенсатор помех демонстрирует лучшие показатели качества в терминах меньшего коэффициента усиления шума, большего коэффициента подавления помех, а также большей доли направлений, в которых в установившемся режиме возможен приём спутниковых сигналов относительно всех направлений, окружающих ААР. Эти характеристики заметно различаются в рассматриваемых ААР, если число помех меньше значения $M-1$. Данные ААР могут быть также использованы в радиосистемах связи, радиолокации, радиоуправления и в ряде других радиосистем, где принимаемые радиосигналы скрыты шумами радиоприёмников.

Работа проведена в рамках выполнения государственного задания НИЦ «Курчатовский институт».

Литература

1. Maillou R. J. Phased array antenna handbook, 3rd ed. Artech House, Inc., 2017. 506 p.
2. Xiao Z., Zu L., Bai L., Xia X.-G. Array beamforming enabled wireless communications. CRC Press, 2023. 438 p.
3. Vaidyanathan P. P. Signals, systems, and signal processing. Cambridge University Press, 2024. 726 p.
4. Умняшкин С. В. Основы цифровой обработки сигналов: учебное пособие. 8-е изд. М.: Техносфера, 2025. 550 с.
5. Джиган В. И. Адаптивная фильтрация сигналов: теория и алгоритмы. М.: Техносфера, 2013. 528 с.
6. Sarvagya M., Kumar R. V. G. Adaptive filtering – recent advances and practical implementation. LAP Lambert Academic Publishing, 2024. 76 p.
7. Hudson J. E. Adaptive array principles. The Institution of Engineering and Technology, 2007. 253 p.
8. Ратынский М. В. Адаптация и сверхразрешение в антенных решетках. М.: Ленард, 2024. 240 с.
9. Косичкина Т. П., Сперанский В. С. Цифровые сигнальные процессоры и их применение в системах телекоммуникаций и электроники. М.: Горячая линия-Телком, 2024. 315 с.
10. Basak B., Sheet D. K., Santra M. Reconfigurable DSP architectures: FPGA and embedded system approaches. Essay Publication Research and Consultancy, 2025. 140 p.
11. Джиган В. И. Активное шумоподавление: архитектуры, алгоритмы и их параметры // Т-Comm: Телекоммуникации и транспорт. 2025. № 8. С. 3-11.
12. Джиган В. И. Сравнительная эффективность цифровых преобразователей сигналов на базе адаптивных рекурсивных алгоритмов по критерию наименьших квадратов // Т-Comm: Телекоммуникации и транспорт. 2025. № 11. С. 4-14.

13. Джиган В. И. Линейно-ограниченная адаптивная антенная решетка в арифметике действительных чисел для радиоприемников наземной аппаратуры потребителя глобальных навигационных спутниковых систем // Т-Comm: Телекоммуникации и транспорт. 2026. № 3. С. 4-14.
14. Howells P. W. Intermediate frequency side-lobe canceller // US Pattern 3202990. 1959. 8 p.
15. Frost O. L. An algorithm for linearly constrained adaptive array processing // Proceedings of the IEEE. 1972. Vol. 60. No. 8. P. 926-935.
16. Джиган В. И. Цифровая адаптивная антенная решетка для приема полезных сигналов, находящихся под тепловыми шумами // Цифровая обработка сигналов. 2024. № 2. С. 3-10.
17. Hofmann-Wellenhof B., Lichtenegger H., Wasle E. GNSS – Global Navigation Satellite Systems: GPS, GLONASS, Galileo, and more. Springer, 2008. 516 p.
18. Zepernick H.-J., Finger A. Pseudo random signal processing: theory and applications. Wiley, 2013. 672 p.
19. Джиган В. И. Выбор адаптивного алгоритма для цифровой антенной решетки наземной аппаратуры систем глобальной спутниковой навигации // Цифровая обработка сигналов. 2025. № 2. С. 27-35.
20. Reyland J. M. Software defined radio: theory and practice. Artech House, 2023. 796 p.
21. Григорьев Л. Н. Цифровое формирование диаграммы направленности в фазированных антенных решетках. М.: Радиотехника, 2010. 144 с.
22. Плетнева И. Д., Джиган В. И. Моделирование обработки сигналов в цифровых антенных решетках // Исследования в области цифровых систем связи (Межвузовский сборник). – М.: Изд. МИЭТ, 2007. С. 36-43.
23. Джиган В. И., Шериунович А. Ю., Луферчик П. В., Штро П. В. Методика оценки вероятности приема сигналов при использовании адаптивного компенсатора помех в качестве антенны наземной аппаратуры глобальной навигационной спутниковой системы // Успехи современной радиоэлектроники. 2026. Т. 80. № 2. С. 8-19.

COMPARATIVE ANALYSIS OF INTERFERENCE CANCELLER AND LINEARLY-CONSTRAINED ADAPTIVE ANTENNA ARRAY EFFICIENCY IN GROUND-BASED NAVIGATION RECEIVERS

Victor I. Djigan, National Research Center "Kurchatov Institute", National Research University of Electronic Technology, Moscow, Russia,
dzhigan_vi@nrcki.rum

Abstract

Various modern radio systems often operate in interference conditions caused by the radio systems sharing the same frequency band. Separation of the informational signal through the frequency filtering is ineffective in this case, and code separation of the signals is only possible if the informational signal has a special structure. A more universal method for combating the interferences is a spatial filtering when an Adaptive Antenna Array (AAA) is used as the radio system's antenna. The AAA suppresses the interferences by the creating notches in its radiation pattern in the directions of the interference sources. These notches are determined by AAA weights, calculated by adaptive algorithms in real time. This paper examines the architectures and computational procedures of two AAA types: an AAA in the form of an interference canceller and an AAA with linear constraints, both for the usage as the antennas for the ground-based radio receivers of the Global Navigation Satellite Systems (GNSS). The AAA computational procedures are based on the usage of the adaptive Recursive Least Squares filtering algorithms with QR decomposition without square root operations. The AAA simulations under identical operating conditions revealed that, on average, the interference canceller demonstrates a superior performance in terms of smaller noise gain at the AAA output, larger interference suppression ratio, and the larger portion of directions in which the satellite signals can be received in steady-state mode relative to all directions surrounding the AAA. These characteristics differ significantly in these AAAs under consideration if the number of the interference sources to be suppressed is less than the maximal their number that can be suppressed by the AAAs. The AAAs under consideration can be used not only in GNSS radio receivers but also in other radio systems, in which the operate with the radio signals that are obscured by noise.

Keywords: Adaptive antenna array, interference canceller, linear constraints, recursive least-squares algorithm, global navigation satellite systems, noise gain, interference suppression ratio, signal-to-noise ratio maps.

References

- [1] R. J. Maillou, Phased Array Antenna Handbook, 3rd ed. Artech House, Inc., 2017, 506 p.
- [2] Z. Xiao, L. Zu, L. Ba, and X.-G. Xia, Array Beamforming Enabled Wireless Communications. CRC Press, 2023. 438 p.
- [3] P. P. Vaidyanathan, Signals, Systems, and Signal Processing, Cambridge University Press, 2024, 726 p.
- [4] S. V. Umnyashkin, Fundamentals of Digital Signal Processing: Tutorial. 8th ed., Moscow, Russia, Tekhnosfera Publisher, 2025, 550 p.
- [5] V. I. Djigan, Adaptive Filtering: Theory and Algorithms, Moscow, Russia, Technosfera Publisher, 2013, 528 p. (In Russian).
- [6] M. Sarvagya and R. V. G. Kumar, Adaptive Filtering - Resent Advances and Practical Implementation, LAP Lambert Academic Publishing, 2024, 76 p. (In Russian).
- [7] J. E. Hudson, Adaptive Array Principles, The Institution of Engineering and Technology, 2007, 253 p.
- [8] M. V. Ratynski, Adaptation and Supperresolution in Antenna Arrays, Moscow, Russia, Lenard Publisher, 2024, 240 p. (In Russian).

- [9] T. P. Kosichkina and V. S. Speransky, *Digital Signal Processors and their Application in Telecommunication and Electronics Systems*, Moscow: Goryachaya Liniya-Telkom, 2024, 315 p. (In Russian).
- [10] B. Basak, D. K. Sheet, and M. Santra, *Reconfigurable DSP Architectures: FPGA and Embedded System Approaches*, Essay Publication Research and Consultancy, 2025, 140 p.
- [11] V. I. Djigan, "Active noise control: architectures, algorithms and its parameters", *T-Comm*, 2025, no. 8, pp. 3-11. (In Russian).
- [12] V. I. Djigan, "Comparative efficiency of signal digital predistorters based on adaptive Recursive Least Squares algorithms", *T-Comm*, 2025, no. 11, pp. 4-14. (In Russian).
- [13] V. I. Djigan, "Linearly-constrained adaptive antenna array for ground-based radio receivers of global navigation satellite system", *T-Comm*, 2026, vol. 20, no. 3, pp. 4-14. (In Russian).
- [14] P. W. Howells, "Intermediate frequency side-lobe canceller", *US Pattern* 3202990, 1959, 8 p.
- [15] O. L. Frost, "An algorithm for linearly constrained adaptive array processing", *Proceedings of the IEEE*, 1972, vol. 60, no. 8, pp. 926-935.
- [16] V. I. Djigan, "Digital adaptive antenna array for receiving useful signals under thermal noise", *Digital Signal Processing*, 2024, no. 2, pp. 3-10. (In Russian).
- [17] B. Hofmann-Wellenhof, H. Lichtenegger, and E. Wasle, *GNSS – Global Navigation Satellite Systems: GPS, GLONASS, Galileo, and More*, Springer, 2008, 516 p.
- [18] H.-J. Zepernick and A. Finger, *Pseudo Random Signal Processing: Theory and Applications*. Wiley, 2013, 672 p.
- [19] V. I. Djigan, "Selection of an adaptive algorithm for a digital antenna array of ground-based equipment for global satellite navigation systems," *Digital Signal Processing*, 2025, no. 2, pp. 27-35. (In Russian).
- [20] J. M. Reyland, *Software Defined Radio: Theory and Practice*, Artech House, 2023, 796 p.
- [21] L. N. Grigoriev, *Digital Beamforming in Phased Antenna Arrays*, Moscow, Russia, Radiotekhnika Publisher, 2010, 144 p. (In Russian).
- [22] I. D. Pletneva and V.I. Djigan, "Simulation of signal processing in digital antenna arrays", *Research in Field of Digital Communication Systems* (Interuniversity collection), Moscow, Russia, MIET Publisher, 2007, pp. 36-43. (In Russian).
- [23] V. I. Djigan, A. Yu. Shershunovich, P. V. Luferchik, and P. V. Shtro, "Methodology for assessing the probability of signal reception when using an adaptive interference canceller as an antenna for ground-based equipment of a global navigation satellite system," *Achievements of Modern Radioelectronics*, 2026, vol. 80, no. 2, p. 8-19. (In Russian).

Information about author:

Victor I. Djigan, Doctor of Sciences, Principal Researcher of Sub-Department for Chip Design Methodology of Department for Design Problems in Microelectronics of Prospective Microelectronics Center in National Research Center "Kurchatov Institute," Professor, Institute of Microdevices and Control Systems of National Research University of Electronic Technology, Moscow, Russia. ORCID 0000-0001-7485-1623

ПОВЫШЕНИЕ КОНТРАСТНОСТИ ПРИЗНАКОВ РАСПОЗНАВАНИЯ ИЗОБРАЖЕНИЙ НА ОСНОВЕ КРАТНОМАСШТАБНОЙ ОБРАБОТКИ ГИСТОГРАММ РАСПРЕДЕЛЕНИЯ ИХ ЯРКОСТИ

Дворников Сергей Викторович,
Санкт-Петербургский государственный университет
аэрокосмического приборостроения;
Военная академия связи, Санкт-Петербург, Россия

Бестугин Александр Роальдович,
Санкт-Петербургский государственный университет
аэрокосмического приборостроения, Санкт-Петербург, Россия

Киришина Ирина Анатольевна,
Санкт-Петербургский государственный университет
аэрокосмического приборостроения, Санкт-Петербург, Россия

Васильева Дина Владимировна,
Санкт-Петербургский государственный университет
аэрокосмического приборостроения, Санкт-Петербург, Россия,
dolli.dina@mail.ru

Васильев Александр Протальонович,
Московский технический университет связи и информатики,
Москва, Россия

DOI: 10.36724/2072-8735-2026-20-9-18-26

Manuscript received 15 June 2026;
Accepted 20 August 2026

Ключевые слова: аномалии изображений,
разливы нефтепродуктов, векторы признаков,
распознаваемые классы изображений,
гистограммы распределения яркости изображений

Особую актуальность проблема обнаружения аномалий на обрабатываемых изображениях, получаемых с аппаратуры визуального контроля, приобретает в ходе экологического мониторинга акваторий мирового океана при поиске результатов разливов нефтепродуктов. Под аномалиями в представленном исследовании определены фрагменты изображения, характеризующие последствия разливов нефтепродуктов. В качестве научного аппарата исследования определены элементы теории распознавания образов, в соответствии с которыми обрабатываемые изображения представляются в виде векторов признаков. Цель работы – рассмотреть возможные подходы к вейвлет-преобразованиям первичных векторов признаков при помощи декоррелирующей обработки для выбора вейвлета, позволяющего обеспечить наименьшие различия между векторами признаков обрабатываемых изображений, характеризующих один и тот же класс. Проанализированы известные декоррелирующие преобразования для снижения вариативности гистограмм обрабатываемых изображений. Рассмотрены их достоинства и недостатки. Обоснован выбор вейвлета Гаусса 1-го порядка, позволившего обеспечить наименьшие различия между векторами признаков обрабатываемых изображений, характеризующих один и тот же класс. Предложен подход к формированию вторичных признаков распознавания путем усреднения компонентов матриц кратномасштабного преобразования гистограмм распределения яркости изображений вдоль оси их параметра сдвига. Обоснован выбор гистограмм, характеризующих распределение битов яркости обрабатываемых изображений, в качестве первичных векторов признаков. Обоснован переход к выбору избыточного числа масштабов декомпозиции в вейвлет-преобразованиях при формировании вторичных векторов признаков. Правомерность предложенного подхода подтверждена результатами моделирования. Обоснован выбор гистограмм, характеризующих распределение битов яркости обрабатываемых изображений, в качестве первичных векторов признаков. Рассмотрены виды вейвлет-преобразования и различные типы формирующих функций.

Информация об авторах:

Дворников Сергей Викторович, д.т.н., профессор, профессор кафедры проектирования и технологии электронных и лазерных Санкт-Петербургского государственного университета аэрокосмического приборостроения, профессор кафедры радиосвязи Военной академии связи им. Маршала Советского Союза С. М. Буденного, Санкт-Петербург, Россия. <http://orcid.org/0000-0002-4889-0001>

Бестугин Александр Роальдович, директор института радиотехники и инфокоммуникационных технологий Санкт-Петербургского государственного университета аэрокосмического приборостроения, д.т.н., профессор, Санкт-Петербург, Россия. <http://orcid.org/0000-0003-3847-2516>

Киришина Ирина Анатольевна, доцент кафедры проектирования и технологии электронных и лазерных средств Санкт-Петербургского государственного университета аэрокосмического приборостроения, к.э.н., доцент, Санкт-Петербург, Россия. <http://orcid.org/0000-0002-3110-2363>

Васильева Дина Владимировна, инженер по специальности "Химическая технология тугоплавких неметаллических и силикатных материалов", Санкт-Петербургский государственный технологический институт (технический университет), старший преподаватель кафедры радиотехнических систем Института радиотехники и инфокоммуникационных технологий Санкт-Петербургского государственного университета аэрокосмического приборостроения, Санкт-Петербург, Россия. <http://orcid.org/0009-0009-5343-3356>

Васильев Александр Протальонович, Московский технический университет связи и информатики, Москва, Россия

Для цитирования:

Дворников С.В., Бестугин А.Р., Киришина И.А., Васильева Д.В., Васильев А.П. Повышение контрастности признаков распознавания изображений на основе кратномасштабной обработки гистограмм распределения их яркости // Т-Comm: Телекоммуникации и транспорт. 2026. Том 20. №9. С. 18-26.

For citation:

S.V. Dvornikov, A.R. Bestugin, I.A. Kirshina, D.V. Vasilyeva, A.P. Vasilyev, "Contrast enhancement of image recognition features based on multiscale processing of histograms of their brightness Distribution," T-Comm, 2026, vol. 20, no.9, pp. 18-26. (in Russian)

Введение

Все возрастающий поток информации существенно усложняет процедуру ее обработки [1-3]. Особую актуальность данная проблема приобретает при поиске абстрактных объектов в базах данных обрабатываемых изображений [4, 5]. Примером таких объектов являются результаты разливов нефтепродуктов в акваториях морей и океанов, поскольку они не имеют четкой структуры, позволяющей осуществить их единую формализацию, для последующей обработки на основе методов машинного обучения [6, 7]. Кроме того, сложность своевременного обнаружения результатов такого рода техногенных катастроф усугубляется необходимостью мониторинга огромных площадей контролируемых акваторий [8, 9].

Именно поэтому актуальным видится поиск научных подходов, позволяющих решать задачи обработки изображений в интересах своевременного обнаружения последствий техногенных катастроф, в том числе разливов нефтепродуктов, непосредственно по мере их поступления с аппаратуры фото и видео фиксации, устанавливаемой на летно-подъемных платформах.

Вместе с тем удаленность летно-подъемных платформ от пунктов экологического контроля с телекоммуникационным оборудованием предполагает необходимость организации высокопроизводительных каналов передачи информации [10, 11]. Но это не всегда возможно в силу различных обстоятельств. Поэтому рациональным решением в такой ситуации является предварительная селекция информационных данных непосредственно на борту летно-подъемных платформ, что позволило бы существенно снизить объем передаваемого канального трафика, представляемого операторам для принятия решения о результатах обнаружения разлива нефтепродуктов в контролируемых акваториях [12]. Реализация такого подхода видится возможной только при условии наличия высокоэффективных алгоритмов обработки изображений, обеспечивающих автоматическое обнаружение последствий разливов нефтепродуктов. Однако проведенный анализ подходов к организации экологического мониторинга показал, что преимущественно, используется обработка изображений в отложенном режиме на основе методов машинного обучения.

Указанные обстоятельства определяют необходимость поиска и разработки новых методов и способов, позволяющих успешно разрешить рассмотренные проблемы.

Так, в представленной статье демонстрируются результаты исследования по разработке аналитического аппарата для метода вероятностной оценки результатов обработки изображений в процессе поиска последствий техногенных катастроф, в следствии разливов нефтепродуктов в акваториях морей и океанов. А также приводятся данные экспериментов, подтверждающие правомерность разработанного подхода. жениях.

Обоснование подхода к вероятностной оценке результатов обнаружения аномалий в обрабатываемых изображениях

Основополагающим фактором, определяющим достоверность обнаружения объектов на обрабатываемых изображениях, является форма представления исходных данных.

В частности, в [8, 13] обоснована целесообразность представления данных кадров обрабатываемых изображений в виде гистограмм распределения их яркости.

Такой выбор, прежде всего, обусловлен не только компактностью представления исходных данных, но и достаточно высокой их чувствительностью к возникновению аномалий [14, 15]. Под аномалиями, в рамках проведенного исследования, будем понимать изображение разлива нефтепродуктов в контролируемой акватории [8].

В качестве примера на рис. 1 представлены изображения формата *.jpeg с разрешением 640×480 фрагмента акватории Черного моря, без разливов нефтепродуктов (рис. 1а) и с последствиями техногенной катастрофы (рис. 1б).

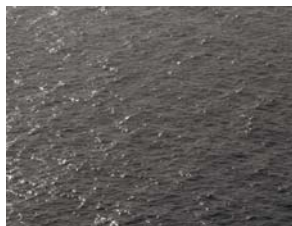


Рис. 1а. Тестовое изображение ТИ-1

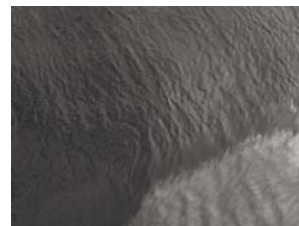


Рис. 1б. Тестовое изображение ТИ-2

На рис. 2 отображены, соответственно, матрицы амплитудных значений тестовых изображений, представленных на рис. 1. Такой формат изображения достаточно неудобный и громоздкий для его обработки и требует существенных вычислительных ресурсов при формировании векторов признаков, имеющих 307 200 элементов. Переход к векторному представлению, за счет усреднения амплитудных значений матриц вдоль одной из осей ведет к снижению информативности такого представления [16, 17].

Вместе с тем в [18] обосновано представление изображений для их последующей обработки в виде векторов распределения яркости.

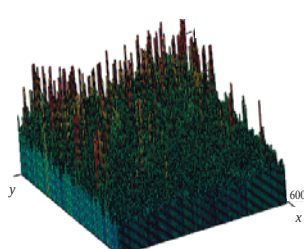


Рис. 2а. Матрица амплитудных значений ТИ-1

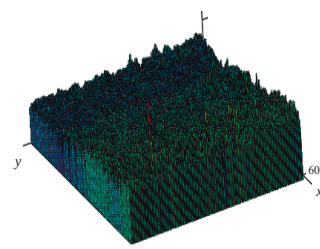


Рис. 2б. Матрица амплитудных значений ТИ-2

Учитывая, что при 8-битовом кодировании яркость представляется всего лишь 256 значениями, то соответственно, и вектор будет иметь такую же размерность.

На рис. 3 представлены гистограммы распределения яркости для тестовых изображений ТИ-1 и ТИ-2 (далее по тексту – гистограммы) в одинаковом масштабе.

Согласно результатам исследования, представленным в [6, 8, 17], векторы признаков, сформированные на основе гистограмм, обеспечивают наилучшую контрастность по отношению к другим форматам представления амплитудных значений матрицы обрабатываемых изображений.

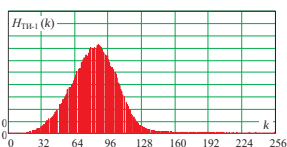


Рис. 3а. Гистограмма тестового изображения ТИ-1

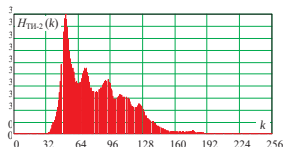


Рис. 3б. Гистограмма тестового изображения ТИ-2

Однако высокая вариативность водной поверхности даже в условиях отсутствия разливов нефтепродуктов приводит к тому, что характеризующие их изображения векторы признаков на основе гистограмм распределения яркости имеют весьма высокую дисперсию, существенно усложняющую получения высоких результатов достоверности выявления аномалий [6]. В частности, для эталонного изображения ТИ-1, рис. 1а, с водной поверхностью не загрязненной разливами нефтепродуктов, были подготовлены кадры с различной освещенностью, при незначительном изменении характера волнения, см. изображения ТИ-11, ТИ-12, ТИ-13 и ТИ-14 на рис. 4. И соответствующие им векторы признаков $H_{ТИ-11}$, $H_{ТИ-12}$, $H_{ТИ-13}$, $H_{ТИ-14}$, (рис. 5).



Рис. 4а. Тестовое изображение ТИ-11



Рис. 4б. Тестовое изображение ТИ-12

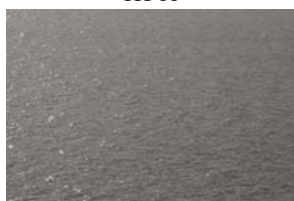


Рис. 4в. Тестовое изображение ТИ-13

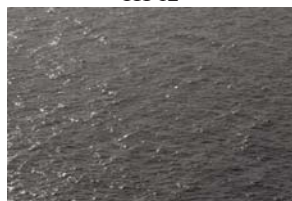


Рис. 4г. Тестовое изображение ТИ-14

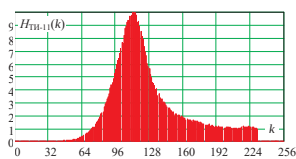


Рис. 5а. Гистограмма тестового изображения ТИ-11

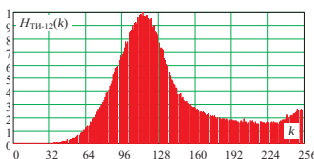


Рис. 5б. Гистограмма тестового изображения ТИ-12

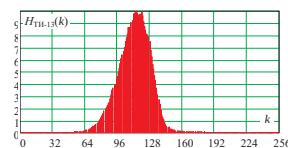


Рис. 5в. Гистограмма тестового изображения ТИ-13

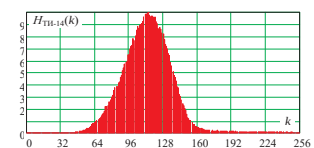


Рис. 5г. Гистограмма тестового изображения ТИ-14

Сопоставительный анализ изображений и соответствующих им гистограмм позволяет сделать следующее заключение. Формирование устойчивых векторов признаков, характеризующих один и тот же класс распознаваемых объектов

(водную акваторию без разливов нефтепродуктов), из-за значительной дисперсии амплитудных значений гистограмм, будет происходить с определенными сложностями.

Для оценки вариативности гистограмм $H_{ТИ-11}$, $H_{ТИ-12}$, $H_{ТИ-13}$, $H_{ТИ-14}$, фактически характеризующих тестовое изображение ТИ-1 в других условиях, построим функцию различий в соответствии с выражением:

$$\Delta r_{0,N}(k) = |H_{ТИ-1}(k) - H_{ТИ-N}(k)| \quad (1)$$

где N – текущий номер гистограммы обрабатываемого изображения, сравниваемый с гистограммой изображения ТИ-1.

На рис. 6 представлены графики функций различий $\Delta r_{0,N}(k)$ гистограмм тестовых изображений ТИ-11, ТИ-12, ТИ-13 и ТИ-14.

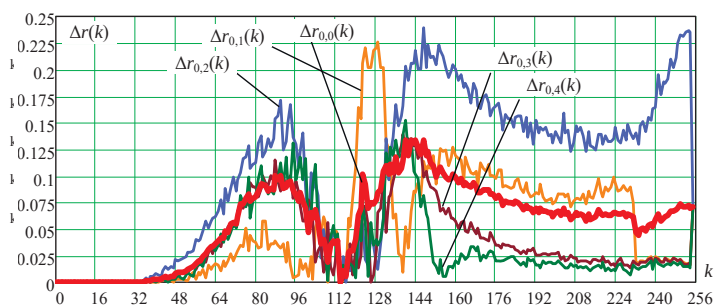


Рис. 6. Графики функций различий гистограмм ТИ-11, ТИ-12, ТИ-13 и ТИ-14 по отношению к эталонному описанию ТИ-1: $\Delta r_{0,1}(k)$ – функция различий между изображением ТИ-1 и ТИ-11; $\Delta r_{0,2}(k)$ – функция различий между изображением ТИ-1 и ТИ-12; $\Delta r_{0,3}(k)$ – функция различий между изображением ТИ-1 и ТИ-13; $\Delta r_{0,4}(k)$ – функция различий между изображением ТИ-1 и ТИ-14; $\Delta r_{0,0}(k)$ – усредненная функция различий.

Все рассчитанные функции различий равны нулю в тех местах, где элементы гистограмм, как эталонного описания, так и текущего равны нулю, а также в позиции их максимальных значений.

На рис. 6 представлена и усредненная функция различий $\Delta r_{0,0}(k)$, ее положение как раз и позволяет оценить параметры вариативности.

Для оценки размаха дисперсии усредненной функции различий, дополнительно были рассчитаны функции, характеризующие ее границы сверху и снизу (рис. 7).

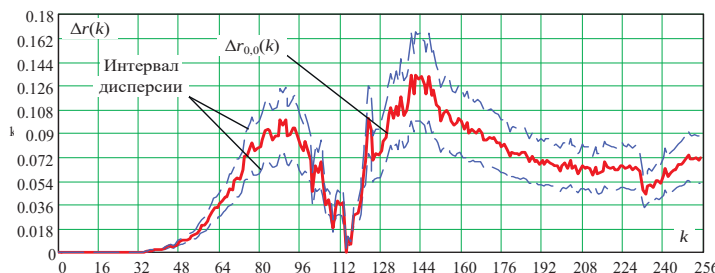


Рис. 7. График усредненной функции различий, с нанесенными границами ее дисперсии

В таблице 1 приведены результаты оценки различий гистограмм тестовых изображения, полученных на основе выделений

$$\Delta r = \left| \overline{H_{\text{ТИ-1}} - H_{\text{ТИ-N}}} \right| \quad (2)$$

и

$$\Delta r [\%] = \left| \overline{H_{\text{ТИ-1}} - H_{\text{ТИ-N}}} \right| \times \left(\sum_{k=0}^{255} H_{\text{ТИ-1}_k} \right)^{-1} \times 100\% \quad (3)$$

Таблица 1

Оценка различий гистограмм тестовых изображений

	$H_{\text{ТИ-1}} - H_{\text{ТИ-11}}$	$H_{\text{ТИ-1}} - H_{\text{ТИ-12}}$	$H_{\text{ТИ-1}} - H_{\text{ТИ-13}}$	$H_{\text{ТИ-1}} - H_{\text{ТИ-14}}$
Δr	1,187	2,031	0,815	0,782
$\Delta r, \%$	2,5	4,3	1,7	1,6

Разница $H_{\text{ТИ-1}} - H_{\text{ТИ-2}}$, характеризующая различия между эталонным изображением ТИ-1 и изображением с разливами нефтепродуктов ТИ-2, рассчитанная в соответствии с выражением (3) составляет 7,6%.

Отметим, что наибольший интервал текущей величины дисперсии у функции различий средних значений согласно табл. 1 для обрабатываемых гистограмм тестовых изображений достигает порядка 4% по отношению к эталонному описанию. Очевидно, что в этом случае, согласно [6], даже потенциальная вероятность правильного обнаружения не превышает значения 0,96. Вместе с тем для усредненной функции различий гистограмм среднее значение Δr равно 1,204. А общая величина дисперсии достигает значения 0,338.

Тогда, если для классификации результата обнаружения разливов нефтепродуктов использовать ранговый алгоритм вида

$$\text{sign}(d_{01} - d_{02}) = \begin{cases} 0 > A_0 \in A_2; \\ 0 \leq A_0 \in A_1, \end{cases} \quad (4)$$

где

$$d_{01(02)} = \frac{1}{M} \sum_{k=1}^M |H_{\text{ТИ-1}_k} - H_{\text{ТИ-N}_k}|. \quad (5)$$

Здесь d_{01} – абсолютная разность между вектором эталонного описания $H_{\text{ТИ-1}}$ класса A_1 , соответствующего изображению с отсутствием разливов нефтепродуктов и одного из обрабатываемых изображений $H_{\text{ТИ-N}}$ этого же класса; d_{02} – разность между вектором эталонного описания $H_{\text{ТИ-1}}$ класса A_1 и одного из обрабатываемых изображений $H_{\text{ТИ-N}}$ класса A_2 , соответствующего изображению с разливами нефтепродуктов. Поскольку в алгоритме принятия решения (4) искомой величиной выступает алгебраическая разность между векторами признаков, то можно заключить, что непосредственное использование гистограмм распределения яркости в качестве конечных векторов признаков распознавания не позволит получить высокую достоверность результатов. Поэтому необходим поиск к снижению дисперсии первичных векторов признаков изображений, характеризующих один и тот же класс распознаваемых изображений.

Выбор декоррелирующего преобразования для снижения вариативности гистограмм обрабатываемых изображений

Одной из причин высокой вариативности гистограмм является изначальная избыточность тестовых изображений, обусловленная достаточно сильной пространственной корреляцией их элементов [15, 19].

Поэтому была выдвинута гипотеза, что применение декоррелирующих преобразований к гистограммам распределения яркости, соответствующим различным изображениям одного и того же класса, позволит снизить их вариативность.

Согласно [20], основной задачей декоррелирующих преобразований является снижение объема представляемых для обработки данных при сохранении их информационного наполнения.

Среди декоррелирующих преобразований широкое распространение получили следующие:

сингулярное преобразование [21];

дискретное преобразование Карунена-Лоэва (ДПКЛ) [22];

дискретное косинусное преобразование (ДКП) [23];

и дискретное вейвлет-преобразование (ДВП) [24].

Сингулярные преобразования (англ. *Singular Value Decomposition* – SVD) относятся к способам факторизации комплексных матриц, с последующим их масштабированием и поворотом. Аналитически операцию факторизации комплексных матриц $\mathbf{M}$ по сингулярным значениям $m \times n$ можно представить в следующем виде:

$$\mathbf{M} = \mathbf{U} \mathbf{\Sigma} \mathbf{V}^{*T}. \quad (6)$$

где $\mathbf{U}$ – комплексная унитарная матрица размерностью $m \times m$; $\mathbf{\Sigma}$ – диагональная прямоугольная матрица (масштабирующая матрица) размером $m \times n$, у которой диагональные значения неотрицательны; $\mathbf{V}^{*T}$ – комплексно сопряженная транспонированная матрица комплексной унитарной матрицы $\mathbf{V}$, размерность которой $n \times n$.

Особенность сингулярного разложения состоит в том, что диагональные элементы $\sigma_i = \Sigma_{ii}$ матрицы $\mathbf{\Sigma}$ представляют собой сингулярные значения матрицы $\mathbf{M}$. Применительно к обработке изображений в качестве матрицы $\mathbf{M}$ выступают битовые значения, определяющие интенсивность яркости. Заметим, что точное разложение по сингулярным значениям применимо лишь к квадратным диагонализуемым матрицам. Это связано с тем, что алгоритмов расчета для других структур матриц не существует. На практике для расчета собственных значений матриц наиболее широкое применение получил итерационный алгоритм Якоби [15]. Вместе с тем существенным ограничением практического применения сингулярного преобразования является значительное количество ненулевых сингулярных значений, что делает их вычисления непрактичными [25]. Поэтому применение сингулярных преобразований для обработки гистограмм изображений нецелесообразно.

Дискретное преобразование Карунена-Лоэва относится к оптимальным преобразованиям набора данных, представленных вектором x_n в базисе ортогональных функций, определяемых ковариационной функцией процесса. То есть преобразование Карунена-Лоэва X_t как бы адаптируется к характеру изменения вектора обрабатываемых данных x_n .

$$X_t = \sum_{k=1}^{N-1} x_k e_k(t). \quad (7)$$

где $e_k(t)$ – вещественные непрерывные функции на интервале $[a, b]$, которые являются попарно ортогональными, т.е. $L^2([a, b])$.

Особенность ДПКЛ состоит в том, что используемые в нем базисные функции ортонормированные. При этом ДПКЛ минимизирует общую среднеквадратичную ошибку преобразования по отношению к любым другим преобразованиям [26].

А дисперсия формируемого результата X_t является суммой дисперсий его составляющих. Кроме того, разложение Карунена-Лоэва обладает свойством минимальной энтропии. Но основные сложности практической реализации ДПКЛ связаны с вычислением собственных векторов линейного оператора, о чем указано в [27].

Дискретное косинусное преобразование (англ. *Discrete Cosine Transform*, DCT) – линейное ортогональное преобразование, являющееся гомоморфизмом векторного пространства преобразования Фурье. В настоящее время известно четыре вида ДКП [28], среди которых наибольшее практическое приложение получило ДКП, обеспечивающее преобразование набора данных $x_0, \dots, x_{N-1}$, в набор:

$$X_k = \sum_{n=0}^{N-1} x_n \cos[k(n+0,5)\pi / n], \quad (8)$$

для $k = 0, \dots, N-1$.

Реализация ДКП предполагает, что x_n четно для $n = -0,5$ и $n = N-0,5$; а X_k четно для $k = 0$ и нечетно для $k = N$.

Анализ результатов исследования [29] показал, что для изображений с высокой динамикой изменения фона, ДКП проигрывают ДВП по многим показателям. При этом согласно результатам исследований представленных в [30], ДВП является наиболее эффективным декоррелирующим преобразованием, обеспечивающим наилучшую контрастность между векторами признаков, характеризующих альтернативные классы при минимизации их размерности. Данный аргумент определил выбор ДВП в качестве инструмента для последующей обработки гистограмм распределения яркости тестируемых изображений.

В общем случае дискретное вейвлет-преобразование представляет собой результат декомпозиции обрабатываемого процесса в базисах вейвлет-функций (англ. *wavelet*) [31]. Свойство масштабируемости вейвлет-функций открывает возможность локализовать различные по характеру изменения исследуемого процесса в различных проекциях масштаба, что позволяет рассматривать их при одинаковом разрешении, независимо от реальных размеров.

В общем случае, реализация ДВП $W(k, n)$ представляет собой свертку входной реализации обрабатываемого процесса $x(t)$ и формирующего материнского вейвлета $\psi(t)$ [32]

$$W(k, n) = \frac{1}{\sqrt{2^k}} \int_0^{N-1} x(t) \psi\left(\frac{t-2^k n}{2^k}\right) dt. \quad (9)$$

где n – переменная сдвига; k – переменная масштаба.

Таким образом, в соответствии с (9), ДВП представляет собой результат декомпозиции входного процесса путем его последовательной ВЧ и НЧ-фильтрации с временным сдвигом n и кратномасштабным изменением амплитудно-частотной характеристики НЧ-фильтра $g[2n-t]$ и ВЧ-фильтра $h[2n-t]$.

При этом результат ВЧ-фильтрации $h(2n-k)$ позволяет получить детализирующие значения, определяемые как вейвлет-коэффициенты $u_{вч}(n)$, а НЧ-фильтрация $g[2n-t]$ ведет к формированию аппроксимирующих коэффициентов $u_{нч}[n]$:

$$y_{нч}(n) = \sum_{t=0}^{N-1} x(t)g(2n-t), \quad y_{вч}(n) = \sum_{t=0}^{N-1} x(t)h(2n-t) \quad (10)$$

Фильтры, описываемые посредством выражений (10), получили название квадратурных зеркальных фильтров [33]. Последовательная реализация процедур, в соответствии с выражениями (10), ведет к формированию двумерной матрицы вейвлет-коэффициентов, которая как раз и представляет собой дискретное вейвлет-преобразование [34].

С представленных позиций ДВП можно рассматривать как распределение энергии (в рассматриваемом случае амплитудных значений гистограмм) в пространстве их масштабных преобразований и временных сдвигов, реализуемых на основе формирующих (порождающих) материнских вейвлетов.

Вместе с тем, анализ выражения (9) показывает, что ДВП обеспечивает кратномасштабный сдвиг, пропорциональный множителю 2^k , в пределах которого как раз и происходит усреднение обрабатываемых значений вдоль оси n . Для исключения потери информативности первичных векторов признаков в результате их усреднения на каждом последующем сдвиге 2^k , без потери общности воспользуемся непрерывной формой кратномасштабного вейвлет-преобразования, определяемого следующим выражением [31]

$$W(k, n) = \frac{1}{\sqrt{k}} \int_0^{N-1} x(t) \psi\left(\frac{t-n}{k}\right) dt. \quad (11)$$

Кратномасштабное преобразование, описываемое выражением (11), также позволяет формировать двумерные матрицы вейвлет-коэффициентов, структура которых будет определяться типом формирующей вейвлет-функции $\psi(*)$.

Обоснование формирующей вейвлет-функции для кратномасштабного преобразования над гистограммами обрабатываемых изображений

Свойства вейвлет-преобразований во многом определяются выбором исходной порождающей функции.

Простейшей формирующей функцией (далее по тексту – материнский вейвлет), используемой для кратномасштабного преобразования, является вейвлет Хаара, описываемый следующим выражением [35]:

$$\psi_x(t) = \begin{cases} 1, & 0 \leq t < 0,5; \\ -1, & 0,5 \leq t < 1; \\ 0, & t \notin [0, 1). \end{cases} \quad (12)$$

А масштабирующая функция для вейвлета Хаара (далее по тексту – отцовский вейвлет) определяется формулой:

$$\phi_x(t) = \begin{cases} 1, & 0 \leq t < 1; \\ 0, & t \notin [0, 1). \end{cases} \quad (13)$$

Здесь и далее нижний индекс $\psi(*)$ определяет первую букву названия функции вейвлета.

В соответствии с аналитической формой представления формирующей функции вейвлета Хаара $\psi_x(t)$ (12), элементы входной реализации группируются по два. А затем для каждой группы вычисляется их сумма и разность. Указанная последовательность процедур реализуется рекурсивно, с образованием нового уровня масштабного преобразования в пространстве временных сдвигов [32]. В результате выполнения указанных процедур получают 2^{k-1} уровней декомпозиций и одну общую сумму в виде ошибки (остатка) аппроксимации. На рисунке 8 показано временное представление отцовского

$\varphi_X(n)$ и материнского $\psi_X(n)$ вейвлетов Хаара. На рисунке 9 показано временное представление вейвлета Добеши.

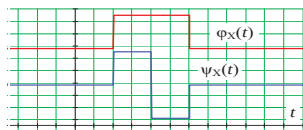


Рис. 8. Временное представление формирующей и масштабирующей функций вейвлета Хаара



Рис. 9. Временное представление формирующей и масштабирующей функций вейвлета Добеши

В ДВП активное применение находят вейвлеты Добеши. Это семейство ортогональных функций, не имеющих аналитического описания, коэффициенты которых рассчитывают на основе итерационных алгоритмов. Представляют их следующим образом [36]:

$$\psi_D(n) = \sqrt{2} \sum_{t=0}^{N-1} g_t \varphi(2n-t) \quad (14)$$

$$\varphi_D(n) = \sqrt{2} \sum_{t=0}^{N-1} h_t \varphi(2n-t) \quad (15)$$

Расчет отцовского вейвлета $\varphi(n)$ происходит при использовании вычисленных значений ортогональных нормированных коэффициентов, получивших название коэффициенты Добеши низких порядков [36]. В частности, на рис. 9 представлены формирующий и масштабирующий вейвлеты Добеши $\psi_D(n)$ и $\varphi_D(n)$ 4-го порядка.

в дальнейшем будем рассматривать лишь формирующие материнские функции основных видов вейвлетов.

Для дискретных данных достаточно часто используют, так называемые, дуальные (двойные) вейвлеты, синтезируемые на основе передаточных характеристик низкочастотных и высокочастотных фильтров, или R -функций, связанных следующим условием [37]:

$$\psi_{Дв}(n) = \varphi_{Дв}(n) + z \varphi_{Дв}(2n) \quad (16)$$

где z – некоторое комплексное число.

Формирующая (порождающая) функция для двойного вейвлета демонстрируется на рис. 10. На рис. 11 показана масштабирования функция биортогонального вейвлета, коэффициенты которой аппроксимированы посредством В-сплайна.

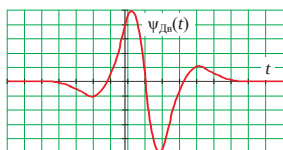


Рис. 10. Временное представление формирующей функции двойного вейвлета

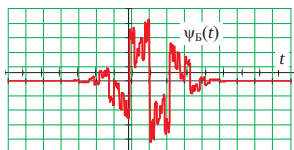


Рис. 11. Временное представление формирующей функции биортогонального вейвлета B-spline

В обработке изображений широкое применение находят биортогональные вейвлеты (B -сплайновые (англ. *B-spline*) вейвлеты) [38]. Главная особенность биортогональных вейвлетов заключается в том, что формируемый на их основе кратномасштабный базис не является ортогональным [37].

Однако применение биортогональных вейвлетов допускает больше возможных степеней свободы для построения симметричных вейвлет-функций. В преобразованиях на основе биортогональных вейвлетов применяют две формирующие ψ_B, ψ'_B и две масштабирующие φ_B, φ'_B функции.

Аналитически это можно представить следующим образом:

$$\varphi_B(n) = \sum_{t=0}^{N-1} p_t \varphi(2n-t) \quad \varphi'_B(n) = \sum_{t=0}^{N-1} p'_t \varphi'(2n-t) \quad (17)$$

$$\psi_B(n) = \sum_{t=0}^{N-1} (-1)^{t+1} p_{-t+1} \varphi(2n-t)$$

$$\psi'_B(n) = \sum_{t=0}^{N-1} (-1)^{t+1} p'_{-t+1} \varphi'(2n-t) \quad (18)$$

В (17) и (18) коэффициенты p_t и p'_t являются расчетными, значения которых представлены в [39].

Вместе с тем наиболее широкое применение в кратномасштабном анализе получили вейвлеты Гаусса различных порядков [26, 28, 30, 31, 40]. В аналитическом виде вейвлет Гаусса n -го порядка можно представить на основе следующего выражения:

$$\psi_{Гn}(t) = (-1)^n \frac{d^n}{dt^n} \exp(-t^2/2) \quad (19)$$

где n – порядок дифференцирования.

Согласно выражению (19), повышение порядка дифференцирования ведет к увеличению количества осцилляций функции на длительности ее элемента. Поэтому учитывая, что гистограммы представляют собой векторы дискретных значений, применение вейвлет-функций выше 2-го порядка не целесообразно.

С учетом такого ограничения, аналитические выражения, описывающие вейвлеты Гаусса 1-го и 2-го порядка, будут иметь вид:

$$\psi_{Г1}(t) = -t \exp(-t^2/2) \quad (20)$$

$$\psi_{Г2}(t) = (1-t^2) \exp(-t^2/2) \quad (21)$$

Временное представление вейвлетов Гаусса 1-го и 2-го порядков демонстрируется, соответственно, на рис. 12 и 13.

Заключение

Полученные результаты позволяют сделать вывод о достижении целевой установки, направленной на разработку аналитического аппарата, позволяющего получить вероятностную оценку результатов обработки изображений.

Несмотря на то, что объектом исследования являлись результаты разливов нефтепродуктов в акваториях морей и океанов, разработанный аналитический аппарат может быть применен и в других областях науки и техники, связанных с обработкой кадров фото и видеоизображений.

В работе обосновано понятие аномалий, под которыми, в широком аспекте, авторы понимают элементы фона (фрагменты) графического изображения, возникающие в

результате внешних факторов. В частности, последствий разливов нефтепродуктов.

В категориях теории распознавания образов предложен подход к формированию первичных и вторичных векторов признаков обрабатываемых изображений на основе гистограмм распределения яркости, с последующей их кратномасштабной обработкой.

Доказана целесообразность выбор вейвлета Гаусса 1-го порядка в декоррелирующей обработке векторов признаков.

Практическую направленность полученных результатов исследований авторы связывают с их применением для поиска заданных фрагментов изображений в больших базах данных.

Литература

1. Козловский А. В. Распараллеливание обработки информации при формировании составных изображений // Известия ЮФУ. Технические науки. 2025. № 1(243). С. 92-103. DOI 10.18522/2311-3103-2025-1-92-103. EDN XEQNBN.
2. Ушенкин В. А. Метод высокоточной обработки топографических интерферометрических изображений с привлечением низкодетальной стереофотограмметрической информации от оптико-электронных систем ДЗЗ // Цифровая обработка сигналов. 2025. № 1. С. 59-66. EDN YWPXNC.
3. Хомяков О. О., Панищев В. С., Титов В. С., Ватутин Э. И. Математическая модель и параллельный алгоритм обработки изображений, содержащих символическую информацию // Труды МАИ. 2024. № 137. EDN YCKKNY.
4. Копейкин Р. Е., Перов А. Ю., Соловьев М. А. Алгоритмы поляризации селекционной информации для обработки информации в изображениях, полученных с помощью радара с синтезированной апертурой // Научно-технический вестник Поволжья. 2023. № 12. С. 304-308. EDN BMDKLN.
5. Хомколов И. В., Медиевский А. В., Зотин А. Г. и др. Информационное обеспечение обработки и анализа изображений медицинского эксперимента // Информатизация и связь. 2023. № 3. С. 35-42. DOI 10.34219/2078-8320-2023-14-3-35-42. EDN YDRKUW.
6. Дворников С. В., Васильева Д. В. Повышение достоверности выявления аномалий на изображениях при формировании их векторов признаков в базах вейвлетов // Информатика и автоматизация. 2024. Т. 23, № 6. С. 1698-1729. DOI 10.15622/ia.23.6.5. EDN JDDWJS.
7. Abbasov I. B. Image Recognition in Agriculture and Landscape Protection // International Journal of Science and Research. 2020. Vol. 9, No. 12. P. 757-763. DOI 10.21275/SR201212144831
8. Васильева Д. В., Дворников С. В. Структуризация первичных признаков изображений систем мониторинга окружающей среды // Радиотехника. 2024. Т. 88, № 8. С. 5-17. DOI 10.18127/j00338486-202408-01. EDN CXJBKD.
9. Пахлян И. А., Омельянюк М. В. Разработка технических и технологических решений для очистки акваторий от разлива нефтепродуктов // Нефтяное хозяйство. 2025. № 8. С. 94-97. DOI 10.24887/0028-2448-2025-8-94-97. EDN WXLKXU.
10. Маценко С. В., Блиновская Я. Ю. Показатели эффективности мероприятий по ликвидации разливов нефти и нефтепродуктов на морских акваториях: организационные и технологические аспекты // Защита окружающей среды в нефтегазовом комплексе. 2023. № 2(311). С. 5-11. DOI 10.33285/2411-7013-2023-2(311)-5-11. EDN FKLXYZ.
11. Vasilyeva D. V., Dvornikov S. V., Yakushenko S. A., Dvornikov S. S. Automation of detection procedures based on the results of processing images from video surveillance systems. Collection of abstracts of reports of the III International Forum "Mathematical methods and models in high-tech production". 08.11.2023 Ch. 1. GUAP. 2023.
12. Петрова Е. Г. Чрезвычайные ситуации с аварийным выбросом опасных химических веществ // Экологические системы и приборы. 2023. № 5. С. 47-53. DOI 10.25791/esip.5.2023.1372. EDN PSFNPU.
13. Подрезов Р. В., Райфельд М. А. Непараметрический метод определения числа классов в задаче сегментации изображений // Автоматизация. 2020. Т. 56, № 3. С. 81-90. DOI 10.15372/AUT20200309. EDN KEDMMS.
14. Андриянов Н. А., Гаврилина Ю. Н. Исследование алгоритма обнаружения детерминированных аномалий на сложных изображениях с помощью дважды стохастической модели // Ural Radio Engineering Journal. 2020. Т. 4, № 1. С. 18-32. DOI 10.15826/urej.2020.4.1.002. EDN HAHHGA.
15. Li J., Li X., Li H. et al. Fabrication of the Cu/AgCuTi/Nb composite for superconducting radio-frequency material under extreme service conditions based on electroplating additive manufacturing // Materials & Design. 2024. Vol. 244. P. 113220. DOI 10.1016/j.matdes.2024.113220. EDN UHVEUS.
16. Белов А. М., Денисова А. Ю. Классификация изображений ДЗЗ с использованием алгоритма комплексирования данных различных сенсоров // Компьютерная оптика. 2020. Т. 44, № 4. С. 627-635. DOI 10.18287/2412-6179-CO-735. EDN LJANDA.
17. Vasilieva D. V., Dvornikov S. V., Yakushenko S. A., Dvornikov S. S. Automation of forest fire detection procedures based on the results of video processing // Scientific and analytical journal "Bulletin of the St. Petersburg University of the State Fire Fighting Service of the Ministry of Emergency Situations of Russia". 2023. № 4. С. 47-58. DOI 10.61260/2218-130X-2024-2023-4-47-58. EDN TQYCOI.
18. Суринов В. А. Об обработке зашумленных контрастных изображений // Вестник Южно-Уральского государственного университета. Серия: Математика. Механика. Физика. 2021. Т. 13, № 1. С. 14-21. DOI 10.14529/mmph210102. EDN WHOFR.
19. Велигоша А. В., Малышко Н. Н., Струков Р. И. Применение искусственных нейронных сетей для снижения избыточности дискретного вейвлет-преобразования изображений // Теория и техника радиосвязи. 2020. № 2. С. 5-12. EDN QVMKEW.
20. Чебачев В. О. Обработка сигналов в системах радиочастотной идентификации с помощью декоррелирующих преобразований // Вопросы радиоэлектроники. 2016. № 9. С. 35-40. EDN WJTJXR.
21. Булатов Ю. Н. Теорема о норме сингулярных J-псевдодифференциальных операторов Киприянова-Катрахова // Вестник Воронежского государственного университета. Серия: Физика. Математика. 2025. № 2. С. 38-48. EDN MMYKYK.
22. Осанов В. А. Сравнительный анализ методов декорреляции сетевого трафика // Вестник Рязанского государственного радиотехнического университета. 2025. № 92. С. 25-38. DOI 10.21667/1995-4565-2025-92-25-38. EDN APDMUC.
23. Семенов В. И. Сравнение алгоритма кратномасштабного анализа изображений в частотной области с алгоритмом Малла // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2022. № 59. С. 73-82. DOI 10.17223/19988605/59/8. EDN VMBXDU.
24. Sheikh H. R., Bovik A. C. Image information and visual quality // IEEE Transactions on Image Processing. 2006. Vol. 15, iss. 2. P. 430-444. doi: 10.1109/tip.2005.859378
25. Bisgard James. Analysis and Linear Algebra: The Singular Value Decomposition and Applications. Student Mathematical Library (1st ed.). AMS. 2021. ISBN 978-1-4704-6332-8.
26. Dvornikov S. V., Dvornikov S. S., Konoplev M. A. Algorithm of radio communication signals recognition on the basis of symmetric matrices // Information technologies. 2010. № 9. С. 75-77.
27. Karhunen-Loeve transform (KLT) Archived 2016-11-28 at the Wayback Machine, Computer Image Processing and Analysis (E161) lectures, Harvey Mudd College. URL: <https://pages.hmc.edu/ruye/MachineLearning/lectures/ch8/node4.html> (дата обращения 02.10.2025)
28. Ahmed Nasir, Natarajan R. Raj, Rao K.R. Discrete Cosine Transform // IEEE Transactions on Computers. C-23 (1). IEEE Computer Society. January 1974, pp. 90-93. doi:10.1109/T-C.1974.223784.
29. Loeffler C., Ligtenberg A. and Moschytz G. Practical Fast 1-D DCT Algorithms with 11 Multiplications // Proc. Int'l. Conf. on Acoustics, Speech, and Signal Processing 1989 (ICASSP '89), pp. 988-991.

30. Dvornikov S. V., Stepynin D. V., Dvornikov A. S., Bukareva A. P. Formation of signal feature vectors from wavelet coefficients of their frame transformations // *Information Technologies*. 2013. № 5. С. 46-49. EDN QAQWUF.

31. Flores Soriano M. Solar radio bursts impact on the International GNSS Service Network during Solar Cycle 24 // *Journal of Space Weather and Space Climate*. 2024. DOI 10.1051/swsc/2024034. EDN TWLMJI.

32. Васильева Д.В. Исследование чувствительности векторов признаков, сформированных на основе кратномасштабных преобразований обрабатываемых изображений. Известия высших учебных заведений России. Радиоэлектроника. 2024. Т. 27. № 3. С. 20-29. doi.org/10.32603/1993-8985-2024-27-3-20-29.

33. Singh U., Dua A., Kumar N., Guizani M. QoS Aware Uplink Scheduling for M2M Communication in LTE/LTE-A Network: A Game Theoretic Approach // *IEEE Transactions on Vehicular Technology*. 2022. Vol. 71, No. 4, pp. 4156-4170. – DOI 10.1109/TVT.2021.3132535. EDN CCFNBN.

34. Сай С. В., Зинкевич А. В., Фомина Е. С. Сравнение дискретного косинус и вейвлет-преобразований в системах сжатия RAW-изображений // *Компьютерная оптика*. 2022. Т. 46, № 6. С. 929-938. DOI 10.18287/2412-6179-CO-1094.

35. Велигоша А. В., Малышко Н. Н., Струков Р. И. Применение искусственных нейронных сетей для снижения избыточности

дискретного вейвлет-преобразования изображений // *Теория и техника радиосвязи*. 2020. № 2. С. 5-12.

36. Грищенко А. Ю., Коровкин Н. В., Островский Д. П. Вейвлеты Эрмита-Гаусса: синтез дискретных форм и исследование свойств // *Научно-технический вестник информационных технологий, механики и оптики*. 2025. Т. 25, № 4. С. 789-796. DOI 10.17586/2226-1494-2025-25-4-789-796. EDN SDEFFM.

37. Charles K., Chui An. Introduction to Wavelets (Wavelet Analysis & Its Applications), 1992, Academic Press, San Diego, ISBN 0-12-174584-8.

38. Семенов В. И., Чумаров С. Г. От конструирования вейвлетов на основе производных функции Гаусса к синтезу фильтров с конечной импульсной характеристикой // *Научно-технический вестник информационных технологий, механики и оптики*. 2024. Т. 24, № 2. С. 306-313. DOI 10.17586/2226-1494-2024-24-2-306-313. EDN KLQTXM.

39. Беляев А. А., Кононов Д. П., Кудрин А. М. Исследование выбора материнской вейвлет-функции при вибродиагностике тепловозных двигателей // *Известия Петербургского университета путей сообщения*. 2025. Т. 22, № 2. С. 290-299. DOI 10.20295/1815-588X-2025-2-290-299. EDN FDGUUM.

40. Vasilieva D. V., Dvornikov S. S., Tolstukha Y. E. et al. Formation of feature vectors for video surveillance systems // *Voprosy radioelektroniki. Series: Television Technics*. 2023. № 4. С. 62-68.

INCREASING THE CONTRAST OF IMAGE RECOGNITION FEATURES BASED ON MULTISCALE PROCESSING OF HISTOGRAMS OF THEIR BRIGHTNESS

Sergey V. Dvornikov, Saint-Petersburg State University of Aerospace Instrumentation;
Military Telecommunications Academy, St Petersburg, Russia

Alexander R. Bestugin, Saint-Petersburg State University of Aerospace Instrumentation, St Petersburg, Russia

Irina A. Kirshina, Saint-Petersburg State University of Aerospace Instrumentation, St Petersburg, Russia

Dina V. Vasilieva, Saint-Petersburg State University of Aerospace Instrumentation, St Petersburg, Russia, dolli.dina@mail.ru

Alexander P. Vasiliev, Moscow Technical University of Communications and Informatics, Moscow, Russia

Abstract

The problem of detecting anomalies in processed images obtained from visual monitoring equipment becomes particularly relevant during environmental monitoring of the world's oceans when searching for the results of oil product spills. Under the anomalies in the presented study, fragments of the image characterizing the consequences of oil product spills are identified. As a scientific research tool, the elements of the pattern recognition theory are defined, according to which the processed images are represented as feature vectors. The purpose of the work – to consider possible approaches to wavelet transformations of primary feature vectors using decorrelating processing to select a wavelet that allows for the smallest differences between feature vectors of processed images characterizing the same class. An approach to the formation of secondary recognition features is proposed by averaging the components of multiscale transformation matrices of histograms of the brightness distribution of images along the axis of their shift parameter. The choice of histograms characterizing the distribution of brightness bits of processed images as primary feature vectors is justified. The transition to the choice of an excessive number of decomposition scales in wavelet transformations during the formation of secondary feature vectors is justified. To consider possible approaches to wavelet transformations of primary feature vectors using decorrelating processing to select a wavelet that allows for the smallest differences between feature vectors of processed images characterizing the same class. The validity of the proposed approach is confirmed by the simulation results. The choice of histograms characterizing the distribution of brightness bits of processed images as primary feature vectors is justified. The types of wavelet transformations and various types of generating functions are considered.

Keywords: image anomalies, oil product spills, feature vectors, recognizable image classes, histograms of image brightness distribution

References

- [1] A. V. Kozlovsky, "Parallelization of Information Processing in the Generation of Composite Images," *Izvestiya SFedU. Engineering Sciences*. 2025. No. 1(243), pp. 92-103. DOI 10.18522/2311-3103-2025-1-92-103.
- [2] V. A. Ushenkin, "Method for high-precision processing of topographic interferometric images incorporating low-detail stereophotogrammetric information from optoelectronic Earth remote sensing systems," *Digital Signal Processing*. 2025. No. 1, pp. 59-66.
- [3] O. O. Khomyakov, V. S. Panishchev, V. S. Titov, E. I. Vatutin, "Mathematical model and parallel algorithm for processing images containing symbolic information," *Proceedings of MAI*. 2024. No. 137.
- [4] R. E. Kopeikin, A. Yu. Perov, M. A. Solov'yev, "Polarization selection algorithms for information processing in synthetic aperture radar images," *Scientific and Technical Bulletin of the Volga Region*. 2023. No. 12, pp. 304-308.
- [5] I. V. Khomkolov, A. V. Medievsky, A. G. Zotin et al., "Algorithmic support for processing and analyzing medical experiment images," *Informatization and Communication*. 2023. No. 3, pp. 35-42. DOI 10.34219/2078-8320-2023-14-3-35-42.

- [6] S. V. Dvornikov, D. V. Vasilyeva, "Improving the reliability of anomaly detection in images through feature vector generation in wavelet bases," *Informatics and Automation*. 2024. Vol. 23, No. 6, pp. 1698-1729. DOI 10.15622/ia.23.6.5.
- [7] I. B. Abbasov, "Image Recognition in Agriculture and Landscape Protection," *International Journal of Science and Research*. 2020. Vol. 9, No. 12, pp. 757-763. DOI 10.21275/SR201212144831
- [8] D. V. Vasilyeva, S. V. Dvornikov, "Structuring of primary image features for environmental monitoring systems," *Radiotekhnika*. 2024. Vol. 88, No. 8, pp. 5-17. DOI 10.18127/r00338486-202408-01.
- [9] I. A. Pakhlyan, M. V. Omelyanyuk, "Development of technical and technological solutions for cleaning water areas following petroleum product spills," *Neftyanoe Khozyaystvo (Oil Industry)*. 2025. No. 8, pp. 94-97. DOI 10.24887/0028-2448-2025-8-94-97.
- [10] S. V. Matsenko, Ya. Yu. Blinovskaya, "Performance indicators for oil and petroleum product spill response operations in marine waters: organizational and technological aspects," *Environmental Protection in the Oil and Gas Complex*. 2023. No. 2(311), pp. 5-11. DOI 10.33285/2411-7013-2023-2(311)-5-11.
- [11] D. V. Vasilyeva, S. V. Dvornikov, S. A. Yakushenko, S. S. Dvornikov, "Automation of detection procedures based on the results of processing images from video surveillance systems," *Collection of abstracts of reports of the III International Forum "Mathematical methods and models in high-tech production"*. 08.11.2023 Ch. I. SUAI. 2023.
- [12] E. G. Petrova, "Emergency situations involving the accidental release of hazardous chemical substances," *Ecological Systems and Devices*. 2023. No. 5, pp. 47-53. DOI 10.25791/esip.5.2023.1372.
- [13] R. V. Podrezov, M. A. Raifeld, "Non-parametric method for determining the number of classes in image segmentation," *Avtometriya*. 2020. Vol. 56, No. 3, pp. 81-90. DOI 10.15372/AUT20200309.
- [14] N. A. Andriyanov, Yu. N. Gavrilina, "Investigation of an algorithm for detecting deterministic anomalies in complex images using a doubly stochastic model," *Ural Radio Engineering Journal*. 2020. Vol. 4, No. 1, pp. 18-32. DOI 10.15826/urej.2020.4.1.002.
- [15] J. Li, X. Li, H. Li et al., "Fabrication of the Cu/AgCuTi/Nb composite for superconducting radio-frequency material under extreme service conditions based on electroplating additive manufacturing," *Materials & Design*. 2024. Vol. 244. P. 113220. DOI 10.1016/j.matdes.2024.113220.
- [16] A. M. Belov, A. Yu. Denisova, "Classification of Earth remote sensing images using a multi-sensor data fusion algorithm," *Computer Optics*. 2020. Vol. 44, No. 4, pp. 627-635. DOI 10.18287/2412-6179-CO-735.
- [17] D. V. Vasilieva, S. V. Dvornikov, S. A. Yakushenko, S. S. Dvornikov, "Automation of forest fire detection procedures based on the results of video processing," *Bulletin of the St. Petersburg University of the State Fire Fighting Service of the Ministry of Emergency Situations of Russia*. 2023. No. 4, pp. 47-58. DOI 10.61260/2218-130X-2024-2023-4-47-58.
- [18] V. A. Surin, "On the processing of noisy high-contrast images," *Bulletin of South Ural State University. Series: Mathematics. Mechanics. Physics*. 2021. Vol. 13, No. 1, pp. 14-21. DOI 10.14529/mmph210102.EDN WHOFR.
- [19] A. V. Veligosh, N. N. Malyshko, R. I. Strukov, "Application of artificial neural networks to reduce the redundancy of the discrete wavelet transform of images," *Theory and Technique of Radio Communication*. 2020. No. 2, pp. 5-12.
- [20] V. O. Chebachev, "Signal processing in radio-frequency identification systems using decorrelating transforms," *Issues of Radio Electronics*. 2016. No. 9, pp. 35-40.
- [21] Yu. N. Bulatov, "Theorem on the norm of singular Kipriyanov-Katrakhov J-pseudodifferential operators," *Proceedings of Voronezh State University. Series: Physics. Mathematics*. 2025. No. 2, pp. 38-48.
- [22] V. A. Osanov, "Comparative analysis of network traffic decorrelation methods," *Bulletin of Ryazan State Radio Engineering University*. 2025. No. 92, pp. 25-38. DOI 10.21667/1995-4565-2025-92-25-38.
- [23] V. I. Semenov, "Comparison of a frequency-domain multiscale image analysis algorithm with Mallat's algorithm," *Tomsk State University Journal of Control, Computer Science and Informatics*. 2022. No. 59, pp. 73-82. DOI 10.17223/19988605/59/8.
- [24] H. R. Sheikh, A. C. Bovik, "Image information and visual quality," *IEEE Transactions on Image Processing*. 2006. Vol. 15, iss. 2, pp. 430-444. doi: 10.1109/tip.2005.859378
- [25] Bisgard James, "Analysis and Linear Algebra: The Singular Value Decomposition and Applications," *Student Mathematical Library* (1st ed.). AMS. 2021. ISBN 978-1-4704-6332-8.
- [26] S. V. Dvornikov, S. S. Dvornikov, M. A. Konoplev, "Algorithm of radio communication signals recognition on the basis of symmetric matrices," *Information technologies*. 2010. No. 9, pp. 75-77.
- [27] Karhunen-Loeve transform (KLT) Archived 2016-11-28 at the Wayback Machine, Computer Image Processing and Analysis (EI61) lectures, Harvey Mudd College.URL: <https://pages.hmc.edu/ruey/MachineLearning/lectures/ch8/node4.html> (date of access 02.10.2025).
- [28] Ahmed Nasir, Natarajan T. Raj, Rao K.R., "Discrete Cosine Transform," *IEEE Transactions on Computers*. C-23 (1). IEEE Computer Society, 1974, pp. 90-93. doi:10.1109/T-C.1974.223784.
- [29] C. Loeffler, A. Ligtenberg, and G. Moschytz, "Practical Fast 1-D DCT Algorithms with 11 Multiplications," *Proc. Int'l. Conf. on Acoustics, Speech, and Signal Processing 1989 (ICASSP '89)*, pp. 988-991.
- [30] S. V. Dvornikov, D. V. Stepynin, A. S. Dvornikov, A. P. Bukareva, "Formation of signal feature vectors from wavelet coefficients of their frame transformations," *Information Technologies*. 2013. No. 5, pp. 46-49.
- [31] Soriano M. Flores, "Solar radio bursts impact on the International GNSS Service Network during Solar Cycle 24," *Journal of Space Weather and Space Climate*. 2024. DOI 10.1051/swsc/2024034.
- [32] D. V. Vasilyeva, "Investigation of the sensitivity of feature vectors generated based on multiscale transforms of processed images," *Journal of the Russian Universities. Radioelectronics*. 2024, no. 27(3), pp. 20-29. doi.org/10.32603/1993-8985-2024-27-3-20-29.
- [33] U. Singh, A. Dua, N. Kumar, M. Guizani, "QoS Aware Uplink Scheduling for M2M Communication in LTE/LTE-A Network: A Game Theoretic Approach," *IEEE Transactions on Vehicular Technology*. 2022. Vol. 71, No. 4, pp. 4156-4170. DOI 10.1109/TVT.2021.3132535.
- [34] S. V. Sai, A. V. Zinkevich, E. S. Fomina, "Comparison of discrete cosine and wavelet transforms in RAW image compression systems," *Computer Optics*. 2022. Vol. 46, No. 6, pp. 929-938. DOI 10.18287/2412-6179-CO-1094.
- [35] A. V. Veligosh, N. N. Malyshko, R. I. Strukov, "Application of artificial neural networks to reduce the redundancy of the discrete wavelet transform of images," *Theory and Technique of Radio Communication*. 2020. No. 2, pp. 5-12.
- [36] A. Yu. Grishentsev, N. V. Korovkin, D. P. Ostrovsky, "Hermite-Gaussian wavelets: synthesis of discrete forms and investigation of properties," *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*. 2025. Vol. 25, No. 4, pp. 789-796. DOI 10.17586/2226-1494-2025-25-4-789-796.
- [37] K. Charles, An. Chui, "Introduction to Wavelets (Wavelet Analysis & Its Applications)," 1992, Academic Press, San Diego, ISBN 0-12-174584-8.
- [38] V. I. Semenov, S. G. Chumarov, "From Wavelet Design Based on Gaussian Function Derivatives to FIR Filter Synthesis," *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*. 2024. Vol. 24, No. 2, pp. 306-313. DOI 10.17586/2226-1494-2024-24-2-306-313.
- [39] A. A. Belyaev, D. P. Kononov, A. M. Kudrin, "Investigation of mother wavelet function selection for vibration-based diagnostics of diesel locomotive engines," *Proceedings of Petersburg State Transport University*. 2025. Vol. 22, No. 2, pp. 290-299. DOI 10.20295/1815-588X-2025-2-290-299.
- [40] D. V. Vasilieva, S. S. Dvornikov, Y. E. Tolstukha et al., "Formation of feature vectors for video surveillance systems," *Voprosy radioelektroniki. Series: Television Technics*. 2023. No. 4, pp. 62-68.

Information about authors:

Sergey V. Dvornikov, Dr Sci. (Eng.), Professor of the Radio Communication Department of the Military Telecommunications Academy, Professor of the Department of Radioengineering and Fiber-optic Complexes of the Saint Petersburg State University of Aerospace Instrumentation, St. Petersburg, Russia. <https://orcid.org/0000-0002-4889-0001>

Aleksandr R. Bestugin, Director of the Institute of Radio Engineering and Information Communication Technologies at the St. Petersburg State University of Aerospace Instrumentation, Doctor of Technical Sciences, Professor, St. Petersburg, Russia. <http://orcid.org/0000-0003-3847-2516>

Irina A. Kirshina, associate professor of the Department of Design and Technology of Electronic and Laser Devices at St. Petersburg State University of Aerospace Instrumentation, Candidate of Economic Sciences, associate professor, St. Petersburg, Russia. <http://orcid.org/0000-0002-3110-2363>

Dina V. Vasilieva, St. Petersburg State Technical University (TI), chemical technology of refractory non-metallic and silicate materials, senior lecturer at the Department of Radio Engineering Institute of Radio Engineering and Infocommunication Technologies, St. Petersburg, Russia. <http://orcid.org/0009-0009-5343-3356>

Aleksandr P. Vasiliev, Moscow Technical University of Communications and Informatics, Moscow, Russia

ОБЕСПЕЧЕНИЕ БЕЗОПАСНОСТИ БЕСПРОВОДНЫХ УСТРОЙСТВ ИОТ-ИНФРАСТРУКТУРЫ С ПОМОЩЬЮ ПОЛУПРОВОДНИКОВЫХ ФИЗИЧЕСКИ НЕКЛОНИРУЕМЫХ ФУНКЦИЙ

DOI: 10.36724/2072-8735-2026-20-9-27-35

Manuscript received 28 May 2026;

Accepted 07 August 2026

Мосин Сергей Геннадьевич,*Казанский (Приволжский) федеральный университет,
Казань, Россия;**ООО "ЛабСистемс", г. Владимир, Россия;**Рязанский государственный радиотехнический университет
им. В.Ф. Уткина, Рязань, Россия, smosin@ieee.org***Ключевые слова:** Интернет вещей (IoT), беспроводные IoT-устройства, безопасность, физически неклонируемая функция (PUF), схемотехническое моделирование, XOR-вентиль, реконфигурируемый кольцевой генератор

В работе рассмотрена проблема обеспечения безопасности беспроводной инфраструктуры Интернета вещей (IoT – Internet of Things) в контексте Индустрии 4.0. Критическим вызовом является ограниченность ресурсов беспроводных IoT-устройств (автономное энергопотребление, вычислительная мощность, объем памяти), что не позволяет применять сложные криптографические алгоритмы в реальном времени. Для решения этой задачи предложено использование полупроводниковых физически неклонируемых функций (PUF – Physically Unclonable Functions), генерирующих уникальные псевдослучайные последовательности на основе случайных вариаций технологического процесса производства интегральных схем. Представлена системная модель IoT-инфраструктуры в терминах теории множеств, позволяющая выполнять структурный и параметрический синтез в ходе проектирования. Основным научным вкладом является разработка схемотехнической модели XOR-вентилей, оптимизированной для синтеза PUF на основе реконфигурируемых кольцевых генераторов. Модель адаптирована под КМОП-технологии TSMC 180 нм и позволяет на ранних этапах проектирования беспроводных IoT-устройств по заказной технологии выбирать рациональные конфигурации PUF с учетом фактических ограничений. Проведено численное моделирование методом Монте-Карло с учетом коррелированных и некоррелированных технологических вариаций. Реализованы две тестовые обвязки для накопления битовых последовательностей. Полученные результаты верифицированы и продемонстрировали соответствие теоретическим ожиданиям, подтверждая эффективность предложенного подхода для синтеза псевдослучайных битовых последовательностей для целей идентификации и аутентификации беспроводных IoT-устройств. Дальнейшие исследования направлены на оценку надежности и устойчивости к атакам с использованием машинного обучения.

Информация об авторах:

Мосин Сергей Геннадьевич, доктор технических наук, профессор кафедры прикладной математики и искусственного интеллекта, Казанский (Приволжский) федеральный университет, г. Казань, Россия; заместитель генерального директора ООО "ЛабСистемс", г. Владимир, Россия; профессор кафедры САПР ВС, Рязанский государственный радиотехнический университет им. В.Ф. Уткина, г. Рязань, Россия. ORCID 0000-0003-1389-2602

Для цитирования:

Мосин С. Г. Обеспечение безопасности беспроводных устройств IoT-инфраструктуры с помощью полупроводниковых физически неклонируемых функций // T-Comm: Телекоммуникации и транспорт. 2026. Том 20. №9. С. 27-35.

For citation:

S. G. Mosin, "The security ensuring of IoT-infrastructure's wireless devices using semiconductor physically unclonable functions", T-Comm, 2026, vol. 20, no.9, pp. 27-35. (in Russian)

Введение

Переход к четвертой промышленной революции (Индустрия 4.0) обусловлен разработкой и активным внедрением киберфизических систем (КФС, *CPS – Cyber-Physical System*) в различные секторы экономики с целью организации умного производства, под которым понимают взаимодействие между умной фабрикой и умной продукцией. Используемые в работе определения основаны на принятых в российских и международных стандартах понятиях (табл. 1). Умное производство можно охарактеризовать использованием автоматизированных систем управления, взаимодействующих между собой без непосредственного участия человека с применением технологий искусственного интеллекта. А Индустрию 4.0 – как концептуальную модель создания единого цифрового пространства, где физические объекты и нематериальные активы представлены виртуально и могут взаимодействовать друг с другом на протяжении своего жизненного цикла.

Таблица 1

Перечень стандартов, определяющих понятия предметной области

ГОСТ Р 71531-2024 «Системы киберфизические. Термины и определения».
ГОСТ Р 71576-2024 «Системы киберфизические. Общие положения».
ГОСТ Р 59799-2021 «Умное производство. Модель эталонной архитектуры индустрии 4.0 (RAMI 4.0)».
ГОСТ Р 71842-2024 «Цифровая промышленность. Умное производство. Часть 1. Термины и определения»
ISO/IEC 20924-2021 <i>Information technology – Internet of Things (IoT) – Vocabulary</i> .

Каркасом Индустрии 4.0 выступает стек ключевых технологий, который включает: искусственный интеллект (*AI – Artificial Intelligence*) и машинное обучение (*ML – Machine Learning*), большие данные и предиктивную аналитику (*Big Data*), цифровые двойники (*Digital Twins*), интернет вещей (*IoT – Internet of Things*), распределенные и высокопроизводительные вычисления (*HPC – High-Performance Computing*), роботизированные и мехатронные агенты, дополненная и виртуальная реальность (*AR/VR – Augmented/Virtual Reality*) и др.

Технология *IoT* в представленном стеке определяет инфраструктуру взаимосвязанных сущностей, систем и информационных ресурсов, а также служб, позволяющих обрабатывать информацию о физическом и виртуальном мире и реагировать на нее. Такая инфраструктура представлена сетью датчиков, сенсоров, контрольно-измерительных узлов и устройств (КИУУ), собирающих данные о состоянии объектов и параметрах процессов, преобразующих их к удобному формату и передающих в единую систему обработки, которые обладают интерфейсом подключения к проводным сетям или беспроводным каналам связи. Причем беспроводные *IoT*-устройства получают все большее распространение в силу повышения гибкости систем, их мобильности и масштабируемости, а также простоты установки и обслуживания.

Обеспечение кибербезопасности *IoT*-инфраструктуры КФС требует комплексной защиты каналов связи, конечных устройств и обрабатываемых данных. Переход на беспроводные интерфейсы расширяет поверхность атак за счет уязвимостей физического уровня (перехват, глушение, подмена пакетов), что критически повышает риски компрометации по

сравнению с проводными технологиями. Применение недоверенной электронной компонентной базы и несанкционированное подключение сторонних устройств выступают источниками злонамеренных ошибок в виде сторонних атак, провоцирования отказов и сбоев в обслуживании, нелегального доступа к данным, нарушения их целостности и достоверности, что негативно сказывается на безопасности и защищенности *IoT*-инфраструктуры КФС.

Объектом предлагаемого исследования выступает безопасность беспроводных *IoT*-устройств. Проблема – ограниченные ресурсы и возможности беспроводных *IoT*-устройств (слабые процессоры/микроконтроллеры, мало памяти, узкий канал связи, лимит энергопотребления и др.) для использования сложных защитных механизмов в реальном времени. Предмет исследования – подходы к внутрисхемной реализации уникальных защитных механизмов с учетом фактических ограничений беспроводных *IoT*-устройств.

Цель предложенной работы – реализация и исследование генераторов уникальных псевдослучайных битовых последовательностей на основе полупроводниковых физически неклонируемых функций для обеспечения идентификации и аутентификации беспроводных *IoT*-устройств.

Основные результаты:

1. Предложена системная модель *IoT*-инфраструктуры в терминах теории множеств.
2. Определена совокупность требований к беспроводным *IoT*-устройствам и обозначено перспективное направление формирования уникальных псевдослучайных последовательностей на основе полупроводниковых физически неклонируемых функций (*PUF – Physically Unclonable Functions*).
3. Предложена схмотехническая модель *XOR*-вентиля, которая позволяет синтезировать, моделировать и выбирать рациональные конфигурации *PUF* для реализации в специализированных интегральных схемах (*ASIC*).
4. Проведено численное моделирование реконфигурируемых кольцевых генераторов (*RO*), выступающих в качестве полупроводниковой *PUF*, и исследование характеристик сформированных ими псевдослучайных битовых последовательностей. Полученные результаты продемонстрировали соответствие теоретическим ожиданиям и результатам других авторов, но для базиса программируемых логических интегральных схем (*ПЛИС*).

1 Архитектура и особенности *IoT*-инфраструктуры

В общем случае *IoT*-инфраструктуру можно рассматривать как систему гетерогенных систем, которые различаются видом взаимодействия с окружающими системами, степенью связности, топологией и технологией связи, назначением, техническими и функциональными ограничениями и др. (рис. 1). Математически архитектуру *IoT*-инфраструктуры представим кортежем множеств физических *IoT*-устройств – **E**, сетей связи – **N**, форматов накапливаемых данных – **D**, вычислительных ресурсов – **C**, приложений – **App**, технологий обеспечения комплексной безопасности – **S** и технических / функциональных требований и ограничений – **R**.

$$\text{Arch}_{\text{IoT}} = [\text{E}, \text{N}, \text{D}, \text{C}, \text{App}, \text{S}, \text{R}]. \quad (1)$$

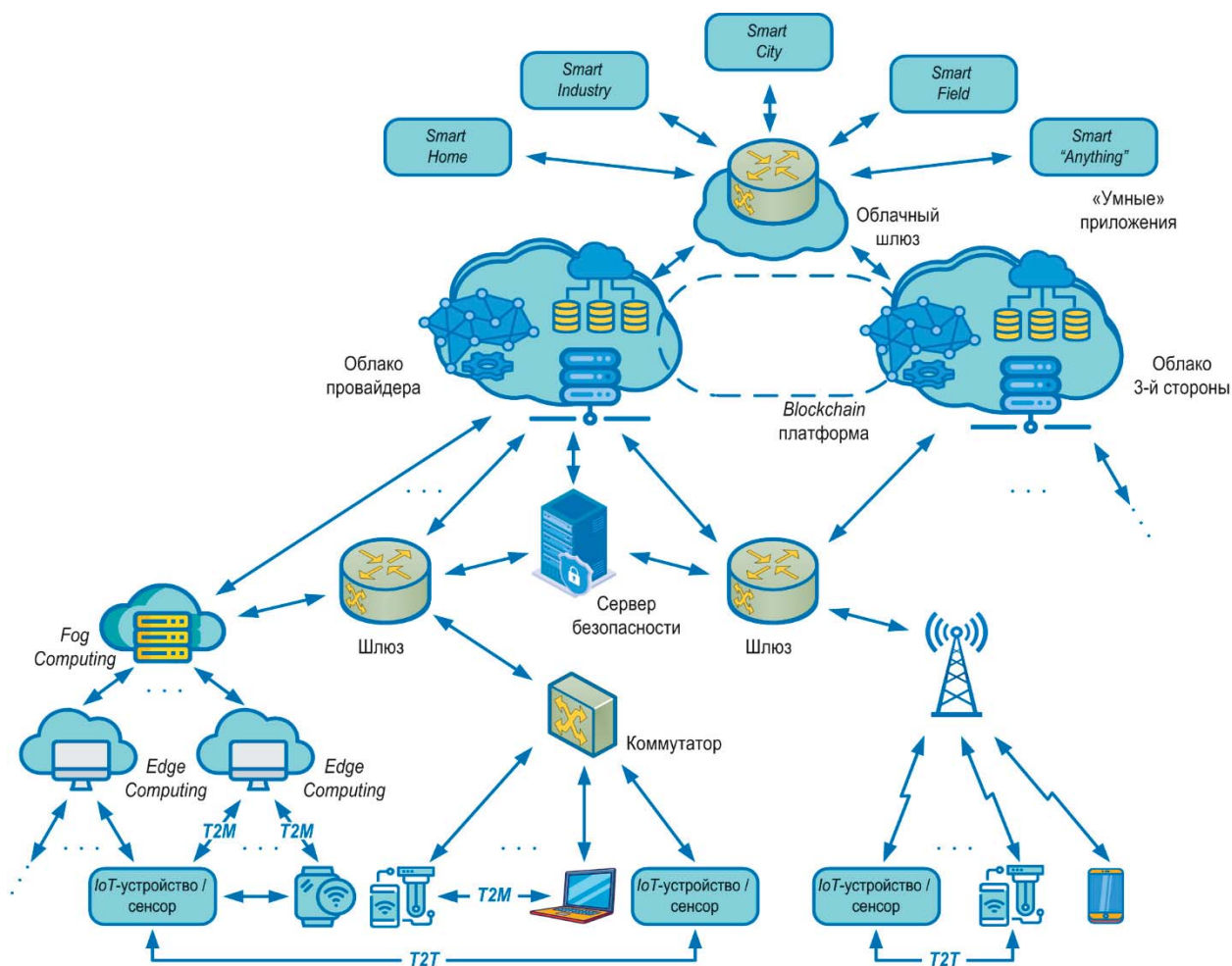


Рис. 1. Архитектура IoT-инфраструктуры

Множество IoT-устройств (**E**) включает элементы сбора данных (датчики/сенсоры, КИУУ и т.п.) и актуаторы – элементы, выполняющие действия над объектами управления (регуляторы, переключатели, сервоприводы, клапаны и т. д.) и отвечающие за взаимодействие киберпространства IoT с физическим миром.

Множество сетей связи (**N**) включает стандарты и протоколы передачи информации с использованием проводных и беспроводных технологий, а также устройства, обеспечивающие сетевое взаимодействие – коммутаторы, шлюзы, точки доступа, базовые станции и др. Правила организации и функционирования сетей связи определяет стек протоколов от физического уровня до прикладного. Взаимодействие IoT-устройств с другими элементами инфраструктуры рассматривают как межмашинную коммуникацию (*M2M* – *Machine-to-Machine*) в рамках которой в зависимости от цели и уровня интеграции выделяют две модели:

- *Things-to-Machine (T2M)*, которая описывает сценарий взаимодействия оконечных устройств сбора данных (*Things*) с интеллектуальной системой (*Machine*) в виде контроллера, компьютера/сервера или облачного сервиса (ЦОД – центр обработки данных) для постоянной или периодической передачи накопленной информации. Как правило, этот сценарий предусматривает односторонний обмен данными.

- *Things-to-Things (T2T)*, которая описывает сценарий взаимодействия IoT-устройств между собой без прямого участия человека для совместного решения задач или создания сложной, динамической и, зачастую, самоорганизующейся сети – двухсторонний обмен.

Множество форматов данных (**D**) включает правила представления накапливаемой и передаваемой информации, которые непосредственно связаны с используемыми протоколами передачи (табл. 2).

Множество вычислительных ресурсов (**C**) включает централизованные или распределенные информационно-вычислительные средства обработки данных, поступающих от оконечных устройств. Простейшие вычисления могут выполняться на микроконтроллере IoT-устройства (при наличии). Интеллектуальный анализ данных выполняют на специальных вычислительных ресурсах согласно выбранной вычислительной модели:

- Граничные вычисления (*Edge Computing*) – модель распределенных вычислений, в которой обработка информации происходит в непосредственной близости к месту, где данные были получены и будут использованы в дальнейшем (на «границе» сети) [1, 2].

- Облачные вычисления (*Cloud Computing*) – модель централизованных вычислений в ЦОД и дата-центрах на удаленных

серверах и/или высокопроизводительных вычислительных системах, которые предоставляет провайдер облачных сервисов [3]. Возможно использование коммерческих облачных сервисов по подписке или сервисов с открытым доступом, которые могут быть предоставлены не только ключевым провайдером, но и сторонним поставщиком. Интеграция облачных вычислений и технологии *Blockchain* позволяет организовать распределенную доверенную вычислительную сеть, узлами которой выступают «облака». Использование распределенного реестра устраняет единую точку отказа, повышает прозрачность операций и защищает данные от фальсификации [4].

• Туманные вычисления (*Fog Computing*) – модель децентрализованных вычислений, в которой вычислительные ресурсы, хранилища данных и сетевые сервисы размещены ближе к конечным устройствам и источникам данных, формируя промежуточный слой между ними и облаком для балансировки вычислительной нагрузки и пропускной способности сети [5].

Таблица 2

Форматы данных и протоколы передачи в IoT-инфраструктуре

Название	Источник
Форматы данных	
<i>Текстовые форматы для хранения и обмена структурированными данными</i>	
XML (eXtensible Markup Language)	RFC 3023: XML Media Types. January 2001. 39 p. https://datatracker.ietf.org/doc/html/rfc3023
JSON (JavaScript Object Notation)	ISO/IEC 21778:2017 Information technology – The JSON data interchange syntax
CSV (Comma-Separated Values)	RFC 4180: Common Format and MIME Type for Comma-Separated Values (CSV) Files. October 2005. 6 p. https://datatracker.ietf.org/doc/html/rfc4180
<i>Двоичные форматы для хранения и обмена структурированными данными</i>	
Protobuf (Protocol Buffers)	Protocol Buffers v35.1 https://github.com/protocolbuffers/protobuf/releases/tag/v35.1
Apache Avro	Avro 1.12.1 https://avro.apache.org/blog/releases/
Apache Parquet	Parquet 2.13.0 https://parquet.apache.org/blog/
Протоколы передачи	
CoAP (Constrained Application Protocol)	RFC 7252: The Constrained Application Protocol (CoAP). June 2014, IETF Trust and the Document Authors. 112 p. http://www.rfc.fr/rfc/en/rfc7252.pdf
MQTT (Message Queuing Telemetry Transport)	ISO/IEC 20922:2016 Information technology – Message Queuing Telemetry Transport (MQTT) v3.1.1. ГОСТ Р 58603–2019 Информационные технологии. Интернет вещей. Протокол организации очередей доставки телеметрических сообщений MQTT. Вер. 3.1.1.
Data Distribution Service	Object Management Group. Data distribution service for real-time systems, 2004.
AMQP (Advanced Message Queuing Protocol)	ISO/IEC 19464:2014 Information technology – Advanced Message Queuing Protocol (AMQP) v1.0 specification
XMPP (Extensible Messaging and Presence Protocol)	RFC 3920 & 3921: Extensible Messaging and Presence Protocol (XMPP): Core. October 2004. 91 p. http://www.rfc.fr/rfc/en/rfc3920.pdf

Множество приложений (**App**) включает программные решения под различное аппаратное обеспечение (смартфоны, сенсорные панели, настольные/переносные компьютеры и серверы, и др.), представленные в плоскости пользователя и плоскости разработчика. В первом случае это мобильные приложения, web-страницы, программы визуализации и обработки запросов, доступные в использовании даже для неквалифицированных пользователей. Во втором случае это библиотеки программных компонентов, программные интерфейсы приложений (*API – Application Programming Interface*), БД и СУБД, т.е. специализированные средства, предназначенные для разработчиков программного обеспечения. Можно выделить иерархию приложений от уровня отдельных пользователей до бизнес-уровня и уровня отраслей промышленности с учетом масштаба востребованности и универсальности.

Множество технологий обеспечения комплексной безопасности (**S**) включает методы и инструменты, обеспечивающие защиту всех элементов IoT-инфраструктуры, предотвращение уязвимостей и несанкционированный доступ при их использовании в разных сценариях взаимодействия [6]. На уровне обеспечения безопасности IoT-устройств важными средствами выступают: идентификация – распознавание устройства и его свойств по уникальному признаку [7, 8], аутентификация – проверка подлинности пользователя, устройства или системы [9-11] и авторизация – предоставление прав на осуществление определенных действий и доступа к ресурсам. Аутентификация использует открытые / закрытые ключи и специализированные протоколы [12, 13]. В основе аутентификации как процесса установления, что устройство (пользователь) действительно является тем, за кого себя выдает, лежат различные решения:

• Факторная аутентификация (*FA – Factor Authentication*) – метод подтверждения подлинности устройства (пользователя), при котором требуется использование нескольких независимых факторов из разных категорий. Распространена двухфакторная аутентификация (*2FA*) – частный случай многофакторной аутентификации (*MFA*), использующий только два фактора.

• Взаимная аутентификация (*MA – Mutual authentication*) – метод, при котором подлинность подтверждают обе стороны соединения до начала обмена данными.

• Аутентификации через отдельный канал (*OObA – Out-of-Band Authentication*) – метод подтверждения подлинности с использованием отдельного независимого канала связи, отличного от основного канала передачи данных.

Множество технических / функциональных требований и ограничений (**R**) включает разрешенные или допустимые диапазоны для параметров системы и используемых ресурсов, обязательные и опциональные возможности, нормативную базу и т.п. Например технические требования задают энергоэффективность, пропускную способность, совместимость, устойчивость, возможности вертикального и горизонтального масштабирования; функциональные требования определяют сценарии сбора и передачи данных, логику их обработки и формирования управляющих воздействий для актуаторов, аспекты визуализации и интерфейса пользователя для приложений; нормативная база регламентирует обработку персональных данных, взаимодействие с критическими инфраструктурами (здравоохранение, правоохранительные

службы, энергетика и др.), разрешения на использование определенных технологий и сервисов, договорные отношения между провайдером и пользователем.

В ходе проектирования *IoT*-инфраструктуры осуществляют структурный и параметрический синтез, а также анализ и оптимизацию получаемых решений, которые на ранних стадиях основаны на совокупности теоретико-множественных операций над элементами кортежа Arch_{IoT} (1), что позволяет определять состав необходимых компонентов, распределять нагрузку и осуществлять балансировку между шлюзами и ресурсами граничных, туманных и облачных вычислений, моделировать связи и отношения один-ко-многим / многие-ко-многим, масштабировать и управлять подсистемами и др.

2 Безопасность *IoT*-устройств с использованием полупроводниковых *PUF*

К беспроводным *IoT*-устройствам в отличие от проводных предъявляют особые требования к энергопотреблению и длительности автономной работы, производительности встроенных вычислительных ресурсов, к дальности и скорости передачи данных, а также к безопасности, что связано с уязвимостями и атаками – возможностью перехвата радиосигнала, подмены данных и несанкционированной интеграцией стороннего устройства в систему.

Оценка энергопотребления в течение цикла эксплуатации *IoT*-устройства может быть выполнена на ранних стадиях проектирования *IoT*-инфраструктуры с учетом особенностей используемой технологии беспроводной связи [14-16]. Минимальная производительность встроенных вычислителей также выступает ограничением на использование надежных, но вычислительно сложных технологий шифрования данных и аутентификации. Поэтому для обеспечения безопасности беспроводных *IoT*-устройств рассматривают варианты с учетом обозначенных ограничений. Одно из перспективных направлений формирования уникальных псевдослучайных

последовательностей и цифровых ключей с их использованием – внутрисхемная реализация полупроводниковых физически неклонировуемых функций (*PUF* – *Physically Unclonable Functions*), в основе которых лежат случайные вариации технологического процесса производства интегральных схем [17, 18]. В общем случае *PUF* определена множеством пар запрос-ответ (*CPR* – *Challenge-Response Pair*), которые непредсказуемы и не поддаются клонированию. Применение реконфигурируемых схем *PUF* позволяет на ограниченном множестве электронных компонентов, занимающих минимально необходимую площадь кристалла ИС, формировать семейство потенциально полезных *CPR*, где запросы (*C*) активируют конкретную конфигурацию, которой соответствует уникальный ответ (*R*).

Вентиль «Исключающее-ИЛИ» (*XOR*) – эффективный базис реализации реконфигурируемых *PUF* с точки зрения функциональных возможностей и минимизации занимаемой площади на кристалле (рис. 2). Если входы *XOR*-вентиля рассматривать как управляющий (*Cx*) и информационный (*Ix*), то в силу функциональных особенностей в зависимости от значения управляющего входа вентиль можно рассматривать как инвертор ($Cx = 1$) или как буфер ($Cx = 0$). Это обстоятельство позволяет формировать структуры, реконфигурируемые с помощью значений управляющих входов, совокупность которых определяет уникальный запрос *PUF* (*C*).

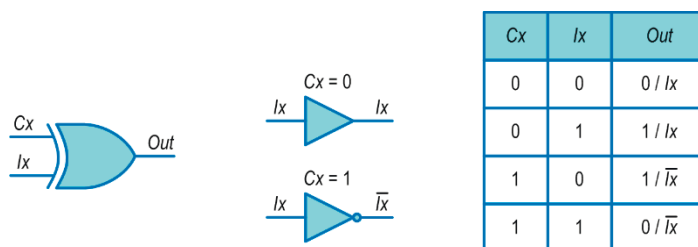


Рис. 2. *XOR*-вентиль и варианты реконфигурирования

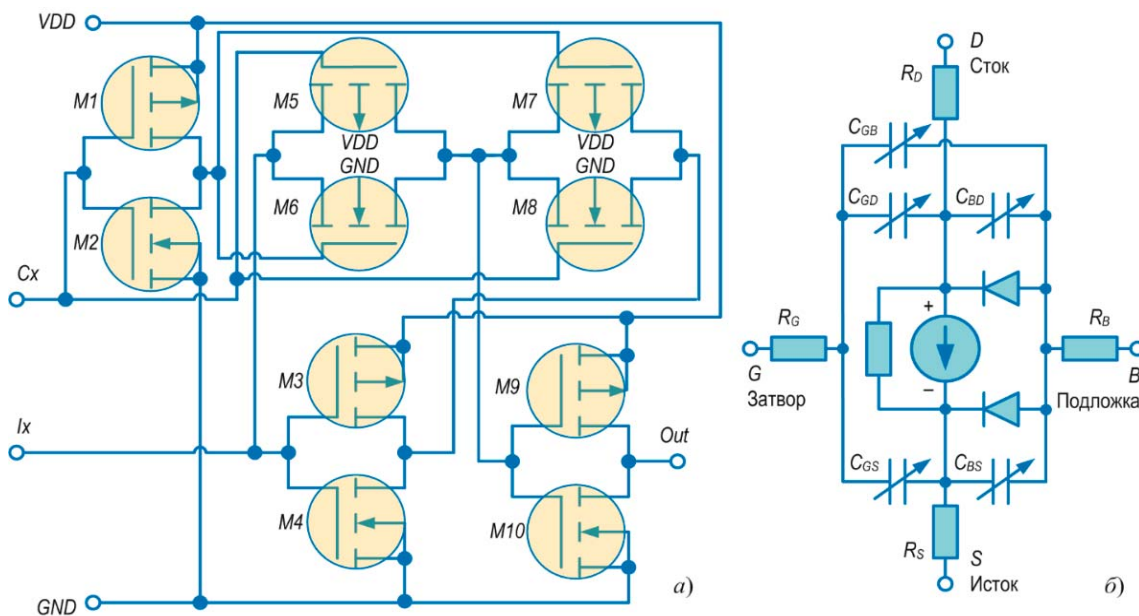


Рис. 3. Схема электрическая: *XOR*-вентиль (а), модель транзистора (б)

В работе предложена схмотехническая модель *XOR*-вентилля, которая учитывает особенности используемой интегральной технологии при производстве ИС беспроводных *IoT*-устройств и позволяет уже на ранних стадиях проектирования (до создания топологии) синтезировать, моделировать и выбирать рациональные конфигурации *PUF*, обеспечивающие формирование полезных *CRP*. Модель *XOR*-вентилля представлена списком соединений в формате *PSpice*, соответствующим схемной реализации в базе КМОП-транзисторов (рис. 3). Модель позволяет задавать параметры вариаций технологического процесса, необходимые для последующих статистических испытаний методом Монте-Карло, в ходе которых оцениваются динамические характеристики как самого вентилля, так и реконфигурируемых схем *PUF* на его основе.

В качестве таких схем рассмотрены реконфигурируемые кольцевые генераторы (*RO* – *Ring Oscillator*) с изменяемым количеством каскадов (рис. 4) [19-21]. Обязательное требование к *RO* – конфигурирование нечетного числа инверторов в схеме. Для *n*-стадийного каскада возможно формирование 2^{n-1} уникальных схем кольцевых генераторов.

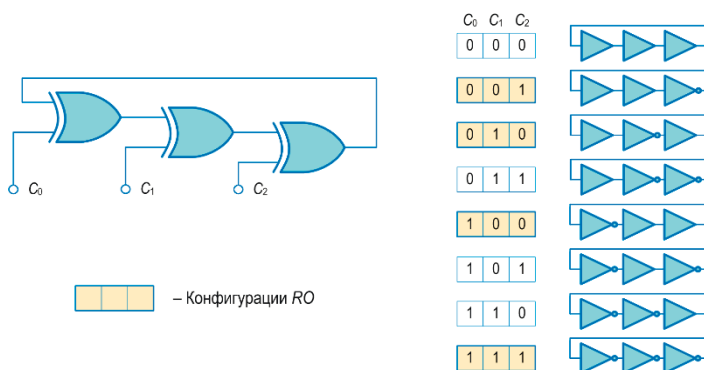


Рис. 4. Структура кольцевого генератора на трех *XOR*-вентиллях и варианты реконфигурирования

3 Экспериментальные исследования

Для подтверждения достижения цели работы – реализации и исследования генераторов уникальных псевдослучайных битовых последовательностей на основе полупроводниковых физически неклонируемых функций (*PUF*), предназначенных для идентификации и аутентификации беспроводных *IoT*-устройств, – проведена серия численных экспериментов. В ходе исследований выполнена верификация схмотехнической модели *XOR*-вентилля. Реконфигурируемые кольцевые генераторы (*RO*), выступающие в качестве полупроводниковой *PUF*, реализованы на четырех *XOR*-вентиллях. Моделирование *RO* выполнено методом Монте-Карло во временной области с использованием EDA-среды *PSpice A/D* (*CADENCE*) при 3 000 итерациях. Параметры модели *XOR*-вентилля настроены под КМОП-технологии *TSMC* 180 нм; случайные вариации технологического процесса заданы независимым разбросом параметров (3% по нормальному закону от номиналов) и коррелированным разбросом в партии (5% по равномерному закону). Расчеты проводились на вычислительной системе с процессором *Intel Core i5-7000* (3.4 ГГц, x64) и 32 Гб ОЗУ.

Для накопления псевдослучайных последовательностей предложены и реализованы две тестовые обвязки – счетчик импульсов на выходе *RO* (разрядность 10 бит) и сдвиговый

регистр с варьируемой длиной от 8 до 16 разрядов для накопления отсчетов выходного сигнала *RO* с прореживанием внутри временного окна. В процессе верификации подтверждены статические и динамические характеристики *XOR*-вентилля.

По итогам статистических испытаний построены гистограммы числовых значений, накопленных в счетчике и сдвиговом регистре, которые отражают широкий диапазон изменения частоты *RO* при вариации технологических параметров, что свидетельствует о чувствительности предложенной структуры к технологическим флуктуациям (рис. 5).

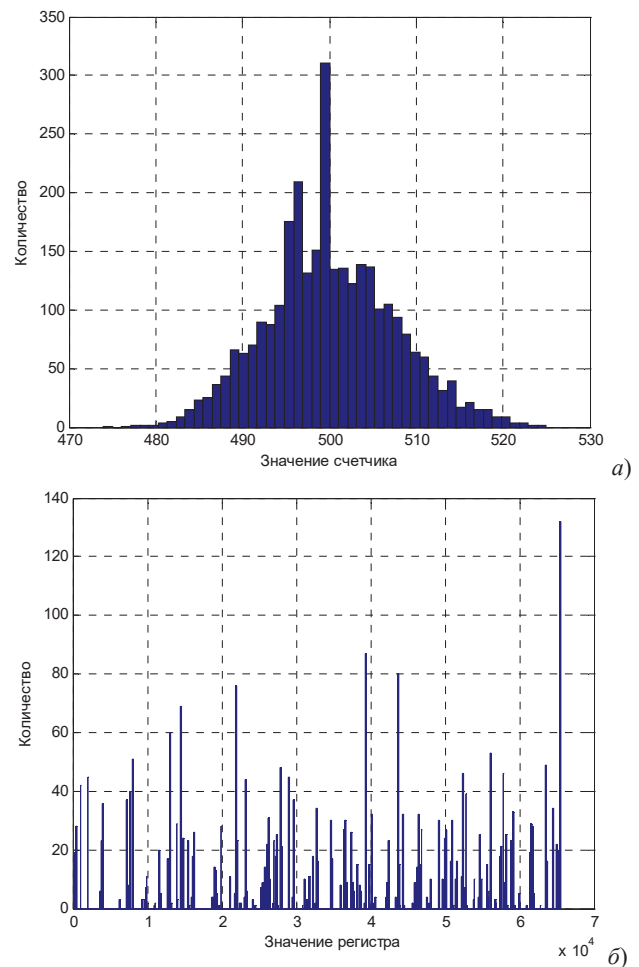


Рис. 5. Гистограмма накопленных значений при длине временного окна 150 нс: в счетчике (а), в сдвиговом 16-разрядном регистре (б)

На основе полученных выборок рассчитаны ключевые характеристики *PUF* – уникальность (U_q – *Uniqueness*), однородность (U_f – *Uniformity*) и общая минимальная энтропия (E) – для обеих обвязок при длинах сдвигового регистра 8, 10, 12, 14 и 16 разрядов (табл. 3). Результаты для сдвигового регистра демонстрируют соответствие ожидаемым теоретическим значениям: U_q и U_f стремятся к 0.5, а E – к 1, что подтверждает статистическую независимость и равномерность выходных последовательностей. Эти данные коррелируют с результатами других авторов для *PUF* на основе программируемых логических интегральных схем (ПЛИС), что дополнительно валидирует сходимость предложенной модели и реконфигурируемых схем на ее основе (табл. 4).

Таблица 3

Полученные характеристики PUF

Характеристика	Устройство					
	Счетчик	Сдвиговый регистр (разрядность, бит)				
		8-	10-	12-	14-	16-
Уникальность	0.3105	0.4962	0.4963	0.4965	0.4967	0.4968
Однородность	0.6413	0.5442	0.5437	0.5424	0.5409	0.5403
Общая минимальная энтропия	0.5037	0.8780	0.8793	0.8827	0.8866	0.8884

Таблица 4

Сравнение результатов с референсными работами

Источник	Уникальность	Однородность	Энтропия
[19]	0.4220	0.4674	–
[20]	–	–	0.9870
[21]	0.5018	0.4755	–
[22]	0.4805	0.5091	0.7825
[23]	0.4200	0.5697	–
[24]	0.4974	0.5466	–
[25]	0.4479	0.4974	–
Предложенное решение:			
сдвиговый регистр 16-бит	0.4968	0.5403	0.8884

Ключевые характеристики для счетчика уступают теоретическим ожиданиям в силу ограниченного диапазона накапливаемых значений (рис. 5, а). Прямое использование этих значений в качестве уникальных ответов PUF нецелесообразно по сравнению со значениями из сдвигового регистра. Близость ключевых характеристик ответов сдвигового регистра различной длины подтверждают возможность внутрисхемного синтеза уникальных битовых последовательностей требуемой длины с минимальными затратами площади кристалла.

4 Заключение

В работе предложена системная модель IoT-инфраструктуры, представленная кортежем конечных множеств, позволяющая выполнять структурный и параметрический синтез в ходе проектирования. Обозначены источники уязвимостей беспроводных IoT-устройств и совокупность функциональных и технологических ограничений, которые осложняют обеспечение безопасности. Определена совокупность требований к беспроводным IoT-устройствам и представлено перспективное направление формирования уникальных псевдослучайных последовательностей на основе полупроводниковых PUF. Предложена схематическая модель XOR-вентилля, которая позволяет уже на ранних стадиях проектирования IoT-устройств по заказной интегральной технологии осуществлять выбор рациональной конфигурации PUF, адаптируя параметры обвязки под требования конкретного приложения. Выполнено численное моделирование реконфигурируемых кольцевых генераторов, реализованных на основе XOR-вентилей и выступающих в качестве полупроводниковой PUF. Проведено исследование характеристик получен-

ных псевдослучайных битовых последовательностей. Экспериментальные результаты продемонстрировали соответствие теоретическим ожиданиям и результатам других авторов, что позволяет судить о достижении поставленной в данной работе цели.

Дальнейшее развитие исследований видится в расширении возможностей модели XOR-вентилля для оценки надежности (*reliability*) реконфигурируемых PUF, а также в реализации средств анализа устойчивости к атакам с использованием машинного обучения (ML-атаки).

Литература

1. Kumar A., Singh D. Reinforcement learning based security method for IoT-Edge computing // *Wireless Networks*, 2026, vol. 32, pp. 1547–1567, doi: 10.1007/s11276-026-04123-5.
2. Jinchuan W., Ming N., Qiaojin G., Zhiwei Y., Suhang L. Design of an Intelligent IoT Platform Based on the Synergy of Edge Computing and Cloud Computing // *Journal of Big Data and Computing*, 2026, vol. 4, no. 1, pp. 27–33, doi: 10.62517/jbdc.202601104.
3. Andriulo F.C., Fiore M., Mongiello M., Traversa E., Zizzo V. Edge Computing and Cloud Computing for Internet of Things: A Review // *Informatics*, 2024, vol. 11, no. 4: 71. 22 p., doi: 10.3390/informat-ics11040071.
4. Jhariya M.K., Dehalwar V., Bharti J. Blockchain-enabled secure, lightweight self-sovereign identification and authentication framework for access management of IoT devices // *Cluster Computing*, 2026, vol. 29, Paper ID 291, 25 p., doi: 10.1007/s10586-026-06005-z.
5. Apat H.K., Goswami V., Sahoo B. et al. An Adaptive Queueing Model for Optimizing IoT Application Performance in Fog-Cloud Computing Environment // *SN Computer Science*, 2026, vol. 7, Paper ID 486, 15 p., doi: 10.1007/s42979-026-04997-4.
6. Минаев В.А., Швырев Б.А., Ромашкин Т.Р. Безопасность интернета вещей: основные решения // *Информация и безопасность*. 2023. Т. 26. Вып. 2. С. 163–168. doi: 10.36622/VSTU.2023.26.2.001.
7. Dou Y. Current Status and Challenges of IoT Device Identity Authentication Technology // *Computers and Artificial Intelligence*, 2026, vol. 3, no. 2, pp. 65–77, doi: 10.70267/cai.26v3n2.6577.
8. Weibing Li, Shenggang Wu, Wei Han. Blockchain-based IoT device identity authentication and data traceability // *International Journal of Computer Information Systems and Industrial Management Applications*, 2025, vol. 17, p. 17, doi: 10.70917/ijcisim-2025-0316.
9. Бойко А. А., Быков М. Ю., Куцев С. С., Перегудов М. А. Аутентификация в сетях сотовой связи: обзор математических моделей // *Системы управления, связи и безопасности*. 2025. № 1. С. 187–219. doi: 10.24412/2410-9916-2025-1-187-219.
10. Миронов А. М. Новая математическая модель протоколов аутентификации и основанный на ней метод верификации // *Интеллектуальные системы. Теория и приложения*. 2018. Т. 22, № 4. С. 79–98.
11. Kang J.H., Seo M. Enhanced Authentication for Decentralized IoT Access Control Architecture // *Cryptography*, 2023, vol. 7, no. 3, Paper ID 42, 17 p., doi: 10.3390/cryptography7030042.
12. Chan Suet Yan S., Tang Su Wei A., Bong J. H., Teh Q. L., Sivalingam S., Khoo S. Y., Nafy T. M. Authentication of IoT device with the enhancement of One-time Password (OTP) // *Journal of IT in Asia*, 2021, vol. 9, no. 1, pp. 29–40, doi: 10.33736/jita.3841.2021.
13. Malik A., Kushwah R. A Secure Authentication and Key Establishment Protocol for IoT Devices Using ECCMD // *SN Computer Science*, 2025, vol. 6, Paper ID 803, 16 p., doi: 10.1007/s42979-025-04354-x.
14. Mosin S. A Model of LoRaWAN Communication in Class A for Design Automation of Wireless Sensor Networks Based on the IoT Paradigm // *2018 IEEE East-West Design & Test Symposium (EWDTS)*, Kazan, 2018, pp. 1–6. doi: 10.1109/EWDTS.2018.8524618.
15. Mosin S. A Model for Univariant and Multivariant Simulating the Battery Lifetime of UE NB-IoT // *2024 International Russian Automation*

Conference (RusAutoCon), Sochi, Russian Federation, 2024, pp. 13–18, Paper ID 10694326, doi: 10.1109/RusAutoCon61949.2024.

16. Мосин С.Г. Поиск оптимальной конфигурации радиоресурсов линии вверх для снижения энергопотребления UE NB-IoT при передаче данных // Динамика сложных систем. 2025. Т. 19. № 2. С. 40–49. doi: 10.18127/j19997493-202502-06.

17. Lata K., Cenkeramaddi L.R. FPGA-Based PUF Designs: A Comprehensive Review and Comparative Analysis // Cryptography, 2023, vol. 7, no. 4, Paper ID 55, 30 p., doi: 10.3390/cryptography7040055.

18. Karmakar M., Naz S.F., Shah A.P. Fault-Tolerant Reversible Logic Gate-Based RO-PUF Design // Memories - Materials, Devices, Circuits and Systems, 2023, vol. 4, Paper ID 100055, 6 p., doi: 10.1016/j.memori.2023.100055.

19. Tanamoto T., Yasuda S., Takaya S., Fujita S. Physically Unclonable Function Using an Initial Waveform of Ring Oscillators // IEEE Transactions on Circuits and Systems II: Express Briefs, 2017, vol. 64, no. 7, pp. 827–831, doi: 10.1109/TCSII.2016.2602828.

20. Rojas-Muñoz L.F., Sánchez-Solano S., Martínez-Rodríguez M.C., Brox P. True Random Number Generation Capability of a Ring Oscillator PUF for Reconfigurable Devices // Electronics, 2022, vol. 11, Paper ID 4028, 24 p., doi: 10.3390/electronics11234028.

21. Zulfikar Z., Soin N., Wan Muhamad Hatta S.F., Abu Talib M.S., Jaafar A. Routing Density Analysis of Area-Efficient Ring Oscillator Physically Unclonable Functions // Applied Sciences, 2021, vol. 11, no. 20, Paper ID 9730, 15 p., doi: 10.3390/app11209730.

22. Gu C., Chang CH., Liu W. et al. A large-scale comprehensive evaluation of single-slice ring oscillator and PicoPUF bit cells on 28-nm Xilinx FPGAs // Journal of Cryptographic Engineering, 2021, vol. 11, pp. 227–238, doi: 10.1007/s13389-020-00244-5.

23. Zulfikar Z., Walidainy H., Rahman A., Muchtar K. Reliability Improvement of 28 nm Intel FPGA Ring Oscillator PUF for Chip Identification // Cryptography, 2025, vol. 9, Paper ID 36, 16 p., doi: https://doi.org/10.3390/cryptography9020036.

24. Casado-Galán A., Núñez J., Tena-Sánchez E., Potestad-Ordóñez F.E., Acosta A.J. Assessment of an FPGA Implementation of a Hybrid PUF Based on a Configurable Transient Effect Ring Oscillator and Ring Oscillator (TERORO-PUF) // Electronics, 2026, vol. 15, no. 3, Paper ID 661, pp. 1–23, doi: https://doi.org/10.3390/electronics15030661.

25. Cui Y., Li J., Chen Y., Wang C., Gu C., O'neill M., Liu W. An efficient Ring Oscillator PUF using programmable delay units on FPGA // ACM Transactions on Design Automation of Electronic Systems, 2023, vol. 29, no. 1, Article No.: 2, pp. 1–20, doi: 10.1145/3593807.

THE SECURITY ENSURING OF IOT-INFRASTRUCTURE'S WIRELESS DEVICES USING SEMICONDUCTOR PHYSICALLY UNCLONABLE FUNCTIONS

Sergey G. Mosin, Kazan Federal University, Kazan, Russia; LabSystems LLC, Vladimir, Russia;
Ryazan State Radio Engineering University, Ryazan, Russia, smosin@ieee.org

Abstract

The paper examines the security of wireless IoT-infrastructure in the context of Industry 4.0. A critical challenge is the limited resources of wireless IoT-devices (power consumption, computing power, and memory capacity), which precludes the use of complex cryptographic algorithms in real time. The use of semiconductor physically unclonable functions (PUFs) that generate unique pseudorandom sequences based on random variations in the integrated circuit manufacturing process is proposed to address this issue. A system model of IoT-infrastructure is presented in terms of set theory, enabling structural and parametric synthesis during design. The main scientific contribution is the development of a circuit model of an XOR-gate optimized for PUF synthesis based on reconfigurable ring oscillators. The model is adapted for TSMC 180 nm CMOS technology and enables the selection of rational PUF configurations at the early stages of designing wireless IoT-devices using custom integrated technology taking into account actual constraints. The Monte Carlo simulations were conducted considering correlated and uncorrelated technological variations. Two testbenches for a pseudorandom bit sequence accumulation were implemented. The obtained results were verified and demonstrated compliance with theoretical expectations, confirming the effectiveness of the proposed approach for synthesizing pseudorandom bit sequences for identification and authentication of wireless IoT-devices. Further research aims to assess the reliability and resilience to machine learning attacks.

Keywords: Internet of Things (IoT), wireless IoT-devices, security, physically unclonable function (PUF), circuit simulation, XOR-gate, reconfigurable ring oscillator.

References

- [1] A. Kumar, D. Singh, "Reinforcement learning based security method for IoT-Edge computing," *Wireless Networks*, 2026, vol. 32, pp. 1547–1567, doi: 10.1007/s11276-026-04123-5.
- [2] W. Jinchuan, N. Ming, G. Qiaojin, Y. Zhiwei, L. Suhang, "Design of an Intelligent IoT Platform Based on the Synergy of Edge Computing and Cloud Computing," *Journal of Big Data and Computing*, 2026, vol. 4, no. 1, pp. 27–33, doi: 10.62517/jbdc.202601104.
- [3] F.C. Andriulo, M. Fiore, M. Mongiello, E. Traversa, V. Zizzo, "Edge Computing and Cloud Computing for Internet of Things: A Review," *Informatics*, 2024, vol. 11, no. 4: 71. 22 p., doi: 10.3390/informatics11040071.

- [4] M.K. Jhariya, V. Dehalwar, J. Bharti, "Blockchain-enabled secure, lightweight self-sovereign identification and authentication framework for access management of IoT devices," *Cluster Computing*, 2026, vol. 29, Paper ID 291, 25 p., doi: 10.1007/s10586-026-06005-z.
- [5] H.K. Apat, V. Goswami, B. Sahoo et al., "An Adaptive Queueing Model for Optimizing IoT Application Performance in Fog-Cloud Computing Environment," *SN Computer Science*, 2026, vol. 7, Paper ID 486, 15 p., doi: 10.1007/s42979-026-04997-4.
- [6] V.A. Minaev, B.A. Shvyrev, T.R. Romashkin, "Internet of things security: key solutions," *Informacija i bezopasnost'*, 2023. vol. 26, no. 2. pp. 163-168. doi: 10.36622/VSTU.2023.26.2.001 (in Russian).
- [7] Y. Dou, "Current Status and Challenges of IoT Device Identity Authentication Technology," *Computers and Artificial Intelligence*, 2026, vol. 3, no. 2, pp. 65-77, doi: 10.70267/cai.26v3n2.6577.
- [8] Li Weibing, Wu Shenggang, Han Wei, "Blockchain-based IoT device identity authentication and data traceability," *International Journal of Computer Information Systems and Industrial Management Applications*, 2025, vol. 17, p. 17, doi: 10.70917/ijcism-2025-0316.
- [9] A. A. Boyko, M. Yu. Bykov, S. S. Kushev, M. A. Peregodov, "Authentication in cellular networks: overview of mathematical models," *Systems of Control, Communication and Security*, 2025, no. 1, pp. 187-219, doi: 10.24412/2410-9916-2025-1-187-219 (in Russian).
- [10] A. M. Mironov, "New mathematical model of authentication protocols and verification method based on this model," *Intelligent systems. Theory and applications*, 2018, vol. 22, no. 4, pp. 79-98 (in Russian).
- [11] J.H. Kang, M. Seo, "Enhanced Authentication for Decentralized IoT Access Control Architecture," *Cryptography*, 2023, vol. 7, no. 3, Paper ID 42, 17 p., doi: 10.3390/cryptography7030042.
- [12] S. Chan Suet Yan, A. Tang Su Wei, J. H. Bong, Q. L. Teh, S. Sivalingham, S. Y. Khoo, T. M. Nafy, "Authentication of IoT device with the enhancement of One-time Password (OTP)," *Journal of IT in Asia*, 2021, vol. 9, no. 1, pp. 29-40, doi: 10.33736/jita.3841.2021.
- [13] A. Malik, R. Kushwah, "A Secure Authentication and Key Establishment Protocol for IoT Devices Using ECCMD," *SN Computer Science*, 2025, vol. 6, Paper ID 803, 16 p., doi: 10.1007/s42979-025-04354-x.
- [14] S. Mosin, "A Model of LoRaWAN Communication in Class A for Design Automation of Wireless Sensor Networks Based on the IoT Paradigm," in *2018 IEEE East-West Design & Test Symposium (EWDTS)*, Kazan, 2018, pp. 1-6. doi: 10.1109/EWDTS.2018.8524618.
- [15] S. Mosin, "A Model for Univariate and Multivariate Simulating the Battery Lifetime of UE NB-IoT," in *2024 International Russian Automation Conference (RusAutoCon)*, Sochi, Russian Federation, 2024, pp. 13-18, Paper ID 10694326, doi: 10.1109/RusAutoCon61949.2024.
- [16] S. G. Mosin, "Search of optimal uplink radio resources configuration for reducing power consumption of UE NB-IoT at data transmission," *Dynamics of complex systems*, 2025, vol. 19, no. 2, pp. 40-49, doi: 10.18127/j19997493-202502-06 (in Russian).
- [17] K. Lata, L.R. Cenkeramaddi, "FPGA-Based PUF Designs: A Comprehensive Review and Comparative Analysis," *Cryptography*, 2023, vol. 7, no. 4, Paper ID 55, 30 p., doi: 10.3390/cryptography7040055.
- [18] M. Karmakar, S.F. Naz, A.P. Shah, "Fault-Tolerant Reversible Logic Gate-Based RO-PUF Design," *Memories - Materials, Devices, Circuits and Systems*, 2023, vol. 4, Paper ID 100055, 6 p., doi: 10.1016/j.memori.2023.100055.
- [19] T. Tanamoto, S. Yasuda, S. Takaya, S. Fujita, "Physically Unclonable Function Using an Initial Waveform of Ring Oscillators," *IEEE Transactions on Circuits and Systems II: Express Briefs*, 2017, vol. 64, no. 7, pp. 827-831, doi: 10.1109/TCSII.2016.2602828.
- [20] L.F. Rojas-Munoz, S. Sanchez-Solano, M.C. Martinez-Rodriguez, P. Brox, "True Random Number Generation Capability of a Ring Oscillator PUF for Reconfigurable Devices," *Electronics*, 2022, vol. 11, Paper ID 4028, 24 p., doi: 10.3390/electronics11234028.
- [21] Z. Zulfikar, N. Soin, S.F. Wan Muhamad Hatta, M.S. Abu Talip, A. Jaafar, "Routing Density Analysis of Area-Efficient Ring Oscillator Physically Unclonable Functions," *Applied Sciences*, 2021, vol. 11, no. 20, Paper ID 9730, 15 p., doi: 10.3390/app11209730.
- [22] C. Gu, CH. Chang, W. Liu et al., "A large-scale comprehensive evaluation of single-slice ring oscillator and PicoPUF bit cells on 28-nm Xilinx FPGAs," *Journal of Cryptographic Engineering*, 2021, vol. 11, pp. 227-238, doi: 10.1007/s13389-020-00244-5.
- [23] Z. Zulfikar, H. Walidainy, A. Rahman, K. Muchtar, "Reliability Improvement of 28 nm Intel FPGA Ring Oscillator PUF for Chip Identification," *Cryptography*, 2025, vol. 9, Paper ID 36, 16 p., doi: <https://doi.org/10.3390/cryptography9020036>.
- [24] A. Casado-Gal?n, J. N??ez, E. Tena-S?nchez, F.E. Potestad-Ord??ez, A.J. Acosta, "Assessment of an FPGA Implementation of a Hybrid PUF Based on a Configurable Transient Effect Ring Oscillator and Ring Oscillator (TERORO-PUF)," *Electronics*, 2026, vol. 15, no. 3, Paper ID 661, pp. 1-23, doi: <https://doi.org/10.3390/electronics15030661>.
- [25] Y. Cui, J. Li, Y. Chen, C. Wang, C. Gu, M. O'neill, W. Liu, "An efficient Ring Oscillator PUF using programmable delay units on FPGA," *ACM Transactions on Design Automation of Electronic Systems*, 2023, vol. 29, no. 1, Article No.: 2, pp. 1-20, doi: 10.1145/3593807.

Information about author:

Sergey G. Mosin, Kazan Federal University, Professor, Kazan, Russia; LabSystems LLC, Deputy Director General, Vladimir, Russia; Ryazan State Radio Engineering University, Professor, Ryazan, Russia. ORCID 0000-0003-1389-2602

ВЗАИМОСВЯЗЬ МОЩНОСТИ И ЭНЕРГИИ ПРИ ПЕРЕДАЧЕ В РЕЖИМЕ СЛУЧАЙНОГО МНОЖЕСТВЕННОГО ДОСТУПА С МИНИМАЛЬНОЙ СРЕДНЕЙ ЗАДЕРЖКОЙ

DOI: 10.36724/2072-8735-2026-20-9-36-46

Manuscript received 11 June 2026;

Accepted 14 August 2026

Бурков Артём Андреевич,

Санкт-Петербургский государственный университет
аэрокосмического приборостроения,
г. Санкт-Петербург, Россия, a.burkov@guap.ru

Статья подготовлена при финансовой поддержке Российского
научного фонда, проект № 22-19-00305-П <https://rscf.ru/project/22-19-00305/> "Пространственно-временные стохастические
модели беспроводных сетей с большим числом пользователей"

Тюрликов Андрей Михайлович,

Санкт-Петербургский государственный университет
аэрокосмического приборостроения,
г. Санкт-Петербург, Россия, turlikov@guap.ru

Ключевые слова: случайный множественный доступ, АЛОХА,
отношение сигнал/шум, энергия на бит, средняя задержка,
стабильность, марковская цепь, задача минимизации

Долгое время исследования в области случайного множественного доступа велись независимо в двух направлениях: теория информации множественного доступа и теория случайного множественного доступа. В последние годы стали появляться работы рассматривающие системы случайного множественного доступа совмещающие оба направления. Однако в существующих работах как правило совместно не рассматриваются вопросы взаимосвязи между мощностью и затратами энергии при наличии повторных передач и минимизации средней задержки в системе. В настоящей работе исследуются такая взаимосвязь. В работе рассматривается система случайного множественного доступа на основе алгоритма АЛОХА с гауссовским каналом множественного доступа. Предполагается, что в системе имеется потенциально неограниченное число абонентских устройств. Для такой системы описан подход для определения области стабильной работы, в которой обеспечивается конечность средней задержки. Предложен способ вычисления верхней и нижней границ для средней задержки с любой заданной точностью. Сформулирована и численно решена оптимизационная задача для минимизации средней задержки в системе за счет выбора длительности сообщения абонентов при фиксированном числе бит в сообщении и заданной мощности передачи. На основе решения оптимизационной задачи предложен подход, позволяющий получить зависимость энергозатрат от мощности при передаче в режиме случайного множественного доступа с минимальной средней задержкой. Показано, что эта зависимость не является монотонно возрастающей функцией. Существует значение мощности, которое минимизирует энергозатраты. Проведено сравнение с системой, в которой имеется только один источник с неограниченной очередью для хранения сообщений и один получатель, коллизии в такой системе отсутствуют. Поэтому для такой системы зависимость энергозатрат от мощности является монотонно возрастающей функцией. При этом энергозатраты для системы со случайным доступом выше, чем для системы с одним источником при любом значении мощности.

Информация об авторах:

Бурков Артём Андреевич, доцент кафедры инфокоммуникационных технологий и систем связи Федерального государственного автономного образовательного учреждения высшего образования "Санкт-Петербургский государственный университет аэрокосмического приборостроения", Санкт-Петербург, Россия, к.т.н., ORCID0000-0002-0920-585X

Тюрликов Андрей Михайлович, заведующий кафедрой инфокоммуникационных технологий и систем связи Федерального государственного автономного образовательного учреждения высшего образования "Санкт-Петербургский государственный университет аэрокосмического приборостроения", Санкт-Петербург, Россия, профессор, д.т.н., ORCID0000-0001-7132-094X

Для цитирования:

Бурков А.А., Тюрликов А.М. Взаимосвязь мощности и энергии при передаче в режиме случайного множественного доступа с минимальной средней задержкой // Т-Comm: Телекоммуникации и транспорт. 2026. Том 20. №9. С. 36-46.

For citation:

A.A. Burkov, A.M. Turlikov, "Relationship between power and energy in random multiple access transmission with minimum average delay," T-Comm, 2026, vol. 20, no.9, pp. 36-46. (in Russian)

Введение

Долгое время исследования в области случайного множественного доступа велись независимо в двух направлениях [1,2]: теория информации множественного доступа [3] и теория случайного множественного доступа [4]. В рамках данных направлений исследовались разные аспекты систем с множеством источников и одним получателем.

В рамках теории случайного множественного доступа рассматривается случайное число абонентских устройств, передающих данные по общему каналу связи. В рамках этого направления исследуются процессы случайного поступления сообщений в систему, проводится анализ стабильности систем и задержек, определяемых коллизией сообщений. При этом не учитывается влияние шумов в канале связи.

В направлении теории информации множественного доступа не рассматривается процесс случайного появления сообщений в системе. Традиционно предполагается, что число абонентских устройств, передающих по общему каналу связи, фиксировано. В рамках этого направления исследуется влияние шумов в канале связи или наложения сигналов пользователей на скорость передачи в системе.

В последние года стали появляться работы рассматривающие системы случайного множественного доступа совмещающие как теорию информации множественного доступа, так и теорию случайного множественного доступа. Первыми работами в этом направлении можно считать [5, 6].

Так в работе [6] авторы рассматривают гауссовский канал множественного доступа (Gaussian multiple access channel) для сценария, где число абонентских устройств фиксировано и известно. При этом их число может быть большим, но передают сообщения только некоторое подмножество из них. Вероятность успешной доставки задается как требование к системе. Рассматривается несколько алгоритмов множественного доступа, включая АЛОХА. Исследуются границы достижимости отношения энергии на бит к спектральной плотности мощности шума. Описанные в работах модели не допускают повторной передачи сообщений. Как следствие в системе присутствуют потери сообщений, и не возникает проблемы, связанной с обеспечением стабильной работы системы.

В работе [7] рассматривается система случайного множественного доступа с алгоритмом АЛОХА и гауссовским каналом множественного доступа. Решается задача минимизации среднего возраста информации за счет выбора оптимальной длины блока передаваемой информации. Для вычисления вероятности ошибки используется приближенное выражение для формулы из работы [6].

В рассмотренных ранее работах исследуется алгоритм АЛОХА без учета применения к конкретным технологиям. Однако в настоящее время активно развивается технология беспроводных сетей 5G, в рамках которой исследуются перспективы применения 2х шаговой процедуры случайного доступа наряду со стандартной 4х шаговой процедурой. Предполагается, что применение 2х шаговой процедуры позволит уменьшить объем передаваемой служебной информации, что в свою очередь может позволить снизить задержки в системе. Как правило, в качестве алгоритма, работающего на первом этапе этих процедур, используется алгоритм типа АЛОХА. Так в работе [8] исследуются различные варианты перехода от 2х шаговой процедуры к 4х шаговой, и рассматривается

вопрос изменения пропускной способности и средней длительности процедуры разрешения коллизии для четырех вариантов перехода. Но в работах аналогичных [8] как правило, не исследуется влияние характеристик шумов в канале связи на работу алгоритма.

В рамках данной работы будет рассмотрена система случайного множественного доступа на основе алгоритма АЛОХА с гауссовским каналом множественного доступа. Предполагается наличие потенциально неограниченного числа абонентских устройств. Будет рассматриваться задача минимизации средней задержки в системе, и исследоваться взаимосвязь мощности и энергии при передаче с минимальной средней задержкой.

1 Модель системы со случайным доступом в канал

В работе рассматривается система с гауссовским каналом множественного доступа [6]. Предполагается, что имеется одна базовая станция и потенциально неограниченное число абонентских устройств (пользователей). Для передачи данных по общему ресурсу канала используется алгоритм случайного множественного доступа АЛОХА. Помимо коллизий при передаче сообщений на успешную передачу сообщения также влияет наличие в канале аддитивного белого гауссовского шума (АБГШ). Введем ряд допущений для рассматриваемой системы (похожая система допущений приводилась в работах [9,10]).

Допущение 1. Каждое сообщение, передаваемое абонентским устройством, содержит k бит информации.

Допущение 2. Предполагается, что используется некоторая кодово-модуляционная схема A , которая на основе k бит информации формирует сигнал, содержащий n отсчетов.

Допущение 3. Время работы системы разделено на окна одинаковой длительности, принятой за единицу времени. Каждое окно позволяет передавать n отсчетов, то есть одно сообщение, содержащее k бит информации. Число сообщений, появляющихся в системе за одно окно, распределено по закону Пуассона с параметром Λ [сообщений/окно].

Допущение 4. Рассматривается потенциально неограниченное количество абонентских устройств (в таком случае понятия абонентское устройство и сообщение можно считать эквивалентными). Абонентское устройство с готовым к отправке сообщением будет называться активным.

Допущение 5. Абонентские устройства и базовая станция имеют синхронизацию, как по отсчетам, так и по окнам. Временные интервалы пронумерованы, и окно с номером t соответствует временному интервалу $[t, t + 1)$.

Допущение 6. Рассматривается дискретный по времени канал связи с аддитивным белым гауссовским шумом, определяемый как

$$Y = \sum_{i=1}^{K_t} X_i + Z$$

где Y – сигнал на входе приемника базовой станции; K_t – количество абонентских устройств, которые передавали в окне с номером t , являющееся случайной величиной; X_i – сигнал i -го пользовательского устройства; Z – аддитивный белый гауссовский шум, а $Z \sim \mathcal{N}(0,1)$. Сигналы, передаваемые пользовательскими устройствами, содержат n отсчетов. Максимальная энергия каждого сигнала ограничена $X_i = nP$.

Допущение 7. Предполагается, что используются такие кодово-модуляционные схемы A , что при наложении пользовательских сигналов ни одно из сообщений в смеси не может быть успешно декодировано.

Допущение 8. При отсутствии наложения сообщений от разных абонентских устройств существует вероятность ошибки декодирования P_e , вызванная наличием в канале АБГШ. В таком случае с вероятностью P_e сообщение абонента не будет принято успешно.

Допущение 9. В обратном канале не возникает ошибок (или они всегда корректно исправляются), поэтому все абонентские устройства гарантированно и достоверно получают информацию от базовой станции.

Допущение 10. Абонентским устройствам и базовой станции достоверно известно число активных устройств $M_{акт}$. Каждое устройство независимо принимает решение о передаче в каждом очередном окне с вероятностью $p = \frac{1}{M_{акт}}$.

Примечание 1. Так как дисперсия шума в рассматриваемой модели равна 1 (см. допущение 6), то отношение сигнал/шум (SNR) в размах равно $P/1$. Таким образом, численное значение мощности сигнала P в ваттах совпадает с отношением сигнал/шум в размах.

Примечание 2. Все абоненты расположены на одинаковом расстоянии от базовой станции, передают с одинаковой мощностью P и находятся в одинаковых условиях радиоканала. Поэтому мощность сигнала от каждого устройства на приемнике одинакова и равна отношению сигнал/шум.

На рисунке 1 представлена схема, описывающая рассматриваемую систему случайного множественного доступа.

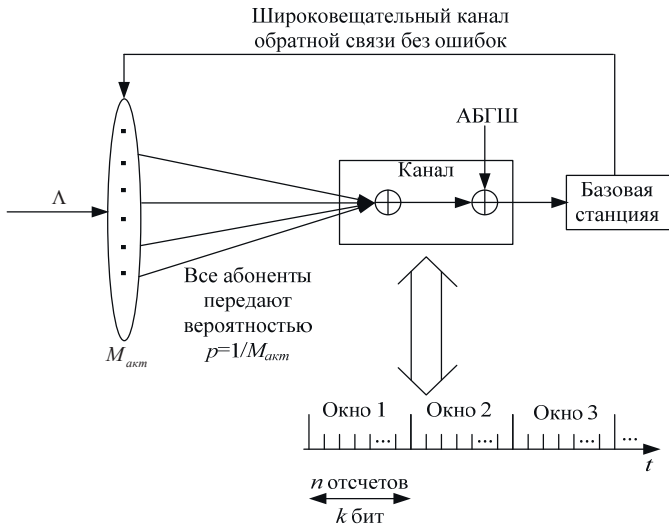


Рис. 1. Схема рассматриваемой системы

Описанная модель системы можно характеризовать следующей набором параметров:

- входная интенсивность Λ [сообщения/слот],
- число передаваемых бит информации k [биты],
- число отсчетов для передачи сигнала n [отсчёты],
- отношение сигнал/шум γ [разы],
- отношение энергии сигнала, приходящейся на один бит передаваемого сообщения E_b , к спектральной плотности шума N_0 – E_b/N_0 [энергия/бит].

Также введем в рассмотрение понятие спектральной эффективности ρ для системы с дискретным по времени каналом и интенсивностью входного потока Λ , определяющей интенсивность входного потока в битах на один отсчет. Взаимосвязь между Λ и ρ с учетом кодово-модуляционной схемы определяется следующим выражением [11]:

$$\rho[\text{бит/отсчет}] = \frac{\Lambda k}{n}. \quad (1)$$

Для систем Интернета вещей значение ρ лежит в диапазоне от доли бит на отсчет, до нескольких бит на отсчет [6]. В данной работе мы будем демонстрировать результаты для случая, когда $\rho=0.1$.

Из работы [12] следует, что существует код длины n отсчетов, содержащий k информационных символов, для которого вероятность ошибки декодирования при заданном значении отношения сигнал/шум γ не превышает величину P_e :

$$P_e(n, k, \gamma) = Q \left(\frac{n \frac{1}{2} \log_2(1 + \gamma) + \frac{1}{2} \log_2 n - k}{\sqrt{n \frac{\gamma}{2} \frac{\gamma + 2}{(\gamma + 1)^2} \log_2 e}} \right). \quad (2)$$

Ключевыми показателями для рассматриваемого типа систем являются [10]:

- среднее число активных абонентов в системе $\bar{N}$,
- средняя задержка $\bar{D}$,
- критическая интенсивность входного $\Lambda_{кр}$ – интенсивность входного потока, до которой система работает стабильно, а средняя задержка конечна [4]:

$$\Lambda_{кр} = \sup_{\Lambda} \{ \Lambda : \bar{D}(\Lambda) < \infty \}. \quad (3)$$

При $\Lambda < \Lambda_{кр}$ средняя задержка $\bar{D}$ и среднее число активных абонентов в системе $\bar{N}$ связаны через теорему Литтла [13]:

$$\bar{D} = \frac{\bar{N}}{\Lambda}. \quad (4)$$

2 Определение области стабильной работы системы

Рассматриваемая система работает стабильно (т.е. средняя задержка в системе конечна) при условии, что $\Lambda < \Lambda_{кр}$.

Для системы случайного доступа на основе алгоритма АЛОХА значение критической интенсивности совпадает с вероятностью того, в насыщенной системе сообщение успешно доставлено до получателя.

Наличие шумов в канале приводит к тому, что даже при отсутствии коллизии в канале с некоторой отличной от нуля вероятностью сообщение не будет принято успешно (см. допущение 8). Вероятность успешной передачи для рассматриваемой системы с учетом шумов в канале, вероятностью передачи p и вероятностью ошибки P_e может быть записана как:

$$\begin{aligned} \Lambda_{кр} &= \Pr \{ \text{успех} \mid M_{акт} \} \Pr \left\{ \begin{array}{l} \text{успешное} \\ \text{декодирование} \end{array} \middle| \text{успех} \right\} = \\ &= M_{акт} p (1 - p)^{M_{акт} - 1} (1 - P_e), \end{aligned}$$

При потенциально неограниченном числе абонентских

устройств с учетом допущения 10, критическая интенсивность для системы с учетом (2) можно вычислить по следующей формуле:

$$\Lambda_{kp} = \lim_{M_{акт} \rightarrow \infty} \left(1 - \frac{1}{M_{акт}}\right)^{M_{акт}-1} (1 - P_e(n, k, \gamma)) = e^{-1} (1 - P_e(n, k, \gamma)) \quad (5)$$

Интенсивность входного потока Λ при заданном значении спектральной эффективности ρ с учетом (1) может быть выражена как:

$$\Lambda = \frac{n\rho}{k} \quad (6)$$

Пусть изначально для рассматриваемой системы заданы следующие требуемые параметры:

- значение спектральной эффективности ρ ,
- число бит передаваемой в сообщении информации k ,
- отношение сигнал/шум γ .

И рассматривается задача выбора числа отсчетов n за которое будет передаваться сообщение. Существует набор значений $n \in (n_{min}, n_{max})$, при которых система работает стабильно, а средняя задержка конечна.

Область стабильной работы, а также значения n_{min} и n_{max} можно определить с использованием выражений (5) и (6). Система стабильна, если соблюдается условие $\Lambda < \Lambda_{kp}$ то есть:

$$\frac{n\rho}{k} < e^{-1} (1 - P_e(n, k, \gamma)) \quad (7)$$

На рисунке 2 представлен пример определения области стабильности системы для значений $k=100$ бит, $\gamma=1$ и $\rho=0.1$. Из рисунка 2 видно, что для заданного набора параметров неравенство (7) выполняется на отрезке от 194 отсчетов до 367 отсчетов. То есть $n_{min}=194$ и $n_{max}=367$, а система стабильна для всех $n \in (194, 367)$.

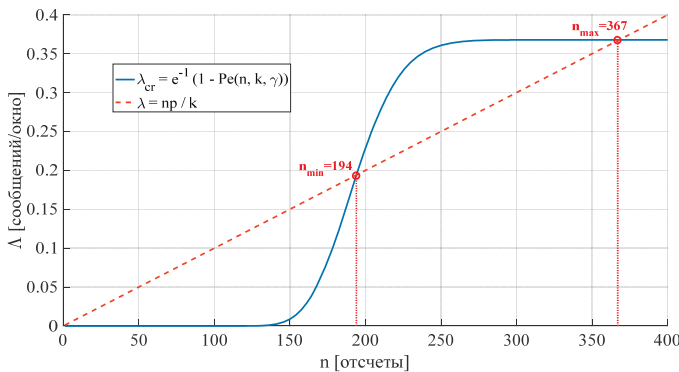


Рис. 2. Определение области стабильности для системы при $k=100$, $\gamma=1$ и $\rho=0.1$

Проведя аналогичные вычисления для различных значений ρ при фиксированных значениях числа информационных бит k и отношения сигнал/шум можно показать, что чем ниже требуемое значение ρ , тем больший диапазон области стабильной работы системы по параметру числа отсчетов n .

В результате, проведенный анализ области стабильности рассматриваемой системы, показано, что существует

некоторый набор значений n , при которых система работает стабильно, а средняя задержка конечна. Однако число отсчетов n влияет как на вероятность ошибки, вызванную шумами в канале (с увлечением числа отсчетов вероятность такой ошибки уменьшается [12]), так и на входную интенсивность. Увеличение числа отсчетов повышает входную интенсивность, что в свою очередь увеличивает среднее число активных абонентов в системе и повышает вероятность коллизии. Можно предположить, что существует такое значение числа отсчетов n , при котором для заданных параметров системы средняя задержка будет минимальной.

3 Анализ средней задержки в системе

Как отмечалось в предыдущем разделе выбор числа отсчетов n , за которое передается сообщение, может влиять на среднюю задержку в системе. Для того, что бы рассмотреть задачу снижения средней задержки за счет выбора оптимального параметра n необходимо рассмотреть способы вычисления или оценки данной характеристики системы.

Система случайного доступа на основе алгоритма АЛОХА с потенциально неограниченным числом абонентских устройств может быть описана марковской цепью со счетным числом состояний [4]. Состояние данной марковской цепи отражает число активных абонентов в системе $M_{акт}$. Если система стабильна, то марковская цепь эргодична и для нее может быть найдено стационарное распределение $\bar{\pi} = \{\pi_0, \pi_1, \pi_2, \dots\}$ где π_i вероятность того, что в системе имеется i активных абонентов. Тогда среднее число абонентов в системе можно вычислить как:

$$\bar{N} = \sum_{i=0}^{\infty} i \pi_i.$$

И далее на основе теоремы Литтла получить среднюю задержку. Однако нахождение стационарного распределения для рассматриваемой системы является достаточно сложной задачей.

В работе [14] для системы с алгоритмом АЛОХА при отсутствии шумов в канале предложен подход для вычисления верхней и нижней границы средней задержки. Идея основана на введении двух вспомогательных систем, первая из которых позволяет вычислить нижнюю границу, а вторая верхнюю. В методе используется параметр K , влияющий на точность получаемых оценок, и с его увеличением интервал между границами уменьшается, что позволяет получить оценку средней задержки системы с требуемой точностью [14].

На основе работы [14] покажем способ вычисления верхней и нижней границ для средней задержки в рассматриваемой системе при учете наличия АБГШ канале связи.

Основным отличием рассматриваемой системы от системы из [14] является то, что вероятность успеха определяется

$$\text{как } \left(1 - \frac{1}{M_{акт}}\right)^{M_{акт}-1} (1 - P_e(n, k, \gamma)).$$

Параметр, влияющий на точность получаемых оценок, определим как K (аналогично работе [14]). Идея заключается в том, что для состояний марковской цепи начиная с некоторого состояния K , меняется вероятность передачи сообщения абонентами, что влияет на вероятность события успех.

Вероятность передачи задается таким образом, что начиная с состояния K вероятность успеха остается постоянной. Для системы, задающей верхнюю границу, эта вероятность ниже, чем у исходной, а для системы, задающей нижнюю границу, выше, чем в исходной.

Так для системы, определяющей верхнюю границу, вероятность успеха будет определяться по следующему правилу:

$$\Pr\{\text{успех} | i\} = \begin{cases} \left(1 - \frac{1}{i}\right)^{i-1} (1 - P_e(n, k, \gamma)), & i \leq K \\ e^{-1} (1 - P_e(n, k, \gamma)), & i > K \end{cases}$$

Для системы, определяющей нижнюю границу, вероятность успеха определяется по следующему правилу:

$$\Pr\{\text{успех} | i\} = \begin{cases} \left(1 - \frac{1}{i}\right)^{i-1} (1 - P_e(n, k, \gamma)), & i \leq K \\ \left(1 - \frac{1}{K}\right)^{K-1} (1 - P_e(n, k, \gamma)), & i > K \end{cases}$$

Далее с учетом введенных вероятностей успеха формируется система линейных уравнений, содержащая $K+1$ уравнение.

В общем виде получение системы уравнений может быть представлено в виде следующих шагов:

Шаг 1. Составляется $K+1$ стандартных уравнений для стационарных вероятностей $\pi_0, \pi_1, \pi_2, \dots, \pi_K$;

Шаг 2. Предпоследнее уравнение заменяется на условие нормировки: $\pi_0 + \pi_1 + \pi_2 + \dots + \pi_K = 1$;

Шаг 3. Последнее уравнение заменяется в зависимости от того какая вспомогательная система рассматривается:

– Для первой вспомогательной системы (верхняя граница):

$$(1 - P_e(n, k, \gamma)) \left(\sum_{k=1}^K \left(1 - \frac{1}{k}\right)^{k-1} \pi_k + e^{-1} \pi_K \right) = \Lambda.$$

– Для второй вспомогательной системы (нижняя граница):

$$(1 - P_e(n, k, \gamma)) \left(\sum_{k=1}^K \left(1 - \frac{1}{k}\right)^{k-1} \pi_k + \left(1 - \frac{1}{K}\right)^{K-1} \pi_K \right) = \Lambda.$$

Решив полученную систему линейных уравнений, получим набор стационарных вероятностей для соответствующей вспомогательной системы.

Далее повторим рассуждения из работы [14]. Для рассматриваемой в данной работе системы можно показать, что верхняя граница для средней задержки может быть получена как:

$$D_U = \frac{\psi_U(\Lambda, K)}{2\Lambda(e^{-1}(1 - P_e(n, k, \gamma)) - \Lambda)} + 0.5, \quad (8)$$

где

$$\psi_U(\Lambda, K) = \Lambda^2 + \Lambda - (1 - P_e(n, k, \gamma)) \left(2 \sum_{k=1}^K k \left(1 - \frac{1}{k}\right)^{k-1} \pi_k + 2e^{-1} \sum_{k=1}^K k \pi_k + (1 - 2\Lambda) \left(\sum_{k=1}^K \left(1 - \frac{1}{k}\right)^{k-1} \pi_k + e^{-1} \pi_K \right) \right)$$

Формула для вычисления нижней границы средней задержки вычисляется по формуле:

$$D_L = \frac{\psi_L(\Lambda, K)}{2\Lambda \left(\left(1 - \frac{1}{K}\right)^{K-1} (1 - P_e(n, k, \gamma)) - \Lambda \right)} + 0.5, \quad (9)$$

где

$$\psi_L(\Lambda, K) = \Lambda^2 + \Lambda - (1 - P_e(n, k, \gamma)) \left(2 \sum_{k=1}^K k \left(1 - \frac{1}{k}\right)^{k-1} \pi_k + 2 \left(1 - \frac{1}{K}\right)^{K-1} \sum_{k=1}^K k \pi_k + (1 - 2\Lambda) \left(\sum_{k=1}^K \left(1 - \frac{1}{k}\right)^{k-1} \pi_k + \left(1 - \frac{1}{K}\right)^{K-1} \pi_K \right) \right)$$

Скорость схождения верхней и нижней границ для средней задержки от параметра K аналогична, представленной в работе [14].

А рисунке 3 приведены зависимости для средней задержки в системе, полученные следующими способами: имитационное моделирование, расчет верхней границы по формуле (8) и расчет нижней границы по формуле (9). Представленные результаты получены в предположении, что $P_e(n, k, \gamma) = 0.4$.

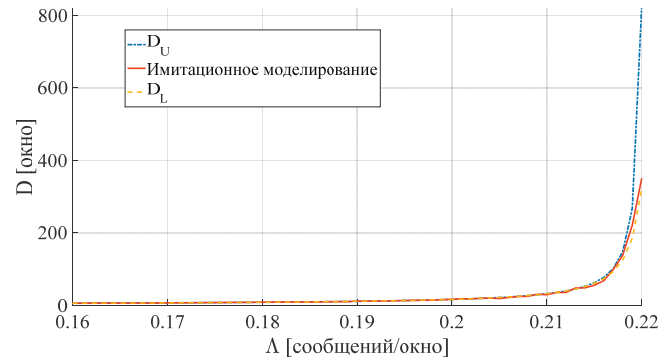


Рис. 3. Зависимость для средней задержки, полученная разными методами при $P_e(n, k, \gamma) = 0.4$

Оценка средней задержки, полученная с помощью имитационного моделирования, при выборе достаточного объема выборки (или времени моделирования) может давать результат с любой требуемой точностью. Однако такой подход является ресурсозатратным, поэтому не подходит в качестве способа, применимого при решении задачи минимизации средней задержки. Верхняя и нижняя границы при подобранном параметре K дают хорошее приближение средней задержки в системе. В дальнейшем при формулировке и решении оптимизационной задачи минимизации средней задержки за счет выбора оптимального значения параметра системы n будет использоваться верхняя граница. Это обусловлено тем, что в случае рассмотрения некоторого ограничения на значение средней задержки можно гарантировать, то, что в рассматриваемой системе её значение не будет выше этой величины.

Далее в работе все результаты будут представлены для параметра оценки $K=100$.

4 Задача минимизации средней задержки

Далее используя верхнюю оценку для средней задержки (см. формулу (8)) сформулируем оптимизационную задачу для минимизации средней задержки в системе по числу отсчетов n .

Предполагается, что изначально заданы: k, ρ, γ , а также подобран параметр оценки K . Средняя задержка, полученная с помощью выражения (9) имеет размерность в окнах. Для формулировки оптимизационной задачи значения средней задержки переводятся в отсчеты путем умножения значения в окнах на число отсчетов в окне:

$$D_U[\text{отсчеты}] = n \left(\frac{\psi_U(\Lambda, K)}{2\Lambda(e^{-1}(1 - P_e(n, k, \gamma)) - \Lambda)} + 0.5 \right),$$

Тогда оптимизационная задача может быть сформулирована следующим образом:

Заданы: k, ρ, γ, K

Минимизировать: $D_U(k, \rho, \gamma, K, n)$ по n

При условиях: $\Lambda = \frac{n\rho}{k}$,

$$P_e(n, k, \gamma) = Q \left(\frac{n \frac{1}{2} \log_2(1 + \gamma) + \frac{1}{2} \log_2 n - k}{\sqrt{n \frac{\gamma}{2} \frac{\gamma + 2}{(\gamma + 1)^2} \log_2 e}} \right),$$

$$n \in (n_{\min}, n_{\max}).$$

Значения $n_{\min}$ и $n_{\max}$ вычисляются для набора k, ρ и γ по методике, описанной в разделе 2.

Решением сформулированной оптимизационной задачи является значение n_0 , которое при заданных k, ρ и γ минимизирует среднюю задержку в системе.

Пример решения данной оптимизационной задачи приведен на рисунке 4 при $k=100, \rho=0.1$ и $\gamma=1$. Средняя задержка на графике представлена в отсчетах. В результате решения сформулированной оптимизационной задачи минимальное значение средней задержки достигается при числе отсчетов $n_0=236$.

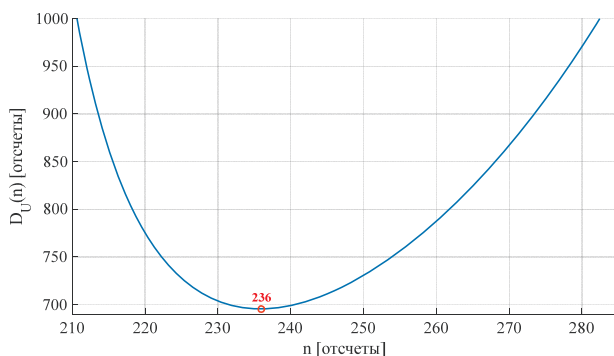


Рис. 4. Решение оптимизационной задачи при $k=100, \rho=0.1$ и $\gamma=1$

5 Взаимосвязь средней задержки и E_b/N_0

Далее рассмотрим взаимосвязь между средней задержкой и затратами энергии на бит E_b/N_0 .

Взаимосвязь между отношением сигнал/шум γ и энергией на бит E_b/N_0 для системы с повторными передачами можно записать в следующем виде [11]:

$$\gamma = \frac{2k \frac{E_b}{N_0}}{Sn}, \quad (10)$$

где S – среднее число передач сообщения. Из работы [11] следует, что данную величину можно вычислить как:

$$S = \frac{1 - \pi_0}{\Lambda}. \quad (11)$$

Как следует из формулы (11) для расчета S необходимо вычислить распределение стационарных вероятностей марковской цепи. В свою очередь для построения матрицы переходных вероятностей этой цепи необходимо вычислить вероятность ошибки в канале $P_e(n, k, \gamma)$, вызванной наличием АБГШ в канале. А для ее вычисления нужно знать значение γ , которое не получить без расчета S .

Для решения возникшей проблемы предлагается подход, который можно описать следующим алгоритмом:

Шаг 1. Задается отношение сигнал/шум γ .

Шаг 2. Определяется область стабильной работы системы для получения значений $n_{\min}$ и $n_{\max}$.

Шаг 3. В пределах области стабильности решается оптимизационная задача для поиска n_0 , которое минимизирует среднюю задержку. В процессе решения оптимизационной задачи для всех $n \in (n_{\min}, n_{\max})$ в том числе и n_0 вычисляется стационарное распределение марковской цепи. В таком случае для марковской цепи, соответствующей значению n_0 имеется набор стационарных вероятностей.

Шаг 4. Используя этот набор стационарных вероятностей для марковской цепи при значении n_0 , вычисляем среднее число передач сообщения по формуле (11).

Шаг 5. Рассчитываем значение энергии на бит E_b/N_0 с учетом результатов решения оптимизационной задачи и формулы (10) как:

$$\frac{E_b}{N_0} = \frac{n_0 S \gamma}{2k}. \quad (12)$$

Используя предложенный алгоритм можно получить зависимость между E_b/N_0 и D_U для системы случайного доступа с потенциально неограниченным числом абонентских устройств. Такую систему можно охарактеризовать как передачу от многих источников к одному получателю.

Далее рассмотрим систему с одним источником и одним получателем, что бы провести их сравнительный анализ и рассмотреть влияние случайного доступа на характеристики системы.

6 Система с одним источником и одним получателем

Рассмотрим подход для получения взаимосвязи между средней задержкой и затратами энергии на бит для системы с одним источником и одним получателем. В такой система абонент всегда передает с вероятностью равной единице. Так как коллизий в такой системе нет, то вероятность успешной передачи зависит только от вероятности ошибки, возникающей из-за присутствия в канале АБГШ.

Для начала покажем, как для такой системы вычислить значение средней задержки. Система допущений в этом случае аналогична допущениям для системы со случайным доступом, за исключением допущения 4 и 10, так как в системе с 1 абонентом их можно опустить.

С точки зрения расчета средней задержки такую систему можно рассмотреть, как систему массового обслуживания с Пуассоновским входным потоком с параметром Λ и с временем обслуживания, распределенным по геометрическому закону с параметром $1 - P_e(n, k, \gamma)$. Средняя задержка для такой системы вычисляется на основе формулы Поллачека-Хинчина для системы M|G|1 [15]. Тогда с учетом того, что время обслуживания распределено по геометрическому закону и процесс передачи может происходить только в начале окна, то средняя задержка в окнах вычисляется по следующей формуле:

$$\bar{D} = \frac{2 - \Lambda}{2((1 - P_e(n, k, \gamma)) - \Lambda)} + 0.5.$$

С учетом системы допущений и формулы (1) данное выражение можно записать как:

$$\bar{D}(n, k, \rho, \gamma) = \frac{2 - \frac{n\rho}{k}}{2\left((1 - P_e(n, k, \gamma)) - \frac{n\rho}{k}\right)} + 0.5.$$

Данная формула дает значение средней задержки в окнах. Для получения средней задержки в отсчетах необходимо умножить ее на n :

$$\bar{D}(n, k, \rho, \gamma)[\text{отсчеты}] = n \left(\frac{2 - \frac{n\rho}{k}}{2\left((1 - P_e(n, k, \gamma)) - \frac{n\rho}{k}\right)} + 0.5 \right).$$

Для решения оптимизационной задачи, аналогично подходу для системы с множеством источников и одним получателем, необходимо определить область стабильной работы системы.

Вероятность успеха или критическая интенсивность входного потока для рассматриваемой системы определяется следующим выражением:

$$\Lambda_{кр} = \Pr\{\text{успешное декодирование}\} = 1 - P_e(n, k, \gamma).$$

Входная интенсивность определяется аналогично предыдущей системе по формуле (6).

Система работает стабильно, когда $\Lambda < \Lambda_{кр}$ то есть выполняется неравенство:

$$\frac{n\rho}{k} < 1 - P_e(n, k, \gamma).$$

Если построить графики этих функций от n , то мы получим две точки пересечения, между которыми система работает стабильно (аналогично системе с множеством источников и одним получателем). Как и раньше назовем их n_{min} и n_{max} , а решение задачи минимизации задержки будем рассматривать для $n \in (n_{min}, n_{max})$.

На рисунке 5 представлено графическое определение области стабильности для системы с одним источником и

одним получателем при $k=100$, $\gamma=1$ и $\rho=0.1$. Из рисунка видно, что для заданного набора параметров указанное ранее неравенство выполняется на отрезке от 171 отсчета до 999 отсчетов. То есть $n_{min}=171$ и $n_{max}=999$, а система стабильна для всех $n \in (171, 999)$. Уменьшение значения ρ увеличивает диапазон, в котором система работает стабильно.

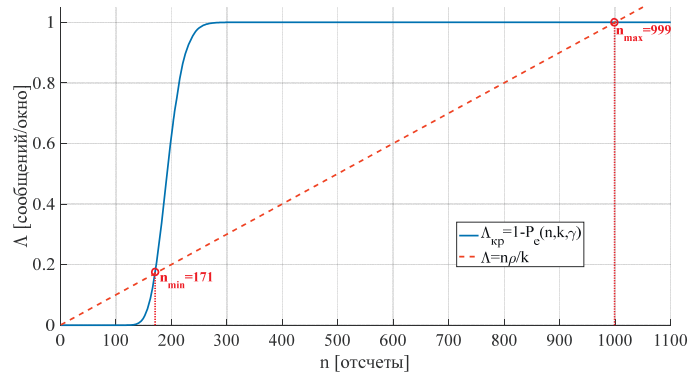


Рис. 5. Определение области стабильности для системы содним источником и одним получателем при $k=100$, $\gamma=1$ и $\rho=0.1$

Как и для предыдущей системы, предполагается, что существует такое значение числа отсчетов n_0 , при котором средняя задержка в системе минимальна.

Сформулируем оптимизационную задачу для данной системы:

Заданы : k, ρ, γ

Минимизировать: $\bar{D}(k, \rho, \gamma, n)$ по n

При условиях : $\Lambda = \frac{n\rho}{k}$,

$$P_e(n, k, \gamma) = Q \left(\frac{n \frac{1}{2} \log_2(1 + \gamma) + \frac{1}{2} \log_2 n - k}{\sqrt{n \frac{\gamma}{2} \frac{\gamma + 2}{(\gamma + 1)^2} \log_2 e}} \right),$$

$$n \in (n_{min}, n_{max}).$$

На рисунке 6 представлен пример решения оптимизационной задачи для $k=100$, $\rho=0.1$ и $\gamma=1$. Средняя задержка на графике представлена в отсчетах. В результате решения сформулированной оптимизационной задачи минимальное значение средней задержки достигается при числе отсчетов $n_0=231$.

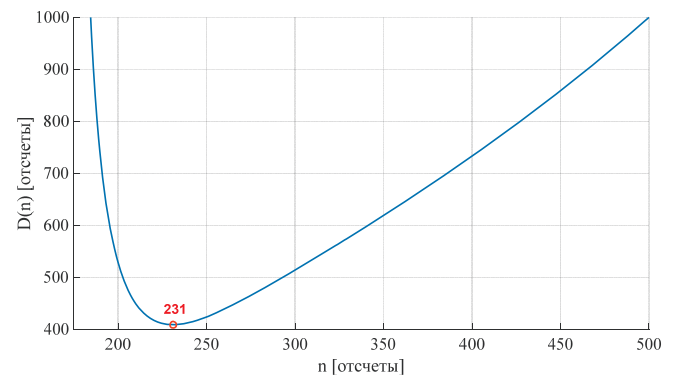


Рис. 6. Решение оптимизационной задачи для системы с одним источником и одним получателем при $k=100$, $\rho=0.1$ и $\gamma=1$

Для получения значений энергии на бит E_b/N_0 будет использоваться методика, описанная в разделе 5. Для этого нужно вычислить среднее число передач сообщения S до успешной отправки.

Покажем, как можно вычислить среднее число передач для системы с одним источником и одним получателем. Число передач, которое понадобится абоненту, чтобы передать сообщение распределена по геометрическому закону с параметром $1 - P_e(n, k, \gamma)$. Следовательно, среднее число передач вычисляется по следующей формуле:

$$S = \frac{1}{1 - P_e(n, k, \gamma)}.$$

Таким образом, мы получили все необходимые величины для сравнения системы с одним источником и одним получателем и системы, рассматриваемой в работе, с множеством источников и одним получателем.

8 Сравнение систем

На рисунке 7 представлена зависимость средней задержки от значения отношения сигнал/шум γ , полученная при решении оптимизационной задачи для обеих систем. Результаты представлены для $k=100$ и $\rho=0.1$. Из графиков видно, что с увеличением значения отношения сигнал/шум средняя задержка в системе уменьшается для обеих систем. Система с множеством источников требует больших значений отношения сигнал/шум для обеспечения эквивалентного значения средней задержки. С увеличением отношения сигнал/шум разница в средней задержке систем уменьшается.

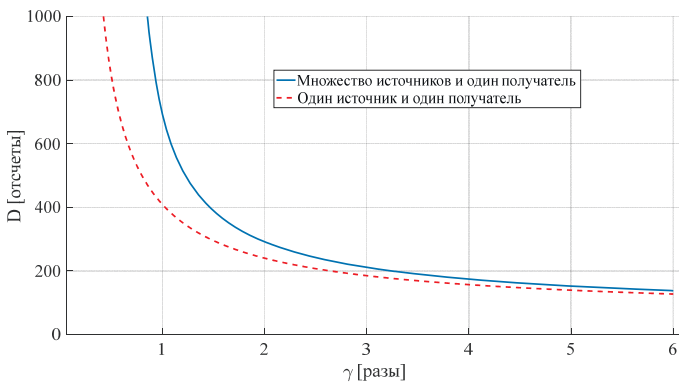


Рис. 7. Зависимость средней задержки от γ при $k=100$ и $\rho=0.1$

На рисунке 8 приведена зависимость значения n_0 от значения отношения сигнал/шум γ для обеих систем. Результаты представлены для $k=100$ и $\rho=0.1$. Значения числа отсчетов, обеспечивающие минимум средней задержки для обеих систем, имеют близкие значения. Для системы с одним источником и одним получателем оптимальное значение числа отсчетов меньше, чем для системы с множеством источников. Наибольшая разница в числе отсчетов при малых значениях отношения сигнал/шум и составляет порядка 4 отсчетов. С ростом отношения сигнал/шум эта разница уменьшается и составляет один отсчет.

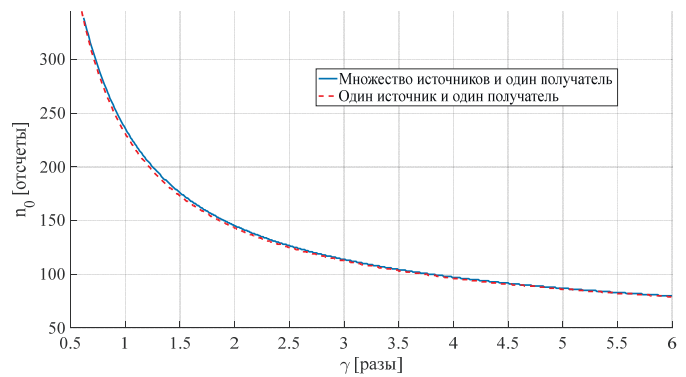


Рис. 8. Зависимость n_0 от γ при $k=100$ и $\rho=0.1$

Далее рассмотрим взаимосвязь между значением отношения сигнал/шум γ и энергии на бит E_b/N_0 для обеих систем при $k=100$ и $\rho=0.1$. Данная зависимость приведена на рисунке 9.

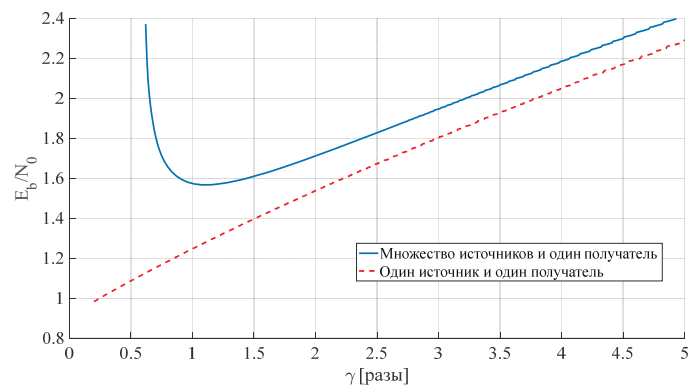


Рис. 9. Зависимость E_b/N_0 от γ для двух систем при минимизации средней задержки ($k=100$ и $\rho=0.1$)

Как видно из рисунка 9 для системы с одним источником и одним получателем затраты энергии на бит оказываются ниже, чем для системы со случайным множественным доступом. При этом зависимость является монотонно возрастающей функцией, что обусловлено отсутствием коллизий в системе. Для системы со случайным доступом (множество источников и один получатель) с ростом значения отношения сигнал/шум сначала убывает, а затем начинает возрастать. То есть имеется точка минимума по значению энергии на бит E_b/N_0 .

Поясним такое поведение на качественном уровне. При малых отношениях сигнал/шум для успешной передачи сообщения даже в отсутствии коллизий требуется большое число отсчетов. Чем больше отсчетов в одном окне передачи, тем выше входная интенсивность сообщений на окно при фиксированном значении спектральной эффективности. Вследствие чего в системе высокая вероятность коллизии и соответственно большее значение среднего числа передач. Все это в совокупности дает высокое значение затрат энергии на бит для успешной передачи сообщения. С ростом отношения сигнал/шум уменьшается, как длина окна в отсчетах, так и вероятность коллизии со средним числом передач. Однако после некоторого момента уменьшение числа отсчетов начинает давать незначительное снижение среднего числа передач, так как вероятность коллизии практически не уменьшается.

При этом затраты энергии на бит продолжают возрастать вследствие повышения мощности сигнала. Все это в совокупности приводит к тому, что зависимость энергозатрат от мощности не является монотонно возрастающей.

На рисунке 10 приведена взаимосвязь между энергией на бит и средней задержкой в системе. Для системы с одним источником и одним получателем с увеличением средней задержки уменьшается и затраты энергии на бит. Для системы с множественным доступом с увеличением задержки энергия на бит сначала уменьшается, а потом снова начинает возрастать, то есть эта зависимость имеет точку минимума по энергии на бит. Этот эффект обусловлен поведением зависимостей представленных на рисунке 9.

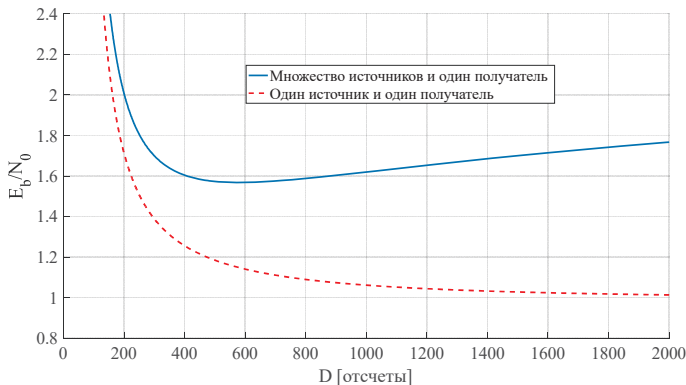


Рис. 10. Взаимосвязь между E_b/N_0 и D для двух систем при минимизации D ($k=100$ и $\rho=0.1$)

Таким образом возник нетривиальный эффект для системы с множеством источников и одним получателем: при определенных условиях рост отношения сигнал/шум сопровождается снижением затрат энергии на бит, которое затем сменяется их увеличением. Аналогичный эффект возникает и при увеличении значений средней задержки. Таким образом существует минимальное значение энергии на бит при котором имеется некоторое значение средней задержки. При необходимости уменьшения средней задержки необходимо увеличивать затраты энергии на бит. При этом ослабление требований по средней задержке (возможность ее увеличения) после точки минимума не позволяет уменьшить затраты энергии на бит.

7

8 Заключение

В работе рассмотрена модель системы случайного множественного доступа на основе алгоритма АЛОХА с гауссовским каналом множественного доступа. В которой все абоненты передают сообщения, содержащие k бит, с помощью сигналов одинаковой длительности, содержащих n отсчетов. Отношение сигнал/шум для всех абонентов одинаковое и равно некоторому значению γ . Эта модель отражает основные особенности современных беспроводных систем, в которых большое число абонентских устройств передают данные на базовую станцию в режиме случайного доступа и используют одну и ту же кодово-модуляционную схему. Основной особенностью модели является наличие потенциально неограниченного числа абонентских устройств, что может приводить к нестабильной работе системы.

Был описан метод определения области стабильной работы системы. Сформулирована и решена оптимизационная задача по минимизации средней задержки по числу отсчетов n в рамках области стабильной работы системы. Было показано, что при фиксированных отношении сигнал/шум γ , числе информационных бит k и спектральной эффективности ρ существует единственное значение числа отсчетов n_0 , при котором средняя задержка минимальна.

Описана проблема для получения зависимости между затратами энергии на бит и средней задержкой. Для её решения предложен алгоритм, позволяющий определить взаимосвязь между этими характеристиками. В рамках данного алгоритма также можно получить взаимосвязь между мощностью и затратами энергии.

При исследовании взаимосвязи между мощностью и затратами энергии был получен нетривиальный эффект: при определенных условиях рост отношения сигнал/шум сопровождается снижением затрат энергии на бит, которое затем сменяется увеличением.

При исследовании взаимосвязи между энергией на бит и средней задержкой в системе было отмечено, что существует такое значение средней задержки, при котором затраты энергии на передачу минимальные. Если в системе требуется уменьшить значение средней задержки, то в таком случае необходимо увеличить энергозатраты.

Для сравнения была исследована система с одним источником и одним получателем, и было показано, что в такой системе отсутствуют особенности, выявленные для системы с множеством источников и одним получателем. При этом затраты энергии и средняя задержка в такой системе всегда ниже, чем для системы со случайным доступом к каналу.

Полученные в работе результаты позволяют выработать рекомендации по выбору режимов работы абонентских устройств систем Интернета Вещей, в которых имеется ограничения на среднее время доставки сообщения от абонента до базовой станции, а сами устройства имеют автономные источники питания, что делает актуальной задачу уменьшения затрат энергии на передачу.

Финансовая поддержка

Статья подготовлена при финансовой поддержке Российского научного фонда, проект № 22-19-00305-П <https://rscf.ru/project/22-19-00305/> "Пространственно-временные стохастические модели беспроводных сетей с большим числом пользователей".

Литература

1. Mathys P. A class of codes for a t active users out of n multiple-access communication system // IEEE Transactions on Information Theory, 1990, vol. 36, no. 6, pp. 1206–1219.
2. Gallager R. G. A perspective on multiaccess channels // IEEE Transactions on Communications, 1985 vol. 31, pp. 124 – 142.
3. El Gamal A., Kim Y. H. Network information theory // Cambridge, UK: Cambridge university press, 2011, pp. 509–611.
4. Tsybakov B. Survey of USSR contributions to random multiple-access communications // IEEE transactions on information theory, 1985, T. 31, № 2, pp. 143–165.
5. Y. Polyanskiy A perspective on massive random-access // 2017 IEEE International Symposium on Information Theory (ISIT). IEEE, 2017, pp. 2523–2527.

6. Ordentlich O., Polyanskiy Y. Low complexity schemes for the random access gaussian channel // in 2017 IEEE International Symposium on Information Theory (ISIT). IEEE, 2017, pp. 2528-2532.
7. Kim Y. Minimizing age of information in slotted ALOHA with short-packet communications //IEEE Access, 2024, T. 12. C. 186779-186793.
8. Рачугин Р. О. Анализ вариантов возврата от 2-шаговой к 4-шаговой процедуре случайного доступа // Т-Comm: Телекоммуникации и транспорт, 2025, №11. С. 33-40.
9. Burkov A. et al. Achievability Bounds for Massive Random Access in the Gaussian MAC with Delay Constraints //2019 XVI International Symposium "Problems of Redundancy in Information and Control Systems"(REDUNDANCY), IEEE, 2019, pp. 224-227.
10. Burkov A. A. Signal power and energy-per-bit optimization problems in mMTC systems // Информационно-управляющие системы, 2021, Т. 5, pp. 51-58.
11. Burkov A., Shneer S., Turlikov A. An achievability bound of energy per bit for stabilized massive random access Gaussian channel // IEEE Communications Letters, 2020, T. 25, №. 1, pp. 299-302.
12. Polyanskiy Y., H. V. Poor S. Verd'u. Channel coding rate in the finite blocklength regime // IEEE Transactions on Information Theory, 2010, T. 56, № 5, pp. 2307-2359.
13. Br'emaud P. Probability theory and stochastic processes // Springer, 2020.
14. Burkov A. A., Shneer S., Turlikov A. M. Arbitrarily Accurate Approximation of Numerical Characteristics of Stationary ALOHA Channels // 2021 Wave Electronics and its Application in Information and Telecommunication Systems (WECONF), IEEE, 2021, pp. 1-8.
15. Клейнрок Л. Вычислительные системы с очередями: Пер. с англ. М.: Мир, 1979.

RELATIONSHIP BETWEEN POWER AND ENERGY IN RANDOM MULTIPLE ACCESS TRANSMISSION WITH MINIMUM AVERAGE DELAY

Artem A. Burkov, Saint Petersburg State University of Aerospace Instrumentation, Saint Petersburg, Russia, a.burkov@guap.ru
Andrey M. Turlikov, Saint Petersburg State University of Aerospace Instrumentation, Saint Petersburg, Russia, turlikov@guap.ru

Abstract

For a long time, research into random multiple access proceeded along two independent lines: multiple access information theory and random multiple access theory. In recent years, studies combining these two approaches have begun to appear. However, existing literature typically fails to jointly address the relationship between power and energy expenditure (accounting for retransmissions) and the minimization of average delay in system. This paper investigates that relationship. The study examines a random multiple access system based on the ALOHA algorithm utilizing a Gaussian multiple-access channel. The system is assumed to have a potentially unlimited number of user devices. An approach is described for determining the stable operating region where the average delay remains finite. A method is proposed for calculating upper and lower bounds on the average delay with any desired degree of accuracy. An optimization problem is formulated and numerically solved to minimize average system delay by selecting the user message duration, given a fixed number of bits per message and a specified transmission power. Based on the solution to this optimization problem, an approach is proposed to establish the relationship between power and energy expenditure for random multiple access transmission with minimal average delay. It is shown that the dependence of energy expenditure on power is not a monotonically increasing function and there exists a specific power level that minimizes energy expenditure. A comparison is made with a system comprising a single source with an infinite message queue and a single receiver – a setup where collisions do not occur. Consequently, in that system, energy expenditure is a monotonically increasing function of power. Furthermore, energy expenditure in the random access system is higher than in the single source system at any power level.

Keywords: Random multiple access, ALOHA, signal-to-noise ratio, energy per bit, average delay, stability, Markov chain, minimization problem

References

- [1] P. Mathys, "A class of codes for a t active users out of n multiple-access communication system," *IEEE Transactions on Information Theory*, 1990, vol. 36, no. 6, pp. 1206-1219.
- [2] R. G. Gallager, "A perspective on multiaccess channels," *IEEE Transactions on Communications*, 1985 vol. 31, pp. 124-142.
- [3] A. El Gamal, Y. H. Kim, "Network information theory," Cambridge, UK: Cambridge university press, 2011, pp. 509-611.

- [4] B. Tsybakov, "Survey of USSR contributions to random multiple-access communications," *IEEE transactions on information theory*, 1985, Vol. 31, no. 2, pp. 143-165.
- [5] Y. Polyanskiy, "A perspective on massive random-access," *2017 IEEE International Symposium on Information Theory (ISIT)*. IEEE, 2017, pp. 2523-2527.
- [6] O. Ordentlich, Y. Polyanskiy, "Low complexity schemes for the random access gaussian channel," in *2017 IEEE International Symposium on Information Theory (ISIT)*. IEEE, 2017, pp. 2528-2532.
- [7] Y. Kim, "Minimizing age of information in slotted ALOHA with short-packet communications," *IEEE Access*, 2024. Vol. 12, pp. 186779-186793.
- [8] R.O. Rachugin, "Analysis types of backoff from 2-step to 4-step random access procedure," *T-Comm*, 2025, vol. 19, no. 11, pp. 33-40. (in Russian)
- [9] Burkov A. et al., "Achievability Bounds for Massive Random Access in the Gaussian MAC with Delay Constraints," *2019 XVI International Symposium "Problems of Redundancy in Information and Control Systems" (REDUNDANCY)*, IEEE, 2019, pp. 224-227.
- [10] A. A. Burkov, "Signal power and energy-per-bit optimization problems in mMTC systems," *Information and Control Systems*, 2021, no. 5, pp. 51-58.
- [11] A. Burkov, S. Shneer, A. Turlikov, "An achievability bound of energy per bit for stabilized massive random access Gaussian channel," *IEEE Communications Letters*, 2020. Vol. 25, no. 1, pp. 299-302.
- [12] Y. Polyanskiy, H. V. Poor, S. Verd'u, "Channel coding rate in the finite blocklength regime," *IEEE Transactions on Information Theory*, 2010. Vol. 56, no. 5, pp. 2307-2359.
- [13] P. Br'emaud, "Probability theory and stochastic processes," Springer, 2020.
- [14] A. A. Burkov, S. Shneer, A. M. Turlikov, "Arbitrarily Accurate Approximation of Numerical Characteristics of Stationary ALOHA Channels," *2021 Wave Electronics and its Application in Information and Telecommunication Systems (WECONF)*, IEEE, 2021, pp. 1-8.
- [15] L. Kleinrock, "Computer applications," *Queueing systems*, 1976, no. 3.

Information about authors:

Artem A. Burkov, PhD, associate professor of Department of Infocommunication Technologies and Communication Systems in Saint Petersburg State University of Aerospace Instrumentation, Saint Petersburg, Russia. ORCID 0000-0002-0920-585X

Andrey M. Turlikov, professor, Doctor of Technical Sciences, head of the Department of Infocommunication Technologies and Communication Systems in Saint Petersburg State University of Aerospace Instrumentation, Saint Petersburg, Russia. ORCID 0000-0001-7132-094X

МЕТОДИКА ПРИМЕНЕНИЯ АРГМОР ДЛЯ ПРОГРАММНОГО РАЗДЕЛЕНИЯ СЕТЕВЫХ РЕСУРСОВ В РАСПРЕДЕЛЕННОЙ ВЫЧИСЛИТЕЛЬНОЙ СЕТИ

DOI: 10.36724/2072-8735-2026-20-9-47-58

Петров Алексей Юрьевич,

Санкт-Петербургский государственный университет
аэрокосмического приборостроения,
Санкт-Петербург, Россия, palexex1998@yandex.ru

Оленев Валентин Леонидович,

Санкт-Петербургский государственный университет
аэрокосмического приборостроения,
Санкт-Петербург, Россия, valentin.olenev@guap.ru

Тюрликов Андрей Михайлович,

Санкт-Петербургский государственный университет
аэрокосмического приборостроения,
Санкт-Петербург, Россия, turlikov@guap.ru

Manuscript received 30 May 2026;

Accepted 02 August 2026

Статья подготовлена при финансовой поддержке Российского научного фонда, проект № 22-19-00305-П <https://rscf.ru/project/22-19-00305/> "Пространственно-временные стохастические модели беспроводных сетей с большим числом пользователей"

Ключевые слова: программное разделение сетевых ресурсов; распределённая вычислительная сеть; АРГМОР; многокритериальная оптимизация; Парето-архив; правило распределения; управляющий узел; узел-исполнитель; динамическое распределение ресурсов; обратная связь

Предметом исследования является методика применения Адаптивного Роево-Генетического Метода Оптимизации Распределения ресурсов (АРГМОР) для программного разделения сетевых ресурсов в распределённой вычислительной сети. Цель работы состоит в формализации перехода от результатов многокритериальной оптимизации к практическому управлению узлами-исполнителями с учётом изменяющегося состояния вычислительной среды. Область применения – централизованно управляемые распределённые вычислительные сети, в которых требуется динамическое программное распределение задач и ресурсов. Методология исследования основана на представлении текущего состояния сети $S(t)$, включающего сведения о задачах, доступности, занятости и загрузке узлов, временных характеристиках и действующих ограничениях и параметрах, определяющих допустимость текущих вариантов распределения ресурсов, а также на использовании Парето-архива A , формируемого АРГМОР по двум критериям – средней задержке и производительности алгоритма распределения ресурсов. Предложена архитектура, разделяющая управляющий и исполнительный уровни, и определена последовательность выбора рабочего решения x^* , его преобразования в правило распределения и далее в адресованные управляющие команды. Сформирован замкнутый управляющий цикл: сбор состояния сети, запуск АРГМОР, формирование Парето-архива, выбор решения, распределение подзадач, выполнение, получение результатов и актуализация состояния $S(t+1)$. Установлено, что повторный запуск оптимизации целесообразно выполнять при событиях, способных изменить применимость действующего правила: поступлении новой задачи, изменении доступности или загрузки исполнителей, перегрузке, превышении допустимого времени ожидания либо ошибке выполнения. Методика обеспечивает многокритериальность, адаптацию к динамике сети, поддержку неоднородных исполнителей и независимость от конкретной аппаратной платформы и протокола передачи данных. Практическое применение методики проиллюстрировано на четырёхузловой модели распределённой вычислительной сети при последовательном поступлении задач оценки числа ? методом Монте-Карло. В серии из 50 запусков результаты АРГМОР сопоставлены с PSO, GA и решением, выбранным из точного Парето-фронта, сформированного полным перебором.

Информация об авторах:

Петров Алексей Юрьевич, старший преподаватель, Санкт-Петербургский государственный университет аэрокосмического приборостроения, Санкт-Петербург, Россия, <https://orcid.org/0009-0005-0231-1544>

Оленев Валентин Леонидович, заведующий кафедрой аэрокосмических компьютерных и программных систем, кандидат технических наук, доцент, Санкт-Петербургский государственный университет аэрокосмического приборостроения, Санкт-Петербург, Россия, <https://orcid.org/0000-0002-1817-2754>

Тюрликов Андрей Михайлович, заведующий кафедрой инфокоммуникационных технологий и систем связи, доктор технических наук, профессор, Санкт-Петербургский государственный университет аэрокосмического приборостроения, Санкт-Петербург, Россия, <https://orcid.org/0000-0001-7132-094X>

Для цитирования:

Петров А. Ю., Оленев В. Л., Тюрликов А. М. Методика применения АРГМОР для программного разделения сетевых ресурсов в распределённой вычислительной сети // Т-Comm: Телекоммуникации и транспорт. 2026. Том 20. №9. С. 47-58.

For citation:

A. Yu. Petrov, V. L. Olenov, A. M. Tyurlikov, "Methodology for Applying ARG MOR to Software-Based Partitioning of Network Resources in a Distributed Computing Network," T-Comm, 2026, vol. 20, no. 9, pp. 47-58. (in Russian)

Введение

Развитие распределённых вычислительных систем сопровождается увеличением числа одновременно выполняемых задач, усложнением их структуры и повышением требований к эффективности использования доступных ресурсов [1]. В условиях распределённой сети назначение вычислительной работы отдельным узлам не может рассматриваться как статическая процедура, выполняемая один раз в начале работы. Состав доступных исполнителей, степень их занятости, текущая загрузка и характеристики взаимодействия изменяются во времени, а в систему продолжают поступать новые задачи. Вследствие этого эффективность программного разделения сетевых ресурсов определяется способностью системы учитывать текущее состояние вычислительной среды и своевременно корректировать принятое распределение.

Ограниченность вычислительных и сетевых ресурсов не позволяет во всех случаях резервировать для каждой задачи или потока объём ресурсов, соответствующий его максимальным требованиям. Кроме того, статическое закрепление ресурсов при изменяющейся интенсивности обмена способно приводить к их недоиспользованию. В связи с этим актуальным является динамическое перераспределение ресурсов с учётом текущего состояния сети и изменения характеристик обслуживаемых потоков [2, 3]. Современные исследования показывают, что аналогичная задача возникает в распределённых, периферийных и спутниковых вычислительных средах, где решение должно учитывать изменение нагрузки, неоднородность вычислительных узлов и ограниченность доступных ресурсов.

Для оценки качества распределения в работе используются два основных параметра – средняя задержка и производительность алгоритма распределения ресурсов. Задержка отдельной задачи характеризует временные издержки между её поступлением в систему и началом выполнения на выбранном узле-исполнителе. В условиях динамически изменяющегося состояния вычислительной сети величина задержки отдельной задачи зависит от текущей загрузки и доступности узлов, состава поступивших задач и принятого варианта распределения и поэтому рассматривается как случайная величина. При сравнении вариантов распределения используется оценка её среднего значения. Производительность алгоритма распределения ресурсов характеризует количество задач, успешно завершённых системой за единицу времени. Совместное рассмотрение средней задержки и производительности позволяет оценивать как временные издержки распределения, так и итоговую результативность функционирования распределённой вычислительной сети.

Задачи динамического распределения вычислительных и сетевых ресурсов с учётом временных характеристик и текущего состояния системы рассматриваются и в современных распределённых вычислительных средах [4-6]. При этом существенное значение имеют неоднородность вычислительных узлов, изменяющаяся загрузка и ограниченность доступных вычислительных ресурсов [7-9].

Для решения многокритериальной задачи распределения ресурсов ранее разработан Адаптивный Роево-Генетический Метод Оптимизации Распределения ресурсов – АРГМОП [10-12]. В рассматриваемой методике АРГМОП используется как компонент, формирующий множество недоминируемых

вариантов распределения. Использование популяционных методов и Парето-подхода является распространённым способом решения многокритериальных оптимизационных задач [13-15].

Наличие оптимизационного метода само по себе не определяет порядок фактического распределения задач. Из множества недоминируемых решений необходимо выбрать рабочий недоминируемый вариант распределения, соответствующий текущим требованиям; затем выбранный вариант должен быть преобразован в правило распределения, пригодное для исполнения; на его основе должны быть сформированы управляющие команды, адресованные конкретным узлам-исполнителям; после выполнения задач должны быть получены результаты и актуализированы сведения о состоянии сети. Эти операции связывают результаты многокритериальной оптимизации с практическим управлением сетью.

В исследованиях посвящённых распределению сетевых ресурсов также рассматриваются механизмы динамического регулирования выделяемых ресурсов, справедливого распределения нескольких типов ресурсов и программного разделения функций управления и непосредственного исполнения [16-18].

Целью статьи является разработка и описание методики применения АРГМОП для программного разделения сетевых ресурсов в распределённой вычислительной сети. Для достижения поставленной цели необходимо определить входные данные, архитектуру взаимодействия компонентов, порядок выбора рабочего решения, формирования правила распределения и организации обратной связи.

Постановка задачи

Рассмотрим распределённую вычислительную сеть, в которой множество вычислительных узлов используется для выполнения поступающих задач или их отдельных подзадач. В каждый момент времени часть узлов может быть свободна, часть – выполнять ранее назначенную работу, а часть – быть временно недоступной для использования. Одновременно в систему могут поступать новые задачи, требующие назначения доступным ресурсам. Следовательно, процедура распределения должна учитывать не только состав вычислительной работы, но и текущие характеристики среды исполнения.

Вариант распределения, формируемый АРГМОП, далее обозначается через x . Рассматриваются только варианты распределения, удовлетворяющие действующим ограничениям и соответствующие текущему состоянию вычислительной сети.

Каждый допустимый вариант распределения x оценивается по двум показателям качества – средней задержке D и производительности P [6, 8, 19]:

$$D \rightarrow \min, P \rightarrow \max \quad (1)$$

Для отдельной задачи i задержка D_i определяется интервалом времени между поступлением задачи в систему и началом её выполнения на назначенном вычислительном узле:

$$D_i = t_i^{start} - t_i^{in}, \quad (2)$$

где t_i^{in} – момент поступления задачи в систему, t_i^{start} – момент начала выполнения задачи на назначенном узле.

Величина D_i зависит от текущего состояния вычислительной сети, занятости и загрузки узлов, состава поступающих

задач и сформированного варианта распределения. Для оценки качества конкретного варианта распределения используется оценка её среднего значения. Для N наблюдаемых задач средняя задержка определяется выражением:

$$\bar{D} = \frac{\sum_{i=1}^N D_i}{N}. \quad (3)$$

Значение D_i характеризует задержку отдельной задачи, тогда как $\bar{D}$ используется при сравнении вариантов распределения. Единицей измерения задержки и средней задержки являются секунды или миллисекунды.

Производительность алгоритма распределения ресурсов определяется количеством задач, успешно завершённых системой за единицу времени:

$$P = \frac{N_{done}}{\Delta t}, \quad (4)$$

где N_{done} – количество задач, успешно завершённых за интервал наблюдения Δt . Единицей измерения производительности алгоритма является задач/с.

На основании сформированных входных данных и текущего состояния сети выполняется запуск АРГМОП, результатом которого является Парето-архив A – множество недоминируемых решений [14, 15]. Формирование архива A завершает оптимизационный этап, однако для практического применения полученного решения требуется определить порядок его преобразования в правило распределения и управляющие команды узлам-исполнителям. Из архива необходимо выбрать рабочее решение x^* , после чего преобразовать его в правило распределения и управляющие команды.

Дальнейшая задача состоит в переходе от полученного оптимизационного результата к практическому управлению сетевыми ресурсами.

Входные данные и состояние сети

Для применения АРГМОП управляющий контур должен располагать актуальными сведениями о задачах и доступных ресурсах. Эти сведения относятся к определённому моменту времени и изменяются по мере выполнения назначений, освобождения узлов, изменения их состояния и поступления новых задач. Текущее состояние распределённой вычислительной сети обозначим $S(t)$.

Под состоянием сети $S(t)$ далее понимается совокупность актуальных сведений о задачах, доступных узлах, временных характеристиках и действующих ограничениях, что схематично представлено на рисунке 1 [3, 7, 20].

К блоку «Задачи» относятся исходные вычислительные задачи и, при необходимости, подзадачи, ожидающие распределения между узлами сети.

Блок «Узлы сети» описывает множество вычислительных узлов, которые могут участвовать в текущем цикле распределения. Для каждого узла учитываются его доступность, занятость и загрузка. Под доступностью понимается способность узла принять назначенную подзадачу и корректно вернуть результат. Занятость отражает факт выполнения узлом ранее назначенной работы в рассматриваемый момент времени. Загрузка характеризует степень текущего использования вычислительного ресурса и может определяться по выбранным в реализации показателям [4, 8].



Рис. 1. Состав состояния сети $S(t)$

Блок «Временные характеристики» включает параметры, используемые для оценки текущего режима функционирования сети: время отправки задания, время получения ответа и фактическое время обработки подзадачи. Эти сведения позволяют оценивать временные затраты распределения и уточнять текущее состояние сети [1, 3].

Блок «Ограничения» задаёт условия допустимого распределения. К ним относятся допустимое время выполнения, допустимое время ожидания, условия корректного завершения и условия повторного назначения подзадачи. Совокупность указанных ограничений определяет, какие варианты распределения могут рассматриваться как допустимые в текущих условиях функционирования сети [5, 21].

Помимо параметров текущего состояния сети, при формировании входных данных используются заданные параметры оптимизационного цикла АРГМОР.

Архитектура применения АРГМОР

Архитектура применения АРГМОР основывается на разделении процессов принятия решения и непосредственного вычислительной работы. На управляющем уровне выполняется подготовка задачи, формируется текущее состояние сети $S(t)$, запускается модуль АРГМОР и формируется Парето-архив A . Из архива выбирается рабочее решение x^* , на основе которого формируется правило распределения. Далее средствами сетевой диспетчеризации правило преобразуется в управляющие команды, адресуемые соответствующим узлам-исполнителям. На исполнительном уровне узлы принимают назначения, выполняют соответствующие подзадачи и возвращают результаты и служебные сведения, используемые для последующего обновления состояния сети. Указанная иерархия представлена на рисунке 2 [9, 18].

В архитектуре можно выделить несколько функциональных компонентов. Компонент подготовки задачи определяет параметры исходной вычислительной задачи и, если её структура допускает декомпозицию, формирует набор подзадач, предназначенных для последующего распределения. Компонент формирования текущего состояния сети объединяет сведения о задачах, узлах-исполнителях, их состоянии, временных характеристиках и действующих ограничениях в составе $S(t)$. Оптимизационный модуль АРГМОР формирует Парето-архив A , содержащий множество недоминируемых вариантов распределения [7]. Из него выбирается рабочее решение x^* , соответствующее текущим требованиям. Выбранное решение преобразуется в правило распределения, после чего сетевой диспетчер формирует адресованные управляющие команды для узлов-исполнителей.

Представленная архитектура не привязана к конкретной аппаратной платформе или физическому интерфейсу передачи данных. Существенными для методики являются функциональные этапы подготовки задачи, формирования состояния сети, многокритериальной оптимизации, выбора рабочего решения, формирования правила распределения и передачи управляющих воздействий узлам-исполнителям [22]. После выполнения назначенных подзадач результаты и сведения о фактическом состоянии исполнителей могут использоваться для актуализации $S(t)$ и последующего цикла распределения [19, 20].



Рис. 2. Архитектура применения АРГМОР в распределённой вычислительной сети

В конкретной программно-аппаратной реализации функции управляющего уровня могут выполняться одним компонентом либо распределяться между несколькими компонентами. Например, подготовка и декомпозиция исходной задачи могут выполняться контролирующим устройством, тогда как функции сетевой диспетчеризации, передачи команд и получения результатов – master-узлом. Такое разделение носит функциональный характер и не изменяет общей последовательности применения АРГМОР.

Формирование правила распределения

Центральным этапом методики является преобразование результата многокритериальной оптимизации в прикладное правило распределения. Парето-архив A содержит несколько недоминируемых вариантов, поэтому первым шагом после завершения работы АРГМОР становится выбор рабочего решения x^* [14, 15]. Выбор осуществляется в соответствии с текущими требованиями цикла распределения: приоритет может отдаваться минимизации средней задержки, максимизации производительности алгоритма распределения ресурсов или компромиссному варианту.

Рабочее решение x^* не является непосредственной командой узлу-исполнителю, оно описывает выбранный вариант распределения в обозначениях АРГМОР. Для практического применения его необходимо преобразовать в правило распределения. Правило распределения определяет, какие подзадачи каким исполнителям передаются, в каком объёме, в каком порядке и при каких условиях контроля.



Рис. 3. Переход от результата оптимизации к управлению исполнителями

Таким образом, необходимо чётко различать три уровня представления результата. Первый уровень — оптимизационное решение x^* , выбранное из Парето-архива. Второй уровень — правило распределения, которое интерпретирует выбранное решение с учётом текущего состояния сети. Третий уровень — управляющие команды, представляющие собой адресованные узлам-исполнителям сообщения, посредством которых правило реализуется на практике. Такое разграничение позволяет избежать смешения математического результата оптимизации и сетевого механизма его исполнения.

Правило распределения включает указание исполнителя, назначаемую подзадачу, объём вычислительной работы, параметры выполнения, порядок передачи, допустимое время

ожидания результата, а также условия контроля и повторного назначения. Конкретный состав и способ представления этих элементов могут уточняться в зависимости от архитектуры и программно-аппаратной реализации вычислительной сети.

На основе сформированного правила подготавливаются управляющие команды для узлов-исполнителей. Команда должна содержать, по крайней мере, идентификатор или адрес исполнителя, идентификатор подзадачи, объём работы, необходимые параметры вычисления и служебные признаки контроля. Конкретный формат управляющей команды определяется используемым протоколом и программно-аппаратной реализацией сети [18, 20].

Последовательность применения методики

Последовательность применения АРГМОР представляет собой совокупность взаимосвязанных этапов, охватывающих путь от постановки вычислительной задачи до получения результатов и актуализации состояния сети. Она описывает взаимодействие управляющего и исполнительного уровней сети, тогда как внутренние процедуры многокритериального оптимизационного поиска выполняются внутри модуля АРГМОР.



Рис. 4. Последовательность применения АРГМОР

Полный цикл начинается с постановки вычислительной задачи управляющим узлом. Затем выполняются инициализация сети и определение множества потенциальных исполнителей. На следующем этапе собираются актуальные параметры вычислительной среды и формируется текущее состояние сети $S(t)$, включающее сведения о задачах, доступности, занятости и загрузке узлов, временных характеристиках и действующих ограничениях. Если структура исходной задачи допускает декомпозицию, она разделяется на подзадачи, после чего формируются входные данные для АРГМОР [3, 4, 7].

В результате оптимизационного цикла АРГМОП формируется Парето-архив A , содержащий множество недоминируемых вариантов распределения. Из него выбирается рабочее решение x^* , соответствующее текущим требованиям по средней задержке и производительности алгоритма распределения ресурсов. Выбранное решение преобразуется в правило распределения, на основании которого формируются управляющие команды, адресованные узлам-исполнителям.

После передачи команд исполнители выполняют назначенные подзадачи и возвращают результаты и служебные сведения. На их основе рассчитываются фактические параметры работы и актуализируется состояние сети $S(t+1)$. Обновлённые сведения используются при последующем цикле распределения, что обеспечивает обратную связь между результатами исполнения и очередным запуском АРГМОП.

Управляющий цикл и обратная связь

Программное разделение сетевых ресурсов в динамической вычислительной среде не может ограничиваться однократным формированием распределения. В процессе выполнения задач меняются доступность, занятость и загрузка узлов, могут поступать новые задания, часть исполнителей может выходить из числа доступных ресурсов, а отдельные назначения могут завершаться ошибкой. Поэтому применение АРГМОП должно рассматриваться как замкнутый управляющий цикл, включающий обратную связь [1, 3, 4].

Общий цикл можно описать последовательностью: состояние сети $S(t)$ – запуск АРГМОП – Парето-архив A – выбор рабочего решения x^* – правило распределения – управляющие команды – результаты выполнения – обновлённое состояние сети $S(t+1)$.

Обратная связь выполняет две основные функции. Во-первых, она позволяет фиксировать успешность или неуспешность текущих назначений и получать собственно вычислительные результаты. Во-вторых, она служит источником сведений для актуализации состояния сети. После завершения или срыва исполнения изменяются характеристики доступности, занятости и загрузки ресурсов; кроме того, могут уточняться временные характеристики взаимодействия. Всё это используется при подготовке следующего цикла применения метода.

Новый цикл следует запускать только при изменениях, способных повлиять на действующее правило распределения. К таким событиям относятся поступление новой задачи, изменение доступности исполнителя, существенное изменение загрузки, перегрузка, превышение допустимого времени ожидания, ошибка выполнения либо отсутствие корректного результата. При возникновении одного из указанных событий обновлённое состояние $S(t+1)$ используется как вход для нового запуска АРГМОП. Если же значимых изменений не произошло, система может продолжать исполнение уже сформированного правила или перейти в режим ожидания.

Свойства и ограничения методики

Предлагаемая методика обладает рядом свойств, непосредственно вытекающих из её структуры.

Первым свойством является многокритериальность: формирование распределения опирается на Парето-архив, в котором одновременно учитываются средняя задержка и производительность алгоритма распределения ресурсов. Это

позволяет не сводить задачу к единственной интегральной оценке, а сохранять несколько недоминируемых вариантов и выбирать из них рабочее решение в соответствии с текущими приоритетами [14, 15 19].

Вторым свойством является адаптивность методики к изменению состояния вычислительной сети. Состояние сети актуализируется по мере поступления новых данных, а правило распределения может формироваться повторно при изменении условий функционирования. Результаты выполнения используются для обновления состояния сети и подготовки следующего цикла АРГМОП [1, 3, 4].

Третьим важным свойством является разделение оптимизационного и исполнительного уровней. АРГМОП отвечает за формирование множества недоминируемых решений, после чего выбранный вариант преобразуется в правило распределения и далее – в управляющие команды. Поэтому методика не зависит от конкретного физического интерфейса передачи данных и может применяться в сетях с различной аппаратной организацией. Ещё одним свойством является возможность работы с неоднородными исполнителями: различия между узлами учитываются в составе текущего состояния сети и не требуют изменения общей последовательности применения метода [7-9].

Наряду со свойствами необходимо выделить и ограничения методики. Во-первых, её применение требует достаточно актуального и достоверного состояния сети. Если сведения о доступности, занятости или загрузке существенно расходятся с фактическими условиями, выбранное решение может потерять актуальность ещё до начала исполнения. Во-вторых, повторная оптимизация требует вычислительных затрат, поэтому слишком частый запуск АРГМОП способен сам по себе увеличивать служебные временные издержки. В-третьих, методика ориентирована на наличие централизованного управляющего контура, что упрощает принятие согласованных решений, но предъявляет повышенные требования к производительности и устойчивости управляющего уровня. Наконец, качество практического применения существенно зависит от корректности подготовки подзадач и от полноты входных данных.

Демонстрационный пример применения методики

Для количественной иллюстрации её применения далее рассматривается четырёхузловой экспериментальный сценарий, позволяющий проследить переход от текущего состояния сети к конкретному распределению вычислительной нагрузки, а также сопоставить АРГМОП с PSO, GA и решением, выбранным из точного Парето-фронта, сформированного полным перебором.

Для пояснения практического применения предложенной методики рассмотрим задачу оценки числа π методом Монте-Карло [10]. Геометрической основой данной задачи является определение площади четверти единичного круга. Площадь этой области может быть представлена выражением:

$$\pi = 4 \int_0^1 \sqrt{1-x^2} dx. \quad (5)$$

При оценке числа π методом Монте-Карло генерируются независимые случайные точки (x, y) , равномерно распределённые внутри единичного квадрата:

$$0 \leq x \leq 1, \quad 0 \leq y \leq 1. \quad (6)$$

Для каждой сгенерированной точки проверяется условие принадлежности четверти единичной окружности:

$$x^2 + y^2 \leq 1. \quad (7)$$

Следовательно отдельное испытание представляет собой генерацию одной случайной точки и проверку условия её принадлежности четверти единичной окружности. Поскольку результаты отдельных испытаний не зависят друг от друга, общий объём вычислений может быть разделён на независимые части и распределён между вычислительными узлами.

Для демонстрации работы методики принимаются следующие допущения.

1. Исходная вычислительная задача содержит 50 000 независимых испытаний Монте-Карло.

2. Минимальной единицей вычислительной работы, передаваемой отдельному узлу, является вычислительный квант. Под вычислительным квантом далее понимается группа из 250 независимых испытаний Монте-Карло.

3. В рассматриваемой вычислительной сети используются четыре узла-исполнителя.

При указанных допущениях общее число вычислительных квантов M составляет:

$$M = \frac{50\,000}{250} = 200. \quad (8)$$

Следовательно, задача распределения состоит в назначении 200 вычислительных квантов четырём узлам-исполнителям. Объём вычислений, назначаемый каждому узлу, определяется непосредственно в процессе формирования решения.

Вариант распределения вычислительной нагрузки между K узлами представим вектором:

$$z = (z_1, z_2, \dots, z_K), \quad (9)$$

где z_i – число вычислительных квантов, назначаемых i -му узлу.

Для допустимого варианта распределения выполняются условия

$$\sum_{i=1}^K z_i = M. \quad (10)$$

$$z_i \in \mathbb{Z}, \quad z_i \geq 1, \quad i = 1, \dots, K. \quad (11)$$

В рассматриваемом четырёхузловом примере $K = 4$, $M = 200$, поэтому:

$$z = (z_1, z_2, z_3, z_4), \quad (12)$$

а допустимые варианты распределения удовлетворяют условиям:

$$z_1 + z_2 + z_3 + z_4 = 200 \quad (13)$$

Поскольку один вычислительный квант содержит 250 испытаний, фактический объём вычислительной работы N_i , назначаемой i -му узлу, составляет:

$$N_i = 250 * z_i. \quad (14)$$

При любом допустимом варианте распределения выполняется

$$\sum_{i=1}^4 N_i = 50\,000 \quad (15)$$

то есть общий объём исходной вычислительной задачи остаётся неизменным.

Распределение вычислительных квантов между узлами представляет собой известную комбинаторную задачу [23] определения числа положительных целочисленных решений уравнения:

$$z_1 + z_2 + \dots + z_K = M. \quad (16)$$

При условии $z_i \geq 1$ число различных вариантов распределения N_{var} определяется биномиальным коэффициентом:

$$N_{var}(M, K) = \binom{M-1}{K-1} = \frac{(M-1)!}{(K-1)!(M-K)!} \quad (17)$$

Для $M = 200$ квантов и $K = 4$ узлов число допустимых вариантов составляет

$$N_{var}(200, 4) = \binom{199}{3} = \frac{199!}{3!196!} = 1\,293\,699 \quad (18)$$

Таким образом, даже при использовании всего четырёх вычислительных узлов и однородных вычислительных квантов полный перебор должен проверить более 1,29 млн вариантов распределения.

Поскольку при увеличении объёма вычислительной задачи число распределяемых квантов существенно превышает число вычислительных узлов $M \gg K$, для оценки роста числа вариантов распределения целесообразно использовать приближение биномиального коэффициента при $n \gg k$. При $k = o(n)$, то есть $k \rightarrow 0$, биномиальный коэффициент имеет вид [24]:

$$\left(\frac{n}{k}\right) \sim \left(\frac{ne}{k}\right)^k \frac{1}{\sqrt{2\pi k}} \exp\left(-\frac{k^2}{2n}(1 + o(1))\right). \quad (19)$$

Для рассматриваемой задачи $n = M - 1$, $k = K - 1$. Поэтому число допустимых вариантов распределения может быть оценено как:

$$N_{var}(M, K) = \binom{M-1}{K-1} \sim \left(\frac{e(M-1)}{K-1}\right)^{K-1} \frac{1}{\sqrt{2\pi(K-1)}} \exp\left(-\frac{(K-1)^2}{2(M-1)}(1 + o(1))\right). \quad (20)$$

Полученная оценка показывает, что при увеличении числа распределяемых квантов и вычислительных узлов количество допустимых вариантов распределения быстро возрастает, вследствие чего применение полного перебора становится вычислительно значительно более затратным.

Рассмотрим отдельный цикл функционирования распределённой вычислительной сети. Помимо приведённых выше общих допущений, для данного цикла задаётся конкретное начальное состояние узлов.

В момент поступления новой вычислительной задачи все четыре узла уже выполняют ранее назначенную работу. Остаточное время до освобождения узлов составляет соответственно 2,9846; 1,9214; 1,6115; 0,6010 с. Поступившая задача не вытесняет ранее выполняемые вычисления и не выполняется параллельно с ними на одном вычислительном узле. Каждый узел сначала завершает ранее назначенную работу, после чего начинает выполнение назначенных ему квантов новой задачи.

Вследствие этого алгоритм распределения должен определять такой вектор

$$z = (z_1, z_2, z_3, z_4) \quad (21)$$

который учитывает различное остаточное время занятости вычислительных узлов.

Для заданного начального состояния в первом рассматриваемом цикле АРГМОП формирует вектор распределения:

$$z^{(1)} = (2, 44, 57, 97) \quad (22)$$

Верхний индекс (1) обозначает номер рассматриваемого цикла распределения, а элементы вектора определяют число квантов, назначенных соответственно первому, второму, третьему и четвёртому вычислительным узлам.

При размере вычислительного кванта 250 испытаний указанному вектору соответствует распределение - (500; 11 000; 14 250; 24 250).

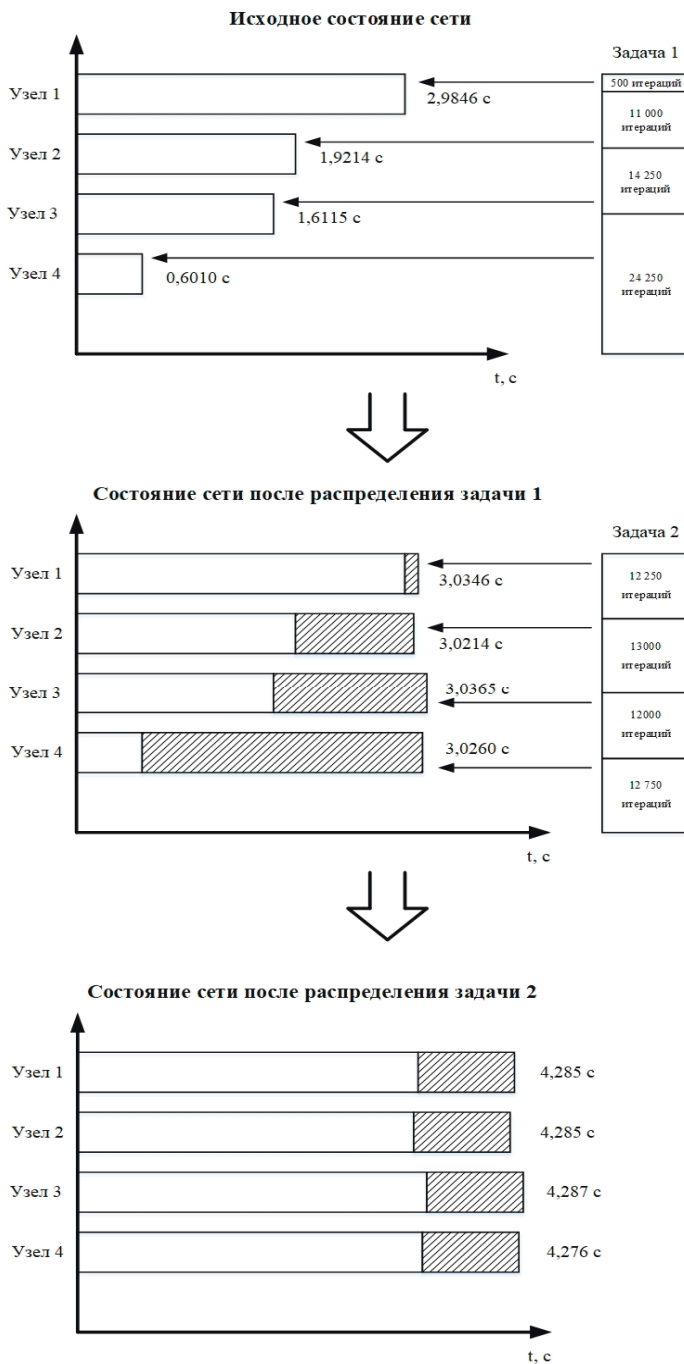


Рис. 5. Пример последовательного распределения двух вычислительных задач между четырьмя узлами

Наиболее занятый первый вычислительный узел получает только два кванта, то есть 500 новых испытаний, тогда как четвёртый узел, имеющий минимальное остаточное время занятости, получает 97 квантов, соответствующих 24 250 испытаниям. Тем самым сформированное распределение учитывает текущее состояние сети и компенсирует исходную неравномерность занятости исполнителей.

На рисунке 5 показан последовательный характер применения методики. Для первой поступившей задачи распределение формируется исходя из остаточной занятости вычислительных узлов. После назначения вычислительных квантов состояние сети изменяется. Поэтому при поступлении следующей задачи используется уже не исходное, а обновлённое состояние вычислительной сети.

В рассматриваемом примере после первого назначения прогнозируемая занятость узлов составляет:

$$S(t_1) = (3,035; 3,021; 3,037; 3,026) \quad (23)$$

При поступлении второй вычислительной задачи АРГМОП запускается повторно уже для этого обновлённого состояния. Поскольку после первого распределения занятость узлов практически выровнена, следующая однотипная задача может распределяться близко к равномерному варианту. Таким образом, правило распределения не является заранее зафиксированным: оно формируется заново при поступлении каждой новой задачи с учётом состояния сети в соответствующий момент времени.

Для количественной оценки методики эксперимент проводился серией из 50 воспроизводимых запусков. В каждом запуске АРГМОП, PSO, GA и полный перебор получали одинаковое исходное состояние четырёх узлов, одинаковый объём поступающей вычислительной задачи и одинаковые случайные отклонения фактического времени исполнения.

На рисунке 6 представлены значения задержки, полученные в каждом из 50 экспериментальных запусков. Каждая точка соответствует одному состоянию вычислительной сети и результату распределения поступившей задачи при использовании соответствующего метода. Пунктирная линия соответствует рабочему решению, выбранному из точного Парето-фронта, сформированного полным перебором, и используется в качестве эталона для рассматриваемой четырёхузловой модели.

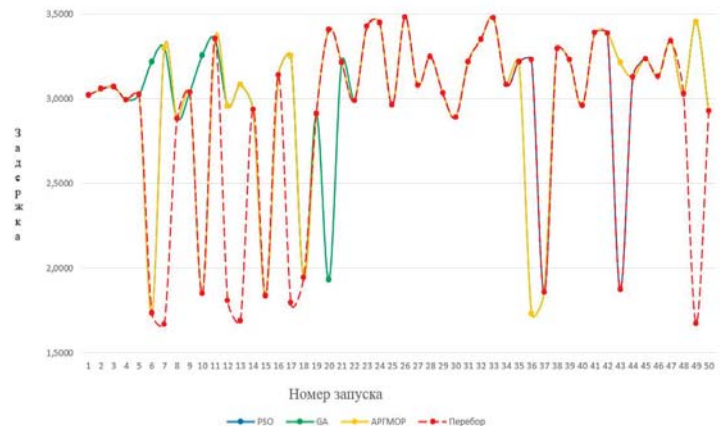


Рис. 6. Задержка алгоритмов распределения ресурсов для серии из 50 запусков

На рисунке 7 представлена производительность распределённой системы при использовании четырёх рассматриваемых способов формирования распределения.

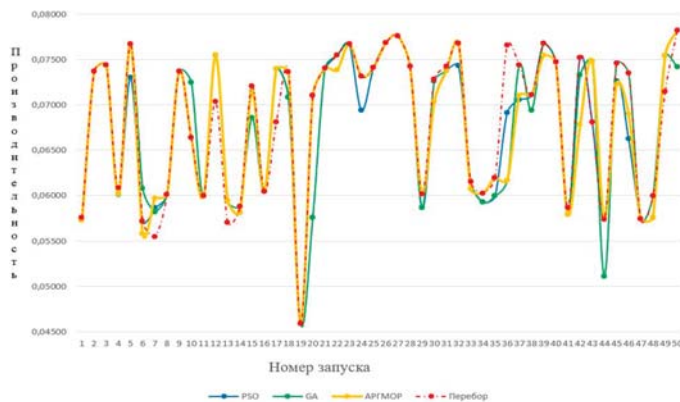


Рис. 7. Производительность алгоритмов распределения ресурсов для серии из 50 запусков

По результатам серии из 50 запусков были рассчитаны средние значения задержки и производительности для АРГМОП, PSO, GA и полного перебора. Полученные значения представлены в таблице 1.

Таблица 1

Средние значения задержки и производительности по результатам 50 запусков

Метод	Средняя задержка D, с	Средняя производительность Р, задач/с
АРГМОП	3,00506	0,067877
PSO	3,00825	0,067732
GA	3,03338	0,067850
Полный перебор	2,86006	0,068230

В рассматриваемой четырёхузловой модели пространство решений сравнительно невелико, а задача распределения в четырёхузловой модели может быть достаточно эффективно решена всеми сравниваемыми эвристическими методами. Поэтому различия между АРГМОП, PSO и GA в данном эксперименте имеют ограниченную величину. С увеличением числа вычислительных узлов и размерности пространства распределений возрастает количество допустимых вариантов назначения вычислительной нагрузки и усложняется поиск сбалансированного решения. В этих условиях преимущество гибридного механизма АРГМОП может проявляться более заметно за счёт сочетания роевого поиска, внешнего Парето-архива и генетической коррекции при стагнации.

Дополнительным направлением исследования является переход от распределения однородных вычислительных квантов к задаче назначения неоднородных вычислительных блоков. В такой постановке отдельные подзадачи могут иметь различную вычислительную трудоёмкость, а узлы — различную производительность, остаточную занятость и временные характеристики взаимодействия. Решение должно определять не только объём нагрузки каждого исполнителя, но и состав конкретных подзадач, назначаемых каждому узлу. Такое усложнение существенно увеличивает размер и неоднородность пространства решений и позволяет более полно оценить эффективность адаптивной гибридизации АРГМОП по сравнению с отдельным применением PSO и GA.

Заключение

В статье предложена методика применения ранее разработанного Адаптивного Роево-Генетического Метода Оптимизации Распределения ресурсов АРГМОП в распределённой вычислительной сети. Методика определяет порядок включения АРГМОП в контур управления распределённой вычислительной сетью и переход от текущего состояния сети и результатов оптимизации к правилу распределения и управляющим командам.

Для описания методики используются представление допустимого решения x , два целевых параметра — средняя задержка $\bar{D}$ и производительность P , обозначение текущего состояния сети $S(t)$, Парето-архива A и выбранного рабочего решения x^* . На их основе определена последовательность перехода от текущего состояния сети к правилу распределения и управляющим воздействиям.

Определены состав текущего состояния сети, архитектура применения АРГМОП, порядок выбора рабочего решения из Парето-архива, формирование правила распределения, построение управляющих команд и общий управляющий цикл с обратной связью. Показано, что в рассматриваемой централизованной архитектуре правило распределения может быть функционально связано как с подготовкой вычислительной задачи, так и с сетевой диспетчеризацией исполнителей. Тем самым устанавливается связь между многокритериальным оптимизационным алгоритмом и практическим контуром управления распределённой вычислительной сетью.

Для иллюстрации применения методики рассмотрена работа распределённой вычислительной сети из четырёх узлов в условиях различной остаточной занятости исполнителей и последовательного поступления новых вычислительных задач. Показано, что решение АРГМОП непосредственно определяет объём вычислительной работы, назначаемый каждому узлу, а после формирования распределения состояние сети обновляется и используется при поступлении следующей задачи. На малой четырёхузловой модели результаты метода сопоставлены с PSO, GA и решением, выбранным из точного Парето-фронта, сформированного полным перебором. Проведение серии из 50 запусков позволило непосредственно оценить два принятых в работе показателя — среднюю задержку и производительность алгоритма распределения ресурсов. Полученные результаты подтверждают эффективную практическую применимость предложенной методики для динамического формирования правил распределения в условиях изменяющегося состояния вычислительной сети.

Ограниченная размерность рассматриваемой модели приводит к относительно небольшим различиям между АРГМОП и отдельным применением PSO и GA. Дальнейшее исследование целесообразно направить на оценку масштабируемости метода при увеличении числа вычислительных узлов, поскольку рост размерности существенно расширяет пространство допустимых распределений.

Финансовая поддержка

Статья подготовлена при финансовой поддержке Российского научного фонда, проект № 22-19-00305-П <https://rscf.ru/project/22-19-00305/> "Пространственно-временные стохастические модели беспроводных сетей с большим числом пользователей".

Литература

1. Оленев В.Л. Эволюция коммуникационных технологий через автоматизацию процессов разработки протоколов // Известия Самарского научного центра Российской академии наук, т. 28, № 2, 2026, С. 185-195
2. Суворова Е. А. Динамическое распределение сетевых ресурсов в сети SpaceFibre на основе мультиагентных самообучающихся систем // Наукоемкие технологии в космических исследованиях Земли. 2025. Т. 17, № 4. С. 15–27.
3. Суворова Е. А. Оценка состояния сети SpaceFibre для динамического распределения ресурсов буферного пространства сетевых устройств между потоками данных // Системный анализ и логистика. 2026. № 1 (49). С. 3-13.
4. Zhang Q., Zhang C., Zhao J., Wang D., Xu W. Dynamic Resource Allocation for Multi-Access Edge Computing in Urban Rail Transit // IEEE Transactions on Vehicular Technology. 2025. Vol. 74, No. 2. P. 3296-3310.
5. Huang Z., Wu X., Dong S. Multi-objective task offloading for highly dynamic heterogeneous Vehicular Edge Computing: An efficient reinforcement learning approach // Computer Communications. 2024. Vol. 225. P. 27-43.
6. Liu Y., Yan B., Yu J. Multi objective task resource allocation method based on hierarchical Bayesian adaptive sparsity for edge computing in low voltage stations // IET Cyber-Physical Systems: Theory & Applications. 2024. Vol. 9, No. 1. P. 63-71.
7. Aslanpour M. S., Toosi A. N., Cheema M. A., Chhetri M. B., Salehi M. A. Load balancing for heterogeneous serverless edge computing: A performance-driven and empirical approach // Future Generation Computer Systems. 2024. Vol. 154. P. 266-280.
8. Aminu J., Latip R., Hanafi Z. M., Kamarudin S., Gabi D. Efficient Task Allocation for Energy and Execution Time Trade-Off in Edge Computing Using Multi-Objective IPSO // Computers, Materials & Continua. 2025. Vol. 84, No. 2. P. 2989-3011.
9. Li H., Yu J., Cao L., Zhang Q., Song Z., Hou S. Multi-agent reinforcement learning based computation offloading and resource allocation for LEO Satellite edge computing networks // Computer Communications. 2024. Vol. 222. P. 268-276.
10. Петров А. Ю. Программно-аппаратная архитектура демонстратора алгоритма АРГМОП // Инновационное приборостроение. 2026. Т. 5, № 1. С. 54-58.
11. Петров А. Ю. Гибридный Алгоритм PSO-GA в Вычислительных системах // Аэрокосмическое приборостроение и эксплуатационные технологии: Пятая Международная научная конференция. Сборник докладов. В 2-х частях, СПб. Том Часть 2. Санкт-Петербург: Санкт-Петербургский государственный университет аэрокосмического приборостроения, 2025. С. 220-223.
12. Петров А.Ю., Оленев В.Л. Формализация гибридного метода распределения задач с использованием Сетей Петри // Научно-технический вестник Поволжья. 2025. № 6. С. 228-232.
13. Kennedy J., Eberhart R. Particle swarm optimization // Proceedings of the IEEE International Conference on Neural Networks. Perth: IEEE, 1995. P. 1942-1948.
14. Deb K., Pratap A., Agarwal S., Meyarivan T. A fast and elitist multiobjective genetic algorithm: NSGA-II // IEEE Transactions on Evolutionary Computation. 2002. Vol. 6, No. 2. P. 182-197.
15. Coello Coello C. A., Pulido G. T., Lechuga M. S. Handling multiple objectives with particle swarm optimization // IEEE Transactions on Evolutionary Computation. 2004. Vol. 8, No. 3. P. 256-279.
16. Kelly F. P. Charging and rate control for elastic traffic // European Transactions on Telecommunications. 1997. Vol. 8, No. 1. P. 33-37.
17. Ghodsi A., Zaharia M., Hindman B., Konwinski A., Shenker S., Stoica I. Dominant Resource Fairness: Fair Allocation of Multiple Resource Types // Proceedings of the 8th USENIX Symposium on Networked Systems Design and Implementation. Boston: USENIX Association, 2011. P. 323-336.
18. McKeown N., Anderson T., Balakrishnan H. [et al. OpenFlow: Enabling Innovation in Campus Networks // ACM SIGCOMM Computer Communication Review. 2008. Vol. 38, No. 2. P. 69-74.
19. Nagalakshmi B., Subramanian S. An efficient multi-objective task scheduling in edge computing using adaptive honey badger optimisation // International Journal of Web Engineering and Technology. 2024. Vol. 19, No. 2. P. 110-126.
20. Pei X., Hu Y., Dong Y., Li D. Task Scheduling in Edge Network System Using Resource Interleaving and Reinforcement Learning // IEEE Transactions on Information and Systems. 2025. Vol. E108.D, No. 9. P. 1142-1145.
21. Liu X., Lv J., Kim B.-G. et al. Cooperative Digital Healthcare Task Scheduling and Resource Management in Edge Intelligence Systems // Tsinghua Science and Technology. 2025. Vol. 30, No. 2. P. 926-945.
22. Olenov V., Lavrovskaya I., Korobkov I., Sheynin Yu. Design and Simulation of Onboard SpaceWire Networks // Conference of Open Innovations Association, FRUCT. 2019. №. 24. P. 291-299.
23. Виленкин Н. Я., Виленкин А. Н., Виленкин П. А. Комбинаторика. 6-е изд., стер. М.: МЦНМО: ФИМА, 2017. 400 с.
24. Spencer J. Asymptopia. Providence, RI: American Mathematical Society, 2014. 189 p.

METHODOLOGY FOR APPLYING ARGMOR TO SOFTWARE-BASED PARTITIONING OF NETWORK RESOURCES IN A DISTRIBUTED COMPUTING NETWORK

Alexey Yu. Petrov, Saint Petersburg State University of Aerospace Instrumentation, Saint Petersburg, Russia, palexeq1998@yandex.ru
Valentin L. Olenev, Saint Petersburg State University of Aerospace Instrumentation, Saint Petersburg, Russia, valentin.olenov@guap.ru
Andrey M. Tyurlikov, Saint Petersburg State University of Aerospace Instrumentation, Saint Petersburg, Russia, turlikov@guap.ru

Abstract

This study addresses a methodology for applying the Adaptive Swarm-Genetic Method for Resource Allocation Optimization (ARGMOR) to software-based partitioning of network resources in a distributed computing network. The aim is to formalize the transition from multi-objective optimization results to practical control of worker nodes, taking into account changes in the computing environment. The methodology is intended for centrally managed distributed computing networks requiring dynamic software-based allocation of tasks and resources. The research approach is based on a representation of the current network state $S(t)$, comprising information on tasks, node availability, occupancy and load, timing characteristics, and the applicable constraints and parameters that determine the feasibility of resource allocation options. It also uses a Pareto archive A generated by ARGMOR according to two criteria: average delay and throughput achieved using the resource allocation algorithm. An architecture separating the control and execution layers is proposed, and a sequence is defined for selecting a working solution x^* , converting it into an allocation rule, and subsequently generating control commands addressed to individual nodes. A closed-loop control process is established: collecting network state information, running ARGMOR, generating the Pareto archive, selecting a solution, distributing subtasks, executing them, receiving results, and updating the network state to $S(t+1)$. Optimization is found to be appropriate when triggered by events that may affect the applicability of the current allocation rule, including the arrival of a new task, changes in worker availability or load, overload, excessive waiting time, or an execution error. The methodology supports multiple optimization criteria, adaptation to network dynamics, heterogeneous worker nodes, and independence from any specific hardware platform or data transmission protocol. Its practical application is illustrated using a four-node model of a distributed computing network with successive arrivals of tasks for estimating τ by the Monte Carlo method. Across 50 runs, ARGMOR results are compared with those obtained using PSO, GA, and a solution selected from the exact Pareto front generated by exhaustive search.

Keywords: software-based partitioning of network resources; distributed computing network; ARGMOR; multi-objective optimization; Pareto archive; allocation rule; control node; worker node; dynamic resource allocation; feedback.

References

- [1] V. L. Olenov, "Evolution of communication technologies through automation of protocol development processes," *Izvestia of the Samara Scientific Center of the Russian Academy of Sciences*, vol. 28, no. 2, pp. 185-195, 2026.
- [2] E. A. Suvorova, "Dynamic allocation of network resources in the SpaceFibre network based on multiagent self-learning systems," *H&ES Research*, vol. 17, no. 4, pp. 15-27, 2025.
- [3] E. A. Suvorova, "The SpaceFibre network state evaluation for dynamic allocation of buffer space resources for network devices between data flows," *System Analysis and Logistics*, no. 1 (49), pp. 3-13, 2026.
- [4] Q. Zhang, C. Zhang, J. Zhao, D. Wang, and W. Xu, "Dynamic resource allocation for multi-access edge computing in urban rail transit," *IEEE Trans. Veh. Technol.*, vol. 74, no. 2, pp. 3296-3310, 2025.
- [5] Z. Huang, X. Wu, and S. Dong, "Multi-objective task offloading for highly dynamic heterogeneous vehicular edge computing: An efficient reinforcement learning approach," *Comput. Commun.*, vol. 225, pp. 27-43, 2024.
- [6] Y. Liu, B. Yan, and J. Yu, "Multi objective task resource allocation method based on hierarchical Bayesian adaptive sparsity for edge computing in low voltage stations," *IET Cyber-Phys. Syst., Theory Appl.*, vol. 9, no. 1, pp. 63-71, 2024.
- [7] M. S. Aslanpour, A. N. Toosi, M. A. Cheema, M. B. Chhetri, and M. A. Salehi, "Load balancing for heterogeneous serverless edge computing: A performance-driven and empirical approach," *Future Gener. Comput. Syst.*, vol. 154, pp. 266-280, 2024.
- [8] J. Aminu, R. Latip, Z. M. Hanafi, S. Kamarudin, and D. Gabi, "Efficient task allocation for energy and execution time trade-off in edge computing using multi-objective IPSO," *Comput., Mater. Continua*, vol. 84, no. 2, pp. 2989-3011, 2025.
- [9] H. Li, J. Yu, L. Cao, Q. Zhang, Z. Song, and S. Hou, "Multi-agent reinforcement learning based computation offloading and resource allocation for LEO satellite edge computing networks," *Comput. Commun.*, vol. 222, pp. 268-276, 2024.
- [10] A. Yu. Petrov, "Hardware and software architecture of an ARGMOR algorithm demonstrator," *Innovative Instrumentation*, vol. 5, no. 1, pp. 54-58, 2026.
- [11] A. Yu. Petrov, "A hybrid PSO-GA algorithm in computing systems," in *Proc. 5th Int. Sci. Conf. Aerospace Instrumentation and Operational Technologies*, pt. 2. Saint Petersburg, Russia: Saint Petersburg State University of Aerospace Instrumentation, 2025, pp. 220-223.
- [12] A. Yu. Petrov and V. L. Olenov, "Formalization of a hybrid task allocation method using Petri nets," *Scientific and Technical Volga Region Bulletin*, no. 6, pp. 228-232, 2025.

- [13] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proc. IEEE Int. Conf. Neural Networks*, Perth, WA, Australia, 1995, pp. 1942-1948.
- [14] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," *IEEE Trans. Evol. Comput.*, vol. 6, no. 2, pp. 182-197, 2002.
- [15] C. A. Coello Coello, G. T. Pulido, and M. S. Lechuga, "Handling multiple objectives with particle swarm optimization," *IEEE Trans. Evol. Comput.*, vol. 8, no. 3, pp. 256-279, 2004.
- [16] F. P. Kelly, "Charging and rate control for elastic traffic," *Eur. Trans. Telecommun.*, vol. 8, no. 1, pp. 33-37, 1997.
- [17] A. Ghodsi, M. Zaharia, B. Hindman, A. Konwinski, S. Shenker, and I. Stoica, "Dominant resource fairness: Fair allocation of multiple resource types," in *Proc. 8th USENIX Symp. Netw. Syst. Des. Implement. (NSDI)*, Boston, MA, USA, 2011, pp. 323-336.
- [18] N. McKeown et al., "OpenFlow: Enabling innovation in campus networks," *ACM SIGCOMM Comput. Commun. Rev.*, vol. 38, no. 2, pp. 69-74, 2008.
- [19] B. Nagalakshmi and S. Subramanian, "An efficient multi-objective task scheduling in edge computing using adaptive honey badger optimisation," *Int. J. Web Eng. Technol.*, vol. 19, no. 2, pp. 110-126, 2024.
- [20] X. Pei, Y. Hu, Y. Dong, and D. Li, "Task scheduling in edge network system using resource interleaving and reinforcement learning," *IEICE Trans. Inf. Syst.*, vol. E108-D, no. 9, pp. 1142-1145, 2025.
- [21] X. Liu et al., "Cooperative digital healthcare task scheduling and resource management in edge intelligence systems," *Tsinghua Sci. Technol.*, vol. 30, no. 2, pp. 926-945, 2025.
- [22] V. Olenov, I. Lavrovskaya, I. Korobkov, and Yu. Sheynin, "Design and simulation of onboard SpaceWire networks," in *Proc. 24th Conf. Open Innovations Assoc. (FRUCT)*, 2019, pp. 291-299.
- [23] N. Ya. Vilenkin, A. N. Vilenkin, and P. A. Vilenkin, *Combinatorics*, 6th ed. Moscow, Russia: MCCME and FIMA, 2017.
- [24] J. Spencer, *Asymptopia*. Providence, RI, USA: American Mathematical Society, 2014.

Information about authors:

Alexey Yu. Petrov, Senior Lecturer, Saint Petersburg State University of Aerospace Instrumentation, Saint Petersburg, Russia, <https://orcid.org/0009-0005-0231-1544>

Valentin L. Olenov, Head of department of Aerospace computer and software systems, Candidate of Technical Sciences, Associate Professor, Saint Petersburg State University of Aerospace Instrumentation, Saint Petersburg, Russia, <https://orcid.org/0000-0002-1817-2754>

Andrey M. Tyurlikov, Head of Department of Infocommunication technologies and communication systems, Doctor of Technical Sciences, Professor, Saint Petersburg State University of Aerospace Instrumentation, Saint Petersburg, Russia, <https://orcid.org/0000-0001-7132-094X>

DEVELOPMENT AND DEBUGGING OF SOFTWARE FOR THE IMPLEMENTATION OF A BACKUP RADIO CHANNEL FOR UAV CONTROL

DOI: 10.36724/2072-8735-2026-20-9-59-67

Manuscript received 11 June 2026;

Accepted 15 August 2026

Anastasia Yu. Kudryashova,
Moscow Technical University of Communications and
Informatics, Moscow, Russia, a.i.kudriashova@mtuci.ru

Keywords: UAV, control system, backup radio channel,
DC-DC, control command unit, LFSR, interference-resistant
encoding, binary encoding

The development of unmanned aerial vehicles (UAVs) in recent years has significantly expanded the possibilities for building mobile, rapidly deployable communication systems. One promising area is the use of unmanned aerial vehicle (UAV) as airborne mast, allowing for the rapid raising of relay equipment to a significant height. This article, based on a functional diagram of a fiber-optic communication line (FOCL) control system with power supply via a wire between the control center and the UAV used as an airborne communication mast, implements M-sequence encoding of a backup radio control channel. The following circuits have been developed: an onboard Direct Current to Direct Current (DC-DC) converter, a power supply for the onboard equipment, and an M-sequence unit. The article presents the results of developing a functional diagram of the control and power supply system for an unmanned aerial vehicle. To implement a backup, interference-resistant radio control channel, software is being developed using the Quartus II computer-aided design (CAD) system, which is designed for developing digital devices on programmable logic integrated circuits (FPGAs) and provides M-sequence encoding of data packets.

Information about author:

Anastasia Yu. Kudryashova, Moscow Technical University of Communications and Informatics (MTUCI), Ph.D., Associate Professor, Moscow, Russia

Для цитирования:

Кудряшова А. Ю. Разработка и отладка программного обеспечения для реализации резервного радиоканала управления БПЛА // T-Comm: Телекоммуникации и транспорт. 2026. Том 20. №9. С. 59-67.

For citation:

A. Yu. Kudryashova, "Development and debugging of software for the implementation of a backup radio channel for UAV control," *T-Comm*, 2026, vol. 20, no. 9, pp. 59-67.

Introduction

In previous research, the author developed a functional diagram of a fiber-optic communication line (FOCL) control system and a circuit for providing power via a wire between the control center and the unmanned aerial vehicle (UAV).

Mast devices built on wheeled chassis, with heights ranging from 20 to 50 meters, require 30 minutes to several hours to deploy the system. During this time, it is necessary to:

- stabilize and level the vehicle platform on the outriggers;
- raise the mast sections to the required height;
- install and tension cables to prevent the platform from tipping over in strong winds.
- lay cables from the ground station to the communication devices located on the tower;

The construction of permanent telecommunications masts now takes several weeks to months. This is because their installation requires:

- selecting the correct location for the antenna mast, with a clear line of sight to ensure stable communication with subscribers;
- design and develop the communications tower structure;
- prepare the site;
- manufacture the components and assemble the mast;
- install the communications equipment and lay the cables.

The deployment time of an airborne communications mast is now measured in minutes (from 5 to 15 minutes). Once the unmanned aerial vehicle reaches its destination, it launches. Once it reaches the required altitude, signal relaying begins. This is a significant advantage in situations where every minute is crucial, such as during search and rescue operations or combat situations.

Despite its high potential, the previously discussed technology of using an unmanned aerial vehicle as an airborne communications mast has a number of significant limitations that cannot be ignored.

1. Limited battery capacity, and therefore flight time, are the main technical barrier to implementing this technology.
2. Using radio communications to control an airborne communications mast makes the system vulnerable to interception by an adversary, which can lead to the loss of the drone and the leakage of

classified information. When exposed to electronic warfare, control of the unmanned aerial vehicle systems becomes impossible, resulting in the crash and destruction of the communications node.

3. Strong dependence on weather conditions is also a serious drawback of unmanned platforms. Strong winds, icing, and heavy rainfall can make unmanned aerial vehicle flight impossible or dangerous.

The range of telecommunications devices mounted on traditional antenna towers is highly dependent on terrain and the presence of various obstacles in the signal path.

Mountainous or rugged terrain creates so-called "dead zones," where the base station's signal either fails to reach the tower or is very weak, preventing proper communication between the tower-mounted transceivers and subscriber radios located within the zone. Eliminating this problem requires increasing the number of fixed towers, which significantly increases the cost of the entire communications system.

Using an airborne telecommunications platform offers a significant advantage in this regard. A unmanned aerial vehicle can vary its altitude from a few to several hundred meters, allowing it to cover areas otherwise inaccessible with ground-based towers. Thus, by varying its altitude, it becomes possible to adapt the coverage area to the current needs.

Unmanned platforms have a significant advantage in terms of survivability and vulnerability due to their high mobility and low observability. If threatened, the aircraft can quickly reposition or land. The loss of a single unmanned aerial vehicle is not catastrophic, as it can be quickly replaced by a backup, demonstrating the high fault-tolerance of the entire communications system.

A fiber-optic line is essentially a subtype of a wired communication channel. However, it uses fiber-optic cable as the transmission medium, in which the signal propagates as modulated light pulses [1, 2].

The only viable solution to the first two problems today is to use a wired power supply from a ground station and a fiber-optic communication line as the control channel for the aerial mast. This is the focus of the author's previous research.

The complete functional diagram of the wired power and fiber-optic control system for an unmanned aerial vehicle used as an aerial mast is shown in Figure 1.

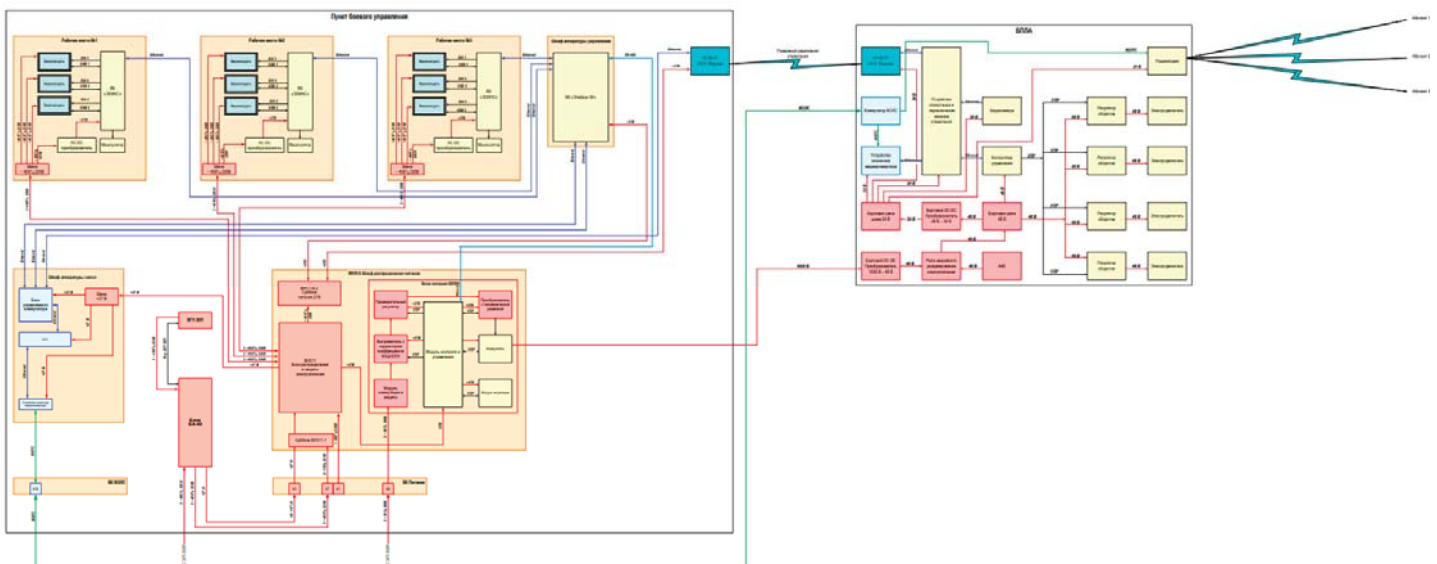


Fig. 1. Functional diagram of the fiber-optic power and control system for an unmanned aerial vehicle

Initial data for the development:

- unmanned aerial vehicle payload: at least 50 kg
- flight altitude: at least 100 m

Power is supplied to the control center from a diesel generator. The generator's output voltage can fluctuate significantly, significantly affecting the operation of the control console systems.

The development involves the use of an airborne mast for communications in special-purpose vehicles. During the implementation of measures to improve the mobility and technological advancement of communications equipment, the decision was made to use an unmanned platform to house the electronic systems onboard.

The aerial mast is connected to the control center via fiber-optic communication lines using a tactical optical cable, bundled with a power cable and wound onto a special reel to ensure the cable can be coiled and deployed during takeoff, landing, and altitude changes.

A type SO-156 connector will be used for connection to the equipment container's input box, while lightweight FC optical connectors are used on the mast side.

To prevent complete loss of control over the unmanned platform in the event of a break or damage to the cable, the mast must be equipped with a backup communication channel.

The optimal solution to this problem is the use of a radio module. In an emergency, this will allow the transmission of information to the subscriber to be terminated, other participants in the broadcast will be alerted to the failure, and an emergency landing can be performed.

A UV-89-07 Wi-Fi module was selected as a device capable of transmitting control commands to the mast, ensuring its mobility, and relaying information from the unmanned aerial vehicle to the complex's subscribers. The main characteristics of this device are summarized in Table 1.

Table 1

Main Characteristics of the UV-89-07

Characteristic name	Value
Operating frequency range	5.1 – 5.35 GHz
Transmitter power	No more than 0.6 W
Bandwidth of the transmitting and receiving channel	No less than 2 Mbps
Communication range	Up to 1 km
Power consumption	No more than 10 W

This module enables data exchange with similar devices using a multipoint-to-point configuration, where the module in multipoint mode is the master, and the device operating in point-to-point mode is the slave. Slave devices cannot communicate with each other.

In this system, the radio module installed onboard the unmanned aerial vehicle (unmanned aerial vehicle) is the slave, while the one located in the control center is the master. Data transfer between the radio device and the switch at the unmanned aerial vehicle control center or onboard the unmanned aerial vehicle occurs via a 10/100 Mbps Ethernet interface.

The unmanned aerial vehicle power supply output is expected to have a DC voltage of approximately 1000 V. This voltage is required to transmit 10 kW of power using a cable with minimal cable thickness.

The development of the onboard unit is subject to strict weight and size requirements. Its primary function is to step down the 1000 V input voltage to 48 V to power the controller and power units. The functional diagram of the step-down Direct Current to Direct Current (DC-DC) converter is shown in Figure 2.

The electrical circuit is based on parallel-connected pulse modules with galvanic isolation. Using five such stages with a rated power of 2000 W each produces a total output of 10 kW.

Given the module's 95% efficiency and conversion factor, the maximum power consumption at 1000 V will be approximately 10,526 W at 10.5 A. Using a 1.5 mm² cable approximately 100 m long, the voltage loss will be approximately 0.6 V per meter, resulting in approximately 940 V being supplied to the onboard DC-DC converter input.

The onboard unit's output voltage can vary from 51 V (at idle) to 45 V (at maximum load), which is acceptable for operating the speed controllers and electric motors. Voltage from the DC-DC converter's output is fed to the emergency power backup relay. A backup lithium-polymer battery (Li-PO) is also connected to it.

When operating from the ground power grid, the relay compares the voltage supplied by the onboard DC-DC converter and the battery. If the ground power fails or drops significantly, the relay automatically switches to the backup battery.

The onboard 48V bus distributes power to all controls operating at this voltage. Non-standard fiber-optic and radio control devices are powered by a DC voltage of 21-29V, requiring the installation of an additional step-down DC-DC converter [7].

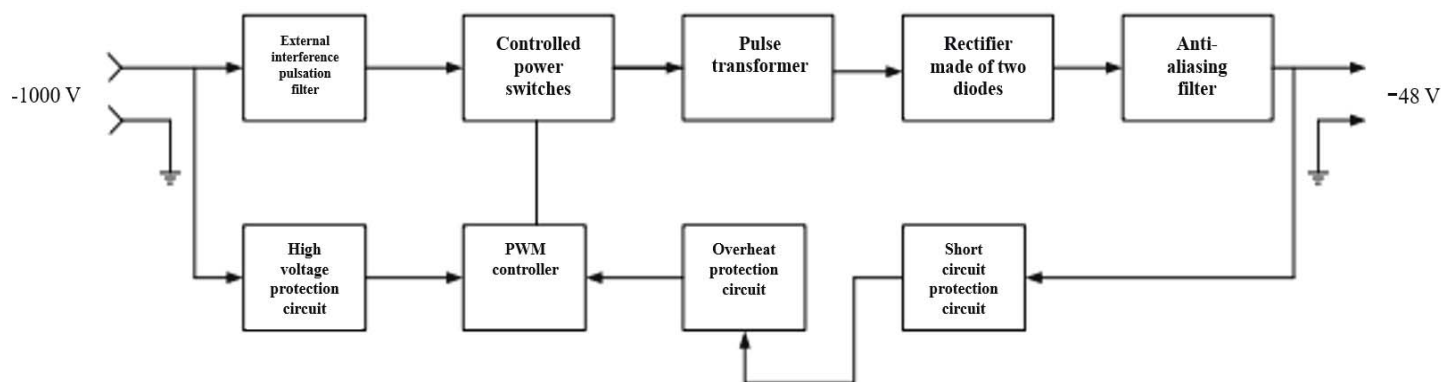


Fig. 2. Functional diagram of the on-board DC-DC converter

Developing the software code involves creating a model based on a block diagram, which reflects the functional diagram of the control module, assembled from basic logic components: and/or gates, multiplexers, triggers, etc. The module is organized into three blocks:

1. control command block
2. -sequence generation block
3. Hamming code generation block

Control Command Block Development

This block is built on a specific set of commands. Each command corresponds to a unique binary combination, which is stored as a sequence written in constant elements. The diagram is shown in Figure 3.

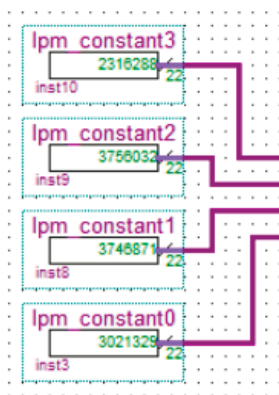


Fig. 3. Constant elements in which code combinations are written

The following code combinations were chosen for control commands:

- gas – 10001101011 (11 bits);
- yaw – 11100101010 (11 bits);
- pitch – 1110010010110000110111 (22 bits).

The selection of a command for transmission to the airmobile mast is carried out by specifying its number on the STATIC_COMMAND element. The selected binary sequence is sent to shift registers, which generate a strobe signal (INFO_STROB).

The active level of such a signal is a logical one. Its leading edge is synchronized with the beginning of the transmission of the information sequence, and the trailing edge is synchronized with its end. The control command block diagram is shown in Figure 4.

When STATIC_COMMAND = 10, the next element receives the INFO_IN information sequence, which corresponds to the "yaw" command, as well as the INFO_STROBE signal. The binary sequence is shown in Figure 5.

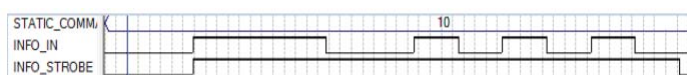


Fig. 5. Binary sequence for the "yaw" command

Changing the STATIC_COMMAND value to 01 changes the code sequence to "pitch" and generates a new strobe signal. The binary sequence is shown in Figure 6.

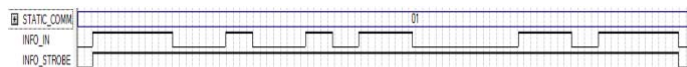


Fig. 6. Binary sequence for the pitch command

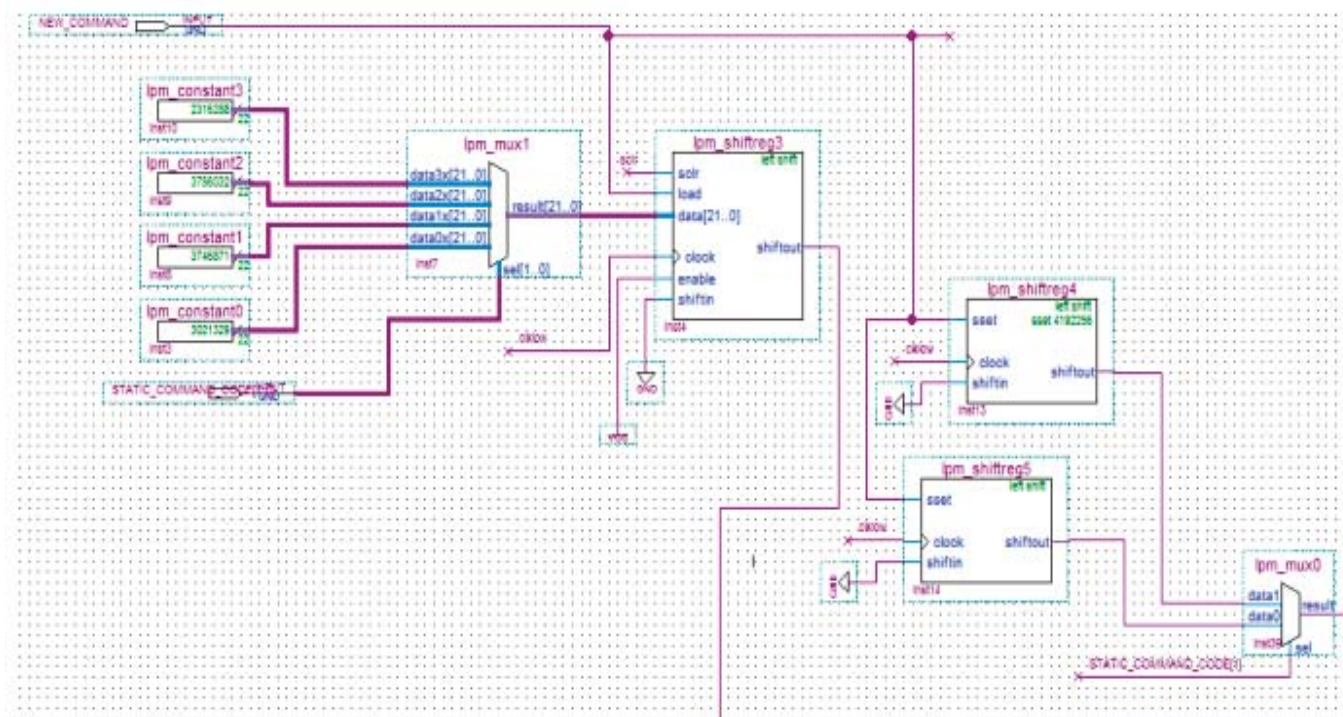


Fig. 4. Control Command Block

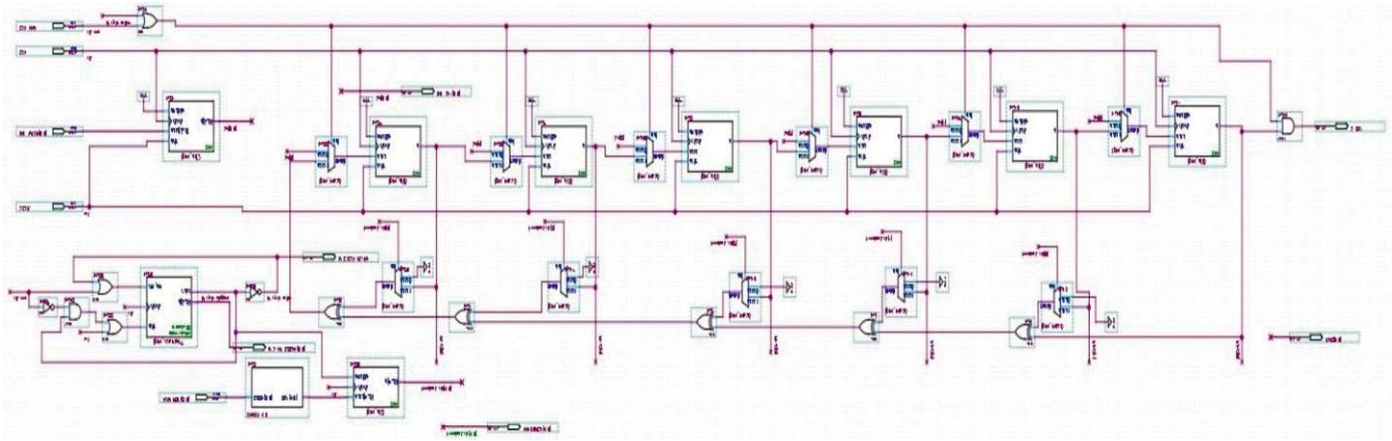


Fig. 7. Functional diagram of the M-sequence block

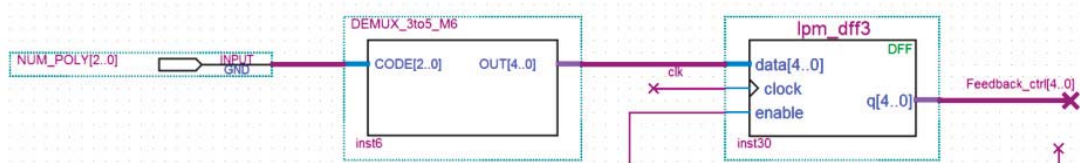


Fig. 8. Feedback configuration selection and fixation node

M-Sequence Generation Block Development

An M-sequence is a pseudo-random binary sequence with the maximum possible repetition period for a given circuit, defined by the formula $N = 2^m - 1$, where m is the bit depth of the linear feedback shift register (LFSR) that generates this type of sequence. Each bit of the shift register can only take two values – one or zero – as it is an element built on binary logic [3-6].

An important property of the M-sequence is that any combination of m symbols occurs no more than once per period. A 6-bit linear feedback shift register was chosen for the implementation, resulting in a sequence period of 63. The basis for constructing this shift register is a generator polynomial whose degree is equal to the register's bit depth. The coefficients of the polynomial determine the feedback structure.

For a 6-bit register, there are six 6th-degree polynomials. The bit depth of a linear-feedback shift register circuit contains m D-flip-flops, the input of each of which (except the last) is connected to the output of the next.

The input of the last (highest) flip-flop is fed with the result of the modulo-2 addition (XOR) operation from the outputs of those flip-flops whose numbers correspond to the non-zero coefficients of the low-order terms of the selected polynomial. The circuit is shown in Figure 7 [8-10].

During the development of this circuit, the ability to select one of the six existing generator polynomials was implemented. To do this, a 3-bit combination corresponding to one of the possible polynomials must be fed to the NUM_POLY[2...0] input before the M-sequence is generated. The polynomials and corresponding combinations are presented in Table 2.

Table 2

Correspondence of Code Combinations to Generator Polynomials

NUM_POLY	Generator polynomial
000	$x^6 + x + 1$
001	$x^6 + x^4 + x^3 + x + 1$
010	$x^6 + x^5 + 1$
011	$x^6 + x^5 + x^2 + x + 1$
100	$x^6 + x^5 + x^3 + x^2 + 1$
101	$x^6 + x^5 + x^4 + x + 1$

The selected NUM_POLY combination is fed to the decoder, which generates a 5-bit control signal. Each bit of this signal controls the on/off of the feedback loops of the D-type flip-flops included in the shift register. The configuration is stored in a special register until the end of the M-sequence generation. The circuit diagram is shown in Figure 8 [11-14].

Feedback control is implemented using lpm_mux0 multiplexers, controlled from the Feedback ctrl bus. The schematic is shown in Figure 9.

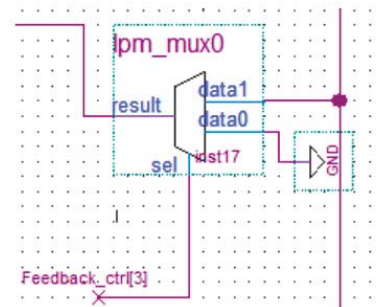


Fig. 9. Feedback connection diagram via multiplexers

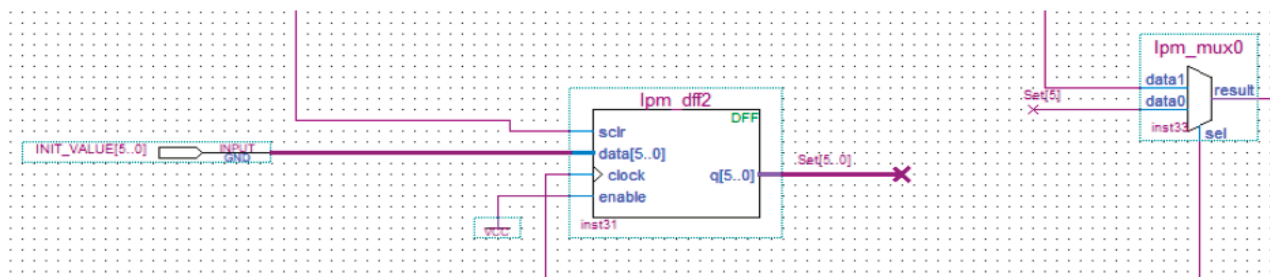


Fig. 10. Schematic diagram of the initial setup unit of the shift register with linear feedback

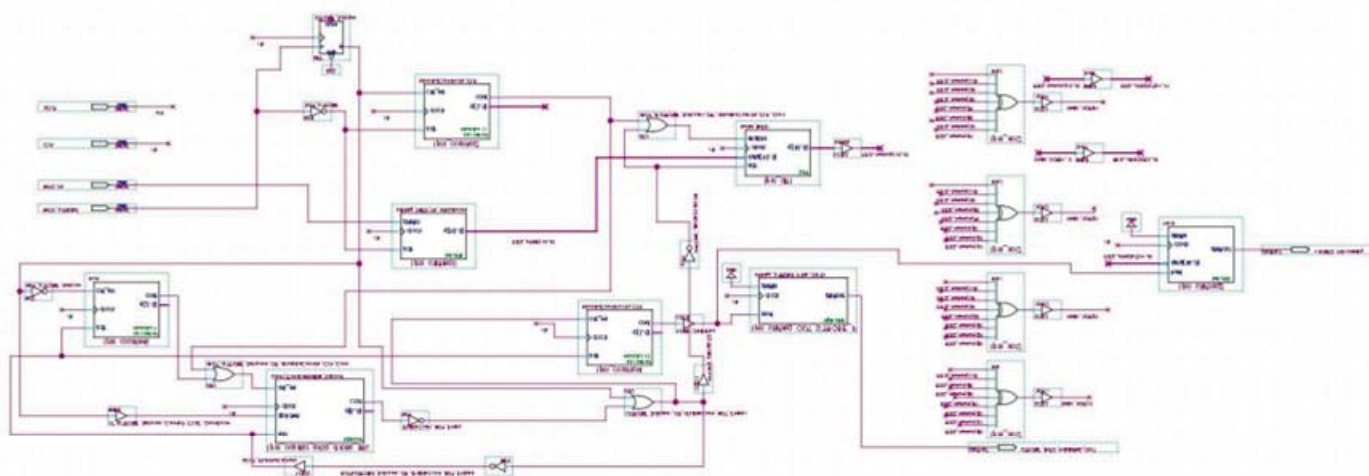


Fig. 11. Functional diagram of the Hamming code complement block

The block operates as follows:

1. INFO_IN information bits are fed sequentially to the input of the lpm_shiftreg0 shift register, which outputs them over the parallel REG shiftin [10..0] bus, 11-bit at a time.
2. The received 11-bit word is stored in the lpm_dff1 register. This allows the receiving shift register to simultaneously receive the next word while the current one is being processed.
3. Each information bit from the buffer register, in accordance with Table 1, is fed to the corresponding inputs of the exclusive-or elements (lpm_xor2), where four check bits are calculated: HMM_XOR[1] = 0, HMM_XOR[2] = 0, HMM_XOR[3] = 1, HMM_XOR[4] = 0.

The multi-input modulo-2 adder was replaced with five two-input adders, which sequentially sum the output of a D-flip-flop either to the output of the next D-flip-flop, when feedback is enabled, or to a logical zero, when the operating system is disabled.

Since a zero state for all flip-flops is prohibited (resulting in the generator freezing), an initial setup mechanism was added to the circuit. An arbitrary non-zero 6-bit value is written to the register via the INIT_VALUE [5...0] input. Additional multiplexers are used to load this value.

During the initialization phase, these multiplexers apply the INIT_VALUE value to the inputs of the D-flip-flops, and during the operating phase, they switch to signals from the feedback circuit and previous bits. The circuit diagram of the unit is shown in Figure 10.

Hamming Code Complement Block

The Hamming code belongs to the class of error-correcting codes capable of detecting and correcting single errors and fixing double errors.

The process of generating check bits (p) for the (15,11) code is described by the following relations, where D0 – D10 are the bits of the original information combination. The logic is presented in Table 3.

Table 3

Logic for Generating Check Bits of the Hamming Code

Initial information sequence	D0	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10
Hamming code											
check bits											
p0											
p1											
p2											
p3											

Explanations for the table using formulas:

$$p_0 = D_0 + D_1 + D_3 + D_4 + D_6 + D_8 + D_{10}$$

$$p_1 = D_0 + D_2 + D_3 + D_5 + D_6 + D_9 + D_{10}$$

$$p_2 = D_1 + D_2 + D_3 + D_7 + D_8 + D_9 + D_{10}$$

$$p_3 = D_4 + D_5 + D_6 + D_7 + D_8 + D_9 + D_{10}$$

For example, let's select the polynomial $x^6 + x + 1$, which corresponds to a NUM_POLY combination of 000 and an initial value of INIT_VALUE [5...0] = 000001. After loading these parameters and receiving the enable signal CLK_ENA, generation begins.

At each clock cycle, the register contents are shifted, the least significant bit is output, and the most significant bit is written to the XOR result of the zeroth (in this polynomial) and first bits. After 63 clock cycles, the register state repeats, and the sequence begins again. The CLK_ENA signal also controls the lock on changes to the initial value and polynomial configuration during generation.

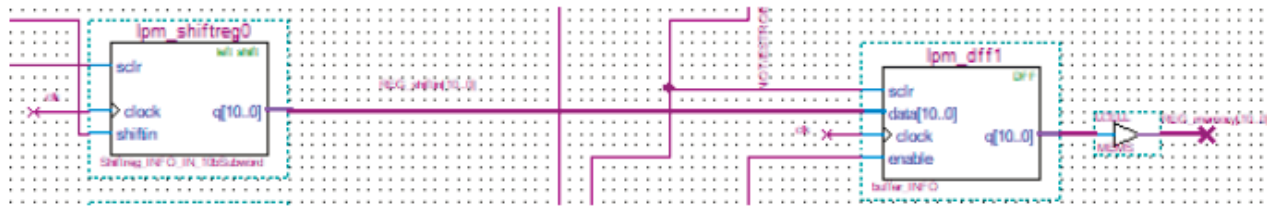


Fig. 12. Conversion of serial code to parallel code

The Hamming code complementation block. The circuit is shown in Figure 11.

The algorithm is clearly shown in Figures 12 and 13.

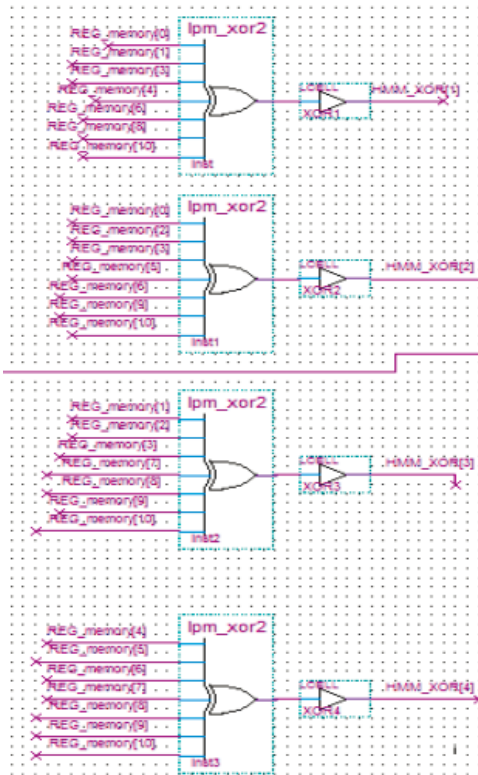


Fig. 13. Generating Check Symbols

1. Using the concatenation operation in inst7 and inst8, four calculated check symbols are appended to the original information bits, forming a 15-bit codeword.

2. The generated sequence of binary symbols is loaded into the output shift register lpm_shiftreg2 via the parallel bus and is output sequentially from the 11bINFO_4bHamming output.

The execution of the operation and the output register are shown in Figures 14 and 15.

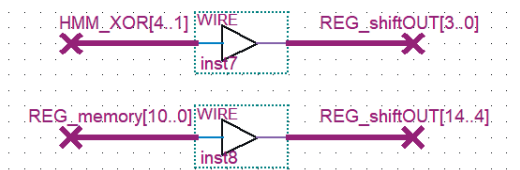


Fig. 14. Scheme of execution of the concatenation operation

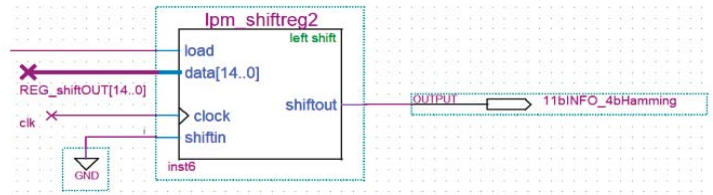


Fig. 15. Output Shift Register

3. To synchronize subsequent blocks, a STROBE_INFO_HAMMING_OUT strobe pulse is generated. It is generated by the lpm_shiftreg_OUT_STROBE_P shift register, indicating the beginning and end of the 15-bit sequence to the M-sequence generation block.

The generation of the strobe signal and its output sequence are shown in Figures 16 and 17.

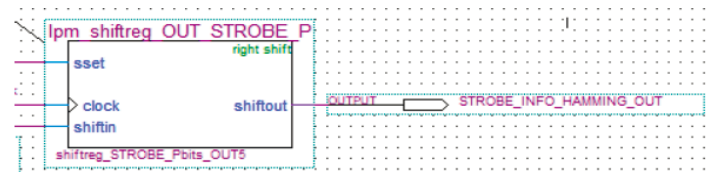


Fig. 16. Formation of a strobe signal



Fig. 17. Output sequence with strobe signal

The resulting 15-bit sequence is encoded using an M-sequence to improve signal noise immunity. This operation is performed using an exclusive-or-not (XNOR) logic element.

The clock frequencies of the blocks are synchronized such that during the transmission of one bit of the information sequence with additional Hamming symbols, the M-sequence block has time to fully output its entire 63-bit period. The STROBE_INFO_HAMMING_OUT strobe triggers the generation of the M-sequence, ensuring the synchronous start of the two sequences. The XNOR element receives the information bit and the M-sequence bits.

If the information symbol is "1," the output displays the normal M-sequence bit values; if the information bit is "0," the output displays an inverted version of the M-sequence. Thus, the circuit outputs a continuous bit stream consisting of alternating direct and inverted M-sequence symbols, which correspond to the bits of the encoded message. The output sequence of the circuit is shown in Figure 18.



Fig. 18. Code sequence at the circuit output

Justification of Development Feasibility

The justification of the feasibility of developing a software product can be conducted based on the results of comparing the new version with an existing equivalent, determining the quantitative values of quality characteristics and their significance [15].

All characteristics selected for analysis and their values are summarized in Table 4.

Table 4

Approximate scale of novelty of proposals

Software Quality Characteristics	Units of measurement	Values of PP quality characteristics		The importance of characteristics
		Analogue	New option	
Suitability	Points	3	4	0,3
Error Resistance	Points	2	3	0,4
Understandability	Points	3	4	0,1
Changability	Points	4	5	0,1
Structured	Points	3	4	0,1

Since the software product being developed is not final, the technical level of the developed version must be linked to the degree of influence it has as a component on the final outcome of the functioning of the system at a higher level of the hierarchy.

The technical level index of software product development will be equal to:

$$J_{tl} = 1,39 * (0,25 + 1) = 1,73 \quad (1)$$

Based on the technical level index (1) of software development, one can conclude that the solutions obtained as results of the evaluated development are superior to those of similar developments conducted previously. This is largely due to the development of personal computers and software products used in the study, as well as a rational approach to planning and development.

The cost-effectiveness of development costs when using its results in scientific applications is $E = 1.48$. A coefficient of $E > 0.5$ corresponds to a science-intensive product and indicates a full return on investment within two years. Therefore, the ongoing development of software for implementing a backup radio control channel for an unmanned aerial vehicle used as an airborne communications mast can be considered cost-effective.

During the development process, the developer's working conditions were assessed based on an analysis of the following production factors: indoor microclimate, vibroacoustic factors (noise), lighting environment, non-ionizing radiation, and work severity and intensity. The assessment revealed that all factors meet established standards.

Conclusion

During the development process, potential operational issues with the wired power supply system and fiber-optic line were taken into account. Therefore, to ensure their reliability, a backup radio control channel and onboard lithium-polymer battery were added to the design. This ensures that, in the event of a failure of the primary systems, the mast can complete its mission and land safely.

To ensure the backup radio channel's noise immunity, software for the control module was developed in the Quartus II environment using a functional logic diagram. The diagram includes three main blocks: a control command block, an M-sequence generator,

and a Hamming code generator. This software is intended to be installed on a programmable logic integrated circuit (PLC) included in the UV-89-07 radio control unit.

References

1. S. F. Gorgadze, A. V. Ermakova, A. Yu. Kudryashova, "Group signals based on symmetric orthogonal matrices and processing of multipath signals," *T-Comm*. 2025. Vol. 19, No. 10, pp. 21-34. DOI 10.36724/2072-8735-2025-19-10-21-34.
2. V. S. Shalavin, A. Yu. Kudryashova, "Modeling the Spectrum of the Output Signal to a Response Signal Generator for a Radio Transmitter in an Inter-Satellite Communication Link," *Synchroinfo Journal*. 2025. Vol. 11, No. 6, pp. 34-39. DOI 10.36724/2664-066X-2025-11-6-34-39.
3. A. Yu. Kudryashova, A. I. Guro, "Design of a Reed-Solomon turbo codec," *DSPA: Application issues of digital signal processing*. 2024. Vol. 14, No. 4, pp. 4-19.
4. Y. Kudryashova and A. S. Adzhemov, "Building an Algorithm for Estimating the Effective Coding of a Source when Converting Signals in Various Metric Spaces," *2018 Wave Electronics and its Application in Information and Telecommunication Systems (WECONF)*, St. Petersburg, Russia, 2018, pp. 1-4, doi: 10.1109/WECONF.2018.8604380.
5. S. Adzhemov and A. Y. Kudryashova, "Development of a General Method for Evaluating the Effectiveness of Primary Coding," *2022 Wave Electronics and its Application in Information and Telecommunication Systems (WECONF)*, St. Petersburg, Russian Federation, 2022, pp. 1-5, doi: 10.1109/WECONF55058.2022.9803744.
6. A. S. Adzhemov, S. A. Adzhemov, A. Y. Kudryashova, "Methods of comparison of code combinations with message elements from the source," *Systems of Signal Synchronization, Generating and Processing in Telecommunications*. 2023. Vol. 6, No. 1, pp. 6-10. DOI 10.1109/SYNCHROINFO57872.2023.10178629.
7. A. Yu. Kudryashova, A. M. Makeev, "Development of the cable part of the HF path of the telecommand system," *REDS: Telecommunication devices and systems*. 2024. Vol. 14, No. 4, pp. 32-35.
8. G. A. Dolin, A. Y. Kudryashova, V. V. Frisk and V. N. Shakin, "Representation of Algorithms for Schematic Synthesis of Radio Engineering Devices in the Knowledge Base of the Expert System," *2020 International Conference on Engineering Management of Communication and Technology (EMCTECH)*, Vienna, Austria, 2020, pp. 1-5, doi: 10.1109/EMCTECH49634.2020.9261556.
9. G. A. Dolin, A. Yu. Kudryashova, "Synthesis of structural electrical circuits of radio engineering devices in a hybrid production expert system," *Synchroinfo Journal*. 2020. Vol. 6, No. 3, pp. 5-9. DOI 10.36724/2664-066X-2020-6-3-5-9.
10. G. A. Dolin, A. Y. Kudryashova, "Modified methods of circuit simulation of radio engineering devices in the time domain," *Synchroinfo Journal*. 2020. Vol. 6, No. 2, pp. 7-11. DOI 10.36724/2664-066X-2020-6-2-7-11.
11. A. Y. Kudryashova, T. I. Semyonova, V. V. Frisk, V. N. Shakin, "Study of the effectiveness of Scilab software means for solving optimization problems," *2020 Wave Electronics and its Application in Information and Telecommunication Systems, WECONF 2020*, Saint-Petersburg, June 01-05, 2020. Vol. 3. Saint-Petersburg, 2020. P. 9131166. DOI 10.1109/WECONF48837.2020.9131166.
12. V. N. Shakin, T. I. Semyonova, A. Y. Kudryashova, V. V. Frisk, "Comparison of computer modeling of RC filter in matlab and Scilab environments," *2020 Wave Electronics and its Application in Information and Telecommunication Systems, WECONF 2020*, Saint Petersburg, June 1-5, 2020. Vol. 3. Saint Petersburg, 2020. P. 9131473. DOI 10.1109/WECONF48837.2020.9131473.
13. T. I. Semyonova, V. N. Shakin, V. V. Frisk, A. Yu. Kudryashova, "Analysis of the efficiency of Scilab package tools in solving optimization problems," *Infocommunication and radioelectronic technologies*. 2020. Vol. 3, No. 2, pp. 149-161.
14. T. I. Semenova, V. N. Shakin, V. V. Frisk, A. Yu. Kudryashova, "Study of the efficiency of Scilab package tools in solving optimization problems," *Microwave engineering and telecommunication technologies*. 2020. No. 1-2, pp. 136-137.
15. A. Yu. Kudryashova, V. V. Khoroshun, "Development of automated testing tools for a switching equipment control system," *REDS: Telecommunication devices and systems*. 2024. Vol. 14, No. 4, pp. 21-26.

РАЗРАБОТКА И ОТЛАДКА ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ ДЛЯ РЕАЛИЗАЦИИ РЕЗЕРВНОГО РАДИОКАНАЛА УПРАВЛЕНИЯ БПЛА

Кудряшова Анастасия Юрьевна, Московский технический университет связи и информатики, Москва, Россия,
a.i.kudriashova@mtuci.ru

Аннотация

Развитие беспилотных летательных аппаратов (БПЛА) в последние годы существенно расширило возможности построения мобильных, быстро разворачиваемых систем связи. Одним из перспективных направлений стало использование беспилотного летательного аппарата в качестве аэромобильной мачты, позволяющей оперативно поднимать ретрансляционное оборудование на значительную высоту. В данной статье на основе функциональной схемы системы управления по волоконно-оптическим линиям связи (ВОЛС) с обеспечением электропитания по проводу между пунктом управления и беспилотным летательным аппаратом, используемым в качестве аэромобильной мачты связи реализовано кодирование резервного радиоканала управления с помощью М-последовательности. В том числе разработаны схемы: бортового Direct Current to Direct Current (постоянный ток в постоянный ток, DC-DC) преобразователя, обеспечения электропитания бортового оборудования, блока М-последовательности. Представлены результаты разработки функциональной схемы системы управления и обеспечения электропитания беспилотного летательного аппарата. Для реализации резервного помехозащищенного радиоканала управления разрабатывается программное обеспечение в системе автоматизированного проектирования (САПР) Quartus II, предназначенной для разработки цифровых устройств на программируемых логических интегральных схемах (ПЛИС) и обеспечивающей кодирование информационных посылок с помощью М-последовательности.

Ключевые слова: БПЛА, система управления, резервный радиоканал, DC-DC, блок команд управления, РСЛОС, помехоустойчивое кодирование, двоичное кодирование.

Литература

1. Горгадзе С.Ф., Ермакова А.В., Кудряшова А.Ю. Групповые сигналы на основе симметричных ортогональных матриц и обработка многолучевых сигналов // Т-Comm: Телекоммуникации и транспорт. 2025. Т. 19, № 10. С. 21-34. DOI 10.36724/2072-8735-2025-19-10-21-34. EDN AMTGTV.
2. Шалавин В.С., Кудряшова А.Ю. Моделирование спектра выходного сигнала генератора ответного сигнала для радиопередатчика в межспутниковой линии связи // Synchroinfo Journal. 2025. Т. 11, № 6. С. 34-39. DOI 10.36724/2664-066X-2025-11-6-34-39. EDN YBDSLQ.
3. Кудряшова А. Ю., Гуро А.И. Проектирование турбокодека Риди-Соломона // DSPA: Вопросы применения цифровой обработки сигналов. 2024. Т. 14, № 4. С. 4-19. EDN KADRVV.
4. Kudryashova A. Y., Adzhemov A.S. Building an Algorithm for Estimating the Effective Coding of a Source when Converting Signals in Various Metric Spaces // 2018 Wave Electronics and its Application in Information and Telecommunication Systems (WECONF), St. Petersburg, Russia, 2018, pp. 1-4, doi: 10.1109/WECONF.2018.8604380.
5. Adzhemov S., Kudryashova A. Y. Development of a General Method for Evaluating the Effectiveness of Primary Coding // 2022 Wave Electronics and its Application in Information and Telecommunication Systems (WECONF), St. Petersburg, Russian Federation, 2022, pp. 1-5, doi: 10.1109/WECONF55058.2022.9803744.
6. Adzhemov A. S., Adzhemov S. A., Kudryashova A. Y. Methods of comparison of code combinations with message elements from the source // Systems of Signal Synchronization, Generating and Processing in Telecommunications. 2023. Vol. 6, No. 1, pp. 6-10. DOI 10.1109/SYNCHROINFO57872.2023.10178629. EDN XWJYBA.
7. Кудряшова А. Ю., Макеев А.М. Разработка кабельной части ВЧ-тракта системы телеуправления // REDS: Телекоммуникационные устройства и системы. 2024. Т. 14, № 4. С. 32-35. EDN TYZROX.
8. Dolin G. A., Kudryashova A. Y., Frisk V. V., Shakin V. N. Representation of Algorithms for Schematic Synthesis of Radio Engineering Devices in the Knowledge Base of the Expert System // 2020 International Conference on Engineering Management of Communication and Technology (EMCTECH), Vienna, Austria, 2020, pp. 1-5, doi: 10.1109/EMCTECH49634.2020.9261556.
9. Dolin G. A., Kudryashova A. Y. Synthesis of structural electrical circuits of radio engineering devices in a hybrid production expert system // Synchroinfo Journal. 2020. Vol. 6, No. 3, pp. 5-9. DOI 10.36724/2664-066X-2020-6-3-5-9. EDN BRQDTU.
10. Dolin G. A., Kudryashova A. Y. Modified methods of circuit simulation of radio engineering devices in the time domain // Synchroinfo Journal. 2020. Vol. 6, No. 2, pp. 7-11. DOI 10.36724/2664-066X-2020-6-2-7-11. EDN USYMLF.
11. Kudryashova A. Y., Semyonova T. I., Frisk V. V., Shakin V. N. Study of the effectiveness of Scilab software means for solving optimization problems // 2020 Wave Electronics and its Application in Information and Telecommunication Systems, WECONF 2020, Saint-Petersburg, June 01-05, 2020. Vol. 3. Saint-Petersburg, 2020. P. 9131166. DOI 10.1109/WECONF48837.2020.9131166. EDN KBBBTC.
12. Shakin V. N., Semyonova T. I., Kudryashova A. Y., Frisk V. V. Comparison of computer modeling of RC filter in matlab and Scilab environments // 2020 Wave Electronics and its Application in Information and Telecommunication Systems, WECONF 2020, Saint Petersburg, June 1-5, 2020. Vol. 3. Saint Petersburg, 2020. P. 9131473. DOI 10.1109/WECONF48837.2020.9131473. EDN ZQFUV
13. Семёнова Т.И., Шакин В.Н., Фриск В.В., Кудряшова А.Ю. Анализ эффективности инструментов пакета Scilab при решении задач оптимизации // Инфокоммуникационные и радиоэлектронные технологии. 2020. Т. 3, № 2. С. 149-161. EDN HMXWFS.
14. Семёнова Т.И., Шакин В.Н., Фриск В.В., Кудряшова А.Ю. Исследование эффективности инструментов пакета Scilab при решении задач оптимизации // Микроволновая техника и телекоммуникационные технологии. 2020. № 1-2. С. 136-137. EDN EOBXJ.
15. Кудряшова А.Ю., Хорошун В.В. Разработка автоматизированных инструментов тестирования системы управления коммутационным оборудованием // REDS: Телекоммуникационные устройства и системы. 2024. Т. 14, № 4. С. 21-26. EDN AGCFMD.

Информация об авторе:

Кудряшова Анастасия Юрьевна, Московский технический университет связи и информатики (МТУСИ), к.т.н., доцент., Москва, Россия

INVESTIGATION OF THE COGNITIVE SATURATION THRESHOLD OF THE YOLOV8 MODEL UNDER LIMITED TRAINING DATA

DOI: 10.36724/2072-8735-2026-20-9-68-75

Maria Yu. Karelina,
State University of Management;
Moscow Automobile and Road Construction State Technical University (MADI), Moscow, Russia, karelinamu@mail.ru

Alexey A. Akulov,
State University of Management;
Moscow Automobile and Road Construction State Technical University (MADI), Moscow, Russia, a.akulov.98@mail.ru

Dmitriy S. Taldykin,
Moscow Automobile and Road Construction State Technical University (MADI);
State University of Management, Moscow, Russia,
dima.dima.taldykin@mail.ru

Manuscript received 20 June 2026;

Accepted 27 August 2026

This article was prepared as part of the implementation of research work, implemented at the expense of the federal budget (source of financing - the Ministry of Education and Science of the Russian Federation) on the topic: "Development of software for the agro-industrial complex for monitoring the state of the tops and fruits of greenhouse plants using machine vision and artificial intelligence" (code of the scientific theme FSFM-2025-0002)

Keywords: YOLOv8; tomatoes; weed detection; small datasets; model hallucinations

This study examines the training-set size at which YOLOv8m ceases to benefit from adding visually similar frames in weed detection among tomato plants. The experimental design reflects a common precision-agriculture constraint: RGB images are relatively easy to acquire, whereas agronomist-verified annotation requires substantially more time. The experiment used 600 images with a resolution of 640x640 pixels collected in greenhouse tomato rows during a single growing season. Three classes were annotated: tomato, weed vegetation, and soil-substrate background. Seven training series were used to test the hypothesis that generalization peaks near $N=250$ when the number of unique scenes is below 300. Beyond this point, additional frames acquired under nearly unchanged conditions were accompanied by an increase in high-confidence false detections. The effect was characterized by a hallucination coefficient and an oversaturation index comparing the variance of the weights in the final detection layer with the mean validation loss. At $N=600$, Precision was 41.94% below its peak value and the hallucination coefficient reached 1.80. The practical implications concern annotation planning and detector-capacity selection: increasing the number of near-duplicate frames should not be treated as an unconditional means of improving the reliability of an automated system.

Information about authors:

Maria Yu. Karelina, Doctor of Technical Sciences, Doctor of Pedagogical Sciences, Professor, Head of the Department of Logistics and Transport and Technological Systems, State University of Management, Head of the Department of Machine Parts and Theory of Mechanisms at the Moscow Automobile and Road Construction State Technical University (MADI), Moscow, Russia. ORCID ID: 0000-0003-0335-7550

Alexey A. Akulov, Candidate of Technical Sciences, Associate Professor of the Department of Logistics and Transport and Technological Systems of the State University of Management, Associate Professor of the Department of Machine Parts and Theory of Mechanisms of the Moscow Automobile and Road Construction State Technical University (MADI), Moscow, Russia. ORCID ID: 0000-0001-8952-0237

Dmitriy S. Taldykin, Moscow Automobile and Road Construction State Technical University (MADI), State University of Management, Moscow, Russia. ORCID ID: 0009-0008-5071-426X

Для цитирования:

Карелина М.Ю., Акулов А.А., Талдыкин Д.С. Исследование порога когнитивного насыщения модели YOLOv8 в условиях ограниченной обучающей выборки // Т-Comm: Телекоммуникации и транспорт. 2026. Том 20. №9. С. 68-75.

For citation:

M.Yu. Karelina, A.A. Akulov, D.S. Taldykin, "Investigation of the cognitive saturation threshold of the YOLOv8 model under limited training data," T-Comm, 2026, vol. 20, no. 9, pp. 68-75.

Introduction

A detector output acquires technological significance before any mechanical action on the plants begins. The camera evaluates a section of a tomato row, the recognition module separates weed vegetation from the inter-row space and the soil-substrate surroundings, and the result is then transferred to the actuator system. At this point, an error ceases to be merely a classification error. An incorrectly assigned class may alter the command sent to the working tool and, consequently, the course of a specific agromonic operation.

The visual structure of a greenhouse row offers no single separating feature. Soil, substrate, trellis components, drip tape, plant residues, shadows, and structural elements can occur in the same frame. Against this background, tomato leaves have a complex dissected contour: their orientations differ, adjacent blades overlap, and texture varies even within one plant. Weed color is often close to that of the crop. The boundaries among the classes “tomato - weed - background” must therefore be inferred from a combination of morphological and background cues; no stable single criterion based on color or geometry exists in the scenes considered.

The main data shortage did not arise at the image-acquisition stage. Capturing RGB images required comparatively little time, whereas expert preparation of one frame containing several plants and background regions could take up to 15 minutes. Scenes with small weeds, overlapping leaves, and ambiguous object boundaries were especially time-consuming to annotate. Consequently, the volume of raw photographs substantially exceeded the number of images verified by a specialist. Hundreds of reliably validated scenes, rather than thousands of raw frames, were therefore available for training.

For such a dataset, the number of files must be distinguished from the number of newly observed conditions. A frame acquired with a different plant geometry or a different background state contributes new information to the model. A sequence recorded during one season, in the same row type, and under similar illumination is different. Similar leaf shadows, substrate patterns, trellis directions, drip-tape edges, and reflections on wet surfaces recur from frame to frame. The network receives useful object cues together with a repeated noise structure; under these conditions, increasing N is no longer equivalent to increasing diversity.

Most published studies in agricultural machine vision use YOLO to detect fruit, weeds, plant diseases, and technological objects. With large datasets, attention is naturally concentrated on final accuracy. A small dataset raises a different question: what happens when additional images barely expand the space of observed scenes? Of particular interest is the regime in which aggregate metrics have deteriorated only moderately, whereas high-confidence false bounding boxes have already appeared among the errors.

For an agricultural robot, two predictions assigned to the same wrong class may have different technological consequences. A low-confidence result can be removed by thresholding or sent for repeated analysis. A value above 0.9, by contrast, normally falls within the detector's confidence region. If such a bounding box is false, the actuator loop may treat a weed-free area as a target and perform an unnecessary local operation. In this study, this specific type of error is referred to as a model hallucination.

The study was limited to $N \leq 600$ and to a single task: weed detection in tomato plantings using YOLOv8m. The objective was to locate the boundary beyond which enlarging a structurally similar dataset ceases to improve generalization. A dataset with a fixed annotation scheme was prepared, and training series were carried out for different values of N . For each series, mAP@0.5, Precision, Recall, the hallucination coefficient, the L_2 gradient norm, and the oversaturation index were compared; activation maps were additionally examined for the rational and oversaturated regimes.

In agricultural object detection, YOLO is attractive primarily because of its one-stage prediction scheme: object coordinates and class are determined within a single network. The family is therefore used for plant monitoring, fruit detection, robotic harvesting, and weed detection. The requirements of a mobile platform, however, extend beyond accuracy alone. Computational load and robustness to heterogeneous backgrounds become part of the same engineering problem because the detector must operate within a constrained onboard computing loop.

For tomato crops, the literature more often addresses fruit ripeness, developmental stages, and robotic harvesting. Typical sources of error are familiar: leaf occlusion, complex backgrounds, and variable illumination. Attention modules, lightweight blocks, or other YOLOv8 modifications are commonly proposed in response. Such architectural changes may improve detection, but the number of original annotated scenes remains a separate issue. Architectural complexity cannot create plant or environmental states that are absent from the source material.

The high mAP values reported for multiclass weed detection in cotton using YOLOv8 [4, p. 2] were obtained from a different observation structure. A field dataset covers a broader range of scenes, whereas the present experiment was deliberately confined to a greenhouse row and one growing season. Weed species composition, background, and the range of acquisition conditions all differ. The information content of the two datasets is therefore not directly comparable. The required size of the small tomato dataset was established in a dedicated experimental series rather than by transferring a threshold from the field domain.

Additional ambiguity is introduced by the tomato leaf apparatus itself. For cereal crops, recognized difficulties include similarity in crop and weed color, object overlap, and the absence of an unambiguous size cue. The large dissected tomato leaf adds another class of background repetition: its shadows and characteristic texture fragments recur at different locations along the row and often appear near annotated target objects. In a small dataset, such co-occurrence can reinforce an accidental background combination as a weed-related cue.

The concept of double descent is used here only as a framework for discussing non-monotonic error behavior. The conducted series does not reproduce the classical double-descent curve and is not intended to demonstrate its universality. The relevant point is a possibility recognized in modern learning theory: more examples or a more complex model do not improve generalization monotonically in every regime. In the present formulation, the content of each new image is decisive. Dataset expansion is informative while a frame introduces a previously unrepresented scene state. When additional frames are close realizations of an already observed distribution, the increase in N remains quantitative and may not expand the feature space available to the model.

YOLOv8m has substantial parametric capacity for this regime. According to the Ultralytics documentation, the model contains 25.9 million parameters at an input size of 640 pixels. This expressiveness is sufficient for complex object detection; on several hundred closely related scenes, however, the same capacity also creates an opportunity to memorize incidental background dependencies. Along with tomato and weed morphology, the network may interpolate covariations associated with recurrent image structure.

Changes in the network's internal focus were evaluated using Grad-CAM. The method produces a map of image regions that contribute most strongly to the decision of a convolutional model. Under a rational data regime, activations are expected to concentrate on leaves, stems, and weed vegetation. A shift toward background textures in an oversaturated regime, conversely, indicates weakening of the detector's object-specific focus.

The scientific problem addressed in this study is therefore the identification of a boundary beyond which expansion of a small, weakly variable dataset no longer broadens the model's representation of target objects. Past this boundary, the repeated noise component of the empirical distribution receives additional weight and strengthens memorization. The corresponding state is termed cognitive saturation of the model.

Materials and Methods

The source dataset contained 600 RGB images with a resolution of 640x640 pixels, acquired in tomato rows during one growing season. This deliberately narrow domain was selected to isolate saturation from ordinary domain variability. A change of season, cultivar, camera, or cultivation technology would have introduced an additional covariate shift and complicated interpretation of the saturation effect. The experiment therefore did not aim to estimate maximum domain transferability; it sought an internal limit within a closely related distribution of scenes.

Annotation was performed for three classes. The “tomato plant” class included leaves, distinguishable stem fragments, and other visible tomato parts that could be separated from the surroundings. The “weed vegetation” class covered unwanted plants in the inter-row area and near the root zone. Problematic background fragments - soil, substrate, plant residues, drip tape, shadows, and free background regions - were assigned to the “soil-substrate background” class. The third class was retained specifically because such fragments could provoke false detections.

Training subsets were formed for seven values of N : 100, 200, 250, 300, 400, 500, and 600 images. Class proportions were maintained as the dataset size changed. Source scenes included in validation were excluded from training. This restriction is methodologically important for a small dataset because placing nearly identical frames on opposite sides of the split artificially inflates the estimated performance.

The transformation set was intentionally restricted. Only rotations of ± 15 degrees and horizontal flips were applied; intensive color correction, Mosaic, CutMix, and strong geometric distortions were excluded. This regime made it possible to observe the boundary between useful data expansion and memorization of the noise component of the empirical distribution without masking it by strong augmentation-based regularization.

Training used early stopping with patience=50 epochs, and weight checkpoints were saved every 10 epochs. Final weights, validation loss, the confidence-score distribution, and the L_2 norm

of gradients in the final detection layer were recorded for each series. The stochastic parameter-update rate was held constant across the series. Its potential influence on the location of the saturation threshold was therefore excluded from the cross-regime comparison.

The loss function was represented as the sum of a localization component and a focal component. The latter accounted for imbalance between object and background regions. The basic form was:

$$L_{total} = L_{CIoU} + \alpha \cdot L_{Focal}, \alpha = 0.75 \quad (1)$$

In expression (1), L_{CIoU} characterizes the localization quality of bounding boxes. L_{Focal} reduces the contribution of easily classified examples and reallocates weight toward difficult objects; the focal term was weighted by $\alpha=0.75$. This choice is consistent with the foreground-background imbalance known for one-stage detectors [9, p. 319].

High-confidence false responses were described using a separate hallucination coefficient. It was calculated as the mean logarithmically normalized number of false detections with confidence above 0.9:

$$Khal = (1/M) \cdot \sum_{j=1..M} \ln(1 + FP_{0.9, j}) \quad (2)$$

Here, M is the number of images in the validation set, and $FP_{0.9, j}$ is the number of false bounding boxes in the j -th image at confidence > 0.9 . Logarithmic normalization converts the indicator into a dimensionless form and limits the influence of a single frame containing an anomalously large number of false objects.

Precision, Recall, and $mAP@0.5$ alone were insufficient to describe changes within the final detection layer. The relationship between parameter dispersion and validation error was therefore compared across the training series. For this purpose, the oversaturation index was introduced:

$$Sover = \sigma^2(W_{last}) / \mu(L_{val}) \quad (3)$$

In expression (3), the numerator $\sigma^2(W_{last})$ characterizes the variance of the weights in the final detection layer, whereas $\mu(L_{val})$ in the denominator is the mean validation loss. The *Sover* index combines two quantities that are usually evaluated separately. Within the present experimental series, its increase was considered together with the appearance of locally confident predictions associated with unstable background cues. This interpretation is specific to the adopted formulation and does not define a universal threshold for other architectures or datasets.

The restructuring of the detection head was also monitored through gradients. Their L_2 norm was calculated for the final layer using:

$$G2 = \sqrt{\sum_{i=1..m} (dL / dw_i)^2} \quad (4)$$

Conventional performance assessment was retained through three standard object-detection metrics: Precision, Recall, and $mAP@0.5$. They were calculated as:

$$Precision = TP / (TP + FP) \quad (5)$$

$$Recall = TP / (TP + FN) \quad (6)$$

$$mAP@0.5 = (1/C) \cdot \sum_{c=1..C} AP_c(IoU=0.5) \quad (7)$$

The spatial Grad-CAM comparison was supplemented with a numerical measure of the difference between activation distributions. The Kullback-Leibler divergence was calculated between the rational and oversaturated regimes:

$$DKL = \sum_z P_{250}(z) \cdot \ln(P_{250}(z) / P_{600}(z)) \quad (8)$$

Results

The results of the seven training series were grouped by the type of observed characteristic. Metrics directly related to detection performance are presented in Table 1. Training time, L_{val} , the L_2 norm, and *Sover* are summarized in Table 2. This separation does not alter the data and keeps both tables readable in portrait A4 format.

Table 1

Detection performance metrics for different training-set sizes

N	<i>mAP</i>	<i>Prec.</i>	<i>Rec.</i>	<i>Khal</i>
100	0.67	0.71	0.58	0.05
200	0.83	0.86	0.79	0.12
250	0.91	0.93	0.88	0.18
300	0.88	0.89	0.81	0.45
400	0.80	0.78	0.70	0.89
500	0.71	0.65	0.59	1.35
600	0.62	0.54	0.48	1.80

Table 2

Training time, gradient L_2 norm, and oversaturation index as N changes

N	<i>Time, h</i>	L_{val}	L_2	<i>Sover</i>
100	0.8	0.91	0.40	0.82
200	1.5	0.52	0.52	1.03
250	2.0	0.31	0.64	1.18
300	2.4	0.39	0.91	2.08
400	3.1	0.58	1.36	2.54
500	3.9	0.75	1.82	3.02
600	4.6	0.94	2.30	3.45

The first series represents a shortage of object examples. With 100 images, $mAP@0.5$ was 0.67, Precision was 0.71, and Recall was 0.58. The detector already responded to large objects but frequently missed small weeds and visually ambiguous parts of tomato plants. *Khal* was only 0.05. The small number of high-confidence false responses should not be interpreted as robustness: the model remained insufficiently active in this regime and simultaneously exhibited low recall.

The largest positive shift in the three primary metrics occurred not at the end of the experimental range but after the second hundred images were added. At $N=200$, $mAP@0.5$ reached 0.83, Precision 0.86, and Recall 0.79. The dataset now contained new leaf positions, different planting densities and weed levels, and altered background states. *Khal* increased from 0.05 to 0.12, but this increase was substantially smaller than the gain in detection performance. Over this interval, increasing N was still accompanied by a meaningful broadening of the observed scenes.

The highest combined performance was obtained with 250 images. $mAP@0.5$, Precision, and Recall were 0.91, 0.93, and 0.88, respectively. For the dataset under study, this regime was adopted as the rational saturation level. The number of scenes was already

sufficient for class discrimination, whereas recurrent background details - substrate patterns, drip-tape edges, and leaf shadows - had not yet begun to dominate object decisions according to the results of the series.

The next experiment interrupted the upward trend. An additional 50 images reduced $mAP@0.5$ to 0.88; Precision and Recall fell to 0.89 and 0.81. The decline in aggregate metrics was still modest. *Khal* changed much more strongly: 0.18 at $N=250$ became 0.45 at $N=300$. The first visible restructuring therefore appeared not as an abrupt collapse of mAP but as an accumulation of false bounding boxes to which the network assigned high confidence.

Inspection of predictions from the 400-image series revealed a pronounced spatial localization of errors. False boxes appeared predominantly on furrows, substrate boundaries, elongated stem shadows, and regions containing plant residues. Repeated linear geometry was a common property of these areas. At the same time, $mAP@0.5$ decreased to 0.80, Precision to 0.78, Recall to 0.70, and *Khal* increased to 0.89. Thus, degradation in this series was not spatially uniform and was associated with specific configurations of the background scene.

The final two regimes differed from the preceding points not by an isolated deviation in one metric but by a general deterioration of detection. At $N=500$, $mAP@0.5=0.71$, *Precision*=0.65, *Recall*=0.59, and *Khal*=1.35 were obtained. With 600 images, these values changed to 0.62, 0.54, 0.48, and 1.80. The nature of the error also changed: some false objects had high confidence. Such predictions are more likely to be accepted by the actuator loop as valid. For targeted inter-row treatment, the technologically dangerous transition is specifically the shift from frequent misses to confident false detection of a background region.

To quantify the difference between the regime with maximum Precision and the final experimental point, the relative decrease in this metric was calculated:

$$\Delta P = ((P_{250} - P_{600}) / P_{250}) \cdot 100\% = 41.94\% \quad (9)$$

Substitution of the $N=250$ and $N=600$ values into expression (9) gives 41.94%, or approximately 42%. Over the same range, *Khal* reaches 1.80. These quantities describe different aspects of the same practical risk: positive predictions become less accurate, while the share of high-confidence errors among false decisions increases. A background region may therefore be misinterpreted as a weed more frequently and transferred to the control loop as a basis for actuation.

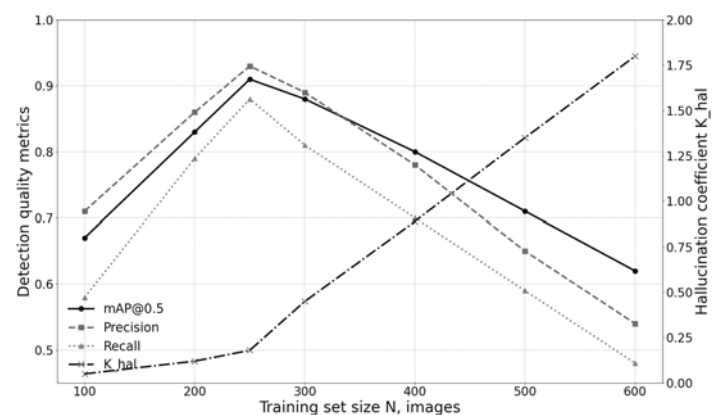


Fig. 1. Changes in detection metrics and hallucination coefficient with increasing training-set size

The four curves in Figure 1 move in the same direction only up to $N=250$. At this point, $mAP@0.5$, *Precision*, and *Recall* reach their maxima; further enlargement of the dataset is accompanied by a decline in all three. *Khal* follows the opposite trajectory, increasing over almost the entire experimental range and more rapidly after $N=300$. The interval between 250 and 300 images therefore cannot be treated as an ordinary plateau. Mean metrics remain close to their maxima, yet the structure of false decisions is already shifting toward high-confidence detections.

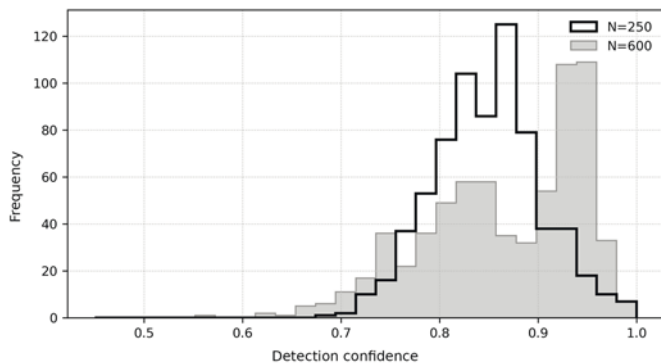


Fig. 2. Detection-confidence distributions for $N=250$ and $N=600$

Figure 2 exposes the difference without reducing it to a single averaged metric. At $N=250$, correct predictions are concentrated mainly around a confidence of 0.85. In the $N=600$ series, the distribution contains two pronounced regions; the second corresponds to false responses at approximately 0.92-0.96. Simply raising the confidence threshold does not resolve the problem. The erroneous objects lie precisely in the part of the scale that is usually considered the most reliable.

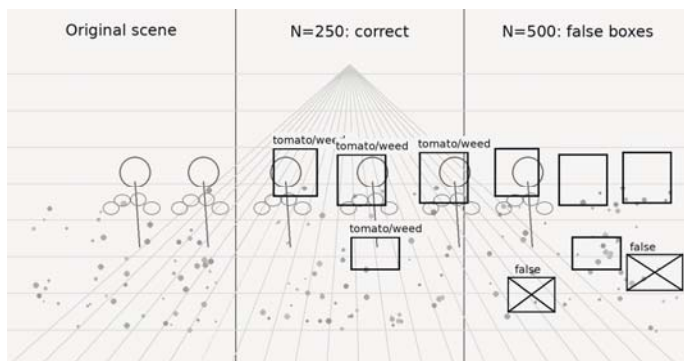


Fig. 3. Comparison of YOLOv8m prediction localization on the tomato scene and background regions

The schematic in Figure 3 separates two fundamentally different types of wrong decision. In the $N=250$ series, bounding boxes correspond to tomato plants and weed vegetation that are actually present. At $N=500$, additional bounding boxes appear on background regions; they are marked with crosses in the schematic. These cases no longer concern an inaccurate object boundary. The target object itself is absent, yet the detector still produces a positive prediction.

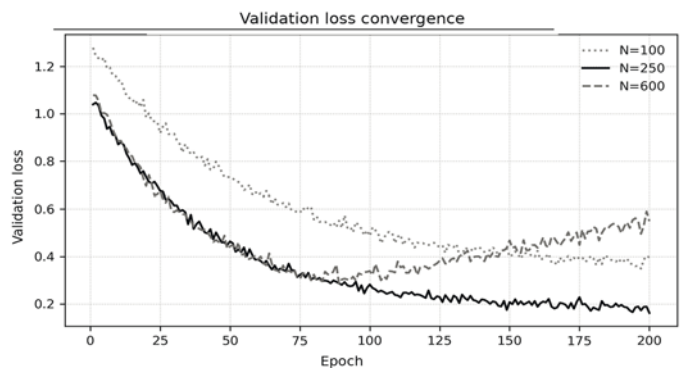


Fig. 4. Validation-loss trajectories for three training regimes

Comparison of the curves in Figure 4 reveals different causes of elevated error. At $N=100$, validation loss remains at a high residual level, which is consistent with a shortage of training examples. The $N=250$ regime behaves differently: the value decreases steadily and then stabilizes. For $N=600$, the initial downward trend is also present, but the direction of the curve later reverses and a secondary increase appears. This is also the series in which memorization of the noise component of the empirical distribution becomes stronger.

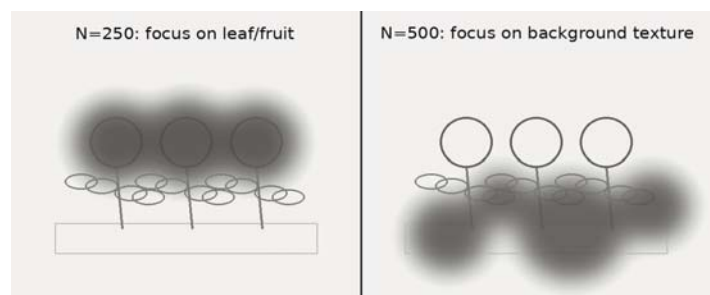


Fig. 5. Comparison of Grad-CAM regions in the rational and oversaturated regimes

The change in the internal basis of the decision can be traced in the Grad-CAM maps in Figure 5. At $N=250$, the strongest contribution is associated with the leaf blade, fruit contour, and fragments immediately adjacent to the object. After the transition to $N=500$, part of the active regions appears on the background texture. The difference between the regimes therefore concerns more than the final magnitude of the error: the spatial set of cues used by the network to form a bounding box also changes.

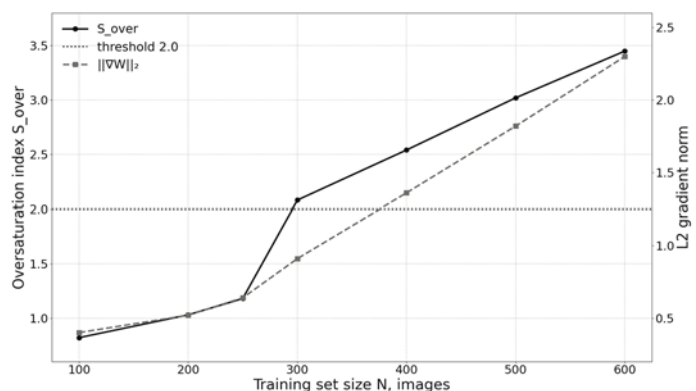


Fig. 6. Dynamics of *Sover* and the L_2 norm of the final detection-layer gradient

Sover exceeds 2.0 after $N=300$. In parallel, the L_2 norm of the final-layer gradient increases from 0.40 to 2.30 across the experimental series. No attenuation of the detection head response is observed. On the contrary, the final-layer parameters continue to undergo active restructuring in the presence of contradictory micro-patterns, and the number of high-confidence false detections increases in the same training regimes.

Discussion

Between 100 and 250 images, enlargement of the dataset coincided with a genuine expansion of the observed scene states. Leaf position, planting density, weed level, soil appearance, and substrate appearance all changed. These differences supplied the network with new combinations of object-related cues; the improvement in mAP@0.5, Precision, and Recall cannot therefore be attributed solely to the number of files. In the initial part of the experimental range, both N and the meaningful variability of the training distribution increased.

The situation changed after the peak regime. New frames continued to enlarge the dataset but increasingly reproduced conditions already represented in the same technological zone. For a highly parameterized YOLOv8m model, this provides additional realizations of local visual dependencies, some of which may belong to the background rather than to stable target morphology. The number of examples then grows faster than the number of independent scene states. This divergence between dataset size and informational breadth forms the basis of our subsequent analysis of oversaturation.

An early sign of the transition is visible when $N=250$ and $N=300$ are compared. mAP@0.5 decreases by only 0.03, from 0.91 to 0.88. Over the same step, *Khal* rises from 0.18 to 0.45, a 2.5-fold increase. If only aggregate accuracy is examined, the change may resemble a moderate fluctuation near the optimum. High-confidence false predictions give a different picture: their accumulation begins before the decline in mAP becomes pronounced.

For a system that controls a working tool, the earlier response of *Khal* relative to mAP is of separate importance. An averaged metric describes overall detection performance but does not indicate how dangerous a particular error is for the subsequent action. A false bounding box with high confidence may pass the trust threshold and enter an actuator command without repeated verification. The hallucination coefficient is therefore used to characterize a technologically significant class of errors rather than as an alternative name for mAP.

We associate the observed behavior with repeated background combinations inside one technological zone. A shadow of a particular shape, the edge of drip tape, or an accumulation of plant residues occurred in many closely related scenes. The parametric capacity of YOLOv8m permits memorization of a statistical association between such local fragments and the target class. This explanation is restricted to the experiment conducted here: with broader domain diversity, the same architecture may distribute its capacity across a larger number of stable visual regularities.

The errors at the two extremes of the experimental range have different origins. At $N=100$, the model lacks examples of object morphology. At $N=600$, the number of files is much larger, but the informational dimensionality of the scene grows disproportionately slowly relative to N . In the latter case, the network repeatedly

receives close realizations of a latent noise structure instead of new independent states of the agrobiocenosis. Data scarcity and an excess of similar data should therefore not be regarded as the same training regime.

The transition in the internal characteristic of the final layer is visible directly in the numerical values: *Sover* is 1.18 at $N=250$ and reaches 2.08 at $N=300$; the index remains above 2.0 thereafter. *Khal* increases over the same range, while Grad-CAM shifts part of the activations toward background regions. For the obtained series, the coincidence of these three features describes the same transition to an oversaturated regime. The present data do not justify treating *Sover* as a universal standalone criterion, but the index can serve as an additional diagnostic signal when deciding whether to continue dataset expansion.

Augmentation can increase the number of variants derived from a single image, but it does not necessarily alter the informational diversity of the scene. Rotation, geometric transformation, or correction of visual characteristics operate on an existing frame. A new plant growth stage or a new environmental condition cannot be introduced in this way. If repeated background structures and random covariations are already present many times in the source set, their transformed versions remain content-related to the same scenes. An increase in the number of augmented examples therefore cannot automatically be equated with the appearance of independent observations.

For the activation distributions, $DKL(P250||P600)=0.37$ and $DKL(P600||P250)=0.52$ were obtained. The asymmetry is particularly important in the reverse direction: the oversaturated regime contains activation states that were not represented at $N=250$. Independently, a second pronounced mode of false responses appears in the confidence-score distribution for $N=600$ at 0.92-0.96. Within the present experiment, part of the additional internal states coincides with the emergence of high-confidence erroneous predictions.

This regularity should be considered before annotation of the next image series begins. After approximately 250 high-quality annotated frames acquired during one season and under similar conditions, the automatic addition of several hundred nearly identical scenes does not guarantee a performance gain. More meaningful expansion requires new acquisition conditions: a different time of day and lighting scheme, another substrate type, or another cultivation area. Growth stage, cultivar, and planting density also alter the observed distribution. Under such expansion, the diversity of states increases rather than N alone.

The influence of architectural capacity can also be tested as a means of reducing the tendency of a small dataset to encourage local texture memorization. YOLOv8n contains fewer trainable parameters than YOLOv8m and may therefore be less able to adapt to idiosyncratic features of individual scenes; for a mobile platform, a smaller model simultaneously reduces computational requirements. This remains a separate experimental hypothesis. The two detectors should be compared not only by final accuracy but also by their robustness to the background structures on which high-confidence false detections occur.

In the transition region $N=250-300$, mAP detects the degradation later than *Khal*. The next stage of the study should therefore address early diagnosis of oversaturation using *Khal*, *Sover*, the dynamics of the gradient L_2 norm, and changes in activation maps jointly. Conventional early stopping is oriented mainly toward validation loss.

For autonomous inter-row treatment, the diagnostic loop should additionally register the moment at which high-confidence false objects begin to appear, because these detections are directly associated with the risk of an erroneous control action.

Conclusions

The maximum performance in the present series occurred at $N=250$: $mAP@0.5=0.91$, $Precision=0.93$, and $Recall=0.88$. Up to this point, dataset growth was accompanied by improved generalization; beyond it, the dependence became non-monotonic. For the scene set considered here, $N=250$ corresponds to the rational cognitive saturation threshold of YOLOv8m.

The final $N=600$ regime differs substantially from the peak regime. Precision decreases to 0.54, $Khal$ increases to 1.80, and the relative loss of Precision is 41.94%. Correct annotation of additional frames alone therefore does not guarantee improved practical reliability: when scene variability is weak, increasing the number of images may be accompanied by deterioration of detector performance.

The dynamics of the final detection layer refine the location of the identified boundary. After $N=300$, $Sover$ exceeds 2.0, while the gradient L_2 norm reaches 2.30 within the experimental range. At the same time, false bounding boxes appear on the greenhouse-row background, soil-substrate surface, shadows, and plant residues. For this series, the combination of these features is associated with an unstable response of the detection head to recurrent background micropatterns.

The obtained value $N=250$ should be regarded only as an experimental optimum within the domain studied. It refers to robotic tomato-care scenes recorded during one growing season and in one technological zone; the result does not establish a universal standard for an arbitrary dataset. Under these conditions, 250 ± 20 annotated frames can be used as a practical reference. Once this range is reached, it is more rational to introduce new observation states – different illumination, growth stage, cultivar, substrate type, planting density, or another acquisition site – rather than to continue adding closely related frames. Such a domain change may also shift the saturation threshold itself.

If heterogeneous scenes cannot be collected, subsequent work should separately test alternative methods for limiting memorization. These include reducing parametric capacity, strengthening regularization, and jointly monitoring $Khal$ and $Sover$. A lighter architecture should not be assumed effective in advance; it must be compared with YOLOv8m in a dedicated series. For an autonomous system, mAP should likewise not remain the sole metric because average accuracy does not express the risk of a confident false response in a region without weeds. In the adopted formulation, $Khal$ and $Sover$ complement the evaluation specifically from this perspective.

Acknowledgements

This article was prepared as part of the implementation of research work, implemented at the expense of the federal budget (source of financing – the Ministry of Education and Science of the Russian Federation) on the topic: “Development of software for the agro-industrial complex for monitoring the state of the tops and fruits of greenhouse plants using machine vision and artificial intelligence” (code of the scientific theme FFSM-2025-0002).

References

1. C. M. Badgujar, A. Poulouse, H. Gan, “Agricultural object detection with You Only Look Once (YOLO) Algorithm: a bibliometric and systematic literature review,” *Computers and Electronics in Agriculture*. 2024. Vol. 223. Article 109090. DOI: 10.1016/j.compag.2024.109090.
2. Y. Fu, J. Chen, X. Li, Y. Zhang, “Multi-stage tomato fruit recognition method based on improved YOLOv8,” *Frontiers in Plant Science*. 2024. Vol. 15. Article 1447263. DOI: 10.3389/fpls.2024.1447263.
3. Z. Yang, H. Zhang, J. Li, X. Wang, “A method for tomato ripeness recognition and detection based on an improved YOLOv8 model,” *Horticulturae*. 2025. Vol. 11. No. 1. Article 15.
4. A. T. Khan, S. M. Jensen, A. R. Khan, “Advancing precision agriculture: a comparative analysis of YOLOv8 for multi-class weed detection in cotton cultivation,” *Artificial Intelligence in Agriculture*. 2025. Article 100598. DOI: 10.1016/j.aiaa.2025.01.013.
5. Y. Liu, F. Zeng, H. Diao, J. Zhu, D. Ji, X. Liao, Z. Zhao, “YOLOv8 model for weed detection in wheat fields based on a visual converter and multi-scale feature fusion,” *Sensors*. 2024. Vol. 24. No. 13. Article 4379. DOI: 10.3390/s24134379.
6. P. Nakkiran, G. Kaplan, Y. Bansal, T. Yang, B. Barak, I. Sutskever, “Deep double descent: where bigger models and more data hurt,” *Journal of Statistical Mechanics: Theory and Experiment*. 2021. No. 12. Article 124003. DOI: 10.1088/1742-5468/ac3a74.
7. Ultralytics. YOLOv8 models: performance metrics and model variants. 2023. URL: <https://docs.ultralytics.com/models/yolov8/> (accessed: 03 July 2026).
8. R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, D. Batra, “Grad-CAM: visual explanations from deep networks via gradient-based localization,” *International Journal of Computer Vision*. 2020. Vol. 128, pp. 336-359. DOI: 10.1007/s11263-019-01228-7.
9. T.-Y. Lin, P. Goyal, R. Girshick, K. He, P. Dollar, “Focal loss for dense object detection,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2020. Vol. 42. No. 2, pp. 318-327. DOI: 10.1109/TPAMI.2018.2858826.
10. M. Belkin, D. Hsu, S. Ma, S. Mandal, “Reconciling modern machine-learning practice and the classical bias-variance trade-off,” *Proceedings of the National Academy of Sciences*. 2019. Vol. 116. No. 32. P. 15849-15854. DOI: 10.1073/pnas.1903070116.
11. H. Sun, Y. Zhou, L. Zhao, Q. Wang, “An improved YOLOv8 model for detecting four stages of tomato ripening and its application deployment in a greenhouse environment,” *Agriculture*. 2025. Vol. 15. No. 9. Article 936.
12. J. Liu, K. Wang, L. Chen, Y. Zhao, “Optimization of a multi-environmental detection model for tomato growing points and flower buds based on improved YOLOv8,” *Scientific Reports*. 2025. Vol. 15. Article 6692.

ИССЛЕДОВАНИЕ ПОРОГА КОГНИТИВНОГО НАСЫЩЕНИЯ МОДЕЛИ YOLOV8 В УСЛОВИЯХ ОГРАНИЧЕННОЙ ОБУЧАЮЩЕЙ ВЫБОРКИ

Карелина Мария Юрьевна, Государственный университет управления; Московский автомобильно-дорожный технический университет (МАДИ), Москва, Россия, karelinamu@mail.ru

Акулов Алексей Андреевич, Государственный университет управления; Московский автомобильно-дорожный технический университет (МАДИ), Москва, Россия, a.akulov.98@mail.ru

Талдыкин Дмитрий Сергеевич, Московский автомобильно-дорожный технический университет (МАДИ), Государственный университет управления, Москва, Россия, dima.dima.taldykin@mail.ru

Аннотация

В работе проверялось, при каком объеме малой размеченной выборки YOLOv8m перестает получать выигрыш от добавления близких по содержанию кадров при детекции сорной растительности в посадках томатов. Постановка связана с обычным для точного земледелия ограничением: RGB-изображения получить сравнительно легко, тогда как проверка разметки агрономом занимает заметно больше времени. Для эксперимента использованы 600 изображений 640x640 пикселей, снятых в тепличных рядах за один вегетационный период; выделялись томат, сорная растительность и фон почвенно-субстратной поверхности. Семь серий обучения позволили проверить предположение о максимуме обобщения в окрестности $N=250$ при числе уникальных сцен менее 300. После этой точки новые кадры, если условия съемки почти не менялись, сопровождалась ростом ложных детекций с высокой уверенностью. Для описания эффекта использованы коэффициент галлюцинаций и индекс перенасыщения, в котором сопоставлены дисперсия весов последнего детектирующего слоя и средняя валидационная потеря. В серии $N=600$ Precision оказалась на 41,94% ниже пикового значения, коэффициент галлюцинаций достиг 1,80. Практический результат относится прежде всего к планированию разметки и выбору емкости детектора: увеличение числа похожих кадров нельзя рассматривать как безусловный способ повышения надежности автоматизированной системы.

Ключевые слова: YOLOv8; томаты; детекция сорняков; малые выборки; галлюцинации модели

Литература

1. Badgujar C. M., Poulouse A., Gan H. Agricultural object detection with You Only Look Once (YOLO) Algorithm: a bibliometric and systematic literature review // Computers and Electronics in Agriculture. 2024. Vol. 223. Article 109090. DOI: 10.1016/j.compag.2024.109090.
2. Fu Y., Chen J., Li X., Zhang Y. Multi-stage tomato fruit recognition method based on improved YOLOv8 // Frontiers in Plant Science. 2024. Vol. 15. Article 1447263. DOI: 10.3389/fpls.2024.1447263.
3. Yang Z., Zhang H., Li J., Wang X. A method for tomato ripeness recognition and detection based on an improved YOLOv8 model // Horticulturae. 2025. Vol. 11. No. 1. Article 15.
4. Khan A. T., Jensen S. M., Khan A. R. Advancing precision agriculture: a comparative analysis of YOLOv8 for multi-class weed detection in cotton cultivation // Artificial Intelligence in Agriculture. 2025. Article 100598. DOI: 10.1016/j.iaia.2025.01.013.
5. Liu Y., Zeng F., Diao H., Zhu J., Ji D., Liao X., Zhao Z. YOLOv8 model for weed detection in wheat fields based on a visual converter and multi-scale feature fusion // Sensors. 2024. Vol. 24. No. 13. Article 4379. DOI: 10.3390/s24134379.
6. Nakkiran P., Kaplan G., Bansal Y., Yang T., Barak B., Sutskever I. Deep double descent: where bigger models and more data hurt // Journal of Statistical Mechanics: Theory and Experiment. 2021. No. 12. Article 124003. DOI: 10.1088/1742-5468/ac3a74.
7. Ultralytics. YOLOv8 models: performance metrics and model variants. 2023. URL: <https://docs.ultralytics.com/models/yolov8/> (accessed: 03 July 2026).
8. Selvaraju R. R., Cogswell M., Das A., Vedantam R., Parikh D., Batra D. Grad-CAM: visual explanations from deep networks via gradient-based localization // International Journal of Computer Vision. 2020. Vol. 128. P. 336-359. DOI: 10.1007/s11263-019-01228-7.
9. Lin T.-Y., Goyal P., Girshick R., He K., Dollar P. Focal loss for dense object detection // IEEE Transactions on Pattern Analysis and Machine Intelligence. 2020. Vol. 42. No. 2. P. 318-327. DOI: 10.1109/TPAMI.2018.2858826.
10. Belkin M., Hsu D., Ma S., Mandal S. Reconciling modern machine-learning practice and the classical bias-variance trade-off // Proceedings of the National Academy of Sciences. 2019. Vol. 116. No. 32. P. 15849-15854. DOI: 10.1073/pnas.1903070116.
11. Sun H., Zhou Y., Zhao L., Wang Q. An improved YOLOv8 model for detecting four stages of tomato ripening and its application deployment in a greenhouse environment // Agriculture. 2025. Vol. 15. No. 9. Article 936.
12. Liu J., Wang K., Chen L., Zhao Y. Optimization of a multi-environmental detection model for tomato growing points and flower buds based on improved YOLOv8 // Scientific Reports. 2025. Vol. 15. Article 6692.

Информация об авторах:

Карелина Мария Юрьевна, д.т.н., д.п.н., профессор, заведующий кафедрой логистики и транспортно-технологических систем, Государственный университет управления, Заведующий кафедрой "Детали машин и теория механизмов" Московский автомобильно-дорожный технический университет (МАДИ), Москва, Россия. ORCID ID: 0000-0003-0335-7550

Акулов Алексей Андреевич, к.т.н., доцент кафедры логистики и транспортно-технологических систем, Государственный университет управления, Доцент кафедры "Детали машин и теория механизмов" Московский автомобильно-дорожный технический университет (МАДИ), Москва, Россия. ORCID ID: 0000-0001-8952-0237

Талдыкин Дмитрий Сергеевич, аспирант, ассистент кафедры "Детали машин и теория механизмов", Московский автомобильно-дорожный государственный технический университет (МАДИ), младший научный сотрудник, Государственный университет управления, Москва, Россия. ORCID ID: 0009-0008-5071-426X

ORGANIZERS:

RUSSIA SECTION TEM/GRS/ITSS JOINT CHAPTER
INSTITUTE OF RADIO AND INFORMATION SYSTEMS (IRIS)

INTERNATIONAL CONFERENCE

"2027 Systems of signals generating and processing in the field of on board communications"

From 10 to 12 March 2027, Moscow, Russia
Avia Plaza, Aviamotornaya str., 10/2

IEEE Conference

Conference will produce a publication
(IEEE Conference Publication Program (CPP)) – **IEEE Explore**,

Possibility of indexing in **Scopus**

The main areas of the conference «Systems of signals generating and processing in the field of on board communications» include modern digital transportation systems design and technical operation, radio waves propagation, transmitting, receiving and processing signals in television and radio broadcasting devices, information technologies in transport.

The papers which are discussed at the conference can be divided into the following chapters:

1. Antennas and Radio Waves Propagation
2. Navigation and Mathematical Algorithms of an Object Space Orientation.
3. Radiofrequency Applications.
4. Wire and Optical Communication and Control Systems.
5. Intelligent Transport Systems
6. Digital signal processing in on-board radio systems
7. Electrical and electronic engineering pedagogy

Organising Committee:

111024, Moscow, Aviamotornaya, 8/1, office 323
Tel.: +7(926) 218-82-43, boardconf@media-publisher.ru
media-publisher.ru/en/2026-on-board