

Издательский дом

МЕДИА ПАБЛИШЕР

Научный журнал "Т-Сотт: Телекоммуникации и транспорт"

Журнал включен в перечень периодических научных изданий, рекомендуемый ВАК Минобразования 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 Паблшер"

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 Паблшер". Перепечатка, цитирование, дублирование на сайтах допускаются только с разрешения издателя

© ООО "ИД Медиа Паблшер", 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 Schience);
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 Паблшер"

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 программной

и оргкомитетов WSEAS), Россия, Мексика

Бестугин Александр Роеальдович

(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.

Линейно-ограниченная адаптивная антенная решетка в арифметике действительных чисел для радиоприемников наземной аппаратуры потребителя глобальных навигационных спутниковых систем

4

Бакулин М.Г., Крейнделин В.Б., Резнев А.А.

FTN и NOMA: мифы и реальность

15

СВЯЗЬ

Глушак Е.В., Ключев Д.С., Воловач В.И.

Моделирование многопутевой маршрутизации в гибридных сетях связи с использованием облачных и туманных вычислений

28

ИНФОРМАТИКА

Каменский В.В., Соколов С.В., Лагунова Е.О.

Оптические вычислители нелинейных функций

35

Леохин Ю.Л., Дымкова С.С., Фатхулин Т.Д., Зозуля И.С.

Разработка методов и алгоритмов решения задач управления в организационных системах торговых компаний

43

ПУБЛИКАЦИИ НА АНГЛИЙСКОМ ЯЗЫКЕ

ЭЛЕКТРОНИКА. РАДИОТЕХНИКА

Троегубов А.Ю., Маслов А.А., Себекин Г.В.,

Степанов С.Н., Щурков А.О.

Ключевые аспекты спутниковых систем с высокой пропускной способностью: от стандартизации к управлению качеством обслуживания

53

СВЯЗЬ

Терентьев С.В., Тихвинский В.О., Девяткин Е.В., Фокин Г.А.

Разработка модели данных для управления базовой станцией eNodeB технологии LTE в сети радиодоступа OpenRAN

65

CONTENT

ELECTRONICS. RADIO ENGINEERING

Djigan V.I.
Linearly-Constrained Adaptive Antenna Array for Ground-Based Radio Receivers of Global Navigation Satellite System 4

Bakulin M.G., Kreyndelin V.B., Reznev A.A.
FTN and NOMA: Beyond the Hype - Toward Shannon-Limited Efficiency 15

COMMUNICATIONS

Glushak E.V., Klyuev D.S., Volovach V.I.
Modeling multipath routing in hybrid communication networks using cloud and fog computing 28

COMPUTER SCIENCE

Kamenskij V.V., Sokolov S.V., Lagunova E.O.
Fast optical computers 35

Leokhin Yu.L., Dymkova S.S., Fatkhulin T.D., Zozulya I.S.
Development of Methods and Algorithms for Solving Management Problems in Organizational Systems of Trading Companies 43

PUBLICATIONS IN ENGLISH

ELECTRONICS. RADIO ENGINEERING

A. Yu. Troegubov, A. A. Maslov, G. V. Sebekin, S. N. Stepanov, A. O. Shchurkov
Focused overview of high-throughput satellite systems: from standardization to QoS management 53

COMMUNICATIONS

Terentiev S.V., Tikhvinsky V.O., Devyatkin E.E., Fokin G.A.
Data model development for LTE eNodeB managing in the OpenRAN radio access network 65

T - C o m m

Telecommunications and transport

Volume 20. No. 3-2026

Release date: 20.03.2026

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. Tikhvinsky

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

ЛИНЕЙНО-ОГРАНИЧЕННАЯ АДАПТИВНАЯ АНТЕННАЯ РЕШЕТКА В АРИФМЕТИКЕ ДЕЙСТВИТЕЛЬНЫХ ЧИСЕЛ ДЛЯ РАДИОПРИЕМНИКОВ НАЗЕМНОЙ АППАРАТУРЫ ПОТРЕБИТЕЛЯ ГЛОБАЛЬНЫХ НАВИГАЦИОННЫХ СПУТНИКОВЫХ СИСТЕМ

DOI: 10.36724/2072-8735-2026-20-3-4-14

Manuscript received 12 January 2026;
Accepted 14 March 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: Глобальные навигационные спутниковые системы, рекурсивный алгоритм по критерию наименьших квадратов, линейное ограничение, адаптивная антенная решетка, арифметика комплексных чисел, арифметика действительных чисел, карта отношения сигнал-шум

Антенные решетки сегодня широко используются в качестве антенн современных радиосистем. Возможность управления формой диаграммы направленности (ДН) таких антенн лежит в основе принципа работы адаптивных антенных решеток (ААР). Такие решетки обеспечивают подавление сигналов источников помех, полосы частот которых пересекаются с полосой частот полезного (информационного) сигнала. Одним из актуальных приложений ААР является их использование в наземной аппаратуре потребителя глобальных навигационных спутниковых систем (ГНСС). В такой аппаратуре в основном используются ААР, именуемые компенсаторами помех. Подобно любым ААР, они обеспечивают подавление помех, но не гарантируют прием наперед заданного числа спутниковых сигналов, необходимого для решения навигационной задачи. В работе предлагается использовать разновидность ААР на основе линейно-ограниченного адаптивного алгоритма. В этом алгоритме задается ограничение на значение ДН ААР в направлении зенита. Сигналы со спутников, которые находятся в окрестности этого направления, характеризуются максимальным отношением сигнал-шум при их приеме в силу кратчайшего расстояния от спутников до Земли. Задаваемое ограничением значение ДН не меняется на протяжении всей работы ААР. Оно не зависит от значений её весовых коэффициентов (ВК), вычисляемых в процессе адаптации. Для вычисления ВК в работе использован рекурсивный алгоритм по критерию наименьших квадратов на основе леммы об обращении матрицы, оптимизированный по числу арифметических операций. Вычислительная сложность этого алгоритма меньше по сравнению с его неоптимизированной версией. Уменьшение сложности достигается за счет учёта эрмитовой структуры корреляционной матрицы входных сигналов ААР. Дальнейшее уменьшение сложности алгоритма вычисления ВК достигается в случае, если ААР является симметричной. Практически все операции в такой ААР выполняются в арифметике действительных чисел. Число этих операций в два раза меньше эквивалентного числа действительных операций аналогичной ААР в арифметике комплексных чисел. Кроме того, в рассматриваемой ААР переходный процесс в два раза короче, а подавление помех в установившемся режиме в среднем на 3 дБ больше по сравнению с ААР в арифметике комплексных чисел. Рассмотренная ААР может найти применение не только в аппаратуре ГНСС, но и в аппаратуре других радиосистем.

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 микророботов 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. Линейно-ограниченная адаптивная антенная решетка в арифметике действительных чисел для радиоприемников наземной аппаратуры потребителя глобальных навигационных спутниковых систем // Т-Comm: Телекоммуникации и транспорт. 2026. Том 20. №3. С. 4-14.

For citation:

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)

Введение

Сегодня в качестве антенн оборудования радиосистем широко используются антенные решётки (АР) [1, 2]. С точки зрения архитектуры, АР представляет собой многоканальную антенную систему. Сигналы в каналах АР взвешиваются с помощью комплексных весовых коэффициентов (ВК) перед их суммированием с целью формирования выходного сигнала, что позволяет изменять форму диаграммы направленности (ДН) АР.

Простейшим изменением формы ДН является ориентация её главного лепестка (луча АР) в интересующем направлении. Такое изменение достигается с помощью комплексных ВК с единичными амплитудами и определенными фазами. Такими ВК в аналоговых АР служат фазовращатели.

Однако в АР можно изменять форму и всей ДН. Это свойство способствовало созданию адаптивных АР (ААР) [3, 4]. За счет формирования провалов в ДН осуществляется подавление сигналов источников помех в выходном сигнале ААР. Эти провалы образуются в результате установки в каналах ААР соответствующих комплексных ВК с разными амплитудами и фазами. В качестве таких ВК в аналоговых ААР используются каскадно включенные аттенюаторы и фазовращатели, а в цифровых ААР – просто комплексные числа.

ААР является пространственным фильтром, позволяющим разделять сигналы, находящиеся в одной полосе частот, но различающиеся по пространственному положению их источников. Для вычисления ВК используются различные адаптивные алгоритмы [5, 6], которые обрабатывают сигналы в каналах ААР, её выходной сигнал, а иногда и так называемый требуемый сигнал. Адаптивные алгоритмы не используют информацию о пространственном положении источников помех и энергетических характеристиках их сигналов. В основе работы адаптивных алгоритмов находится безусловная минимизация различных целевых функций. В ААР также может быть использована условная минимизация целевых функций при наложении определенных ограничений, например, на значения ДН в одном или в нескольких интересующих направлениях [7, 8].

ААР сегодня широко применяются для борьбы с помехами в радиосистемах различного назначения. Одним из таких применений является использование ААР в качестве антенн радиоприёмников наземной аппаратуры потребителя (НАП) глобальных навигационных спутниковых систем (ГНСС) [9, 10]. Такая аппаратура часто эксплуатируется в условиях радиопомех, создаваемых природными явлениями; радиоприёмниками, работающими на тех же частотах что и НАП; или же постановщиками преднамеренных помех. Поэтому для обеспечения работоспособности НАП (способности определять координаты ее радиоприёмника) требуется подавлять сигналы этих помех, что обеспечивается с помощью ААР.

В НАП в основном используется разновидность ААР, именуемая компенсатором помех, принцип работы которой аналогичен компенсатору боковых лепестков направленной антенны [11]. Компенсатор помех минимизирует мощность сигнала ошибки между одной из антенн ААР с фиксированным (единичным) ВК и суммой взвешенных сигналов остальных антенн ААР. Так как сигналы, принимаемые со спутников, находятся ниже уровня шумов приёмников ААР, то наличие

этих сигналов никак не сказывается на эффективности подавления помех.

Использовать в НАП ААР, в которых минимизируются различные целевые функции ошибок между требуемым и выходными сигналами, не представляется возможным, так как для решения навигационной задачи необходимо одновременно принимать сигналы от нескольких спутников, находящихся в разных направлениях относительно ААР, а потому один требуемый сигнал не может быть сформирован и использован. Кроме того, сигналы со спутников на входах, как уже было сказано, а значит, и на выходе ААР, находятся под шумами в силу удалённости их источников от Земли. Отношение сигнал-шум (ОСШ) на выходах радиоприёмников НАП для сигналов от спутников, находящихся в зените, примерно оценивается на уровне -20 дБ [9, 10], что также не позволяет использовать в ААР адаптивные алгоритмы с требуемым сигналом, так как даже в отсутствии помех сигналы со спутников на выходе ААР скрыты шумами.

В многолучевых ААР с предварительной корреляционной обработкой сигналов в каналах [12] требуемые сигналы могут быть использованы (по одному на каждый луч/спутник), однако такие ААР являются достаточно сложными, так как являются многолучевыми (от четырех и более лучей). В них общими являются только антенны, радиоприемники и преобразователи частоты, а цифровая часть размножается пропорционально числу лучей. Кроме того, для каждого луча такой ААР требуется свой набор из M корреляторов или согласованных фильтров, где M – это число антенн ААР.

ААР в виде компенсатора помех, как и любая другая разновидность ААР, обеспечивает подавление сигналов от источников всех помех, если их число не превышает значения $M - 1$ [9]. В установившемся режиме в ДН компенсатора образуются глубокие провалы в направлениях на источники помех, но при этом не обеспечиваются какие-либо наперед заданные значения ДН в направлениях на навигационные спутники. Эти значения зависят от ВК, полученных в результате решения задачи подавления этих помех [13], т.е. в целом носят случайный характер. В результате, если число спутников, для которых на выходе ААР ОСШ > -20 дБ, оказывается меньше четырех штук, то навигационная задача становится нерешаемой.

Навигационные спутники, сигналы которых удовлетворяют критерию ОСШ > -20 дБ, в основном располагаются в районе зенита в силу кратчайшего расстояния от спутников до Земли. Поэтому ограничение ДН в направлении зенита до значения её основного лепестка (с учетом ДН антенн) при фазировании ААР в этом направлении является естественным решением рассматриваемой проблемы. Оно позволяет, по меньшей мере, уверенно принимать сигналы со спутников, находящихся в окрестности направления ограничения ДН. Ограничение задается в вычислительных процедурах ААР, в качестве которых могут быть использованы разные линейно-ограниченные адаптивные алгоритмы [5, 7, 8]. Такие алгоритмы обеспечивают заданное ограничение значения ДН в интересующем направлении независимо от значений ВК, вычисляемых в процессе подавления помех.

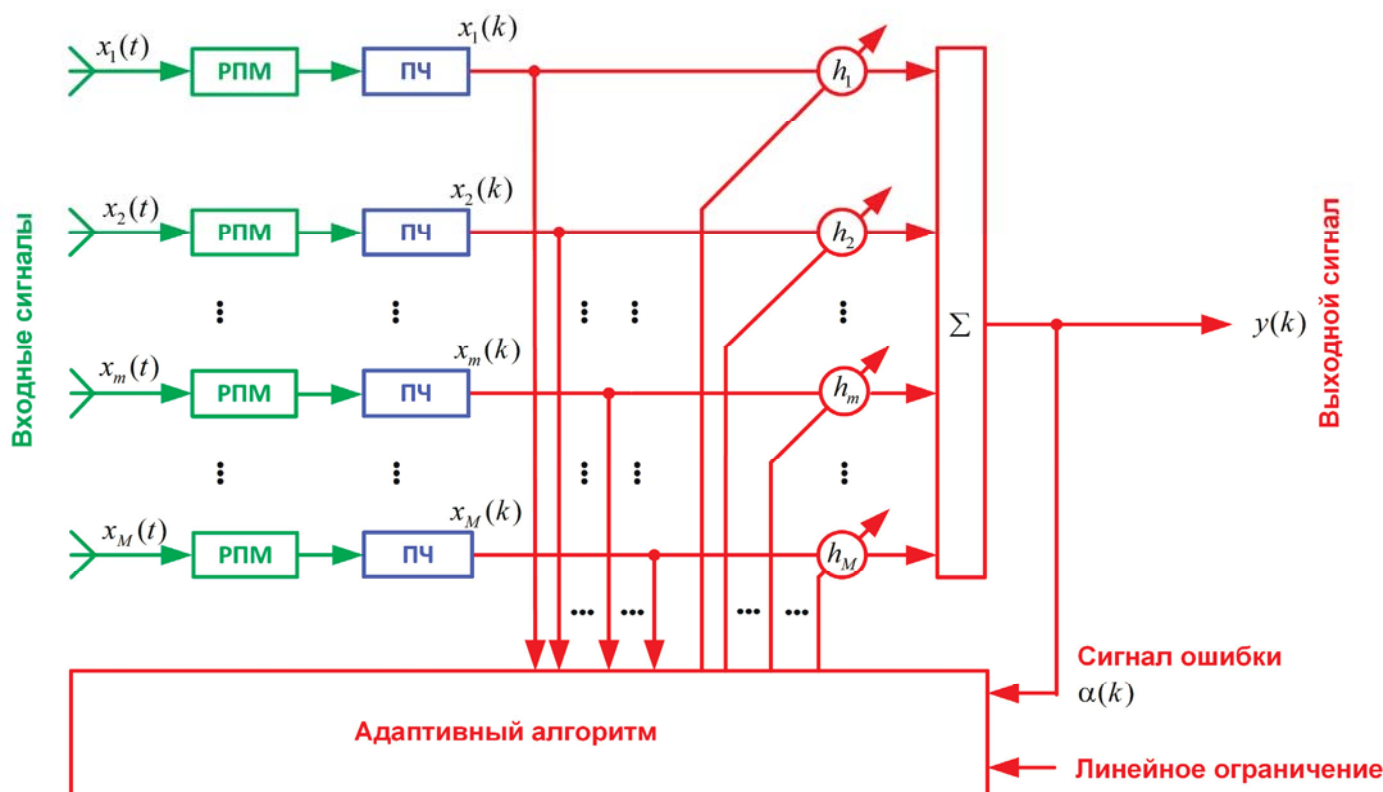


Рис. 1. ААР

Сегодня существует большое разнообразие адаптивных алгоритмов, используемых для вычисления ВК ААР. Это алгоритмы градиентного спуска, нормализованные градиентные алгоритмы, включая алгоритмы аффинных проекций, а также рекурсивные алгоритмы решения систем линейных уравнений по критерию наименьших квадратов (Recursive Least Squares, RLS) [5, 6]. Наиболее эффективными среди них являются RLS-алгоритмы. Критерием эффективности служат длительность переходного процесса, измеряемая числом итераций адаптивного алгоритма, а также минимальное значение целевой функции в установившемся режиме. В связи с этим RLS-алгоритмы сегодня широко применяются не только в ААР, но и в ряде других приложений современной адаптивной обработки сигналов [14, 15].

Однако адаптивные RLS-алгоритмы характеризуются квадратичной вычислительной сложностью как функцией числа ВК, которое в случае ААР совпадает с числом её антенн M . Вычислительную сложность RLS-алгоритмов можно уменьшить за счет учёта эрмитовой структуры корреляционной матрицы входных сигналов ААР, что позволяет вычислять только часть элементов этой матрицы [16]. Также, если ААР симметричная, то почти все её вычисления можно выполнять в арифметике действительных чисел [17-19].

Целью настоящей работы является описание ААР на базе линейно-ограниченного RLS-алгоритма в арифметике действительных чисел, в котором учитывается симметрия ААР и эрмитова структура корреляционной матрицы её входных сигналов.

В следующих разделах работы приводятся математическое описание и результаты моделирования такой ААР, подтверждающие её работоспособность и функциональную эффективность.

ААР в арифметике действительных чисел

Архитектура рассматриваемой линейно-ограниченной ААР приведена на рисунке 1. Она содержит M антенн/каналов и M комплексных ВК h_m , где $m = 1, 2, \dots, M$. ААР принимает непрерывные во времени сигналы $x_m(t)$, где t – это время.

Принятые сигналы $x_m(t)$ после полосовой фильтрации в радиоприёмниках (РПМ) каналов ААР и преобразования по частоте появляются на выходах преобразователей частоты (ПЧ) в виде дискретных комплексных отсчётов

$$x_m(k) = \hat{x}_m(k) + n_m(k),$$

где $\hat{x}_m(k)$ – это дискретные отсчёты сигнала $x_m(t)$, $n_m(k)$ – дискретные отсчеты шума РПМ, а k – номер отсчёта. Номера отсчетов сигналов совпадают с номерами итераций адаптивных алгоритмов, обрабатывающих эти сигналы.

ААР (рисунок 1) является АР с цифровым формированием луча или просто цифровой [20]. Её ПЧ строятся на принципах цифрового радио [21]. Такие ПЧ преобразуют сигналы, принимаемые антеннами ААР, с несущей частоты на нулевую промежуточную частоту. Выходные сигналы ПЧ образуют вектор дискретных отсчётов сигналов

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

Этот вектор используется для вычисления на каждой итерации адаптивного алгоритма выходного сигнала ААР

$$y(k) = \mathbf{h}_M^H(k-1) \mathbf{x}_M(k) = \sum_{m=1}^M h_m^*(k-1) x_m(k)$$

и расчёта вектора ВК ААР

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

который представляет собой набор комплексных чисел.

Здесь и далее, нижний индекс M обозначает число элементов в векторе или в квадратной матрице. В векторе это число равно M , а в квадратной матрице – $M \times M$.

Поскольку в рассматриваемой ААР требуемый сигнал $d(k)$ отсутствует, то сигнал ошибки $\alpha(k)$, минимизируемой в процессе адаптации, совпадает с её выходным сигналом

$$\alpha(k) = d(k) - y(k) = 0 - y(k) = -y(k).$$

С учётом ДН антенн, линейное ограничение в алгоритме работы ААР (вычисления её выходного сигнала и ВК) задается как

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

где вектор \mathbf{c}_M определяется как

$$\mathbf{c}_M = F_{\text{ант}}(\theta_c, \phi_c) [e^{j\psi_1}, e^{j\psi_2}, \dots, e^{j\psi_m}, \dots, e^{j\psi_M}]^T.$$

Здесь $F_{\text{ант}}(\theta_c, \phi_c)$ – это значение 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 направлении ограничения, до каждой из антенн ААР [1, 2].

В качестве алгоритмов, вычисляющих ВК рассматриваемой ААР, могут быть использованы разные линейно-ограниченные RLS-алгоритмы адаптивной фильтрации сигналов [5, 7, 8]. Из них наиболее вычислительно эффективным является RLS-алгоритм на основе леммы об обращении матрицы. Он не содержит сложных арифметических операций, таких как извлечение квадратного корня или деление, число которых в других RLS-алгоритмах обычно пропорционально M . Из сложных операций в рассматриваемом RLS-алгоритме на каждой его итерации используется только одна операция деления.

Во всех RLS-алгоритмах наибольшее число арифметических операций приходится на вычисление обратной корреляционной матрицы $\mathbf{R}_M^{-1}(k)$ входных сигналов ААР $x_m(k)$. Известно, что эта матрица является эрмитовой [5, 6]. Это позволяет примерно в два раза уменьшать сложность обращения такой матрицы за счет вычисления лишь диагональных и под-или над-диагональных её элементов [16].

Все адаптивные алгоритмы, используемые для вычисления ВК ААР, реализуются в арифметике комплексных чисел. Это обусловлено тем, что входные сигналы ААР $x_m(k)$ являются комплексными, а значит, и вычисляемые с помощью адаптивных алгоритмов ВК также являются комплексными.

Однако, если ААР является симметричной, то используя [17–19], почти все вычисления в ней можно выполнять в арифметике действительных чисел. Число действительных операций в такой ААР в два раза меньше эквивалентного числа действительных операций аналогичной ААР в арифметике комплексных чисел.

Симметричной называется ААР, в которой вектор входных сигналов и вектор ВК являются нечетно-симметричными, т.е.

$$\mathbf{x}_M(k) = [x_1(k), \dots, x_{M/2}(k), x_{M/2}^*(k), \dots, x_1^*(k)]^T$$

и

$$\mathbf{h}_M(k) = [h_1(k), \dots, h_{M/2}(k), h_{M/2}^*(k), \dots, h_1^*(k)]^T.$$

Уменьшение вычислительной сложности симметричной ААР достигается следующим образом.

Корреляционная матрица сигналов симметричной ААР $\mathbf{R}_M(k)$ является персимметричной. Согласно [17–19], оценивать такую матрицу можно как

$$\tilde{\mathbf{R}}_M(k) = \frac{1}{2} [\mathbf{R}_M(k) + \mathbf{J}_M \mathbf{R}_M^*(k) \mathbf{J}_M] = \frac{1}{2} \sum_{i=1}^k \lambda^{k-i} \times \\ \times [\mathbf{x}_M(i) \mathbf{x}_M^H(i) + \mathbf{J}_M \mathbf{x}_M^*(i) \mathbf{x}_M^T(i) \mathbf{J}_M],$$

где $1-1/M < \lambda < 1$ – это параметр экспоненциального взвешивания обрабатываемых сигналов, а

$$\mathbf{J}_M = \begin{bmatrix} 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & \dots & 1 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 1 & \dots & 0 & 0 \\ 1 & 0 & \dots & 0 & 0 \end{bmatrix}$$

– перестановочная матрица.

Оценка матрицы $\tilde{\mathbf{R}}_M(k)$ является более точной по сравнению с оценкой матрицы $\mathbf{R}_M(k) = \sum_{i=1}^k \lambda^{k-i} \mathbf{x}_M(i) \mathbf{x}_M^H(i)$ [18]. Однако, вычислительная сложность оценки матрицы $\tilde{\mathbf{R}}_M(k)$ примерно в два раза больше вычислительной сложности оценки матрицы $\mathbf{R}_M(k)$.

В тоже время, используя преобразующие матрицы

$$\mathbf{U}_M = \frac{1}{\sqrt{2}} \begin{bmatrix} \mathbf{I}_{\frac{M}{2}} & \mathbf{J}_{\frac{M}{2}} \\ j\mathbf{J}_{\frac{M}{2}} & -j\mathbf{I}_{\frac{M}{2}} \end{bmatrix},$$

если M – это четное число, или

$$\mathbf{U}_M = \frac{1}{\sqrt{2}} \begin{bmatrix} \mathbf{I}_{\frac{M-1}{2}} & \mathbf{0}_{\frac{M-1}{2}} & \mathbf{J}_{\frac{M-1}{2}} \\ \mathbf{0}_{\frac{M-1}{2}}^T & \sqrt{2} & \mathbf{0}_{\frac{M-1}{2}}^T \\ j\mathbf{J}_{\frac{M-1}{2}} & \mathbf{0}_{\frac{M-1}{2}} & -j\mathbf{J}_{\frac{M-1}{2}} \end{bmatrix},$$

если M – нечетное число, где $j = \sqrt{-1}$ – мнимая единица, матрицу $\tilde{\mathbf{R}}_M(k)$ можно вычислять в арифметике действительных чисел как

$$\bar{\mathbf{R}}_M(k) = \bar{\mathbf{R}}_{M,2}(k) = \lambda \bar{\mathbf{R}}_{M,2}(k-1) + \\ + \bar{\mathbf{x}}_{M,1}(k) \mathbf{x}_{M,1}^T(k) + \bar{\mathbf{x}}_{M,2}(k) \bar{\mathbf{x}}_{M,2}^T(k) = \\ = \bar{\mathbf{R}}_{M,1}(k) + \bar{\mathbf{x}}_{M,2}(k) \bar{\mathbf{x}}_{M,2}^T(k),$$

где

$$\bar{\mathbf{x}}_{M,1}(k) = \text{Re}\{\mathbf{U}_M \mathbf{x}_M(k)\}$$

и

$$\bar{\mathbf{x}}_{M,2}(k) = \text{Im}\{\mathbf{U}_M \mathbf{x}_M(k)\}$$

– это действительная и мнимая составляющие преобразованного комплексного вектора $\mathbf{U}_M \mathbf{x}_M(k)$ входных сигналов ААР.

Из уравнения вычисления матрицы $\bar{\mathbf{R}}_M(k)$ следует, что на каждой k -й итерации работы адаптивного алгоритма обновлять и обращать эту матрицу необходимо дважды. Сначала первый раз, используя векторы сигналов $\bar{\mathbf{x}}_{M,1}(k)$, а затем второй раз, используя векторы сигналов $\bar{\mathbf{x}}_{M,2}(k)$. Поэтому в ААР в арифметике действительных чисел процедура расчета ВК, в которой используется матрица $\bar{\mathbf{R}}_M(k)$, повторяется дважды на каждой k -й итерации, аналогично адаптивным алгоритмам, в которых корреляционная матрица обрабатываемых сигналов оценивается на скользящем окне отсчетов [5]. По этой причине число операций в симметричной ААР в арифметике действительных чисел в два раза больше чем в ААР в арифметике комплексных чисел.

Однако, так как одно комплексное умножение содержит четыре действительные умножения и сложения, то вычислительная сложность симметричной ААР в арифметике действительных чисел примерно в два раза меньше сложности аналогичной ААР в арифметике комплексных чисел.

Кроме того, в ААР в арифметике действительных чисел переходный процесс в два раза короче, а подавление помех в установившемся режиме в среднем на 3 дБ больше (из-за более точной оценки корреляционной матрицы, а значит, и более точного вычисления ВК) по сравнению с ААР в арифметике комплексных чисел.

В настоящей работе представлена математическая модель линейно-ограниченной ААР в арифметике действительных чисел, см. ниже, в которой вычисляются только диагональные и над-диагональные элементы действительной корреляционной матрицы её входных сигналов. Эта модель получена путем использования приемов [5, 16, 17 – 19].

Инициализация : $\mathbf{q}_M(0) = \mathbf{c}_M [\mathbf{c}_M^H \mathbf{c}_M]^{-1}$, $[\mathbf{R}_M(0)]^{-1} = \delta^{-2} \mathbf{I}_M$,

$[\bar{\mathbf{R}}_{M,2}(0)]^{-1} = [\mathbf{R}_M(0)]^{-1}$, $\bar{\mathbf{c}}_M = \mathbf{U}_M \mathbf{c}_M$, $\bar{\mathbf{q}}_M(0) = \mathbf{U}_M \mathbf{q}_M(0)$,

$\bar{\gamma}_M(0) = [\bar{\mathbf{R}}_{M,2}(0)]^{-1} \bar{\mathbf{c}}_M$, $\bar{\mathbf{q}}_{M,2}(0) = \bar{\gamma}(0) [\bar{\mathbf{c}}_M^T \bar{\gamma}(0)]^{-1}$,

$\bar{\mathbf{h}}_{M,2}(0) = \bar{\mathbf{q}}_{M,2}(0) f$, $\beta = \lambda^{-1}$

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

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

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

$\bar{\mathbf{x}}_{M,1}(k) = \text{Re}\{\mathbf{U}_M \mathbf{x}_M(k)\}$

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

$\bar{p}_{n,1}(k) = 0$

For $m = n, n+1, \dots, M$

$\bar{p}_{n,1}(k) = \bar{p}_{n,1}(k) + \bar{R}_{n,m,2}^{-1}(k-1) \bar{x}_{m,1}(k)$

End for m

End for n

For $n = 2, 3, \dots, M$

For $m = 1, 2, \dots, n-1$

$\bar{p}_{n,1}(k) = \bar{p}_{n,1}(k) + \bar{R}_{m,n,2}^{-1}(k-1) \bar{x}_{m,1}(k)$

End for m

End for n

$\bar{s}_1(k) = [\lambda + \bar{\mathbf{x}}_{M,1}^T(k) \bar{\mathbf{p}}_{M,1}(k)]^{-1}$

$\bar{\mathbf{g}}_{M,1}(k) = \bar{\mathbf{p}}_{M,1}(k) \bar{s}_1(k)$

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

For $m = n, n+1, \dots, M$

$\bar{R}_{n,m,1}^{-1}(k) = \beta [\bar{R}_{n,m,2}^{-1}(k-1) - \bar{g}_{n,1}(k) \bar{p}_{m,1}(k)]$

End for m

End for n

$\bar{v}_1(k) = \bar{\mathbf{c}}_M^T \bar{\mathbf{g}}_{M,1}(k)$

$\bar{v}_1(k) = \bar{\mathbf{x}}_{M,1}^T(k) \bar{\mathbf{q}}_{M,2}(k-1)$

$\bar{\mathbf{q}}'_{M,1}(k) = [\bar{\mathbf{q}}_{M,2}(k-1) - \bar{\mathbf{g}}_{M,1}(k) \bar{v}_1(k)] \times$

$\times \left[1 + \frac{\bar{v}_1(k) \bar{v}_1(k)}{1 - \bar{v}_1(k)} \right]$

$\bar{\mathbf{q}}_{M,1}(k) = \bar{\mathbf{q}}'_{M,1}(k) + \bar{\mathbf{q}}_M(0) [1 - \bar{\mathbf{c}}_M^T \bar{\mathbf{q}}'_{M,1}(k)]$

$\bar{\alpha}_1(k) = -\bar{\mathbf{h}}_{M,2}^T(k-1) \bar{\mathbf{x}}_{M,1}(k)$

$\bar{\mathbf{h}}'_{M,1}(k) = \bar{\mathbf{h}}_{M,2}(k-1) + \bar{\mathbf{g}}_{M,1}(k) \bar{\alpha}_1(k)$

$\bar{\mathbf{h}}_{M,1}(k) = \bar{\mathbf{h}}'_{M,1}(k) + \bar{\mathbf{q}}_{M,1}(k) [f - \bar{\mathbf{c}}_M^T \bar{\mathbf{h}}'_{M,1}(k)]$

$\bar{\mathbf{x}}_{M,2}(k) = \text{Im}\{\mathbf{U}_M \mathbf{x}_M(k)\}$

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

$\bar{p}_{n,2}(k) = 0$

For $m = n, n+1, \dots, M$

$\bar{p}_{n,2}(k) = \bar{p}_{n,2}(k) + \bar{R}_{n,m,1}^{-1}(k) \bar{x}_{m,2}(k)$

End for m

End for n

For $n = 2, 3, \dots, M$

For $m = 1, 2, \dots, n-1$

$\bar{p}_{n,2}(k) = \bar{p}_{n,2}(k) + \bar{R}_{m,n,1}^{-1}(k) \bar{x}_{m,2}(k)$

End for m

End for n

$\bar{s}_2(k) = [1 + \bar{\mathbf{x}}_{M,2}^T(k) \bar{\mathbf{p}}_{M,2}(k)]^{-1}$

$\bar{\mathbf{g}}_{M,2}(k) = \bar{\mathbf{p}}_{M,2}(k) \bar{s}_2(k)$

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

For $m = n, n+1, \dots, M$

$\bar{R}_{n,m,2}^{-1}(k) = \bar{R}_{n,m,1}^{-1}(k) - \bar{g}_{n,2}(k) \bar{p}_{m,2}(k)$

End for m

End for n

$\bar{v}_2(k) = \bar{\mathbf{c}}_M^T \bar{\mathbf{g}}_{M,2}(k)$

$\bar{v}_2(k) = \bar{\mathbf{x}}_{M,2}^T(k) \bar{\mathbf{q}}_{M,1}(k)$

$$\begin{aligned} \bar{\mathbf{q}}'_{M,2}(k) &= \left[\bar{\mathbf{q}}_{M,1}(k) - \bar{\mathbf{g}}_{M,2}(k)\bar{v}_2(k) \right] \times \\ &\times \left[1 + \frac{\bar{v}_2(k)\bar{v}_2(k)}{1 - \bar{v}_2(k)\bar{v}_2(k)} \right] \\ \bar{\mathbf{q}}_{M,2}(k) &= \bar{\mathbf{q}}'_{M,2}(k) + \bar{\mathbf{q}}_M(0) \left[1 - \bar{\mathbf{c}}_M^T \bar{\mathbf{q}}'_{M,2}(k) \right] \\ \bar{\alpha}_2(k) &= -\bar{\mathbf{h}}_{M,1}^T(k) \bar{\mathbf{x}}_{M,2}(k) \\ \bar{\mathbf{h}}'_{M,2}(k) &= \bar{\mathbf{h}}_{M,2}(k) + \bar{\mathbf{g}}_{M,2}(k)\bar{\alpha}_2(k) \\ \bar{\mathbf{h}}_{M,2}(k) &= \bar{\mathbf{h}}'_{M,2}(k) + \bar{\mathbf{q}}_{M,2}(k) \left[f - \bar{\mathbf{c}}_M^T \bar{\mathbf{h}}'_{M,2}(k) \right] \\ \mathbf{h}_M(k) &= \mathbf{U}_M^H \bar{\mathbf{h}}_{M,2}(k) \end{aligned}$$

End for k

Здесь $\delta^2 \geq 0.1\sigma_x^2$ – это параметр регуляризации процедуры обращения начальной матрицы $\bar{\mathbf{R}}_M^{-1}(0)$, σ_x^2 – дисперсия входных сигналов ААР $x_m(k)$, \mathbf{I}_M – единичная матрица, $\bar{p}_{n,1}(k)$ и $\bar{p}_{n,2}(k)$ – элементы векторов $\bar{\mathbf{p}}_{M,1}(k)$ и $\bar{\mathbf{p}}_{M,2}(k)$, $\bar{R}_{n,m,1}^{-1}(k)$ и $\bar{R}_{n,m,2}^{-1}(k)$ – элементы матриц $\bar{\mathbf{R}}_{M,1}^{-1}(k)$ и $\bar{\mathbf{R}}_{M,2}^{-1}(k)$, а $\bar{\mathbf{g}}_{M,1}(k)$ и $\bar{\mathbf{g}}_{M,2}(k)$ – векторы коэффициентов Калмана.

В рассматриваемом алгоритме действительные переменные обозначены символами с верхней чертой. Вычислительные операции над этими переменными выполняются в арифметике действительных чисел. И только выходной сигнала ААР $y(k)$ вычисляется в арифметике комплексных чисел. Для этого на каждой k -й итерации требуется преобразовывать вектор действительных ВК $\bar{\mathbf{h}}_{M,2}(k)$ в вектор комплексных ВК

$$\mathbf{h}_M(k) = \mathbf{U}_M^H \bar{\mathbf{h}}_{M,2}(k).$$

В силу структуры матрицы \mathbf{U}_M , указанное преобразование может быть выполнено без использования арифметических операций, а потому на вычислительной сложности ААР в арифметике действительных чисел не сказывается.

Аналогичные (математически эквивалентные) ААР на базе RLS-алгоритмов с вычислением всех элементов обратной корреляционной матрицы, с вычислением только диагональных и под-диагональных её элементов, на базе QR RLS-алгоритмов или RLS-алгоритмов, использующих преобразования Хаусхолдера, могут также быть получены, используя [5, 16-19].

В следующем разделе приводятся результаты компьютерного моделирования, подтверждающие работоспособность и демонстрирующие эффективность рассматриваемой ААР в арифметике действительных чисел.

Результаты моделирования ААР

Для проверки работоспособности ААР в арифметике действительных чисел и сравнения её эффективности с ААР в арифметике комплексных чисел было проведено моделирование адаптивных процессов в этих ААР.

Моделировались две одинаковые квадратные ААР с числом антенн $M = 3 \times 3 = 9$. Ширина ДН каждой антенны (по уровню -3 дБ) составляла $\theta \approx \pm 40^\circ$ в диапазоне углов $\phi = 0^\circ \dots 360^\circ$, а её максимальное значение $|F_{\text{ант}}(\theta, \phi)|_{\text{max}}$ было равно около 6 дБ.

Антенны ААР равномерно располагались на квадратной сетке с расстоянием между соседними парами антенн, равным половине длины радиоволны диапазона L1 (несущая частота 1575.42 МГц) ГНСС Global Positioning System (GPS) [9].

В исходном состоянии каждая ААР размещалась горизонтально, а её главный лепесток и главные лепестки ДН её антенн были направлены в зенит ($\theta_0 = 0^\circ$, $\phi_0 = 0^\circ$). Направление ограничения ДН и направление на один спутник также совпадали с зенитом. Восемь источников помех были размещены в диапазоне углов $\theta = 75^\circ \dots 105^\circ$ и $\phi = 0^\circ \dots 360^\circ$. Расположение источников помех было выбрано случайным. Отношение сигнал/каждая помеха задавалось равным -70 дБ.

Сигнал на выходе ПЧ каждого из каналов ААР определялся как

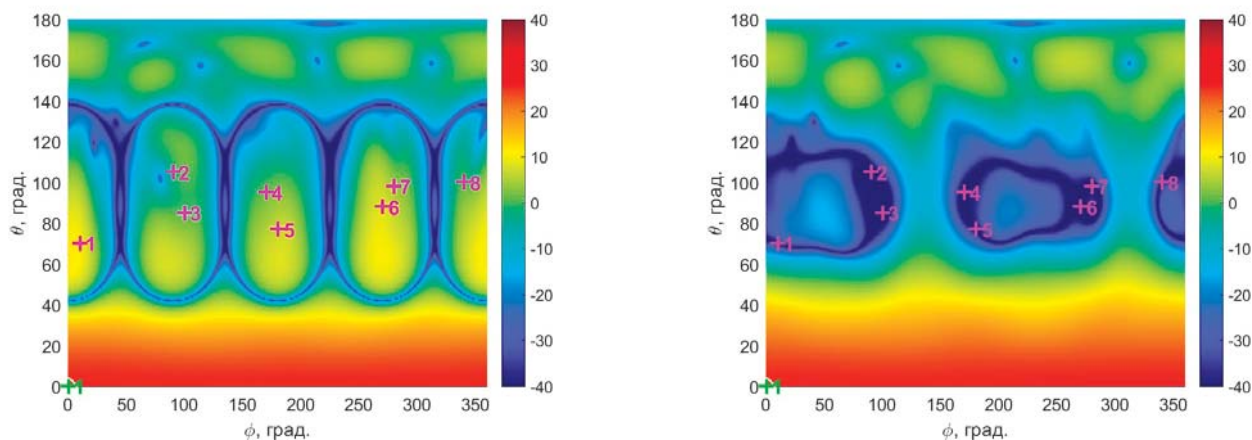
$$\hat{x}_m(k) = \sum_{l=1}^L \tilde{x}_m^{(l)}(k) = \sum_{l=1}^L A^{(l)}(k) e^{j\psi_m^{(l)}},$$

где $A^{(l)}(k)$ – это комплексная огибающая l -го принимаемого сигнала, $\psi_m^{(l)}$ – пространственный набег фазы от источника этого сигнала до m -й антенны относительно фазового центра ААР и $L < M$. Спутниковый сигнал $\tilde{x}_m^{(l)}(k)$ моделировался одной из тридцати двух псевдослучайных последовательностей (ПСП) Голда, используемых в ГНСС GPS [9]. Помехи $\tilde{x}_m^{(l)}(k)$, $l = 2, \dots, L$ моделировались независимыми комплексными реализациями отсчетов белого Гауссова шума. Шумы в каналах ААР $n_m(k)$ также моделировались независимыми реализациями отсчетов белого Гауссова шума с

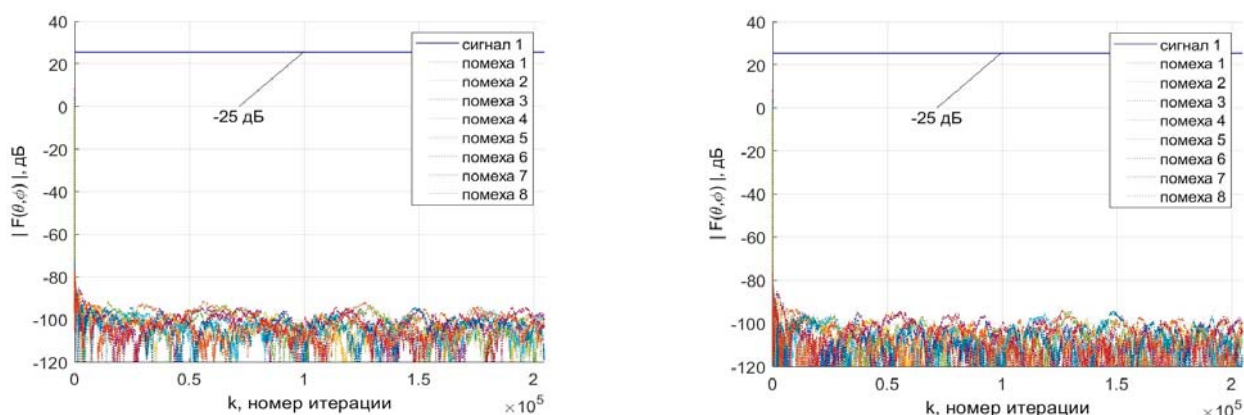
$$\text{ОСШ} = \sigma_{x^{(l)}}^2 / \sigma_n^2 = -20 \text{ дБ},$$

где $\sigma_{x^{(l)}}^2$ – это дисперсия спутникового (полезного) сигнала, а σ_n^2 – дисперсия шумов в каналах РПМ. Если число помех удовлетворяет соотношению $L < M$, то, как уже отмечалось, любая ААР, состоящая из одинаково направленных антенн, может подавить все сигналы этих помех в своём выходном сигнале [3]. Это справедливо и в отношении ААР, рассматриваемых в настоящем разделе.

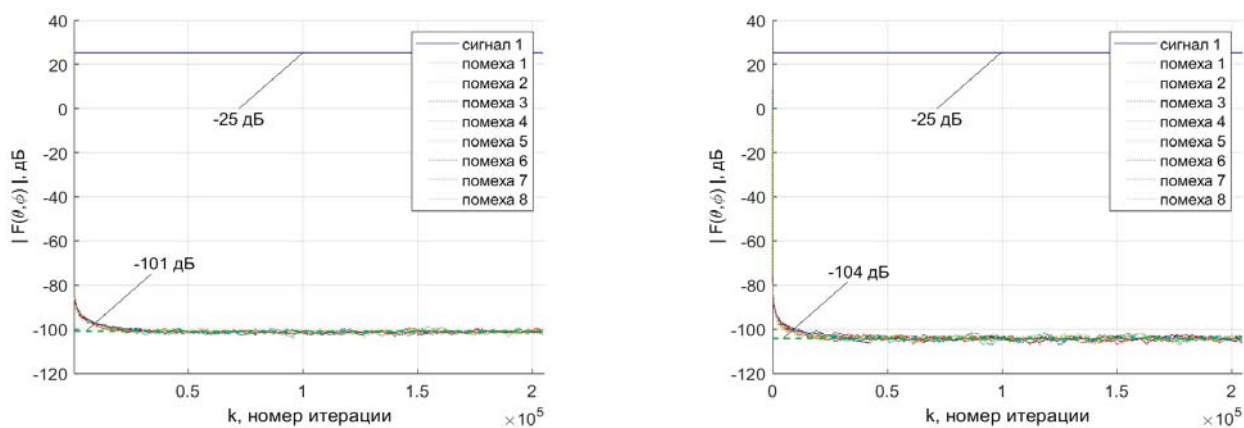
Результаты моделирования ААР приведены на рисунках 2-5. На рисунке 2а и рисунке 2б показаны 3D ДН на плоскости в исходном и в установившемся режимах ААР. Здесь зеленым цветом (см. крестик и цифру) обозначено направление ограничения ДН ААР и направление на источник полезного сигнала. Розовым цветом (см. крестики и цифры) обозначены направления на источники помех и их номера. Цвет самих рисунков обозначает значения ДН в соответствии с цветовой палитрой, расположенной справа от каждого из рисунков.



а) б) **Рис. 2.** ДН ААР: а) – исходная; б) – в установившемся режиме



а) б)



в) г)

Рис. 3. Переходные процессы:

а), в) – ААР в арифметике комплексных чисел; б), г) – ААР в арифметике действительных чисел

В исходном состоянии ВК обеих рассматриваемых ААР одинаковые. Они также практически одинаковые в установившемся режиме. Поэтому их ДН на рис. 2 приведены для случая только одной из двух ААР. Из рисунка 2 следует, что в рассматриваемом эксперименте сигналы всех источников помех подавляются обеими ААР.

На рисунке 3 приведены графики переходных процессов ААР. Здесь $|F(\theta, \phi)|$ – это значения модуля ДН (с учетом ДН антенн) на каждой итерации работы ААР в направлениях источников принимаемых сигналов. Из рисунков 3а и 3б видно,

что в направлениях источников помех в установившемся режиме ААР в её ДН образуются провалы глубиной порядка -100 дБ. Одновременно в направлении зенита выдерживается уровень ДН порядка 25 дБ. Эта величина соответствует значению ограничения ДН ААР

$$20 \log_{10} M + 20 \log_{10} |F(\theta, \phi)|_{\max}.$$

Из рисунков 3а и 3б также видно, что в направлениях источников помех в среднем образуются «более глубокие» провалы в ДН ААР в арифметике действительных чисел, чем в ААР в арифметике комплексных чисел.

Чтобы оценить различие в эффективности подавления помех рассматриваемых ААР необходимо провести статистиче-

ское усреднение результатов моделирования, аналогичных результатам, представленным на рисунках 3а и 3б. При таком моделировании расположение источников сигналов и их энергетические характеристики не менялись. В спутниковом сигнале в каждом эксперименте случайным образом менялась его модуляционная последовательность. Этот сигнал модулировался бинарной фазовой манипуляцией (Binary Phase Shift Keying, BPSK), при которой двадцать ПСП формировали один

бит информации, передаваемой со спутника. Сигналы принимаемых помех и сигналы шумов в каналах ААР также случайным образом менялись в каждом эксперименте. Из усредненных по 50 статистически независимым экспериментам ДН (см. рис. 3в и рис. 3г) следует, что ААР в арифметике действительных чисел, аналогично [17-19], обеспечивает в среднем на 3 дБ большее подавление помех, чем ААР в арифметике комплексных чисел. Из графиков в увеличенном масштабе (рис. 4) также видно, что длительности переходных процессов рассматриваемых ААР, аналогично [17-19], различаются примерно в два раза.

Актуальным, однако, является вопрос, а будут ли приниматься сигналы с других спутников, расположенных в направ-

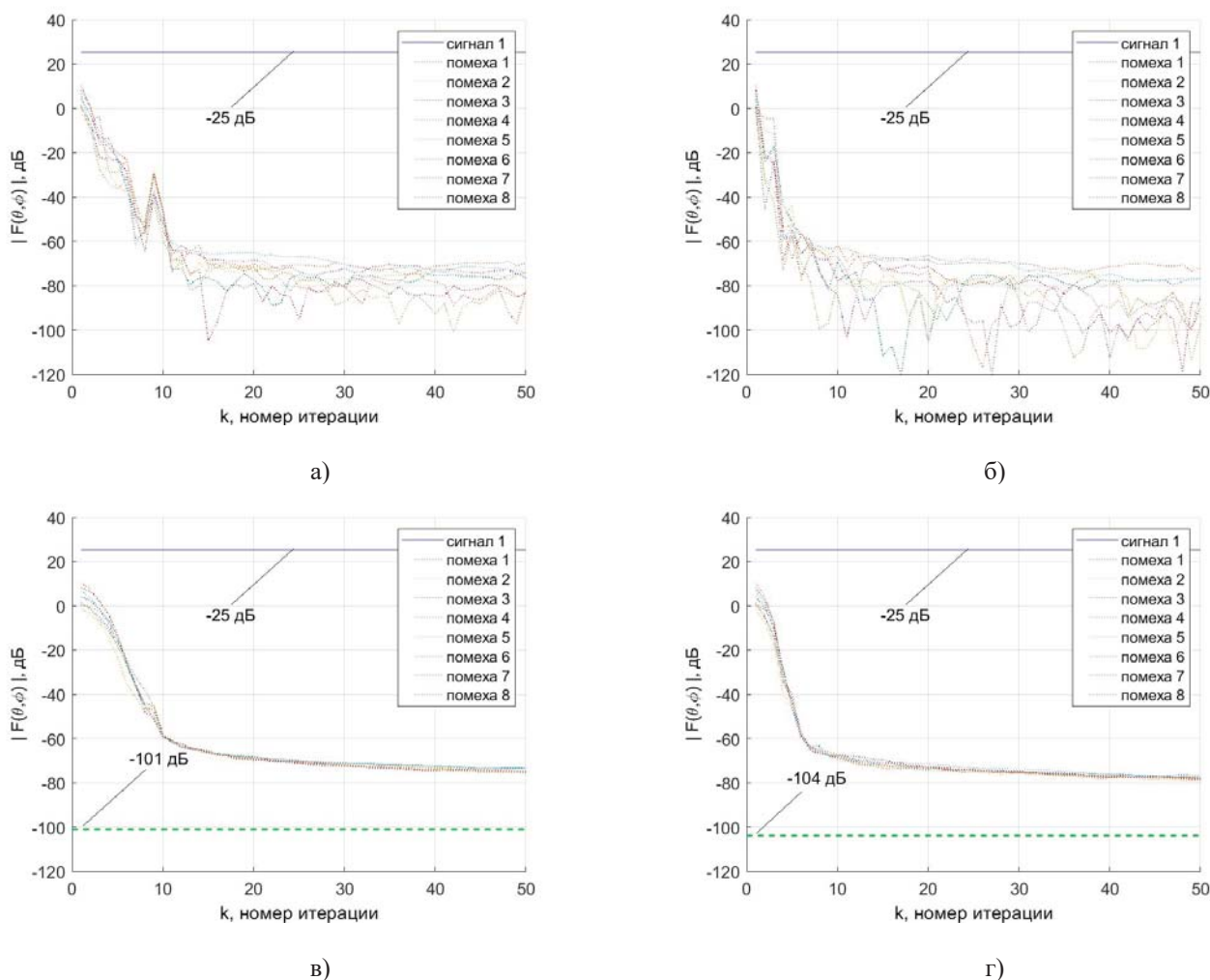


Рис. 4. Переходные процессы (в увеличенном масштабе): а), в) – ААР в арифметике комплексных чисел; б), г) – ААР в арифметике действительных чисел

ских усреднение результатов моделирования, аналогичных результатам, представленным на рисунках 3а и 3б. При таком моделировании расположение источников сигналов и их энергетические характеристики не менялись. В спутниковом сигнале в каждом эксперименте случайным образом менялась его модуляционная последовательность. Этот сигнал модулировался бинарной фазовой манипуляцией (Binary Phase Shift Keying, BPSK), при которой двадцать ПСП формировали один

бит информации, передаваемой со спутника. Сигналы принимаемых помех и сигналы шумов в каналах ААР также случайным образом менялись в каждом эксперименте.

Из усредненных по 50 статистически независимым экспериментам ДН (см. рис. 3в и рис. 3г) следует, что ААР в арифметике действительных чисел, аналогично [17-19], обеспечивает в среднем на 3 дБ большее подавление помех, чем ААР в арифметике комплексных чисел. Из графиков в увеличенном масштабе (рис. 4) также видно, что длительности переходных процессов рассматриваемых ААР, аналогично [17-19], различаются примерно в два раза.

такого сигнала может использоваться ПСП любого из спутниковых сигналов. При этом под ОСШ понимается отношение сигнал шум плюс остаточные подавленные помехи на выходе ААР.

При сильном подавлении (см. рис. 3 и рис. 4), эти помехи не влияют на карты ОСШ. Если при моделировании использовались несколько источников полезных сигналов, то их наличие также не влияет на подавления помех, т.е. на значения вычисляемых ВК, а значит, не влияет и на карты ОСШ, так как полезные сигналы со спутников находятся, как уже отмечалось, ниже уровня тепловых шумов ПРМ в каналах ААР, как минимум, на 20 дБ.

Карты ОСШ позволяют прогнозировать работу ААР в заданной помеховой обстановке, что, в свою очередь, позволяет избежать длительного моделирования ААР при разном числе спутников и их угловых расположениях. Эти карты зависят только от конфигурации ААР, ДН её антенн, используемого адаптивного алгоритма, числа и углового положения источников помех, но не зависят от числа и углового положения источников полезных сигналов (спутников).

Если на картах значения ОСШ, меньшие -20 дБ, проквантовать одним уровнем, то получатся «квантованные» карты

ОСШ (см. рис. 5в и рис. 5г). Эти значения ОСШ на рисунках показаны темно-синим цветом. В направлениях, обозначенных этим цветом, сигналы со спутников не могут быть приняты. Видно, что сигналы со спутников могут приниматься ААР из направлений в диапазоне углов $\theta \approx \pm 40^\circ$ и $\phi = 0^\circ \dots 360^\circ$ относительно направления на зенит. В рассматриваемом эксперименте при отсутствии помех таких направлений около 22% от всех дискретных направлений вокруг неадаптивной ААР.

После подавления помех ААР число направлений с ОСШ > -20 дБ немного увеличилось до 24%. Это обусловлено перераспределением значений исходной ДН из направлений на источники помех в область главного лепестка ДН. Также видно, что подавленные помехи не влияют на прием сигналов со спутников, так как после подавления сигналы помех на выходе ААР находятся ниже уровня шумов РПМ её каналов.

Указанные процентные значения – это отношение площади карты с ОСШ > -20 дБ ко всей площади этой карты. На самом же деле, эти значения можно увеличить в два раза, если учесть, что сигналы со спутников могут приниматься только в верхней полусфере ($\theta = 0^\circ \dots 90^\circ$), т.е. выше линии горизонта,

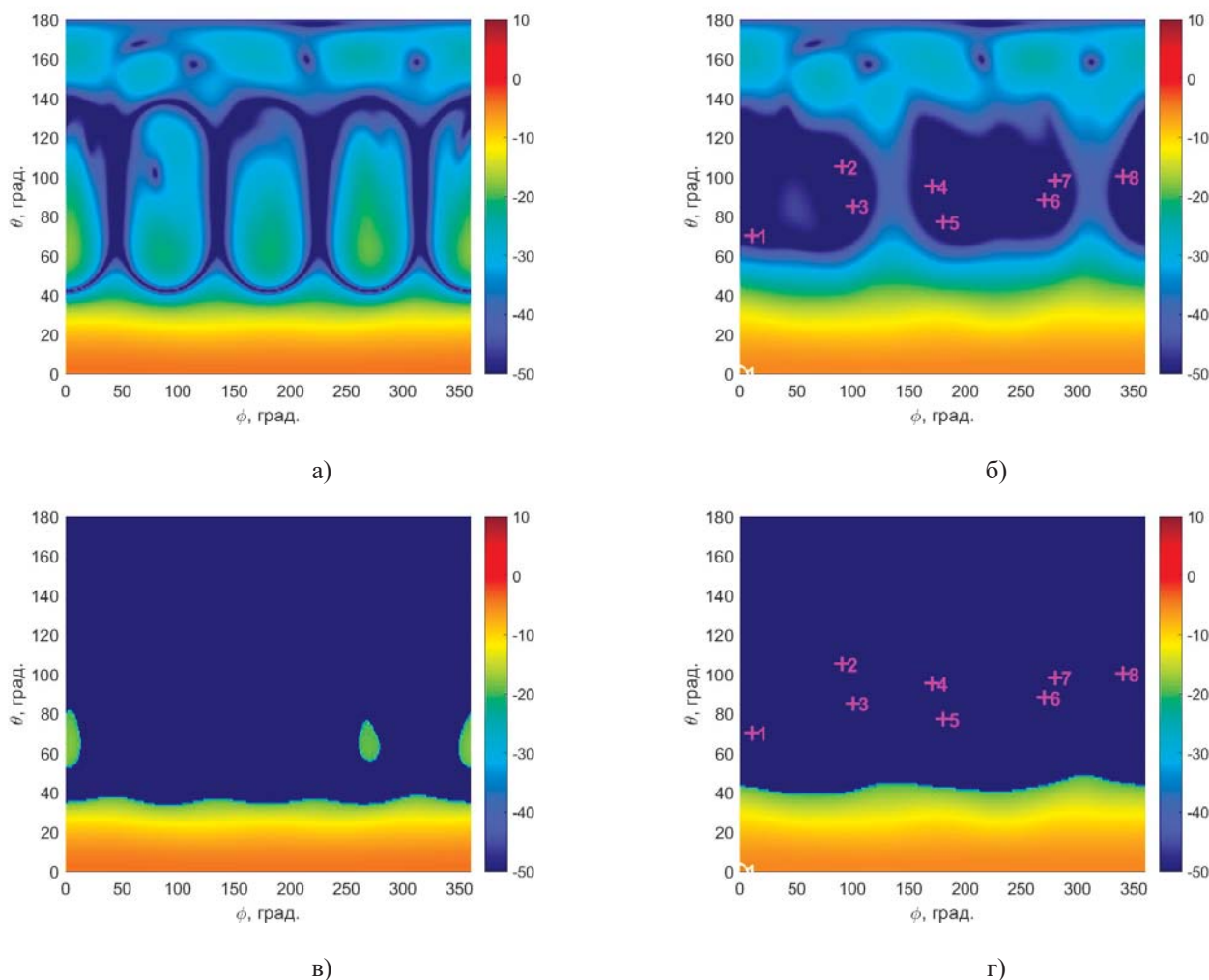


Рис. 5. Карты ОСШ: а) – исходная, не квантованная; б) – в установившемся режиме, не квантованная; в) – исходная, квантованная; г) – в установившемся режиме, квантованная

проходящей через плоскость ААР. Похожие результаты демонстрируются сравнимыми ААР и при других расположениях помех, если их число не превышает значения $M - 1$.

Карты ОСШ, подобно ДН, практически одинаковые для обеих рассматриваемых ААР, так как ВК обеих ААР в установленном режиме примерно одинаковые. В тоже время, ААР в арифметике действительных чисел характеризуется меньшей вычислительной сложностью и более эффективным подавлением помех за счет более точного вычисления ВК по сравнению с ААР в арифметике комплексных чисел.

Заключение

Таким образом, в статье рассмотрена линейно-ограниченная ААР в арифметике действительных чисел. В такой ААР требуемое число действительных арифметических операций в два раза меньше эквивалентного числа действительных операций аналогичной ААР в арифметике комплексных чисел. Переходный процесс ААР в арифметике действительных чисел в два раза короче, а подавление помех в установленном режиме в среднем на 3 дБ больше чем в ААР в арифметике комплексных чисел. Рассмотренная ААР может найти применение не только в аппаратуре ГНСС, но и в аппаратуре других радиосистем, таких как радиосистемы связи, радиолокации и радиоуправления, в которых можно сформулировать и задать ограничение ДН в интересующем направлении, а принимаемые радиосигналы скрыты шумами радиоприемников.

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

Литература

1. Brown A. D., Boeringer D., Cooke T. Electronically scanned arrays. MATLAB® modelling and simulation. CRC Press, 2012, 214 p.
2. Balanis C. A. Antenna theory: analysis and design. 4-th ed. John Wiley & Sons, Inc., 2016, 1095 p.
3. Hudson J. E. Adaptive array principles. The Institution of Engineering and Technology, 2007, 253 p.
4. Ратынский М. В. Адаптация и сверхразрешение в антенных решетках. М.: Ленард, 2024, 240 с.
5. Джиган В. И. Адаптивная фильтрация сигналов: теория и алгоритмы. М.: Техносфера, 2013, 528 с.
6. Diniz P. S. R. Adaptive filtering algorithms and practical implementation, 5-th ed. Springer, 2020, 495 p.
7. Frost O. L. An algorithm for linearly constrained adaptive array processing // Proceedings of the IEEE. 1972. Vol. 60. No. 8. P. 926-935.

8. Resende L. S., Romano J. M. T., Bellanger M. G. A fast least-squares algorithm for linearly constrained adaptive filtering // IEEE Trans. Signal Processing. 1996. Vol. 44. No. 5. P. 1168-1174.

9. Hofmann-Wellenhof B., Lichtenegger H., Wasle E. GNSS – Global Navigation Satellite Systems: GPS, GLONASS, Galileo, and more. Springer, 2008, 516 p.

10. Djigan V. I. Adaptive arrays for GNSS receivers // Proceedings of the IEEE 3rd International Conference on Problems of Informatics, Electronics and Radio Engineering (PIERE). Novosibirsk, Russia, November 15-17, 2024. 4 p.

11. Howells P. W. Intermediate frequency side-lobe canceller // US Patent 3202990. 1959. 8 p.

12. Джиган В. И. Цифровая адаптивная антенная решетка для приема полезных сигналов, находящихся под тепловыми шумами // Цифровая обработка сигналов. 2024. № 2. С. 3-10.

13. Джиган В. И. Выбор адаптивного алгоритма для цифровой антенной решетки наземной аппаратуры систем глобальной спутниковой навигации // Цифровая обработка сигналов. 2025. № 2. С. 27-35.

14. Джиган В. И. Активное шумоподавление: архитектуры, алгоритмы и их параметры // T-Comm: Телекоммуникации и транспорт. 2025. № 8. С. 3-11.

15. Джиган В. И. Сравнительная эффективность цифровых предсказателей сигналов на базе адаптивных рекурсивных алгоритмов по критерию наименьших квадратов // T-Comm: Телекоммуникации и транспорт. 2025. №11. С. 4-14.

16. Djigan V. I. Some tricks of calculations in MIL RLS algorithm // Proceedings of the 23-th International Conference on Digital Signal Processing and its Applications (DSPA-2021). Moscow, Russia, March 24-26, 2021. 4 p.

17. Nitzberg R. Application of maximum likelihood estimation of persymmetric covariance matrices to adaptive processing // IEEE Trans. Aerospace and Electronic Systems. 1980. Vol. 16. No. 1. P. 124-127.

18. Huang K. C., Yen C. C. A unitary transformation method for angle of arrival estimation // IEEE Trans. Acoustic, Speech and Signal Processing. 1991. Vol. 39. No. 4. P. 975-977.

19. Джиган В. И. Нечетная симметрия вектора весовых коэффициентов симметричных антенных решеток с линейными ограничениями // Известия высших учебных заведений. Радиоэлектроника. 2018. № 6. С. 323-335.

20. Григорьев Л. Н. Цифровое формирование диаграммы направленности в фазированных антенных решетках. М.: Радиотехника, 2010. 144 с.

21. Reyland J. M. Software defined radio: theory and practice. Artech House, 2023, 796 p.

22. Джиган В. И., Шериунович А. Ю., Луферчик П. В., Штро П. В. Карты отношения сигнал-шум на выходе адаптивной антенной решетки для оценки возможности приёма сигналов спутниковой навигации // Труды 9-й Всероссийской научно-технической конференции «Системы связи и радионавигации». Россия, г. Красноярск, 16-18 сентября 2025. С. 18-22.

LINEARLY-CONSTRAINED ADAPTIVE ANTENNA ARRAY FOR GROUND-BASED RADIO RECEIVERS OF GLOBAL NAVIGATION SATELLITE SYSTEM

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

Abstract

Antenna arrays are widely used as the antennas of modern radio systems today. The ability to control the Radiation Pattern (RP) of such antennas underlies the operating principle of the Adaptive Antenna Arrays (AAA). These arrays suppress the signals from the interference sources whose frequency bands overlap with the frequency band of the desired (informational) signal. One of the most relevant applications of the AAAs is their usage in the ground-based equipment of the Global Navigation Satellite Systems (GNSS). Such equipment primarily utilizes the AAAs known as the interference cancelers. Like any AAA, they suppress interferences but do not guarantee the receiving of a predetermined number of satellite signals which are necessary for solving the navigation task. This paper proposes to use a version of the linearly-constrained AAA. The constraint ensures a fixed value of the AAA RP towards zenith. The received signals from the satellites, which are located in this direction, provide the maximal signal-to-noise ratio due to the shortest distance from satellites to Earth. The value of the RP specified by the constraint is independent of the values of the AAA weights calculated during the adaptation. To calculate these weights, a recursive least-squares algorithm based on the matrix inversion lemma and optimized for the number of arithmetic operations has been used. The computational complexity of this algorithm is lower than that of its non-optimized version. This complexity reduction is achieved by the taking into account the Hermitian structure of the correlation matrix of the AAA input signals. A further reduction in the computational complexity is achieved if the AAA is symmetric. Almost all operations in such the AAA can be executed using real valued arithmetic. Their number is a half that of the equivalent number of real valued operations in the similar AAA in complex valued arithmetic. Furthermore, compared to the complex valued arithmetic AAA, the transient response of the AAA under the consideration is two times shorter and the interference suppression is on average 3 dB bigger in the steady-state mode. The proposed AAA can be used not only in GNSS equipment but also in other radio systems.

Keywords: Global Navigation Satellite Systems, Recursive Least Squares algorithm, linear constraint, adaptive antenna array, complex-valued arithmetic, real-valued arithmetic, signal-to-noise ratio map.

References

- [1] A. D. Brown, D. Boeringer, and T. Cooke, *Electronically Scanned Arrays. MATLAB® Modelling and Simulation*, CRC Press, 2012, 214 p.
- [2] C. A. Balanis, *Antenna Theory: Analysis and Design*, 4-th ed., John Wiley & Sons, Inc., 2016, 1095 p.
- [3] J. E. Hudson, *Adaptive Array Principles*, The Institution of Engineering and Technology, 2007, 253 p.
- [4] M. V. Ratynski, *Adaptation and Superresolution in Antenna Arrays*, Moscoe, Lenard Publisher, 2024, 240 p.
- [5] V. I. Djigan, *Adaptive Filtering: Theory and Algorithms*, Moscow, Technosphaera Publisher, 2013, 528 p. (In Russian).
- [6] P. S. R. Diniz, *Adaptive Filtering Algorithms and Practical Implementation*, 5-th ed, Springer, 2020, 495 p.
- [7] O. L. Frost, "An algorithm for linearly constrained adaptive array processing", *Proceedings of the IEEE*, 1972, vol. 60, no. 8, pp. 926-935.
- [8] L. S. Resende, J. M. T. Romano, and M. G. Bellanger, "A fast least-squares algorithm for linearly constrained adaptive filtering," *IEEE Trans. Signal Processing*, 1996, vol. 44, no. 5, pp. 1168-1174.
- [9] B. Hofmann-Wellenhof, H. Lichtenegger, and E. Wasle, *GNSS – Global Navigation Satellite Systems: GPS, GLONASS, Galileo, and More*, Springer, 2008, 516 p.
- [10] V. I. Djigan, "Adaptive arrays for GNSS receivers," *Proceedings of the IEEE 3rd International Conference on Problems of Informatics, Electronics and Radio Engineering (PIERE)*, Novosibirsk, Russia, November 15-17, 2024, 4 p.
- [11] P. W. Howells, "Intermediate frequency side-lobe canceller," *US Patent 3202990*, 1959, 8 p.
- [12] V. I. Djigan, "Digital adaptive antenna array for receiving informational signals under thermal noise," *Digital Signal Processing*, Moscow, Russia, 2024, no. 2, pp. 3-10. (In Russian)
- [13] V. I. Djigan, "Selection of adaptive algorithm for digital antenna array of land-based equipment for global satellite navigation systems," *Digital Signal Processing*, Moscow, Russia, 2025, no. 2, pp. 27-35. (In Russian).
- [14] V. I. Djigan, "Active noise control: architectures, algorithms and its parameters", *T-comm - Telecommunications and their Applications in Transport Industry*, Russia, 2025, no. 8, pp. 3-11. (In Russian).
- [15] V. I. Djigan, "Comparative efficiency of signal digital predistorters based on adaptive Recursive Least Squares algorithms," *T-Comm*, Russia, 2025, no. 11, pp. 4-14. (In Russian)
- [16] V. I. Djigan, "Some tricks of calculations in MIL RLS algorithm," *Proceedings of the 23-th International Conference on Digital Signal Processing and its Applications (DSPA-2021)*, Moscow, Russia, March 24-26, 2021, 4 p.
- [17] R. Nitzberg, "Application of maximum likelihood estimation of persymmetric covariance matrices to adaptive processing," *IEEE Trans. Aerospace and Electronic Systems*, 1980, vol. 16, no. 1, pp. 124-127.
- [18] K. C. Huarng and C. C. Yen, "A unitary transformation method for angle of arrival estimation," *IEEE Trans. Acoustic, Speech and Signal Processing*, 1991, vol. 9, no. 4, pp. 975-977.
- [19] V. I. Djigan, "Odd symmetry of weights vector of symmetrical adaptive arrays with linear constraints," *Proceedings of High School. Radioelectronics*, Kiev, Ukraine, 2018, no. 6, pp. 323-335. (In Russian)
- [20] L. N. Grigoriev, *Digital Beamforming in Phased Antenna Arrays*. Moscow, Radiotekhnika Publisher, 2010, 144 p. (In Russian)
- [21] J. M. Reylund, *Software Defined Radio: Theory and Practice*, Artech House, 2023, 796 p.
- [22] V. I. Djigan, A. Yu. Shershunovich, P. V. Lufchik, P. V. Shtro "Signal-to-noise ratio maps at output of adaptive antenna array for assessing possibility of receiving satellite navigation signals," *Proceedings of the 9-th All-Russian Scientific and Practical Conferences on Communication and Global Position Radio Systems*. Krasnoyarsk, Russia, September 16 -18, 2025, pp. 18-22.

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

FTN И NOMA: МИФЫ И РЕАЛЬНОСТЬ

DOI: 10.36724/2072-8735-2026-20-3-15-27

Бакулин Михаил Германович,
Московский технический университет связи и информатики,
Москва, Россия, m.g.bakulin@gmail.com

Крейнделин Виталий Борисович,
Московский технический университет связи и информатики,
Москва, Россия, vitkrend@gmail.com

Резнев Андрей Алексеевич,
Московский технический университет связи и информатики;
АО "НТЦ "Атлас", Москва, Россия,
andrey.reznev@gmail.com

Manuscript received 28 December 2025;
Accepted 05 March 2026

Ключевые слова: спектральная
эффективность, FTN, MIMO, NOMA,
пространственное мультиплексирование,
5G, имитационное моделирование

Технологии Faster-Than-Nyquist (FTN) и Non-Orthogonal Multiple Access (NOMA) широко рассматриваются как перспективные методы повышения спектральной эффективности в системах беспроводной связи 5G и 6G. Обе основаны на преднамеренном нарушении ортогональности: FTN - за счёт уплотнения символов во времени, NOMA - за счёт совместного использования ресурсов несколькими пользователями. В литературе часто утверждается, что это "перегруженные" системы, увеличивающие пропускную способность за счёт контролируемой интерференции. Однако такой взгляд вводит в заблуждение. На основе теоретико-информационного анализа показано, что ни FTN, ни NOMA не позволяют превысить классический предел Шеннона для гауссовского канала. Их реальное преимущество проявляется исключительно при передаче дискретных сигналов (например, QAM/BPSK), которые по своей природе не достигают шенноновской емкости. Таким образом, эффективность этих технологий объясняется не "магией интерференции", а тем, что они компенсируют неоптимальность исходной модуляции, приближая распределение сигнала к гауссовому и тем самым "догружая" изначально недогруженные каналы. Это особенно заметно при использовании низкочастотных модуляций, характерных для IoT и энергоограниченных устройств. Понимание этой природы открывает новые направления для оптимизации FTN- и NOMA-систем: вместо минимизации помех следует максимизировать эффективность использования канала с учётом дискретной структуры передаваемой информации. Выводы подтверждаются имитационным моделированием и анализом взаимной информации для дискретно-непрерывных каналов.

Информация об авторах:

Бакулин Михаил Германович, Московский технический университет связи и информатики (МТУСИ), доцент, к.т.н., Москва, Россия
Крейнделин Виталий Борисович, Московский технический университет связи и информатики (МТУСИ), профессор, д.т.н., Москва, Россия
Резнев Андрей Алексеевич, Московский технический университет связи и информатики (МТУСИ), доцент; АО "НТЦ "Атлас", заместитель генерального директора, к.т.н., Москва, Россия

Для цитирования:

Бакулин М.Г., Крейнделин В.Б., Резнев А.А. FTN и NOMA: мифы и реальность // T-Comm: Телекоммуникации и транспорт. 2026. Том 20. №3. С. 15-27.

For citation:

M.G. Bakulin, V.B. Kreyndelin, A.A. Reznev, "FTN and NOMA: Beyond the Hype - Toward Shannon-Limited Efficiency," T-Comm, 2026, vol. 20, no.3, pp. 15-27. (in Russian)

Введение

Повышение спектральной эффективности и энергетической эффективности – ключевые задачи современных телекоммуникационных систем (5G, 6G, IoT, спутниковая связь). Это обусловлено следующими факторами:

1. Физическими ограничениями (спектр, энергия). Свободные частотные диапазоны почти исчерпаны (особенно ниже 6 ГГц). Новые диапазоны (миллиметровые волны, терагерцовый диапазон) имеют высокие потери и сложны в реализации;

2. Ростом нагрузки (трафик, число устройств). К 2030 году ожидается ~500 млрд подключённых устройств (IoT, умные города, автономные системы). Трафик данных растёт экспоненциально (видео 8K, VR/AR, метавселенные);

3. Экологией и экономикой. На телекоммуникации приходится ~2% глобальных выбросов CO₂ (сопоставимо с авиацией). Энергоэффективные технологии снижают углеродный след. Операторы связи стремятся снизить OPEX и в том числе расходы на энергопотребление, внедряя ESG практики;

4. Требованиями будущих технологий (6G, IoT, метавселенные). 6G ставит цели: спектральная эффективность >100 бит/с/Гц (против ~30 в 5G), энергоэффективность 1 Терабит/Дж.

Поэтому идёт постоянный поиск новых высокоэффективных технологий. Среди таких технологий можно выделить технологию передачи со скоростью больше скорости Найквиста (Faster-Than-Nyquist – FTN) [1-36] и технологию неортогонального множественного доступа (Non-Orthogonal Multiple Acces – NOMA) [37-55].

В статье на основе анализа пропускной способности систем связи с технологиями FTN и NOMA делается вывод о том, что основным фактором повышения эффективности при использовании этих технологий является не столько наличие контролируемой интерференции, сколько тот факт, что для внедрения этих технологий рассматриваются системы с невысокой спектральной эффективностью, т.е. системы с «недогруженными» каналами. Под степенью недогрузки здесь понимается разница между пропускной способностью по Шеннону и реальной пропускной способностью системы связи.

Таким образом, Технологии вроде FTN и NOMA – это не "прорыв за предел Шеннона", а инструменты для его достижения в реальных системах связи, которые сегодня работают далеко от теоретического максимума.

Технология Faster-than-Nyquist. Дискретная модель

FTN-технология представляет собой метод передачи символов со скоростью быстрее, чем скорость Найквиста, определяемой полосой пропускания канала. Основы данной технологии были предложены еще в 1970 году в небольшой работе Lucky [1]. Вскоре последовали работы авторов Salz [2] и Foschini [3], и в 1975 году Мазо [4] показал, что скорость передачи сигналов может быть увеличена до 25% выше скорости Найквиста без снижения минимального евклидова расстояния. Таким образом, была теоретически доказана потенциальная возможность FTN-технологии осуществлять передачу со скоростью выше скорости Найквиста. Однако Мазо изначально использовал идеальные модулирующие

импульсы типа $\frac{\sin(x)}{x}$, но в дальнейшем были получены ре-

зультаты с использованием более удобных для практического использования импульсов, таких как семейство приподнятого-косинуса (raised-cosine pulses) (Liveris и Georghiadis [7]) и другие ограниченные во времени импульсы (Rusek and Anderson [8-10]).

Рассмотрим случай передачи цифровой информации по каналам с ограниченной полосой (рис. 1).

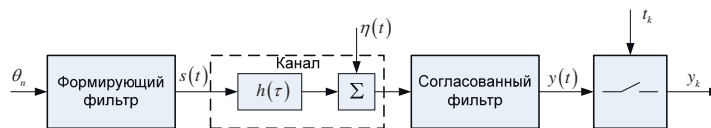


Рис. 1. Упрощённая структура системы передачи дискретной информации по каналам с ограниченной полосой

Последовательность модулированных символов θ_n (отсчётов) проходит через формирующий фильтр с импульсной характеристикой $g_f(\tau)$, который ограничивает спектр излучения сигнала, в результате на входе канала имеем

$$s(t) = \sum_{n=-\infty}^{\infty} \theta_n A g_f(t - nT), \tag{1}$$

где A – амплитуда импульса, которую в дальнейшем будем полагать равной 1.

Мощность сигнала (1) равна

$$P_s = E \left\langle \lim_{\Delta t \rightarrow \infty} \frac{1}{\Delta t} \int_{-\frac{\Delta t}{2}}^{\frac{\Delta t}{2}} \left(\sum_{n=-\infty}^{\infty} \theta_n g_f(t - nT) \right)^2 dt \right\rangle = \frac{1}{T} \int_{-\infty}^{\infty} g_f^2(t) dt = \frac{E_g}{T}, \tag{2}$$

где $E_g = \int_{-\infty}^{\infty} g_f^2(t) dt$ – энергия единичного импульса с формой $g_f(t)$.

Далее сигнал $s(t)$ проходит через линейный канал с импульсной характеристикой $h(\tau)$ и суммируется с аддитивным гауссовским шумом $\eta(t)$. На приемной стороне смесь сигнала с шумом проходит через фильтр, согласованный с формирующим фильтром и каналом. Для простоты анализа рассмотрим случай, когда на протяжении длительности нескольких символов модель канала может быть описана линейной операцией умножения на постоянный множитель h , т.е. канал является достаточно широкополосным и не искажает огибающую символа. В этом случае для непрерывного времени технология передачи информации со скоростью Найквиста описывается следующим уравнением наблюдения:

$$y(t) = h \sum_{n=-\infty}^{\infty} \theta_n g(t - nT) + \eta(t), \tag{3}$$

где $g(\tau) = \int_{-\infty}^{\infty} g_f(\tau - t) g_f(t) dt$ – форма единичного импульса

на выходе согласованного фильтра (СФ). Она, как правило, должна удовлетворять условию отсутствия межсимвольных помех при передаче со скоростью Найквиста по каналам с ограниченной полосой пропускания, например, описываться функцией приподнятого косинуса:

$$g(\tau) = \frac{\sin\left(\frac{\pi\tau}{t}\right) \cos\left(\frac{\alpha\pi\tau}{T}\right)}{\left(\frac{\pi\tau}{t}\right) \left[1 - \left(\frac{2\alpha\pi\tau}{T}\right)^2\right]}, \quad (4)$$

где $\alpha \in [0; 1]$ – параметр скругления. В этом случае, импульсная характеристика формирующего фильтра, известного как «square root raised cosine filter» [50, 51], описывается выражением:

$$g_f(\tau) = \frac{(\sin((1-\alpha)\pi\tau) + 4\alpha\pi\tau \cos((1+\alpha)\pi\tau))}{\pi\tau(1-(4\alpha\pi\tau)^2)}. \quad (5)$$

На выходе СФ наблюдается смесь сигнала и шума

$$y(t) = h \sum_{n=-\infty}^{\infty} \theta_n g(t - nT) + \eta(t) \quad (6)$$

где $\eta(t)$ – аддитивный гауссовский шум с корреляционной функцией $R(\tau) = \frac{N_0}{2} g_N(\tau)$, N_0 – спектральная плотность белого гауссовского шума.

После дискретизации получим уравнение наблюдения в дискретном времени:

$$y(t_n) = h \sum_{k=-\infty}^{\infty} \theta_k g(t_n - kT) + \eta(t_n), \quad (7)$$

или

$$y_n = h \sum_{k=-\infty}^{\infty} \theta_k g((n-k)T) + \eta_n. \quad (8)$$

Пусть ставится задача передача информации с тактовой частотой $F_s = \frac{1}{T_s} > F_N = \frac{1}{T}$, (где F_N – частота Найквиста), по каналу с импульсной характеристикой $g(\tau)$, согласованной с частотой Найквиста, т.е. для непрерывного времени имеем следующее уравнение наблюдения:

$$y(t) = h \sum_{n=-\infty}^{\infty} \theta_n g(t - nT_s) + \eta(t). \quad (9)$$

Так как частота следования символов в данном случае $F_s > F_N$, то такие системы называют системами передачи с частотой больше частоты Найквиста (Faster-than-Nyquist – FTN). Отношение (T_s / T) называют фактором ускорения времени [10Ошибка! Источник ссылки не найден.]. Средняя мощность излучаемого сигнала для такой системы будет определяться следующим выражением:

$$P_{s,FTN} = E \left\langle \lim_{\Delta t \rightarrow \infty} \frac{1}{\Delta t} \int_{-\frac{\Delta t}{2}}^{\frac{\Delta t}{2}} \left(\sum_{n=-\infty}^{\infty} \theta_n g_f(t - nT_s) \right)^2 dt \right\rangle = \frac{1}{T_s} \int_{-\infty}^{\infty} g_f^2(t) dt = \frac{E_g}{T_s}. \quad (10)$$

Отсюда видно, что при увеличении частоты следования импульсов (при уменьшении тактового интервала $T_s < T$) и при условии, что формы этих импульсов не меняются,

мощность сигнала увеличивается $P_{s,FTN} = \frac{E_g}{T_s} > P_s = \frac{E_g}{T}$, что

приводит к увеличению отношения сигнал/шум. Поэтому для корректного сравнения двух систем необходимо ввести нормировку мощности. В этом случае уравнение наблюдения для системы FTN будет следующим:

$$y(t) = h \sqrt{\frac{T}{T_s}} \sum_{n=-\infty}^{\infty} \theta_n g(t - nT_s) + \eta(t). \quad (11)$$

После дискретизации получим

$$y_n = h \sum_{k=-\infty}^{\infty} \theta_k g\left(\left(n - k \frac{T_s}{T}\right)T\right) + \eta_n. \quad (12)$$

Для фильтра, согласованного с частотой Найквиста, характерно

$$g(mT) = \begin{cases} 1, & \text{для } m = 0 \\ 0, & \text{для } m \neq 0, m = \pm 1, \pm 2, \pm 3, \dots \end{cases}. \quad (13)$$

На рис. 2 приведена временная диаграмма сигнала на выходе согласованного фильтра в отсутствии шума (‘синяя линия’) при передаче со скоростью Найквиста ($T_s = T$). Там же изображены индивидуальные сигналы для разных символов и моменты взятия отсчетов. Из этого графика видно, что в моменты взятия отсчетов $t = nT$, значения этих отсчетов совпадают с текущими символами, т.е.

$$s(t_n) = h \sum_{k=-\infty}^{\infty} \theta_k g(t_n - kT) = \theta_k, n = k. \quad (14)$$

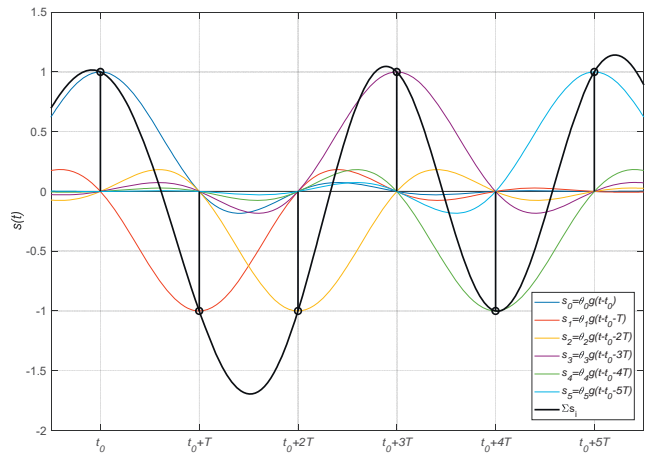


Рис. 2. Временные диаграммы сигналов при частоте следования импульсов равной частоте Найквиста

Пусть по каналу с ограниченной полосой требуется передать N символов. Согласно рассмотренному подходу при передаче со скоростью Найквиста достаточно принять N отсчетов, взятых с временным шагом $T = \frac{1}{F_N}$.

Если рассмотреть данную задачу в дискретном времени, то уравнение наблюдения может быть записано в векторно-матричном виде:

$$y = h \mathbf{I}_N \boldsymbol{\theta} + \boldsymbol{\eta}, \quad (15)$$

где $\mathbf{y} \triangleq [y_1 \ y_2 \ \dots \ y_N]^T$ – вектор отсчётов наблюдения, $\boldsymbol{\theta} \triangleq [\theta_1 \ \theta_2 \ \dots \ \theta_N]^T$ – вектор символов, \mathbf{I}_N – единичная матрица, размером $(N \times N)$, $\boldsymbol{\eta}$ – вектор отсчётов гауссовского шума наблюдения с нулевым математическим ожиданием и корреляционной матрицей $\mathbf{R}_\eta = \sigma_\eta^2 \mathbf{I}_N$.

Уравнение (15) характерно для передачи информации по ортогональным каналам. В данном случае матрица ортогонального преобразования (передачи) описывается единичной матрицей.

Рассмотрим теперь случай, когда требуется передать N_s символов по тому же каналу, т.е. с той же полосой пропускания и за то же время. Причем $N_s > N$, т.е. длительность передаваемых символов будет $T_s = \frac{N}{N_s} T$, соответственно, скорость передачи $R_s = \frac{1}{T_s} = \frac{N_s}{NT} = \frac{N_s}{N} F_N > F_N$. Пример такой передачи для $N_s = 9$ и $N = 6$ иллюстрируется на рис. 3.

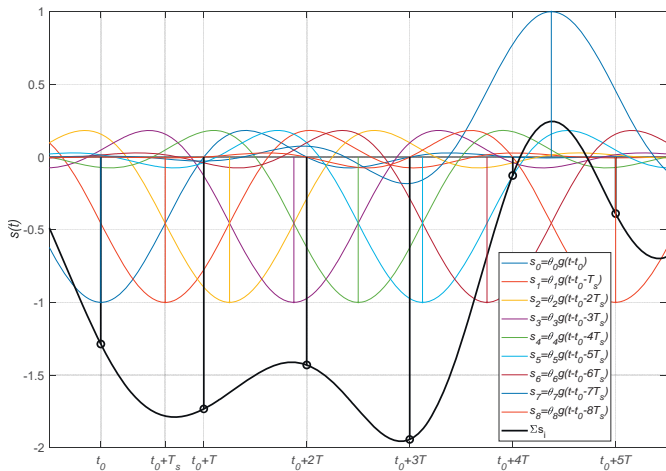


Рис. 3. Временные диаграммы сигналов при частоте следования импульсов больше частоты Найквиста

Здесь хорошо видно, что отсчёты, взятые в моменты $t_n = nT$, не совпадают со значениями самих символов, а, согласно, уравнению (45) содержат сумму разных символов. Этот случай также можно описать в векторно-матричном виде:

$$\mathbf{y} = h \sqrt{\frac{T}{T_s}} \mathbf{B}_{N \times N_s} \boldsymbol{\theta} + \boldsymbol{\eta}, \quad (16)$$

где $\boldsymbol{\theta} \triangleq [\theta_1 \ \theta_2 \ \dots \ \theta_{N_s}]^T$ – N_s -мерный вектор символов, $\mathbf{B}_{N \times N_s}$ – $(N \times N_s)$ -мерная матрица с элементами $b_{nk} = g(nT - kT_s) = g\left(\left(n - k \frac{N}{N_s}\right)T\right)$, для $n = \overline{0, N-1}$, $k = \overline{0, N_s-1}$.

Рассмотренная технология FTN и её дискретный аналог DFTN используется для повышения скорости передачи в

канале во временной области. Аналогичные подходы могут использоваться и в частотной области [18]. В этом случае, уплотнение достигается путем уменьшения частотного интервала между поднесущими OFDM сигнала при неизменной скорости передачи (при этом сужается спектр группового сигнала) или путём увеличения скорости передачи без изменения шага между поднесущими.

Таким образом, технология передачи со скоростью больше скорости Найквиста (FTN) может быть описана в дискретном времени простым векторно-матричным уравнением вида (16). Данный вариант технологии FTN известен как Discrete Time FTN (DTFTN) [30] или её можно назвать дискретная FTN во временной области (Time Domain – Discrete FTN – TD-DFTN). Причем для случая $N_s = N$ матрица преобразования $\mathbf{B}_{N \times N_s}$ становится ортогональной (единичной матрицей).

Аналогичный подход может быть применён и к анализу в частотной области. Рассмотрим систему с модуляцией OFDM с N поднесущими, на каждой поднесущей передаётся модулированный символ длительностью NT . Действительная составляющая комплексного спектра такого сигнала приведена на рис. 4. Как видно из этого рисунка значения спектральных составляющих на частотах поднесущих полностью совпадают со значениями символов передаваемых на этих поднесущих. В отличие от временного описания сигнала здесь мы уже имеем дискретное представление, так как сигналы поднесущих рассматриваются на дискретных значениях частот.

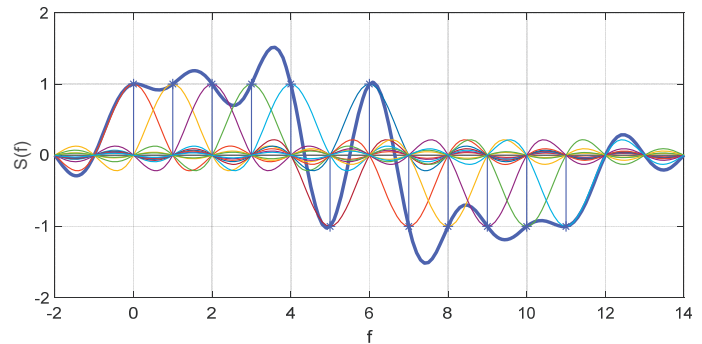


Рис. 4. Действительная составляющая спектра сигнала OFDM при передаче информации со скоростью Найквиста

По аналогии с уравнением во временной области (15) можно записать аналогичное уравнение в частотной области для сигнала OFDM

$$\mathbf{y} = h \mathbf{I}_N \boldsymbol{\theta} + \boldsymbol{\eta}, \quad (17)$$

только в данном случае $\mathbf{y} \triangleq [y_1 \ y_2 \ \dots \ y_N]^T$ – вектор значений смеси сигнала и шума в частотной области, т.е. значения поднесущих. Для канала с частотно-селективными замираниями каналные множители для разных поднесущих могут различаться, в этом случае имеем следующую векторно-матричную модель:

$$\mathbf{y} = \mathbf{H} \mathbf{I}_N \boldsymbol{\theta} + \boldsymbol{\eta}, \quad (18)$$

где \mathbf{H} – диагональная матрица с коэффициентами h_i , $i = \overline{1, N}$, на главной диагонали.

Рассмотрим случай, когда в полосе $N\Delta F$ (где $\Delta F = \frac{1}{NT}$ – дискретный шаг по частоте между ортогональными поднесущими) используются N_s поднесущих с шагом $\Delta F_s = \frac{1}{N_s T}$.

Таким образом, в такой системе за время NT в полосе $\Delta f = \frac{1}{T}$ передаются $N_s > N$ модулированных символов, т.е. мы имеем вариант передачи со скоростью больше скорости Найквиста, но формирование такого сигнала происходит в частотной области. Действительная часть комплексного спектра такого сигнала изображена на рис. 5. Видно, что уменьшение шага между поднесущими приводит к возникновению межканальных помех на поднесущих. Этот вариант также может быть описан векторно-матричным уравнением

$$y = \sqrt{\frac{\Delta F_s}{\Delta F}} \mathbf{H} \mathbf{B}_{N \times N_s} \theta + \eta \quad (19)$$

где $\mathbf{B}_{N \times N_s}$ – $(N \times N_s)$ -мерная матрица с элементами

$$b_{nk} = \gamma(n\Delta F - k\Delta F_s) = \gamma\left(n\frac{1}{NT} - k\frac{1}{N_s T}\right) = \gamma\left(\left(n - k\frac{N}{N_s}\right)\frac{1}{NT}\right) = \gamma\left(\left(n - k\frac{N}{N_s}\right)\Delta F\right)$$

для $n = \overline{0, N-1}$, $k = \overline{0, N_s-1}$, $\gamma(f)$ – комплексная огибающая спектра единичного символа.

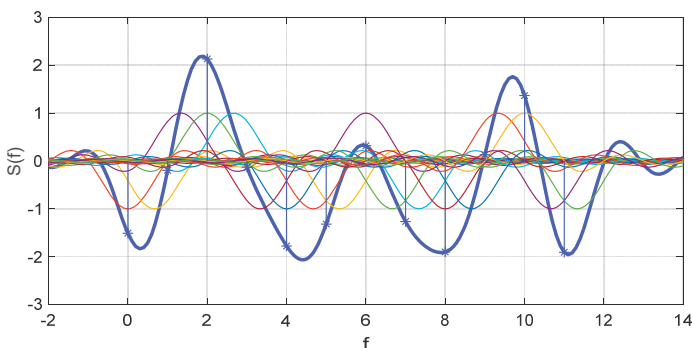


Рис. 5. Действительная составляющая спектра сигнала OFDM при передаче информации со скоростью больше скорости Найквиста на каждой поднесущей

Такой вариант разновидности FTN можно назвать дискретная FTN в частотной области (Frequency Domain Discrete FTN – FD-DFTN).

Анализируя обе разновидности дискретной FTN можно отметить, что для них является общим наличие линейного преобразования N_s -мерного вектора символов в N -мерный вектор отсчётов. Причем и в первом, и во втором случае вектор отсчётов можно рассматривать как результат ортогонального преобразования для сигнала, согласованного с частотно-временными характеристиками канала (с частотой Найквиста), т.е. другими словами говоря, это есть ортогональные

ресурсы, которые позволяют передать согласованный сигнал без потерь. Поэтому дискретные разновидности FTN (DFTN) можно обобщить и рассматривать как передачу N_s модулированных символов по N ортогональным ресурсам. Такая передача может в общем случае описываться векторно-матричным уравнением наблюдения

$$y = \mathbf{W}\theta + \eta, \quad (20)$$

где $\mathbf{W} \triangleq \mathbf{H} \cdot \bar{\mathbf{B}}_{N \times N_s}$ – эквивалентная матрица канала, объединяющая нормированную матрицу преобразования $\bar{\mathbf{B}}_{N \times N_s}$ и матрицу канала \mathbf{H} , нормированная матрица преобразования $\bar{\mathbf{B}}_{N \times N_s} = \beta(N, N_s) \mathbf{B}_{N \times N_s}$ обладает свойством независимости мощности сигнала от соотношения между N и N_s .

Технология NOMA

Принцип неортогонального разделения сигналов абонентов стал развитием и следствием использования алгоритмов обработки квазиортогональных сигналов, который получил название многопользовательский прием (MultiUser Detection – MUD) [42]. Квазиортогональность сигналов была следствием неидеальности каналов, устройств формирования сигналов или использования квазиортогональных последовательностей в системах с кодовым разделением, т.е. последовательностей с невысоким ненулевым уровнем корреляции. Многопользовательский приемник рассматривает все сигналы, приходящие от разных абонентов, как полезные сигналы и демодулирует всю совокупность информационных символов. При этом оказалось возможным достичь характеристик однопользовательской системы даже в условиях межканальных помех.

В системах связи с неортогональным множественным доступом (NOMA – Non-Orthogonal Multiple Access) [37, 38, 40, 48, 53] неортогональность сигналов вводится сознательно для повышения эффективности использования канальных ресурсов.

Технология NOMA имеет множество разновидностей, которые можно разделить на две основные группы [38, 41]: NOMA с разделением по мощности (Power Domain NOMA – PD-NOMA), NOMA с разделением по структуре сигнала. Под разделением по структуре сигнала подразумевается любое различие сигналов пользователей в параметрах сигнала, структуре или параметрах отдельных блоков, используемых при формировании сигнала и т.п. В данной статье не будем останавливаться на описании различных разновидностей NOMA, а ограничимся лишь исследованием NOMA с кодовым разделением (Code Domain NOMA – CD-NOMA). Этот класс NOMA является достаточно широким и к нему относятся такие разновидности, как разделение по шаблону (Pattern Domain Multiple Access – PDMA) [40], разделение с использованием низкоплотных кодовых последовательностей (LDS – Low Density Signatures) [55], разделение с использованием последовательностей Welch (Welch Sequences Multiple Access – WSMA) и многих других [42].

Все разновидности CD-NOMA объединяются одним общим признаком: использование неортогональных сигналов переносчиков для объединения индивидуальных сигналов пользователей. В этом случае, число объединяемых

пользователей больше числа выделенных ортогональных ресурсов.

Модель наблюдения для системы с CD-NOMA зависит от типа линии (восходящая или нисходящая). Для нисходящей линии в общем случае наблюдение, принимаемое k -м пользователем описывается следующим векторно-матричным уравнением:

$$\mathbf{y}^{(k)} = \mathbf{H}^{(k)} \mathbf{B}_{N \times K} \boldsymbol{\theta} + \boldsymbol{\eta}^{(k)}, \quad k = \overline{1, K}, \quad (21)$$

где $\mathbf{y}^{(k)}$ – N -мерный вектор наблюдений для k -го пользователя, N – число ортогональных ресурсов, K – число пользователей, $\boldsymbol{\theta}$ – N_s -мерный вектор модулированных QAM-символов всех пользователей, $\boldsymbol{\eta}^{(k)}$ – вектор комплексных отсчетов гауссовского шума с нулевым математическим ожиданием и корреляционной матрицей \mathbf{R}_η , $\mathbf{H}^{(k)}$ – $(N \times N)$ -мерная диагональная матрица канала нисходящей линии для k -го пользователя, $\mathbf{B}_{N \times K}$ – $(N \times K)$ -мерная кодовая матрица, которая в данном случае является формирующей матрицей, определяемая видом описываемой разновидности NOMA. Для обеспечения постоянства мощности излучаемого сигнала при изменении числа пользователей эта матрица должна удовлетворять следующим условиям

$$P_s = \frac{1}{N} \text{trace}(\mathbf{B}_{N \times K} \mathbf{B}'_{N \times K}). \quad (22)$$

Следует отметить, что выражение (21) описывает разные варианты системы передачи информации. Например, если $K = N$ и $\mathbf{B}_{N \times N}$ – ортogonalная или единичная матрица, то мы имеем обычную систему передачи информации по ортогональным каналам. Если $K > N$, то имеем «перегруженную» систему, а при $K < N$ имеем «недогруженную» систему. Предельным случаем «недогруженной» системы является случай, когда $K = 1$, т.е. один символ распределяется по N ортогональным ресурсам. Это эквивалентно использованию широкополосной модуляции с прямым расширением спектра (DSSS).

Для восходящей линии модель будет другой. Отличия в основном определяются тем, что множители канала для сигналов разных пользователей будут различаться. В общем случае она может быть описана следующим уравнением:

$$\mathbf{y} = \mathbf{H}_\Sigma \odot \mathbf{B}_{N \times K} \boldsymbol{\theta} + \boldsymbol{\eta}, \quad (23)$$

где: $\mathbf{H}_\Sigma \triangleq [\mathbf{h}_1 \dots \mathbf{h}_k \dots \mathbf{h}_K]$ – матрица группового канала, k -й столбец которой содержит канальные множители ресурсов k -го пользователя \mathbf{h}_k , оператор \odot обозначает поэлементное умножение матриц.

Между моделями нисходящей и восходящей линий имеется ещё одно отличие. Так, например, для нисходящей линии мощность суммарного излучаемого сигнала, определяемая матрицей $\mathbf{B}_{N \times K}$, фиксирована и не зависит от числа пользователей. Для восходящей модели, физически, нет такого условия. В ней с появлением каждого нового пользователя суммарная мощность принимаемого сигнала увеличивается. Однако, в данной статье для последующего адекватного сравнения систем NOMA с разной степенью загрузки, а также для

сравнения с системами с ортогональным доступом, будем использовать условие постоянства средней мощности суммарного принимаемого сигнала при любом числе пользователей. Другими словами говоря, и в нисходящей линии, и в восходящей будем считать, что условие (22) выполняется.

Таким образом, для системы NOMA, также как и для системы DFTN, модель описывается следующим уравнением:

$$\mathbf{y} = \mathbf{W} \boldsymbol{\theta} + \boldsymbol{\eta}, \quad (24)$$

где \mathbf{W} – $(N \times K)$ -мерная эквивалентная матрица преобразования, $\mathbf{W} \triangleq \mathbf{H} \odot \mathbf{B}_{N \times K}$ – для восходящей системы, и $\mathbf{W} = \mathbf{W}^{(k)} \triangleq \mathbf{H}^{(k)} \mathbf{B}_{N \times K}$ – для нисходящей системы.

В качестве примера рассмотрим модели для нескольких разновидностей NOMA.

Для системы с разновидностью PDMA используется матрица-шаблон, определяющая правило распределения символов пользователей по ресурсам [41]. Так, для конфигурации 6 пользователей на 4 ресурса (4×6) имеем:

$$\mathbf{B}_{4 \times 6} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 \end{bmatrix}. \quad (25)$$

Здесь единичный элемент определяет номер ресурса (положение в столбце) и номер пользователя в этом ресурсе (положение в строке). Сами по себе символы передаются напрямую в соответствующий ресурс без какого-либо преобразования.

Для системы NOMA разновидности LDS кодовая матрица строится аналогично матрице-шаблону в PDMA, но только вместо единичных элементов используются комплексные множители кодовых последовательностей с низкой плотностью.

Для WSMA кодовая матрица является заполненной матрицей, где в качестве столбцов используются кодовые последовательности, обладающие особыми свойствами (с границей Welch), например, [41, 42]:

$$\mathbf{B}_{4 \times 6} = \begin{bmatrix} -0.2221 + 0.322i & 0.1709 - 0.3679i & 0.4335 - 0.4253i & -0.2877 + 0.4804i \\ -0.4866 + 0.309i & -0.4148 - 0.2589i & 0.5181 + 0.0067i & -0.3439 - 0.2048i \\ 0.0525 - 0.6492i & 0.2786 + 0.2173i & 0.4058 - 0.3688i & -0.0586 - 0.3831i \\ 0.1887 - 0.5138i & -0.5603 + 0.0403i & 0.3714 - 0.066i & 0.3124 + 0.3807i \\ -0.4067 - 0.0166i & 0.5821 - 0.2559i & 0.1316 - 0.231i & 0.5222 - 0.2944i \\ 0.316 + 0.0753i & -0.7029 - 0.1267i & 0.354 - 0.2274i & -0.349 - 0.2925i \end{bmatrix}. \quad (26)$$

Следует отметить, что модель системы связи с разновидностью NOMA вида PD-NOMA, также может быть описана общей моделью (24). Так, например, для системы с двумя пользователями в одном ресурсе для модели с 4-мя ресурсами можно записать эквивалентную матрицу:

$$\mathbf{W} = \begin{bmatrix} \gamma_1 h_{11} & \gamma_2 h_{12} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \gamma_3 h_{23} & \gamma_4 h_{24} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \gamma_5 h_{35} & \gamma_6 h_{36} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \gamma_7 h_{47} & \gamma_8 h_{48} \end{bmatrix}, \quad (27)$$

где γ_k – весовой множитель k -го пользователя, вводимый для организации разделения по мощности, h_{ik} – комплексный множитель канала для сигнала k -го пользователя в i -м ортогональном ресурсе. Здесь, для наглядности, рассмотрены 4 ресурса и 8 пользователей (4×8), но при этом в каждом ресурсе наблюдаются только сигналы двух пользователей, и каждый пользователь использует только один ресурс.

Теоретико-информационный аспект технологий дискретного FTN и NOMA

Технология FTN, а также технология NOMA, в какой-то степени являются сенсационными, так как нарушают некоторые привычные основы. FTN позволяет передавать информацию со скоростью больше скорости Найквиста, а NOMA позволяет объединить сигналы абонентов, число которых больше числа ортогональных каналов. Существует устойчивое мнение на счёт обоснования эффективности технологии FTN, основанное на наличии контролируемой межсимвольной интерференции [14, 32]. Аналогично, эффективность NOMA объясняется неортогональностью используемых сигналов, что приводит также к контролируемым межканальным помехам. Рассмотрим, так ли это на самом деле. Достаточно ли введения контролируемых помех для повышения эффективности с точки зрения теории связи. Или необходимы какие-то дополнительные условия.

Рассмотрим пропускную способность системы передачи информации со скоростью Найквиста и пропускную способность системы с FTN в классическом (Шенноновском) понимании. Для этого воспользуемся векторно-матричной моделью (20) и рассчитаем пропускную способность векторного канала [67]. Для исключения влияния канала на результаты ограничимся рассмотрением канала АБГШ с постоянными одинаковыми канальными множителями для всех отсчётов. Для этой модели пропускная способность векторного канала будет определяться выражением:

$$C_N = \log_2 \det \left(\mathbf{I}_{N_s} + \frac{N}{N_s} \mathbf{B}'_{N \times N_s} \mathbf{R}_\eta^{-1} \mathbf{B}_{N \times N_s} \right) = \log_2 \det \left(\mathbf{I}_{N_s} + \frac{N}{N_s} \rho \mathbf{B}'_{N \times N_s} \mathbf{B}_{N \times N_s} \right), \quad (28)$$

где C_N – пропускная способность (бит/вектор) N -мерного векторного канала, в котором элементами являются временные отсчёты наблюдения, ρ – отношение сигнал/шум на выходе СФ в одном временном отсчёте. В выражении (28) учитывается, что отсчёты наблюдения берутся после согласованной фильтрации с периодом $T = 1 / F_N$, поэтому отсчёты шума будут некоррелированными. Чтобы исключить влияние выбора размерности вектора на пропускную способность рассмотрим нормированную пропускную способность на один отсчёт

$$C = \frac{1}{N} \log_2 \det \left(\mathbf{I}_{N_s} + \frac{N}{N_s} \rho \mathbf{B}'_{N \times N_s} \mathbf{B}_{N \times N_s} \right), \quad (29)$$

Нетрудно показать, что для системы Найквиста ($N = N_s$, $\mathbf{B}_{N \times N_s} = \mathbf{I}_N$) имеем:

$$C = \frac{1}{N} \log_2 \det (\mathbf{I}_N + \rho \mathbf{I}_N) = \log_2 (1 + \rho). \quad (30)$$

Это полностью согласуется с теорией Шеннона о пропускной способности гауссовского канала.

Рассмотрим зависимости пропускной способности векторного канала (29) при разных значениях периода следования импульсов при фиксированной полосе. На рис. 6 приведены зависимости пропускной способности одного на один отсчёт от ОСШ при разных значениях периода следования импульсов. Число отсчётов, взятых с частотой Найквиста $N = 24$, параметр скругления фильтра с приподнятым косинусом $\alpha = 0.3$.

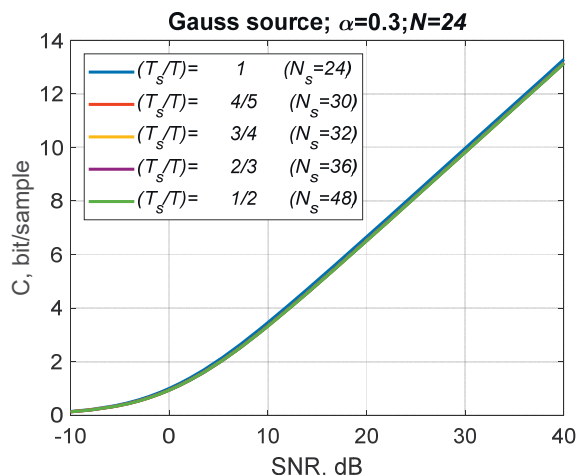


Рис. 6. Зависимости пропускной способности от ОСШ при разных значениях периода следования импульсов

Как следует из этого рисунка, для систем передачи информации с гауссовскими источниками увеличение частоты следования отсчётов не приводит к увеличению пропускной способности по Шеннону.

Рассмотрим теперь пропускную способность при передаче цифровой информации по непрерывному каналу связи, т.е. в качестве сигналов, передаваемых по каналу связи рассмотрим символы с цифровой модуляцией QAM. В данном случае $\theta_n \in \mathbb{C}^M$ – это комплексный символ, принимающий одно из $M = 2^m$ значений, где m – число бит, передаваемых в одном отсчёте сигнала, $n = \overline{1, N_s}$. Этот случай соответствует пропускной способности дискретно-непрерывного канала [61-64, 68, 69, 71]. Здесь под дискретностью понимается дискретный характер значений сигнала. Для этого случая взаимная информация между цифровым источником J с числом состояний 2^{mN_s} , характеризующая пропускную способность, будет определяться следующим выражением [69]:

$$I_{J,Y} = -N - \mathfrak{I}_{J,Y}, \quad (31)$$

где введено следующее обозначение

$$\mathfrak{J}_{J;Y} \triangleq \frac{1}{2^{mN_s}} \sum_{J=0}^{2^{mN_s}-1} \left(\int_y \frac{1}{(2\pi\sigma_\eta^2)^N} \exp\left(-\frac{\|y - \mathbf{W}\theta^{(J)}\|^2}{2\sigma_\eta^2}\right) \times \right. \\ \left. \times \log\left(\frac{1}{2^{mN_s}} \sum_{I=0}^{2^{mN_s}-1} \exp\left(-\frac{\|y - \mathbf{W}\theta^{(I)}\|^2}{2\sigma_\eta^2}\right)\right) dy \right), \quad (32)$$

где $\theta^{(J)}$ – J -ая комбинация N_s -мерного вектора $\theta = [\theta_1 \dots \theta_{N_s}]$.

Для практического использования выражений (31)-(32) можно использовать метод Монте-Карло. для этого его можно преобразовать и записать в следующем виде:

$$\bar{I}_{J;Y} \approx -N - \frac{1}{N_{MC}} \sum_{n=1}^{N_{MC}} \log\left(\frac{1}{2^{mN_s}} \sum_{J=0}^{2^{mN_s}-1} \exp\left(-\frac{\|z_n - \mathbf{W}\Delta\theta^{(I,J)}\|^2}{2\sigma_\eta^2}\right)\right) \quad (33)$$

где N_{MC} – объём статистики метода Монте-Карло, I_n – значение дискретной случайной величины $I \in \{0, 1, 2, \dots, 2^{mN_s} - 1\}$ на n -й реализации вычислений методом Монте-Карло.

Рассмотрим зависимости пропускной способности дискретно-непрерывного канала, приведённой к одному информационному отсчёту $C = \frac{1}{N} \bar{I}_{J;Y}$, от отношения сигнал/шум

при разных соотношениях (T_s / T) при использовании огибающей сигнала в виде приподнятого косинуса (4) с параметром $\alpha = 0.3$. Эти зависимости приведены на рис. 7. Здесь хорошо видно, что в отличие от непрерывного гауссовского источника, при передаче цифровой информации увеличение частоты следования символов (уменьшение соотношения (T_s / T)) приводит к увеличению пропускной способности,

а точнее к реальной скорости передачи информации. На рис. 8 приведены зависимости пропускной способности системы FTN для гауссовского источника и цифрового источника от соотношения $(F_{FTN} / F_N) = 1 / (T_s / T)$ при отношении сигнал/шум 10 дБ. Эти зависимости показывают, что для гауссовского источника увеличение частоты следования информационных отсчётов не увеличивает пропускную способность канала, в то время как для цифрового источника наблюдается рост пропускной способности. Однако, значение пропускной способности при цифровой передаче не превышает пропускную способность гауссовского источника, оно лишь асимптотически стремится к ней.

Эти результаты показывают, что эффект от превышения частоты Найквиста наблюдается только в системах, которые на самом деле неэффективно используют пропускную способность канала. Причём, чем хуже используется канал, тем больше повышается эффективность от использования системы FTN. Таким образом, введение контролируемой межсимвольной интерференции в системах связи с технологией FTN является лишь инструментом для повышения эффективности изначально неэффективных систем.

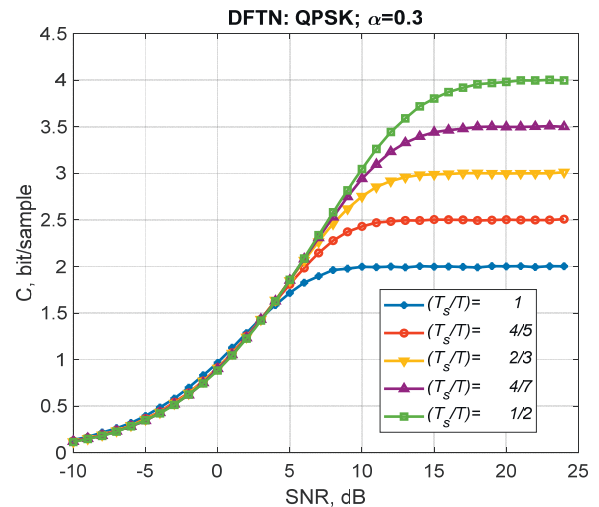


Рис. 7. Зависимости пропускной способности FTN системы при передаче цифровой информации от отношения сигнал/шум при разных соотношениях (T_s / T) .

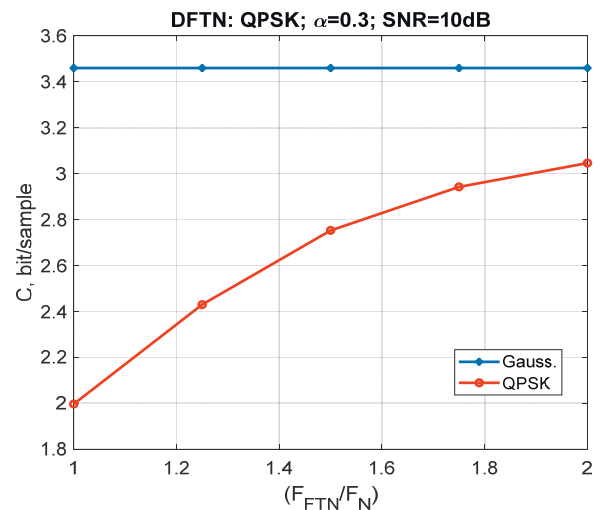


Рис. 8. Зависимости пропускной способности системы FTN для гауссовского источника и цифрового источника от соотношения (F_{FTN} / F_N) при отношении сигнал/шум 10 дБ

Поведём теперь аналогичное исследование пропускной способности системы связи с технологией NOMA. В качестве разновидности NOMA рассмотрим WSMA с числом ресурсов 4 и разным числом пользователей [43]. Поскольку, как было показано в разделе 2, для системы NOMA может использоваться такая же модель, как и для дискретной FTN, то для анализа пропускной способности с гауссовским источником и цифровым в системе NOMA используем те же выражения (28) и (33), соответственно.

На рис. 9 приведены зависимости пропускной способности для гауссовского канала и цифрового канала с модуляцией QPSK для системы связи с NOMA от отношения сигнал/шум при разном числе пользователей. В отличие от аналогичных зависимостей для системы с FTN здесь пропускная способность вычисляется для векторного канала, состоящего из $N_{res} = 4$, т.е. суммарная пропускная способность векторного канала, а не для одного отсчёта.

На рис. 10 приведены зависимости пропускной способности системы NOMA для гауссовского источника и цифрового источника от, так называемого, параметра перегрузки $Overload\ factor = N_{us} / N_{res} \cdot 100\%$ при отношении сигнал/шум 10 дБ, где N_{us} – число пользователей.

Как видно из приведённых рисунков, для гауссовского источника пропускная способность не увеличивается при использовании неортогональных методов множественного доступа (при условии, что суммарная мощность сигнала остается постоянной). В то время, как для систем передачи цифровой информации, использующих модуляцию QPSK, применение неортогональных методов увеличивает суммарную пропускную способность, которая в данном случае выражается в числе пользователей. Однако, при этом суммарная пропускная способность цифровой системы связи с NOMA остается меньше пропускной способности по Шеннону.

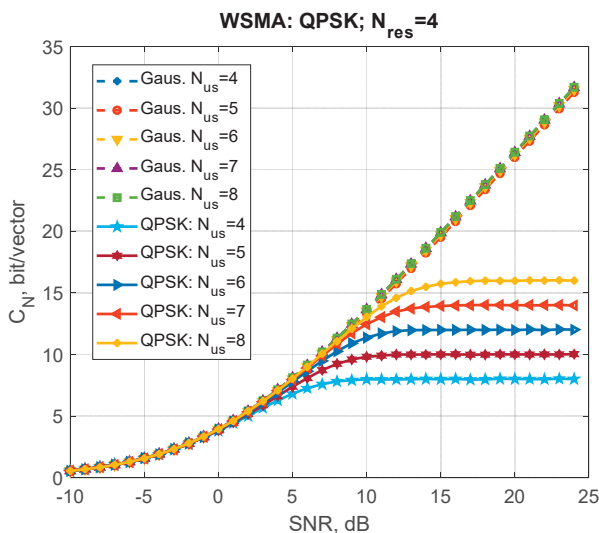


Рис. 9. Зависимости пропускной способности для гауссовского канала и цифрового канала с модуляцией QPSK для системы связи с NOMA от отношения сигнал/шум при разном числе пользователей

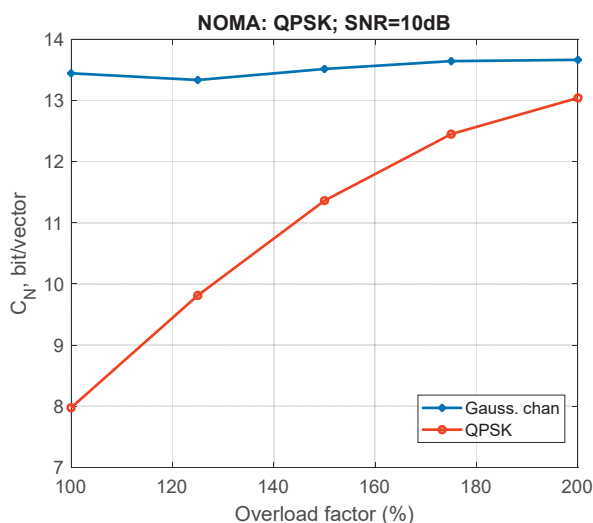


Рис. 10. Зависимости пропускной способности системы NOMA для гауссовского источника и цифрового источника от параметра перегрузки при отношении сигнал/шум 10 дБ

Эти результаты показывают, что использование контролируемой помехи в системах FTN и NOMA позволяет повысить спектральную эффективность и пропускную способность, но только в том случае, когда в системе связи, изначально, свойства канала используются не очень эффективно. Этим, в основном, объясняется, что эти технологии чаще всего используются при модуляциях невысокого порядка и с небольшим числом пользователей.

Следует также обратить внимание на то, что дискретные сигналы в системах с FTN и NOMA имеют больше значений (точек по уровню) чем исходные QAM-сигналы, что приближает их к непрерывным сигналам, и к тому же они становятся практически независимыми. Другими словами говоря, преобразованные сигналы в FTN и NOMA системах приближаются к гауссовским, что приводит к приближению к Шенноновской пропускной способности.

Данные выводы не противоречат и не подвергают сомнению, анализ эффективности систем связи с FTN и NOMA, опубликованные в разных источниках [9, 22, 35, 36, 49, 50], а лишь смещают акцент на истинную природу повышения эффективности и дают направления в поиске критериев для оптимизации, например, использование критериев, учитывающих дискретный характер цифровых сигналов [71-73].

Заключение

Таким образом, проведённый анализ пропускной способности системы связи с новыми технологиями FTN и NOMA показал, что эти технологии, эксплуатирующие введение контролируемых помех путем нарушения ортогональности, не являются «волшебными» технологиями, делающие лучшее лучшим. Они лишь демонстрируют тот факт, что современные системы недостаточно эффективно используют каналы связи, и дают инструмент для улучшения систем связи. Другими словами говоря, они позволяют «догрузить» «недогруженные» системы связи. «Недогруженность» каналов связи возникает из-за использования цифровых методов модуляции низкого порядка, поэтому наибольший эффект от использования FTN и NOMA возникает, именно, в системах с модуляцией типа BPSK и QPSK.

Эффект от использования этих технологий основан на том же, что и применение модуляций высокого порядка, а также многомерных созвездий. Поэтому для оптимизации сигналов систем связи с FTN и NOMA необходимо использовать критерии, учитывающие дискретность значений.

Литература

1. Lucky R. W. Decision feedback and faster-than-Nyquist transmission // Proc. Int. Symp. Inf. Theory (ISIT), Noordwijk, Netherlands, Jun. 1970, p. 7.6.
2. Salz J. Optimum mean-square decision feedback equalization // Bell Syst. Tech. J., vol. 52, pp. 1341-1373, Oct. 1973.
3. Foschini G. J. Contrasting performance of faster binary signaling with QAM // AT&T Bell Lab. Tech. J., vol. 63, pp. 1419-1445, Oct. 1984.
4. Mazo J. E. Faster-than-Nyquist signalling // Bell Syst. Tech. J., vol. 54, pp. 1451-1462, Oct. 1975.
5. Gallager R. A perspective on multiaccess channels // IEEE Trans. Inf. Theory, vol. 31, no. 2, pp. 124-142, Mar. 1985.
6. Gallager R. G. An inequality on the capacity region of multiaccess multipath channels // Communications and Cryptography, R. E. Blahut et

al., Eds. Boston, MA, USA: Springer, 1994, pp. 143-153.

7. *Liveris A. D., Georghiades C. N.* Exploiting faster-than-Nyquist signaling // *IEEE Trans. Commun.*, vol. 51, pp. 1502-1511, Sept. 2003.

8. *Rusek F., Anderson J. B.* Non binary and precoded faster than Nyquist signaling // *IEEE Trans. Commun.*, vol. 56, pp. 808-817, May 2008.

9. *Rusek F., Anderson J. B.* Constrained capacities for faster-than-Nyquist signaling // *IEEE Trans. Inf. Theory*, vol. 55, no. 2, pp. 764-775, Feb. 2009.

10. *Anderson J. B., Rusek F., Öwall V.* Faster-than-Nyquist signaling // *Proc. IEEE*, vol. 101, no. 8, pp. 1817-1830, Aug. 2013.

11. *Fan J., Guo S., Zhou X., Ren Y., Li G. Y., Chen X.* Faster-than-Nyquist signaling: An overview // *IEEE Access*, vol. 5, pp. 1925-1940, 2017.

12. *Hajela D.* On faster than Nyquist signaling: Further estimations on the minimum distance // *SIAM J. Appl. Math.*, vol. 52, no. 3, pp. 900-907, 1992.

13. *Wang Y., Li Q., Li Y., Li L., X. Li.* Linear precoding for faster-than-Nyquist signaling with imperfect channel state information // *Proc. 16th Int. Conf. Wireless Commun. Signal Process. (WCSP)*, Hefei, China, 2024, pp. 962-967.

14. *Belkova S., Cu N. D., Lavrenyuk I.* Joint application of signals with controlled ISI and iterative decision-feedback detection algorithm in DVB-S2 // *Proc. Int. Conf. Electr. Eng. Photonics (EEExPolytech)*, St. Petersburg, Russia, 2024, pp. 134-137.

15. *Matsuyama T., Suzuki T., Saito S., Suganuma H., Maehara F.* Throughput improvement through FTN-based MIMO signaling in fixed wireless access // *Proc. 11th Int. Japan-Africa Conf. Electron., Commun. Comput. (JAC-ECC)*, Alexandria, Egypt, 2023, pp. 79-82.

16. *Fraisse J.-B., Chevalier P., Gautier R., Delaveau F., Traverso S.* Power and second order cyclic covertness of chip-wise direct sequence spread spectrum faster-than-Nyquist signaling // *MILCOM 2023*, Boston, MA, USA, 2023, pp. 522-527.

17. *Hamdar F., Nadal J., Nour C. A., Baghdadi A.* Novel transmission technique based on intentional overlapping for spectral efficiency enhancement in multicarrier systems // *Proc. IEEE PIMRC*, Toronto, ON, Canada, 2023, pp. 1-6.

18. *Hama Y., Ochiai H.* On the achievable spectral efficiency of non-orthogonal frequency division multiplexing // *IEEE Trans. Commun.*, vol. 71, no. 11, pp. 6246-6257, Nov. 2023.

19. *Chiu H.-T. et al.*, Performance comparisons of faster-than-Nyquist signaling considering fixed utilized bandwidth // *Proc. 10th Int. Japan-Africa Conf. Electron., Commun. Comput. (JAC-ECC)*, Alexandria, Egypt, 2022, pp. 109-112.

20. *Zavjalov S., Prokhorov V., Chudnov A.* Comparison of spectral efficiency of FTN signaling with coherent and noncoherent detection // *Proc. Int. Conf. Electr. Eng. Photonics (EEExPolytech)*, St. Petersburg, Russia, 2022, pp. 166-169.

21. *Gunturu A., Sahoo A. K., Chavva A. K. R.* Faster than Nyquist waveform for beyond 5G systems – Evaluation and implementation aspects // *ICC 2022*, Seoul, Korea, 2022, pp. 1859-1864.

22. *Han G., Shamai S.* On sampling continuous-time AWGN channels // *IEEE Trans. Inf. Theory*, vol. 68, no. 2, pp. 782-794, Feb. 2022.

23. *Peleg M., Michaeli T., Shamai S. (Shitz).* On information rates over a binary-input filtered Gaussian channel // *Proc. IEEE COMCAS*, Tel Aviv, Israel, 2021, pp. 330-333.

24. *Baek C.-U., Jung J.-W.* Faster-than-Nyquist signal processing based on unequal error probability for high-throughput wireless communications // *Appl. Sci.*, vol. 9, no. 12, p. 2413, 2019.

25. *Ochiai H.* Peak power properties of band-limited signals: With pulse shaping or windowing // *IEEE Trans. Veh. Technol.*, vol. 73, no. 10, pp. 15200-15213, Oct. 2024.

26. *El Hefnawy M., Taoka H.* Overview of faster-than-Nyquist for future mobile communication systems // *Proc. IEEE VTC Spring*, Dresden, Germany, 2013, pp. 1-5.

27. *Cerci E., Cicek A., Cavus E., Bedeer E., Yanikomeroglu H.* Coded faster-than-Nyquist signaling for short packet communications // *Proc.*

IEEE PIMRC, Helsinki, Finland, 2021, pp. 428-433.

28. *Bedeer E., Yanikomeroglu H.* A novel low complexity faster-than-Nyquist (FTN) signaling detector for ultra high-order QAM // *IEEE Open J. Commun. Soc.*, vol. 2, pp. 2566-2580, 2021.

29. *Che D., Chen X.* Higher-order modulation vs faster-than-Nyquist PAM-4 for datacenter IM-DD optics: An AIR comparison under practical bandwidth limits // *J. Lightw. Technol.*, vol. 40, no. 10, pp. 3347-3357, May 2022.

30. *McGuire M., Sima M.* Discrete time faster-than-Nyquist signaling // *Proc. IEEE GLOBECOM*, Miami, FL, USA, 2010.

31. *Si C., Li S., Liu M., Bai B., Wang J.* An iterative receiver for coded faster-than-Nyquist signaling over Rayleigh fading channels // *Proc. IEEE/CIC ICC Workshops*, Foshan, China, 2022, pp. 472-477.

32. *Jaffal Y., Alvarado A.* Pulses with minimum residual intersymbol interference for faster than Nyquist signaling // *IEEE Commun. Lett.*, vol. 26, no. 11, pp. 2670-2674, Nov. 2022.

33. *Zhang Z., Yuksel M., Yanikomeroglu H.* Faster-than-Nyquist signaling for MIMO communications // *IEEE Trans. Wireless Commun.*, vol. 22, no. 4, pp. 2379-2392, Apr. 2023.

34. *Tong M., Huang X., Zhang J. A.* Faster-than-Nyquist transmission with frame-by-frame decision-directed successive interference cancellation // *IEEE Trans. Commun.*, vol. 71, no. 8, pp. 4851-4861, Aug. 2023.

35. *Yoo Y. G., Cho J. H.* Asymptotic optimality of binary faster-than-Nyquist signaling // *IEEE Commun. Lett.*, vol. 14, no. 9, pp. 788-790, Sep. 2010.

36. *Ganji M., Zou X., Jafarkhani H.* On the capacity of faster than Nyquist signaling // *IEEE Commun. Lett.*, vol. 24, no. 6, pp. 1197-1201, Jun. 2020.

37. *Yuan Y., Yuan Z., Tian L.* 5G non-orthogonal multiple access study in 3GPP // *IEEE Commun. Mag.*, vol. 58, no. 7, pp. 90-96, Jul. 2020.

38. *Al-Dulaimi A. M. K., Al-Dulaimi O. M. K., Al-Dulaimi M. K. H.* Implementing 3GPP code division for non-orthogonal multiple access in 5G // *Proc. ISMSIT*, Ankara, Turkey, 2022, pp. 998-1006.

39. *Бакулин М. Г., Бен Режеб Т. Б. К., Крейнделин В. Б., Панкратов Д. Ю., Смирнов А. Э.* Технология NOMA с кодовым разделением в 3GPP: 5G или 6G? // *T-Comm: Телекоммуникации и транспорт*. 2022. Т. 16, № 1. С. 4-14.

40. *Nikopour H., Baligh H.* Sparse code multiple access // *Proc. IEEE PIMRC*, 2013, pp. 332-336.

41. *Chen S., Ren B., Gao Q., Kang S., Sun S., Niu K.* Pattern division multiple access – A novel nonorthogonal multiple access for fifth-generation radio networks // *IEEE Trans. Veh. Technol.*, vol. 66, no. 4, pp. 3185-3196, 2017.

42. *Li X., Chen H.-H., Qian Y., Rong B., Soleymani M. R.* Welch bound analysis on generic code division multiple access codes with interference free windows // *IEEE Trans. Wireless Commun.*, vol. 8, no. 4, pp. 1603-1607, Apr. 2009.

43. "Signature design for NOMA," 3GPP TSG RAN WG1 Meeting #92, Athens, Greece, Feb. 2018, Doc. R1-1802767.

44. *Бакулин М. Г., Бен Режеб Т. Б. К., Крейнделин В. Б., Панкратов Д. Ю., Смирнов А. Э.* Неортогональный множественный доступ (NOMA) как основа систем связи 5G и 6G. М.: Горячая линия – Телеком, 2024. 264 с.

45. *Luo F.-L. et al.*, *Signal Processing for 5G: Algorithms and Implementations*. Hoboken, NJ, USA: Wiley-IEEE Press, 2016.

46. *Verdú S.* Optimum multiuser asymptotic efficiency // *IEEE Trans. Commun.*, vol. 34, no. 9, pp. 890-897, Sep. 1986.

47. *Бакулин М. Г., Крейнделин В. Б., Шумов А. П.* Вопросы повышения спектральной эффективности систем связи: неортогональная передача // *Цифровая обработка сигналов*. 2013. № 4. С. 55-64.

48. *Lee H., Kim S., Lim J.* Multiuser superposition transmission (MUST) for LTE-A systems // *Proc. IEEE ICC*, 2016, pp. 1-6.

49. *Islam S. M. R., Zeng M., Dobre O. A.* NOMA in 5G systems: Exciting possibilities for enhancing spectral efficiency // *IEEE 5G Tech Focus*, vol. 1, no. 2, Jun. 2017.

50. Wang Y., Ren B., Sun S., Kang S., Yue X. Analysis of non-orthogonal multiple access for 5G // *China Commun.*, vol. 13, no. 2, pp. 52-66, 2016.
51. Shin W., Vaezi M., Lee B., Love D. J., Lee J., Poor H. V. Non-orthogonal multiple access in multi-cell networks: Theory, performance, and practical challenges // *IEEE Commun. Mag.*, vol. 55, no. 10, pp. 176-183, Oct. 2017.
52. Benjebbour A., Saito K., Li A., Kishiyama Y., Nakamura T. Non-orthogonal multiple access (NOMA): Concept, performance evaluation and experimental trials // *Proc. WINCOM*, 2015.
53. Hoshyar R., Wathan F. P., Tafazolli R. Novel low-density signature for synchronous CDMA systems over AWGN channel // *IEEE Trans. Signal Process.*, vol. 56, no. 4, pp. 1616-1626, 2008.
54. Chennakeshu S., Saulnier G. J. Differential detection of $\pi/4$ -shifted-DQPSK for digital cellular radio // *IEEE Trans. Veh. Technol.*, vol. 42, no. 1, pp. 46-57, Feb. 1993.
55. Yan C., Harada A., Benjebbour A., Lan Y., Li A., Jiang H. Receiver design for downlink non-orthogonal multiple access (NOMA) // *Proc. IEEE VTC Spring*, Glasgow, UK, 2015, pp. 1-6.
56. Islam S. M. R., Avazov N., Dobre O. A., Kwak K. Power-domain non-orthogonal multiple access (NOMA) in 5G systems: Potentials and challenges // *IEEE Commun. Surveys Tuts.*, vol. 19, no. 2, pp. 721-742, Second Quart. 2017.
57. Шеннон К. Математическая теория связи // *Bell System Technical Journal*. 1948. С. 379-423, 623-656.
58. Blahut R. E. Principles and Practice of Information Theory. Reading, MA, USA: Addison-Wesley, 1987.
59. Ungerboeck G. Channel coding with multilevel/phase signals // *IEEE Trans. Inf. Theory*, vol. 28, no. 1, pp. 55-67, Jan. 1982.
60. Ng S. X., Hanzo L. On the MIMO channel capacity of multidimensional signal sets // *IEEE Trans. Veh. Technol.*, vol. 55, no. 2, pp. 528-536, 2006.
61. McIllree P. E. Channel capacity calculations for M-ary N-dimensional signal sets // M.S. thesis, School of Electr. Eng., Univ. South Australia, Adelaide, Australia, 1995.
62. Cover T. M., Thomas J. A. Elements of Information Theory, 2nd ed. Hoboken, NJ, USA: Wiley, 2006.
63. Fishman G. S. Monte Carlo: Concepts, Algorithms, and Applications. New York, NY, USA: Springer, 1995.
64. Telatar E. Capacity of multi-antenna Gaussian channels // *Eur. Trans. Telecommun.*, vol. 10, pp. 585-595, 1999.
65. Прокис Дж. Цифровая связь. М.: Радио и связь, 2000.
66. Bakulin M. G., Ben Rejeb T., Kreyndelin V., Pankratov D., Smirnov A. New approach to NOMA optimization based on individual discrete input continuous output memoryless channel capacity // *Alexandria Eng. J.*, vol. 107, pp. 215-224, 2024.
67. Bakulin M. G., Ben Rejeb T. B., Kreyndelin V. B. et al. Individual channel capacity in a communication system with nonorthogonal multiple access // *J. Commun. Technol. Electron.*, vol. 68, Suppl. 2, pp. S198-S211, 2023.
68. Barrueco J., Montalban J., Iradier E., Angueira P. Constellation design for future communication systems: A comprehensive survey // *IEEE Access*, vol. 9, pp. 89778-89797, 2021.
69. Бакулин М. Г., Крейнделин В. Б., Резнёв А. А. Оптимизация матрицы DFTN для повышения помехоустойчивости // *Радиоэлектроника. Наносистемы. Информационные технологии*. 2025. Т. 17, № 4. С. 419-430.
70. Guillén i Fàbregas A., Martínez A., Caire G. Bit-wise mutual information and error exponent for binary-input channels // *IEEE Trans. Inf. Theory*, vol. 54, no. 11, pp. 5140-5154, Nov. 2008.
71. Madhow U. Fundamentals of Digital Communication. Cambridge, U.K.: Cambridge Univ. Press, 2008.
72. Martínez A., Guillén i Fàbregas A. Spectral efficiency of optical intensity channels with bandwidth constraints // *IEEE Trans. Commun.*, vol. 60, no. 10, pp. 2991-3001, Oct. 2012.
73. Ling C., Yeung R. W. On capacity loss due to symbol quantization in discrete-input AWGN channels // *IEEE Commun. Lett.*, vol. 27, no. 5, pp. 1231-1235, May 2023.

FTN AND NOMA: BEYOND THE HYPE – TOWARD SHANNON-LIMITED EFFICIENCY

Mikhail G. Bakulin, Moscow Technical University of Communications and Informatics, Moscow, Russia, m.g.bakulin@gmail.com

Vitaly B. Kreyndelin, Moscow Technical University of Communications and Informatics, Moscow, Russia, vitkrend@gmail.com

Andrey A. Reznev, Moscow Technical University of Communications and Informatics;

JSC NTC Atlas, Moscow, Russia, andrey.reznev@gmail.com

Abstract

Faster-Than-Nyquist (FTN) and Non-Orthogonal Multiple Access (NOMA) are widely regarded as promising techniques for enhancing spectral efficiency in 5G and 6G wireless systems. Both approaches deliberately violate signal orthogonality-FTN by time-domain symbol packing and NOMA by enabling multiple users to share the same resources. The literature often describes them as "overloaded" systems that boost capacity through controlled interference. However, this view is misleading. Based on information-theoretic analysis, we demonstrate that neither FTN nor NOMA can exceed the classical Shannon capacity limit of the Gaussian channel. Their actual gain arises exclusively when transmitting discrete signals (e.g., BPSK, QAM), which inherently operate far below Shannon's bound due to their non-Gaussian nature. Consequently, the performance improvement offered by FTN and NOMA does not stem from "interference exploitation" but from compensating for the inefficiency of low-order modulation schemes-essentially "loading" initially underloaded channels. This effect is most pronounced in systems using low-order constellations, such as IoT or energy-constrained devices. Recognizing this mechanism shifts the optimization paradigm: rather than merely mitigating interference, system design should focus on maximizing channel utilization by leveraging the discrete structure of digital signals. The conclusions are supported by mutual information analysis of discrete-input continuous-output channels and corroborated through Monte Carlo-based simulations. Ultimately, FTN and NOMA are not "magic" breakthroughs beyond fundamental limits-they are practical tools to approach Shannon efficiency in real-world digital communication systems.

Keywords: spectral efficiency, FTN, MIMO, NOMA, spatial multiplexing, 5G, simulation modeling

References

- [1] R. W. Lucky, "Decision feedback and faster-than-Nyquist transmission," in *Proc. Int. Symp. Inf. Theory (ISIT)*, Noordwijk, Netherlands, Jun. 1970, p. 7.6.
- [2] J. Salz, "Optimum mean-square decision feedback equalization," *Bell System Technical Journal*, vol. 52, pp. 1341-1373, Oct. 1973.
- [3] G. J. Foschini, "Contrasting performance of faster binary signaling with QAM," *AT&T Bell Laboratories Technical Journal*, vol. 63, pp. 1419-1445, Oct. 1984.
- [4] J. E. Mazo, "Faster-than-Nyquist signalling," *Bell System Technical Journal*, vol. 54, pp. 1451-1462, Oct. 1975.
- [5] R. Gallager, "A perspective on multiaccess channels," *IEEE Transactions on Information Theory*, vol. 31, no. 2, pp. 124-142, Mar. 1985, doi: 10.1109/TIT.1985.1057022.
- [6] R. G. Gallager, "An inequality on the capacity region of multiaccess multipath channels," in *Communications and Cryptography*, R. E. Blahut, D. J. Costello, U. Maurer, and T. Mittelholzer, Eds. Boston, MA, USA: Springer, 1994, pp. 143-153, doi: 10.1007/978-1-4615-2694-0_14.
- [7] A. D. Liveris and C. N. Georghiades, "Exploiting faster-than-Nyquist signaling," *IEEE Transactions on Communications*, vol. 51, no. 9, pp. 1502-1511, Sep. 2003, doi: 10.1109/TCOMM.2003.816963.
- [8] F. Rusek and J. B. Anderson, "Non binary and precoded faster than Nyquist signaling," *IEEE Transactions on Communications*, vol. 56, no. 5, pp. 808-817, May 2008, doi: 10.1109/TCOMM.2008.060457.
- [9] F. Rusek and J. B. Anderson, "Constrained capacities for faster-than-Nyquist signaling," *IEEE Transactions on Information Theory*, vol. 55, no. 2, pp. 764-775, Feb. 2009, doi: 10.1109/TIT.2008.2009832.
- [10] J. B. Anderson, F. Rusek, and V. Z. wall, "Faster-than-Nyquist signaling," *Proceedings of the IEEE*, vol. 101, no. 8, pp. 1817-1830, Aug. 2013, doi: 10.1109/PROC.2012.2233451.
- [11] J. Fan, S. Guo, X. Zhou, Y. Ren, G. Y. Li, and X. Chen, "Faster-than-Nyquist signaling: An overview," *IEEE Access*, vol. 5, pp. 1925-1940, 2017, doi: 10.1109/ACCESS.2017.2657599.
- [12] D. Hajela, "On faster than Nyquist signaling: Further estimations on the minimum distance," *SIAM Journal on Applied Mathematics*, vol. 52, no. 3, pp. 900-907, 1992, doi: 10.1137/0152052.
- [13] Y. Wang, Q. Li, Y. Li, L. Li, and X. Li, "Linear precoding for faster-than-Nyquist signaling with imperfect channel state information," in *Proc. 16th Int. Conf. Wireless Commun. Signal Process. (WCSP)*, Hefei, China, 2024, pp. 962-967, doi: 10.1109/WCSP62071.2024.10827586.
- [14] S. Belkova, N. D. Cu, and I. Lavrenyuk, "Joint application of signals with controlled ISI and iterative decision-feedback detection algorithm in DVB-S2," in *Proc. Int. Conf. Electr. Eng. Photonics (EExPolytech)*, St. Petersburg, Russia, 2024, pp. 134-137, doi: 10.1109/EEXPOLYTECH62224.2024.10755922.
- [15] T. Matsuyama, T. Suzuki, S. Saito, H. Saganuma, and F. Maehara, "Throughput improvement through FTN-based MIMO signaling in fixed wireless access," in *Proc. 11th Int. Japan-Africa Conf. Electron., Commun. Comput. (JAC-ECC)*, Alexandria, Egypt, 2023, pp. 79-82, doi: 10.1109/JAC-ECC61002.2023.10479615.
- [16] J.-B. Fraisse, P. Chevalier, R. Gautier, F. Delaveau, and S. Traverso, "Power and second order cyclic coherency of chip-wise direct sequence spread spectrum faster-than-Nyquist signaling," in *MILCOM 2023 - IEEE Military Communications Conference*, Boston, MA, USA, 2023, pp. 522-527, doi: 10.1109/MILCOM58377.2023.10356248.
- [17] F. Hamdar, J. Nadal, C. A. Nour, and A. Baghdadi, "Novel transmission technique based on intentional overlapping for spectral efficiency enhancement in multicarrier systems," in *Proc. IEEE 34th Annu. Int. Symp. Pers., Indoor Mobile Radio Commun. (PIMRC)*, Toronto, ON, Canada, 2023, pp. 1-6, doi: 10.1109/PIMRC56721.2023.10293842.
- [18] Y. Hama and H. Ochiai, "On the achievable spectral efficiency of non-orthogonal frequency division multiplexing," *IEEE Transactions on Communications*, vol. 71, no. 11, pp. 6246-6257, Nov. 2023, doi: 10.1109/TCOMM.2023.3300328.
- [19] H.-T. Chiu et al., "Performance comparisons of faster-than-Nyquist signaling considering fixed utilized bandwidth," in *Proc. 10th Int. Japan-Africa Conf. Electron., Commun. Comput. (JAC-ECC)*, Alexandria, Egypt, 2022, pp. 109-112, doi: 10.1109/JAC-ECC56395.2022.10043920.
- [20] S. Zavjalov, V. Prokhorov, and A. Chudnov, "Comparison of spectral efficiency of FTN signaling with coherent and noncoherent detection," in *Proc. Int. Conf. Electr. Eng. Photonics (EExPolytech)*, St. Petersburg, Russia, 2022, pp. 166-169, doi: 10.1109/EEXPOLYTECH56308.2022.9950789.
- [21] A. Gunturu, A. K. Sahoo, and A. K. R. Chavva, "Faster than Nyquist waveform for beyond 5G systems-Evaluation and implementation aspects," in *ICC 2022 - IEEE International Conference on Communications*, Seoul, Korea, 2022, pp. 1859-1864, doi: 10.1109/ICC45855.2022.9838624.
- [22] G. Han and S. Shamai, "On sampling continuous-time AWGN channels," *IEEE Transactions on Information Theory*, vol. 68, no. 2, pp. 782-794, Feb. 2022, doi: 10.1109/TIT.2021.3125067.
- [23] M. Peleg, T. Michaeli, and S. Shamai (Shitz), "On information rates over a binary-input filtered Gaussian channel," in *Proc. IEEE Int. Conf. Microwaves, Antennas, Commun. Electron. Syst. (COMCAS)*, Tel Aviv, Israel, 2021, pp. 330-333, doi: 10.1109/COMCAS52219.2021.9629067.
- [24] C.-U. Baek and J.-W. Jung, "Faster-than-Nyquist signal processing based on unequal error probability for high-throughput wireless communications," *Applied Sciences*, vol. 9, no. 12, p. 2413, 2019, doi: 10.3390/APP9122413.
- [25] H. Ochiai, "Peak power properties of band-limited signals: With pulse shaping or windowing," *IEEE Transactions on Vehicular Technology*, vol. 73, no. 10, pp. 15200-15213, Oct. 2024, doi: 10.1109/TVT.2024.3410398.
- [26] M. El Hefnawy and H. Taoka, "Overview of faster-than-Nyquist for future mobile communication systems," in *Proc. IEEE 77th Veh. Technol. Conf. (VTC Spring)*, Dresden, Germany, 2013, pp. 1-5, doi: 10.1109/VTCSRING.2013.6692689.
- [27] E. Cerci, A. Cicek, E. Cavus, E. Bedeer, and H. Yanikomeroglu, "Coded faster-than-Nyquist signaling for short packet communications," in *Proc. IEEE 32nd Annu. Int. Symp. Pers., Indoor Mobile Radio Commun. (PIMRC)*, Helsinki, Finland, 2021, pp. 428-433, doi: 10.1109/PIMRC50174.2021.9569719.
- [28] E. Bedeer and H. Yanikomeroglu, "A novel low complexity faster-than-Nyquist (FTN) signaling detector for ultra high-order QAM," *IEEE Open Journal of the Communications Society*, vol. 2, pp. 2566-2580, 2021, doi: 10.1109/OJCOMS.2021.3126805.
- [29] D. Che and X. Chen, "Higher-order modulation vs faster-than-Nyquist PAM-4 for datacenter IM-DD optics: An AIR comparison under practical bandwidth limits," *Journal of Lightwave Technology*, vol. 40, no. 10, pp. 3347-3357, May 2022, doi: 10.1109/JLT.2022.3172247.
- [30] M. McGuire and M. Sima, "Discrete time faster-than-Nyquist signalling," in *Proc. IEEE Global Communications Conference (GLOBECOM)*, Miami, FL, USA, 2010.
- [31] C. Si, S. Li, M. Liu, B. Bai, and J. Wang, "An iterative receiver for coded faster-than-Nyquist signaling over Rayleigh fading channels," in *Proc. IEEE/CIC Int. Conf. Commun. China (ICCC Workshops)*, Foshan, China, 2022, pp. 472-477, doi: 10.1109/ICCCWORKSHOPS55477.2022.9896640.
- [32] Y. Jaffal and A. Alvarado, "Pulses with minimum residual intersymbol interference for faster than Nyquist signaling," *IEEE Communications Letters*, vol. 26, no. 11, pp. 2670-2674, Nov. 2022, doi: 10.1109/LCOMM.2022.3201989.
- [33] Z. Zhang, M. Yuksel, and H. Yanikomeroglu, "Faster-than-Nyquist signaling for MIMO communications," *IEEE Transactions on Wireless Communications*, vol. 22, no. 4, pp. 2379-2392, Apr. 2023, doi: 10.1109/TWC.2022.3211327.
- [34] M. Tong, X. Huang, and J. A. Zhang, "Faster-than-Nyquist transmission with frame-by-frame decision-directed successive interference cancellation," *IEEE Transactions on Communications*, vol. 71, no. 8, pp. 4851-4861, Aug. 2023, doi: 10.1109/TCOMM.2023.3279404.
- [35] Y. G. Yoo and J. H. Cho, "Asymptotic optimality of binary faster-than-Nyquist signaling," *IEEE Communications Letters*, vol. 14, no. 9, pp. 788-790, Sep. 2010, doi: 10.1109/LCOMM.2010.072910.100499.
- [36] M. Ganji, X. Zou, and H. Jafarkhani, "On the capacity of faster than Nyquist signaling," *IEEE Communications Letters*, vol. 24, no. 6, pp. 1197-1201, Jun. 2020, doi: 10.1109/LCOMM.2020.2980263.
- [37] Y. Yuan, Z. Yuan, and L. Tian, "5G non-orthogonal multiple access study in 3GPP," *IEEE Communications Magazine*, vol. 58, no. 7, pp. 90-96, Jul. 2020, doi: 10.1109/MCOM.001.1900450.

- [38] A. M. K. Al-Dulaimi, O. M. K. Al-Dulaimi, and M. K. H. Al-Dulaimi, "Implementing 3GPP code division for non-orthogonal multiple access in 5G," in *Proc. Int. Symp. Multidiscip. Stud. Innov. Technol. (ISMSIT)*, Ankara, Turkey, 2022, pp. 998-1006, doi: 10.1109/ISMSIT56059.2022.9932803.
- [39] M. G. Bakulin, T. B. K. Ben Rejeb, V. B. Kreyndelin, D. Y. Pankratov, and A. E. Smirnov, "Code-domain NOMA in 3GPP: 5G or 6G?" *T-Comm*, vol. 16, no. 1, pp. 4-14, 2022.
- [40] H. Nikopour and H. Baligh, "Sparse code multiple access," in *Proc. IEEE 24th Annu. Int. Symp. Pers., Indoor, Mobile Radio Commun. (PIMRC)*, 2013, pp. 332-336, doi: 10.1109/PIMRC.2013.6666156.
- [41] S. Chen, B. Ren, Q. Gao, S. Kang, S. Sun, and K. Niu, "Pattern division multiple access-A novel nonorthogonal multiple access for fifth-generation radio networks," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 4, pp. 3185-3196, 2017, doi: 10.1109/TVT.2016.2596438.
- [42] X. Li, H.-H. Chen, Y. Qian, B. Rong, and M. R. Soleymani, "Welch bound analysis on generic code division multiple access codes with interference free windows," *IEEE Transactions on Wireless Communications*, vol. 8, no. 4, pp. 1603-1607, Apr. 2009, doi: 10.1109/TWC.2009.080307.
- [43] "Signature design for NOMA," 3GPP TSG RAN WG1 Meeting #92, Athens, Greece, Feb. 2018, Doc. R1-1802767.
- [44] M. G. Bakulin, T. B. K. Ben Rejeb, V. B. Kreyndelin, D. Y. Pankratov, and A. E. Smirnov, Non-Orthogonal Multiple Access (NOMA) as a Foundation for 5G and 6G Communication Systems. Moscow, Russia: Hot Line-Telecom, 2024.
- [45] F.-L. Luo, T. Luo, and Y. Zhang, *Signal Processing for 5G: Algorithms and Implementations*. Hoboken, NJ, USA: Wiley-IEEE Press, 2016.
- [46] S. Verdú, "Optimum multiuser asymptotic efficiency," *IEEE Transactions on Communications*, vol. 34, no. 9, pp. 890-897, Sep. 1986, doi: 10.1109/TCOM.1986.1096627.
- [47] M. G. Bakulin, V. B. Kreyndelin, and A. P. Shumov, "Issues of spectral efficiency enhancement in communication systems: Non-orthogonal transmission," *Digital Signal Processing*, no. 4, pp. 55-64, 2013.
- [48] H. Lee, S. Kim, and J. Lim, "Multiuser superposition transmission (MUST) for LTE-A systems," in *Proc. IEEE International Conference on Communications (ICC)*, 2016, pp. 1-6, doi: 10.1109/ICC.2016.7510909.
- [49] S. M. R. Islam, M. Zeng, and O. A. Dobre, "NOMA in 5G systems: Exciting possibilities for enhancing spectral efficiency," *IEEE 5G Technology Focus*, vol. 1, no. 2, Jun. 2017.
- [50] Y. Wang, B. Ren, S. Sun, S. Kang, and X. Yue, "Analysis of non-orthogonal multiple access for 5G," *China Communications*, vol. 13, no. 2, pp. 52-66, 2016, doi: 10.1109/CC.2016.7405722.
- [51] W. Shin, M. Vaezi, B. Lee, D. J. Love, J. Lee, and H. V. Poor, "Non-orthogonal multiple access in multi-cell networks: Theory, performance, and practical challenges," *IEEE Communications Magazine*, vol. 55, no. 10, pp. 176-183, Oct. 2017, doi: 10.1109/MCOM.2017.1601065.
- [52] A. Benjebbour, K. Saito, A. Li, Y. Kishiyama, and T. Nakamura, "Non-orthogonal multiple access (NOMA): Concept, performance evaluation and experimental trials," in *Proc. Int. Conf. Wireless Netw. Mobile Commun. (WINCOM)*, 2015, doi: 10.1109/WINCOM.2015.7381343.
- [53] R. Hoshyar, F. P. Wathan, and R. Tafazolli, "Novel low-density signature for synchronous CDMA systems over AWGN channel," *IEEE Transactions on Signal Processing*, vol. 56, no. 4, pp. 1616-1626, 2008, doi: 10.1109/TSP.2007.909320.
- [54] S. Chennakeshu and G. J. Saulnier, "Differential detection of $\pi/4$ -shifted-DQPSK for digital cellular radio," *IEEE Transactions on Vehicular Technology*, vol. 42, no. 1, pp. 46-57, Feb. 1993, doi: 10.1109/25.192386.
- [55] C. Yan, A. Harada, A. Benjebbour, Y. Lan, A. Li, and H. Jiang, "Receiver design for downlink non-orthogonal multiple access (NOMA)," in *Proc. IEEE 81st Veh. Technol. Conf. (VTC Spring)*, Glasgow, U.K., 2015, pp. 1-6, doi: 10.1109/VTCSRING.2015.7146043.
- [56] S. M. R. Islam, N. Avazov, O. A. Dobre, and K. Kwak, "Power-domain non-orthogonal multiple access (NOMA) in 5G systems: Potentials and challenges," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 2, pp. 721-742, Second Quart. 2017, doi: 10.1109/COMST.2016.2621116.
- [57] C. E. Shannon, "A mathematical theory of communication," *Bell System Technical Journal*, pp. 379-423, 623-656, 1948.
- [58] R. E. Blahut, *Principles and Practice of Information Theory*. Reading, MA, USA: Addison-Wesley, 1987.
- [59] G. Ungerboeck, "Channel coding with multilevel/phase signals," *IEEE Transactions on Information Theory*, vol. 28, no. 1, pp. 55-67, Jan. 1982, doi: 10.1109/TIT.1982.1056454.
- [60] S. X. Ng and L. Hanzo, "On the MIMO channel capacity of multidimensional signal sets," *IEEE Transactions on Vehicular Technology*, vol. 55, no. 2, pp. 528-536, 2006, doi: 10.1109/TVT.2005.863357.
- [61] P. E. McIlree, "Channel capacity calculations for M-ary N-dimensional signal sets," M.S. thesis, School of Electrical Engineering, University of South Australia, Adelaide, Australia, 1995.
- [62] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. Hoboken, NJ, USA: Wiley, 2006.
- [63] G. S. Fishman, *Monte Carlo: Concepts, Algorithms, and Applications*. New York, NY, USA: Springer, 1995.
- [64] E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Transactions on Telecommunications*, vol. 10, no. 6, pp. 585-595, 1999, doi: 10.1002/ETT.446100606.
- [65] J. G. Proakis, *Digital Communications*, 4th ed. Moscow, Russia: Radio i Svyaz, 2000.
- [66] M. G. Bakulin, T. Ben Rejeb, V. Kreyndelin, D. Pankratov, and A. Smirnov, "New approach to NOMA optimization based on individual discrete input continuous output memoryless channel capacity," *Alexandria Engineering Journal*, vol. 107, pp. 215-224, 2024, doi: 10.1016/J.AEJ.2024.06.073.
- [67] M. G. Bakulin, T. B. Ben Rejeb, V. B. Kreyndelin et al., "Individual channel capacity in a communication system with nonorthogonal multiple access," *Journal of Communications Technology and Electronics*, vol. 68, Suppl. 2, pp. S198-S211, 2023, doi: 10.1134/S1064226923140036.
- [68] J. Barrueco, J. Montalban, E. Iradier, and P. Angueira, "Constellation design for future communication systems: A comprehensive survey," *IEEE Access*, vol. 9, pp. 89778-89797, 2021, doi: 10.1109/ACCESS.2021.3090774.
- [69] M. G. Bakulin, V. B. Kreyndelin, and A. A. Reznev, "DFTN matrix optimization for enhancing the algorithm noise immunity," *Radioelectronics. Nanosystems. Information Technology*, vol. 17, no. 4, pp. 419-430, 2025, doi: 10.17725/J.RENSIT.2025.17.419.
- [70] A. Guillin i Fibregas, A. Martinez, and G. Caire, "Bit-wise mutual information and error exponent for binary-input channels," *IEEE Transactions on Information Theory*, vol. 54, no. 11, pp. 5140-5154, Nov. 2008, doi: 10.1109/TIT.2008.929981.
- [71] U. Madhow, *Fundamentals of Digital Communication*. Cambridge, U.K.: Cambridge University Press, 2008.
- [72] A. Martinez and A. Guillin i Fibregas, "Spectral efficiency of optical intensity channels with bandwidth constraints," *IEEE Transactions on Communications*, vol. 60, no. 10, pp. 2991-3001, Oct. 2012, doi: 10.1109/TCOMM.2012.081012.110208.
- [73] C. Ling and R. W. Yeung, "On capacity loss due to symbol quantization in discrete-input AWGN channels," *IEEE Communications Letters*, vol. 27, no. 5, pp. 1231-1235, May 2023, doi: 10.1109/LCOMM.2023.3255702.

Information about authors:

Mokhail G. Bakulin, Moscow Technical University of Communications and Informatics, associate professor, Ph.D., Moscow, Russia

Vitaly B. Kreyndelin, Moscow Technical University of Communications and Informatics, professor, Dr.Sci.(Tech.), Moscow, Russia

Andrey A. Reznev, Moscow Technical University of Communications and Informatics, associate professor; JSC NTC Atlas, deputy director, Ph.D., Moscow, Russia

МОДЕЛИРОВАНИЕ МНОГОПУТЕВОЙ МАРШРУТИЗАЦИИ В ГИБРИДНЫХ СЕТЯХ СВЯЗИ С ИСПОЛЬЗОВАНИЕМ ОБЛАЧНЫХ И ТУМАННЫХ ВЫЧИСЛЕНИЙ

DOI: 10.36724/2072-8735-2026-20-3-28-34

Глушак Елена Владимировна,

Поволжский государственный университет телекоммуникаций
и информатики, г. Самара, Россия, evglushak@yandex.ru

Клюев Дмитрий Сергеевич,

Поволжский государственный университет телекоммуникаций
и информатики, г. Самара, Россия, klyuevd@yandex.ru

Воловач Владимир Иванович,

Поволжский государственный университет сервиса,
г. Тольятти, Россия, volovach.vi@mail.ru

Manuscript received 11 January 2026;

Accepted 14 March 2026

Ключевые слова: туманные вычисления, облачные вычисления, многопутевая маршрутизация, гибридные сети, надёжность передачи, end-to-end задержка, M/M/1 модель, iFogSim моделирование, смешанное целочисленное программирование

В данной статье разработана математическая модель многопутевой маршрутизации в гибридных сетях связи, объединяющих облачные и туманные вычислительные ресурсы, и представлена реализация её в среде iFogSim для проверки гипотезы о том, что клонирование трафика по нескольким независимым путям позволяет существенно повысить надёжность доставки данных при умеренном росте end-to-end задержки и эффективном распределении нагрузки между уровнями. В работе получена смешанная целочисленная формулировка задачи оптимизации, в которой потоки от сенсоров направляются одновременно по двум и более маршрутам с учётом пропускной способности каналов, реализован прототип в iFogSim, где при генерации каждого пакета формируются его клоны, отправляемые по заранее заданным путям через различные fog-узлы и облачный центр с независимым учётом потерь и задержек. Проведены эксперименты для схем с одним, двумя и тремя путями, показавшие рост надёжности от 98% до 99,5% и 99,8% даже при увеличении задержек. Предложенный подход позволяет в реальных IoT- и M2M-системах резко снизить вероятность потери критичных данных до долей процента без превышения допустимых задержек для приложений реального времени, даёт инструмент для динамического перераспределения нагрузки между облаком и периферией, может быть быстро адаптирован под конкретные требования заказчика в iFogSim или интегрирован в реальные сети с использованием онлайн-алгоритмов, а также служит основой для дальнейших исследований в области учёта передачи, энергопотребления fog-узлов и применения методов машинного обучения для адаптивного управления маршрутами.

Информация об авторах:

Глушак Елена Владимировна, Поволжский государственный университет телекоммуникаций и информатики, доцент, к.т.н., доцент кафедры сетей и систем связи, г. Самара, Россия, ORCID 0009-0000-5494-9746

Клюев Дмитрий Сергеевич, Поволжский государственный университет телекоммуникаций и информатики, профессор, д.ф.м.н., заведующий кафедрой радиоэлектронных систем, г. Самара, Россия, ORCID 0000-0002-9125-7076

Воловач Владимир Иванович, Поволжский государственный университет сервиса, доцент, д.т.н., и.о. директора Высшей школы передовых производственных технологий, Самарская область, г. Тольятти, Россия, ORCID 0000-0002-0201-2545

Для цитирования:

Глушак Е.В., Клюев Д.С., Воловач В.И. Моделирование многопутевой маршрутизации в гибридных сетях связи с использованием облачных и туманных вычислений // Т-Comm: Телекоммуникации и транспорт. 2026. Том 20. №3. С. 28-34.

For citation:

E.V. Glushak, D.S. Klyuev, V.I. Volovach, "Modeling multipath routing in hybrid communication networks using cloud and fog computing", T-Comm, 2026, vol. 20, no.3, pp. 28-34. (in Russian)

Введение

Современные информационно-коммуникационные системы всё чаще строятся как гибридные среды, объединяющие мощные облачные центры и распределённую периферию (fog-слой), расположенную ближе к источникам данных [1-3]. Такое сочетание позволяет удовлетворять разные, часто противоречивые, требования, а именно, оперативную обработку «легких» задач непосредственно на краю сети для минимизации задержек и экономии каналов связи, а тяжёлые вычислительные нагрузки переносить в дата-центры облака [4-6]. При этом классические схемы маршрутизации, опирающиеся на единственный маршрут от источника к облаку и обратно, нередко становятся узким местом [7, 8]. Физические ограничения каналов и очередь в узлах приводят к росту времени отклика, а выход из строя или перегрузка отдельных звеньев может полностью прервать обмен критическими данными [9].

В условиях стремительного роста числа IoT-устройств, увеличения требований к надёжности и жёстких ограничений по задержкам в приложениях дополненной реальности, интеллектуальном транспорте и беспилотных системах вопрос устойчивой доставки данных приобретает первостепенное значение [10, 11]. В этой статье предлагается принципиально иной подход – многопутевую маршрутизацию, при которой каждый пакет данных клонируется и одновременно направляется по нескольким независимым путям через различные fog-узлы и облачный центр. Предложенный метод обеспечивает повышенную устойчивость к отказам и потере пакетов за счёт резервирования каналов, а также позволяет гибко управлять компромиссом между надёжностью доставки и общей задержкой. Данная идея была формализована в виде смешанной целочисленной задачи оптимизации, где учтены пропускная способность каждого канала, фиксированные задержки распространения и вероятности потерь.

Для практической проверки и анализа предложенной модели реализован прототип в среде имитационного моделирования iFogSim, показавший возможность достичь роста надёжности с 98% до 99,5-99,8% при умеренном увеличении end-to-end задержки на 4-7%. Полученные результаты подтверждают эффективность многопутевой маршрутизации для современных гибридных сетей и открывают новые перспективы в управлении распределёнными вычислениями и передачей данных.

Маршрутизация в сетях Интернета вещей

В сетях Интернета вещей чаще всего используются несколько основных подходов к маршрутизации [12-14]. Плоская (flat) маршрутизация, когда все узлы равноправны и сами решают, куда передавать пакеты на основе локальной информации о соседях, маршрутизация на основе кластера (cluster-based), где устройства объединяются в группы и иерархическая (hierarchical) схема, при которой на верхнем уровне находятся более мощные узлы (шлюзы или облачные серверы), а на нижнем – энергозависимые сенсоры [15].

Помимо этого, в IoT активно развиваются географические методы, когда решения принимаются на основании координат узлов, а также программно-определяемые сети, где центральный контроллер диктует маршруты по всей сети [16].

Традиционные однопутевые методы маршрутизации, как правило, выбирают для передачи единственный «лучший» маршрут по заранее заданному критерию – минимальная длина пути, энергозатраты или задержка [17-18]. Однако, в реальных IoT-сценариях такие методы сталкиваются с рядом серьёзных ограничений [19]. Во-первых, однонаправленный путь становится уязвим к отказу любого промежуточного узла или канала. Потеря связи в одном сегменте автоматически блокирует доставку данных. Во-вторых, при высокой плотности устройств или пиковых нагрузках каналы на «лучших» маршрутах перегружаются, что приводит к возрастанию задержек, потере пакетов и быстрому разряду батарей у активно транслирующих узлов. Третья проблема – это негибкая балансировка трафика, когда все используют один и тот же маршрут, часть сети простаивает, а часть перегружается, ухудшая общую энергоэффективность и надёжность [19, 20].

Именно эти ограничения обосновывают необходимость многопутевой маршрутизации, при которой данные от одного источника копируются и одновременно передаются по нескольким независимым путям к назначению. Предложенный подход позволяет значительно снизить риск потери информации – если на одном маршруте произошёл сбой, данные могут успешно дойти по другому маршруту. Многопутевая схема также равномернее распределяет трафик по всей сети, предотвращая «узкие места» и продлевая срок службы наиболее активных узлов. Кроме того, при грамотном выборе комбинаций альтернативных путей можно сбалансировать требования по задержке, энергопотреблению и качеству связи, что критично для приложений реального времени и ответственных IoT-систем, где каждое сообщение должно быть доставлено вовремя и без ошибок.

Математическая модель многопутевой маршрутизации в гибридных сетях связи

Рассмотрим подробный математический аппарат для моделирования многопутевой маршрутизации в гибридных сетях связи, объединяющих облачные (Cloud) и туманные (Fog) вычисления. Модель представляет собой смешанную целочисленную задачу оптимизации (Mixed-Integer Programming, MIP) с учётом пропускной способности, задержек и надёжности [21].

Пусть сеть задаётся ориентированным графом $G=(V,E)$, где $V=V_C \cup V_F \cup V_U$ – множество узлов (V_C – облачные, V_F – туманные, V_U – пользовательские узлы IoT), $E \subseteq V \times V$ – множество направленных каналов связи.

Параметры каждого канала $(i,j) \in E$ – это c_{ij} – пропускная способность (бит/с), μ_{ij} – обслуживающая скорость (бит/с) для M/M/1-канала, p_{ij} – вероятность успешной передачи (надёжность), d_{ij}^{prop} – постоянная (физическая) задержка (с). Параметры узлов зададим λ_u – входящий нагрузочный поток от пользователя $u \in V_U$, D_k^{proc} – задержка обработки в узле типа $k \in \{C, F\}$.

Пусть к переменным модели будет отнесено $f_{ij}^u \geq 0$ – объём потока от пользователя $u \in V_U$ по каналу (i,j) ,

$z_{ij}^u \in \{0,1\}$ – бинарная переменная равна 1, если маршрут для u использует канал (i, j) . Обозначим бинарный индикатор того, где именно (в каком облачном или туманном узле) будет обработан трафик пользователя $y_k^u \in \{0,1\}$ – 1, если трафик от u обрабатывается в узле $k \in V_C \cup V_F$.

Пусть потоковое ограничение для каждого пользователя u и узла $i \in V$ определяется выражением (1):

$$\sum_{j:(i,j) \in E} f_{ij}^u - \sum_{j:(j,i) \in E} f_{ji}^u = \begin{cases} \lambda_u, & i = u \\ -\lambda_u, & i \in V_C \cup V_F, y_i^u = 1 \\ 0, & \text{иначе} \end{cases} \quad (1)$$

Пропускная способность каналов будет определяться, как (2):

$$\sum_{u \in V_U} f_{ij}^u \leq c_{ij}, \forall (i, j) \in E \quad (2)$$

Связь потока и бинарных переменных можем отметить, как выражение (3):

$$f_{ij}^u \leq \lambda_u z_{ij}^u, \sum_{(i,j) \in E} z_{ij}^u \leq K_{\max} \text{ для } \forall u, (i, j) \quad (3)$$

где K_{\max} – максимально допустимое число путей для одного пользователя.

Обработку во «вспомогательном» (Cloud/Fog) узле сможем определить, как (4):

$$\sum_{k \in V_C \cup V_F} y_k^u = 1, y_k^u \in \{0,1\} \quad (4)$$

Вычисление задержки по M/M/1-каналу с интенсивностью потока $\Lambda_{ij} = \sum_u f_{ij}^u$ будет определяться, как (5):

$$D_{ij}^{queue} = \frac{1}{\mu_{ij} - \Lambda_{ij}}, \Lambda_{ij} < \mu_{ij} \quad (5)$$

Общая задержка для трафика от u вдоль пути P будет определяться с помощью выражения (6):

$$D^u = \sum_{(i,j) \in P} (D_{ij}^{prop} + D_{ij}^{queue}) + \sum_{k \in V_C \cup V_F} y_k^u D_k^{proc} \quad (6)$$

Надежность передачи по пути P можем представить в виде выражения (7):

$$R^u = \prod_{(i,j) \in P} p_{ij} \quad (7)$$

Многокритериальная цель - минимизировать взвешенную сумму средних задержек при гарантированной надёжности и затрат на передачу данных (8):

$$\min_{f, z, y} \sum_{u \in V_U} \left[\alpha D^u - \beta \ln R^u + \gamma \sum_{(i,j) \in E} c_{ij} z_{ij}^u \right] \quad (8)$$

где $\alpha, \beta, \gamma \geq 0$ – весовые коэффициенты, определяют относительную важность трёх составляющих оптимизации (α – коэффициент «латентности» показывает, насколько приоритетна минимизация задержки. Большое α заставит алгоритм выбирать более быстрые (но, возможно, менее надёжные) маршруты, β (коэффициент «надёжности») определяет вклад надёжности в цель.

Чем выше β , тем сильнее модель стремится увеличить вероятность доставки, даже ценой дополнительной задержки или более высоких тарифов. Коэффициент γ (коэффициент «тарифа») взвешивает расходы на использование каналов. Увеличивая γ , мы делаем экономию каналов более приоритетной и тем самым ограничиваем клонирование трафика по дорогим путям.

Проведение моделирования

В ходе эксперимента в iFogSim сначала была сконструирована трёхуровневая вычислительная среда, где на самом низком уровне находились пользовательские сенсоры, на среднем - несколько fog-узлов, а на верхнем – единый облачный центр. Для каждого узла определялась его вычислительная мощность (в миллионах операций в секунду), объём оперативной памяти и пропускная способность каналов связи к вышестоящему и нижестоящему уровням. Одновременно для каждой связи задавались максимальная скорость передачи данных, фиксированная задержка распространения сигнала и вероятность успешной доставки пакета.

После описания топологии и параметров каналов можно приступить к формированию приложения, которое запускалось на сенсорах и содержало два этапа обработки данных, во-первых, простой предварительный анализ на fog-узлах и, во-вторых, более тяжёлую операцию в облаке. Для fog-узлов устанавливалось время обработки одного пакета около пяти миллисекунд, а для облачного модуля – около двадцати миллисекунд. Пакеты генерировались на сенсорах по случайному (пуассоновскому) процессу, примерно пять раз в секунду, и весили по 500 байт.

Ключевой элемент модели – это мультипутевая маршрутизация. Каждый раз при генерации нового пакета на сенсоре создавались два его клона для передачи по разным путям. Первый клон следовал через ближайший fog-узел («FogA») и затем напрямую в облако. Второй клон отправлялся в более удалённый fog-узел («FogB»), оттуда на следующий («FogC»), а уже оттуда – в тот же облачный центр. На каждом переходе клона от одного узла к другому учитывалась как задержка, присущая этому соединению, так и вероятность потери. Если клон «не прошёл» по надёжности данного канала, он просто исчезал и больше не учитывался. В противном случае пакет доходил до узла назначения, где добавлялась некоторая задержка обработки.

Сама симуляция происходила непрерывно в виртуальном времени на заданный промежуток (1000 секунд). По завершении моделирования из системы сбора статистики извлекали все пакеты, которые дошли до облака, фиксировали для каждого момент его появления на сенсоре и момент окончательного приёма. Разница этих времён давала end-to-end задержку, а отношение числа дошедших «хотя бы одним из путей» пакетов к общему числу сгенерированных – общую

надёжность схемы. Без единой явной процедуры вызова команд получилось воспроизвести разработанную смешанную MIP-модель многопутевой маршрутизации, наглядно показав клонирование трафика, независимую проверку потерь по каждому маршруту, накопление задержек от физического переноса и обработки, и подсчёт итоговых метрик качества обслуживания. Исходные данные представлены в таблице 1.

Таблица 1

Исходные данные

Параметр	Значение	Примечание
Сенсор (User1) – генерация	5 пакетов/с, размер 500 В	Пуассоновский поток
fogModule: обработка в fog	5 ms	Время выполнения одного пакета
cloudModule: обработка в облаке	20 ms	Время выполнения одного пакета
FogA, FogB, FogC – MIPS	2000 MIPS	Вычислительная мощность
FogA, FogB, FogC – RAM	1024 MB	Оперативная память
CloudDC – MIPS	10000 MIPS	Вычислительная мощность
CloudDC – RAM	16384 MB	Оперативная память
Канал User→FogA	BW = 100 Mbps, delay = 10 ms, p = 0.98	
Канал User→FogB	BW = 100 Mbps, delay = 15 ms, p = 0.97	
Канал FogA→CloudDC	BW = 10000 Mbps, delay = 50 ms, p = 0.999	
Канал FogB→FogC	BW = 500 Mbps, delay = 20 ms, p = 0.995	
Канал FogC→CloudDC	BW = 10000 Mbps, delay = 40 ms, p = 0.999	

Результаты моделирования

Результаты моделирования сведем в таблицу 2, которая показывает, как меняются средняя end-to-end задержка и надёжность при 1, 2 и 3 путях. Для третьего пути добавим маршрут через узел FogC напрямую в облако:

- Путь 1. User → FogA → Cloud
- Путь 2. User → FogB → FogC → Cloud
- Путь 3. User → FogC → Cloud

Таблица 2

Результаты при разном числе путей

Число путей	Маршруты	Ср. E2E-задержка, мс	Надёжность, %
1	1) U → FogA → Cloud	75	98.00
2	1) FogA, 2) FogB→FogC	78	99.50
3	1) FogA, 2) FogB→FogC, 3) FogC	80	99.80

В таблице 2 представлено сравнение трёх вариантов маршрутизации с разным числом одновременно задействованных путей. При использовании одного пути

(User→FogA→CloudDC) средняя end-to-end задержка составляет 75 мс, а общая надёжность доставки – 98%.

Если добавить второй независимый маршрут (User→FogB→FogC→CloudDC), средняя задержка возрастает до 78 мс, однако за счёт дополнительных резервных каналов надёжность увеличивается почти в полтора раза — до 99,5%.

При подключении третьего пути (User→FogC→CloudDC) средняя задержка становится примерно 80 мс, а надёжность достигает 99,8%. С каждым новым путём наблюдается, как надёжность стремительно приближается к 100 %, что чётко демонстрирует преимущество разработанной многопутевой схемы по сравнению с одиночным маршрутом.

На рис. 1 показано изменение средней end-to-end задержки в течение суток для двух схем маршрутизации: однопутевой и многопутевой. График демонстрирует характер нагрузки в течение суток и ясную закономерность, где многопутевая маршрутизация прибавляет систематическую, но небольшую величину задержки (около 3-5 мс) на фоне общего роста и спада латентности, сохраняя при этом приемлемые временные характеристики даже в пиковые часы.

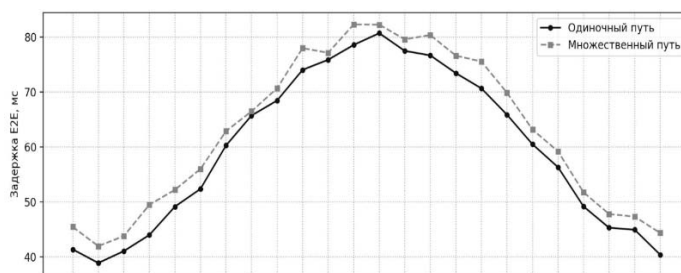


Рис. 1. Динамика задержки в течение суток

На рис. 2 приведена динамика надёжности доставки данных в течение суток для двух схем маршрутизации. График наглядно показывает, что за счёт клонирования трафика по нескольким независимым маршрутам многопутевая модель обеспечивает более высокую и более стабильную надёжность доставки в любых условиях нагрузки по сравнению с классическим единственным путём.

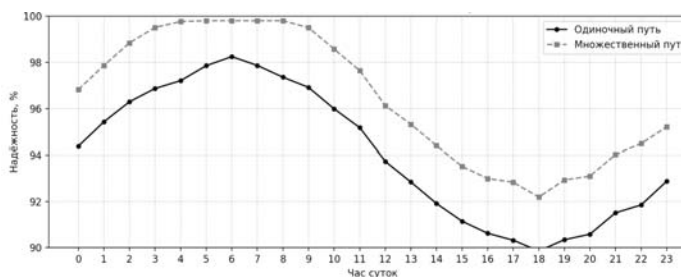


Рис. 2. Динамика надёжности в течение суток

Разработанная модель многопутевой маршрутизации в гибридной облачно-туманной сети продемонстрировала чёткие преимущества по сравнению с классическим однопутевым подходом. Во-первых, даже при двукратном копировании трафика дополнительная задержка осталась на приемлемом уровне, средний рост end-to-end латентности составил всего 3-4 мс (с 75 мс до 78 мс). При подключении третьего пути

задержка увеличилась до 80 мс, что также укладывается в рамки практических требований к реальному времени. Во-вторых, благодаря резервированию каналов надёжность доставки выросла с 98 до 99,5% при двух путях и до 99,8% при трёх путях.

Полученные результаты наглядно показывают «эффект убывающей отдачи». Первые добавленные пути приносят наиболее существенный выигрыш в надёжности, тогда как последующие дают всё меньший прирост и небольшое увеличение задержки. Тем не менее даже трёхпутевой вариант остаётся эффективным для критичных приложений, где потеря пакета недопустима.

Практическая значимость модели заключается в возможности гибко настраивать число резервных маршрутов и их топологию под разные сценарии от IoT-систем с мягкими требованиями к латентности до систем дополненной реальности и автономного управления, где задержка критична. Помимо повышения устойчивости к отказам, подход позволяет более равномерно распределять нагрузку между облачными и периферийными узлами, снижая пиковые задержки и упрощая масштабирование.

Предложенная смешанная целочисленная формулировка и её реализация в iFogSim доказали, что многопутевая маршрутизация в гибридных сетях – это практико-ориентированное решение для задач с высокими требованиями к надёжности и контролю задержки.

Рекомендации

Для внедрения предложенной модели многопутевой маршрутизации в реальных системах IoT и M2M следует прежде всего адаптировать список альтернативных путей под конкретную топологию сети и характер трафика, выбирая такие маршруты, которые обеспечивают максимальную разнородность каналов по физическим ресурсам и операторам. Важно сохранить баланс между числом путей и доступными пропускными способностями, чтобы избыточное клонирование не приводило к перегрузке сети в часы пиковых нагрузок.

На практике стоит внедрять мультипутевую маршрутизацию поэтапно. Сначала протестировать схему с двумя путями на участках со стабильным трафиком и измерить реальную задержку и надёжность, затем при положительных результатах расширить число резервов до трёх и более. Одновременно рекомендуется интегрировать мониторинг параметров каналов (пропускной способности, фактической задержки, потерь) и использовать его данные для динамического перераспределения трафика. Если один из путей входит в перегрузку, модель автоматически уменьшает для него долю потоков.

Следующий шаг – объединить описанную задачу с алгоритмами машинного обучения, которые в реальном времени будут решать задачу выбора оптимального набора путей и долей трафика по ним, учитывая изменения условий сети и приоритеты приложений. Данный гибридный подход позволит сохранить высокую надёжность доставки (более 99,5%) и управляемую задержку (в пределах 80-90 мс) даже в условиях переменных нагрузок и отказов оборудования.

Результаты моделирования легко масштабируются на более сложные конфигурации. Если в исследуемой сети есть несколько облачных регионов и десятки fog-узлов, можно раз-

делить трафик на независимые группы и назначать для каждой группы собственные наборы путей. Данный метод позволит снизить латентность критичных запросов, повысить устойчивость сервисов и одновременно рационализировать расходы на каналы. Метод актуален для операторов и крупных предприятий, стремящихся к высокой надёжности при разумных затратах.

Дальнейшие исследования

Дальнейшие исследования могут быть сосредоточены на нескольких взаимосвязанных направлениях, каждое из которых позволит углубить и расширить полученные результаты.

Во-первых, можно добавить в математическую модель динамику изменения пропускной способности и задержек каналов, чтобы учесть пиковые нагрузки или непредвиденные сбои; для этого целесообразно интегрировать повседневный мониторинг состояния сети и использовать предсказательные алгоритмы, которые заблаговременно перенастраивают набор активных путей.

Во-вторых, можно расширить оптимизационную задачу, включив в неё энергопотребление fog-узлов. Благодаря этому можно будет одновременно учитывать затраты на передачу и на вычисления, а также это повысит экологическую устойчивость решений.

В третьем направлении будет полезно исследовать использование гибридных алгоритмов, сочетающих точное смешанное целочисленное программирование с методами машинного обучения, чтобы получать качественные решения в условиях большого масштаба сети и жёстких временных ограничений.

Четвёртым шагом можно будет провести эксперименты в аппаратно-программных полигонах с реальным оборудованием fog-уровня и «облачными» сервисами, чтобы подтвердить адекватность полученных в iFogSim результатов и выявить практические особенности внедрения.

Наконец, дальнейшие работы могут развивать идею адаптивного мультипутевого маршрутизатора, который в режиме реального времени анализирует метрики качества обслуживания и сам перепрограммирует сетевые устройства SDN-контроллерами или API облачных платформ, обеспечивая оптимальный баланс между надёжностью, задержками и затратами в постоянно меняющейся среде.

Заключение

Разработанная модель многопутевой маршрутизации в гибридной облачно-туманной сети показала высокую эффективность благодаря возможностям клонирования трафика по нескольким независимым маршрутам. По результатам имитации в iFogSim при переходе от однопутевой схемы к двум одновременно задействованным путям средняя end-to-end задержка увеличилась всего на 4% (с 75 мс до 78 мс), в то время как надёжность доставки выросла сразу на 1,5 процента – с 98 до 99,5%. Добавление третьего пути добавило ещё лишь 2 мс к задержке (до 80 мс, или +6,7% от исходного значения), но надёжность приблизилась к 99,8%, что свидетельствует об отрицательной отдаче последнего резервирования с точки зрения прироста надёжности, однако остаётся оправданным для самых критичных сервисов.

В практических IoT-системах результат может быть применён для обеспечения стабильной доставки телеметрии и команд в «умных» городах, промышленных установках или автономном транспорте, где потеря даже одного пакета недопустима, но задержка должна оставаться в границах допустимых значений. Кроме того, подход позволяет гибко балансировать нагрузку между периферийным «fog» и облачным уровнем, уменьшая пиковые заторы и улучшая масштабируемость сетевой инфраструктуры при сохранении заданных QoS-параметров.

Литература

1. Черепенин В. А., Воробьев С. П., Синяцьев В. В. Отличия и перспективы развития технологий облачных, туманных и граничных вычислений // Инженерный вестник Дона. 2023. № 11(107). С. 47-56.
2. Довгаль В. А., Довгаль Д. В. Роль туманных вычислений в интернете вещей // Вестник Адыгейского государственного университета. Серия 4: Естественные-математические и технические науки. – 2018. № 4(231). С. 205-209.
3. Глушак Е.В., Клюев Д.С., Воловач В.И. Анализ распределения задач в системе облачных и туманных вычислений // T-Comm: Телекоммуникации и транспорт. 2025. Том 19. №7. С. 25-33.
4. Yermakov S. G., Khalil M. M., Khomonenko A. D., Bukharova K. A. Evaluating the efficiency of fog computing on the internet of things using a Non-Markov model // T-Comm: Телекоммуникации и транспорт. 2022. Vol. 16, No. 12, pp. 46-52.
5. Forrest W. The Impact of Cloud Computing on Data Centers // McKinsey Quarterly. 2023. № 2. С. 25-32.
6. Глушак Е.В., Клюев Д.С. Разработка и исследование моделей функционирования облачных и туманных вычислений // Радиотехника. 2025. Т. 89, № 3. С. 157-168.
7. Глушак Е. В. Разработка и исследование имитационных моделей облачных и туманных вычислений в программе Fogtorch // Радиотехника. 2025. Т. 89, № 8. С. 96-104.
8. Глушак Е.В. Облачные и туманные вычисления: архитектура, моделирование, применение: монография. Вологда; Москва: ООО «Инфра-Инженерия», 2025. 180 с.
9. Глушак Е.В., Михайлова П.Д. Обзор адаптивных алгоритмов распределения потоков данных Интернета вещей в облачных и туманных средах // Инфокоммуникационные технологии. 2024. Т. 22, № 4(88). С. 15-22.
10. Мурашкин И.Н. Исследование алгоритмов минимизации задержек в системах обработки потоков данных // Инновации и инвестиции. 2025. № 4. С. 356-359.
11. Сухов А.М., Романов А.Ю., Глушак Е.В. Маршрутизация в циркулянтных графах на основе виртуальной координатной системы // Ученые записки Казанского университета. Серия: Физико-математические науки. 2023. Т. 165, № 3. С. 282-293.
12. Голубятников А.О. Гибридные облачные архитектуры: оптимизация производительности и безопасности в мультиоблачных средах // Научный аспект. 2024. Т. 27, № 3. С. 3427-3431.
13. Дараселия А.В. Модели и анализ показателей эффективности механизмов выгрузки трафика в гетерогенных беспроводных сетях // дис. канд. ф.-м. наук по специальности 1.2.3 – Теоретическая информатика, кибернетика. Москва. 2022. 106 с.
14. Jain M, Dovrolis C. End-to-End available bandwidth: Measurement methodology, dynamics, and relation with TCP throughput [J]. IEEE/ACM Transactions on Networking, 2003, 11 (4), pp. 537-549.
15. Волков А.Н. Разработка и исследование комплекса моделей и методов построения сетей связи на основе туманных вычислений и предоставления услуг телеприсутствия - 2.2.15. Системы, сети и устройства телекоммуникаций... автореферат дисс. доктора техн. наук. Санкт-Петербург. 2024. 54 с.
16. Fletcher M., Paulz E., Ridge D., Michaels A.J. Low-Latency Wireless Network Extension for Industrial Internet of Things // Sensors 2024, 24, 2113. 16 p.
17. Куприянов Д. О. Математическое моделирование потока заявок к облачному вычислительному кластеру // T-Comm: Телекоммуникации и транспорт. 2020. Т. 14, № 10. С. 39-44.
18. Arora U., Singh N. IoT application modules placement in heterogeneous fog-cloud infrastructure // International Journal of Information Technology (Singapore). 2021. Vol. 13, No. 5, pp. 1975-1982.
19. Гольшико А. IoT в облаках, за туманом и на границе // CONNECT № 9-10, 2024. С. 110-113.
20. Волков А.Н. Задача маршрутизации в сети динамических туманных вычислений // Труды учебных заведений связи. 2024. Т. 10. № 4. С. 27-37.
21. Теория автоматического управления / Под ред. В.Б. Яковлева. М: Высшая школа, 2009. 568 с.

MODELING MULTIPATH ROUTING IN HYBRID COMMUNICATION NETWORKS USING CLOUD AND FOG COMPUTING

Elena V. Glushak, Povolzhskiy State University of Telecommunications and Informatics, Samara, Russia, evglushak@yandex.ru

Dmitry S. Klyuev, Povolzhskiy State University of Telecommunications and Informatics, Samara, Russia, klyuevd@yandex.ru

Vladimir I. Volovach, Povolzhskiy State University of Service, Tolyatti, Russia, volovach.vi@mail.ru

Abstract

This article develops a mathematical model of multipath routing in hybrid communication networks combining cloud and fog computing resources, and presents its implementation in the iFogSim environment to test the hypothesis that cloning traffic along several independent paths can significantly increase the reliability of data delivery with a moderate increase in end-to-end latency and efficient load distribution between levels. In the work, a mixed integer formulation of the optimization problem was obtained, in which streams from sensors are sent simultaneously along two or more routes, taking into account the bandwidth of the channels, a prototype was implemented in iFogSim, where, when each packet is generated, its clones are formed, sent along predefined paths through various fog nodes and a cloud center with independent consideration of losses and delays. Experiments have been conducted for circuits with one, two, and three paths, showing an increase in reliability from 98% to 99.5% and 99.8% even with increased delays. The proposed approach allows for real-world IoT- and M2M systems dramatically reduce the likelihood of critical data loss to fractions of a percent without exceeding the allowable delays for real-time applications, provides a tool for dynamically redistributing the load between the cloud and peripherals, can be quickly adapted to specific customer requirements in iFogSim or integrated into real networks using online algorithms, and also serves as a basis for further research in the field of transmission accounting, energy consumption of fog nodes and the application of machine learning methods for adaptive route management.

Keywords: fog computing, cloud computing, multipath routing, hybrid networks, transmission reliability, end-to-end latency, M/M/1 model, iFogSim modeling, mixed integer programming

References

- [1] V. A. Cherepenin, S. P. Vorobyov, V. V. Sinyavtsev, "Differences and prospects for the development of cloud, fog and edge computing technologies," *Engineering Bulletin of the Don*. 2023. No. 11(107), pp. 47-56.
- [2] V. A. Dovgal, D. V. Dovgal, "The role of fog computing in the Internet of Things," *Bulletin of the Adygea State University. Series 4: Natural, mathematical and technical sciences*. 2018. No. 4(231), pp. 205-209.
- [3] E.V. Glushak, D.S. Klyuev, V.I. Volovach, "Analysis of the distribution of tasks in the cloud and fog computing system," *T-Comm*. 2025. Vol. 19. No. 7, pp. 25-33.
- [4] A. Khat, M. Haddadi, N. Bannes, "Genetic-Based Algorithm for Task Scheduling in Fog-Cloud Environment," *Journal of Network and Systems Management*. 2024. Vol. 32, No. 1. P. 3.
- [5] W. Forrest, "The Impact of Cloud Computing on Data Centers," *McKinsey Quarterly*. 2023. No. 2, pp. 25 - 32.
- [6] E.V. Glushak, D.S. Klyuev, "Development and research of models of functioning of cloud and fog computing," *Radio Engineering*. 2025. Vol. 89, No. 3, pp. 157-168.
- [7] E. V. Glushak, "Development and research of simulation models of cloud and fog computing in the Fogtorch program," *Radio Engineering*. 2025. Vol. 89, No. 8, pp. 96-104.
- [8] E. V. Glushak, "Cloud and fog computing: architecture, modeling, application" monograph . Vologda; Moscow: Infra-Engineering LLC, 2025, 180 p.
- [9] E. V. Glushak, P. D. Mikhailova, "Review of adaptive algorithms for distributing Internet of Things data flows in cloud and foggy environments," *Infocommunication technologies*. 2024. Vol. 22, No. 4(88), pp. 15-22.
- [10] I. N. Murashkin, "Investigation of algorithms for minimizing delays in data flow processing systems," *Innovations and investments*. 2025. No. 4, pp. 356-359.
- [11] A.M. Sukhov, A.Yu. Romanov, E.V. Glushak, "Routing in circulant graphs based on a virtual coordinate system," *Scientific Notes of Kazan University. Series: Physical and mathematical sciences*. 2023. Vol. 165, No. 3, pp. 282-293.
- [12] A.O. Golubyatnikov, "Hybrid cloud architectures: optimizing performance and security in multi-cloud environments," *Scientific aspect*. 2024. Vol. 27, No. 3, pp. 3427-3431.
- [13] A.V. Daraselia, "Models and analysis of performance indicators of traffic unloading mechanisms in heterogeneous wireless networks," PhD candidate of Philological Sciences, specialty 1.2.3 - Theoretical computer Science, cybernetics. Moscow. 2022. 106 p.
- [14] M. Jain, C. Dovrolis, "End-to-End available bandwidth: Measurement methodology, dynamics, and relation with TCP throughput [J]," *IEEE/ACM Transactions on Networking*, 2003, no. 11 (4), pp. 537-549.
- [15] A.N. Volkov, "Development and research of a set of models and methods for building communication networks based on cloud computing and telepresence services," - 2.2.15. Telecommunications systems, networks and devices ... abstract of the dissertation. Doctors of Technical Sciences. Saint Petersburg. 2024. - 54 p.
- [16] M. Fletcher, E. Paulz, D. Ridge, A. Michaels, "Low-Latency Wireless Network Extension for Industrial Internet of Things," *Sensors*, 2024, 24, 2113, 16 p.
- [17] D.O. Kupriyanov, "Mathematical modeling of the flow of applications to a cloud computing cluster," *T-Comm*. 2020. Vol. 14, No. 10, pp. 39-44.
- [18] U. Arora, N. Singh, "IoT application modules placement in heterogeneous fog-cloud infrastructure," *International Journal of Information Technology (Singapore)*. 2021. Vol. 13, No. 5, pp. 1975-1982.
- [19] A. Golyshko, "IoT in the clouds, beyond the fog and on the border," *CONNECT*. No. 9-10, 2024, pp. 110-113.
- [20] A.N. Volkov, "The routing problem in a network of dynamic fog computing," *Proceedings of educational institutions of communications*. 2024. Vol. 10. No. 4, pp. 27-37.
- [21] "Theory of automatic control" Edited by V.B. Yakovlev. Moscow: Vysshaya shkola Publ., 2009. 568 p.

Information about authors:

Elena V. Glushak, Povolzhskiy State University of Telecommunications and Informatics, Associate Professor, PhD, Associate Professor of the Department of Networks and Communication Systems, Samara, Russia, ORCID 0009-0000-5494-9746

Dmitry S. Klyuev, Povolzhskiy State University of Telecommunications and Informatics, Professor, Ph.D., Head of the Department of Radioelectronic Systems, Samara, Russia, ORCID 0000-0002-9125-7076

Vladimir I. Volovach, Povolzhskiy State University of Service, Associate Professor, Doctor of Technical Sciences, Acting Director of the Higher School of Advanced Production Technologies, Tolyatti, Russia, ORCID 0000-0002-0201-2545

ОПТИЧЕСКИЕ ВЫЧИСЛИТЕЛИ НЕЛИНЕЙНЫХ ФУНКЦИЙ

Каменский Владислав Валерьевич,
Ростовский государственный университет путей
сообщения", г. Ростов-на-Дону, Россия,
kam-vladislav@yandex.ru

DOI: 10.36724/2072-8735-2026-20-3-35-42

Manuscript received 20 December 2025;
Accepted 26 February 2026

Соколов Сергей Викторович,
Московский технический университет связи и
информатики, Москва, Россия,
s.v.s.888@yandex.ru

Лагунова Елена Олеговна,
Ростовский государственный университет путей
сообщения", г. Ростов-на-Дону, Россия,
lagunova@rambler.ru

Ключевые слова: оптическое вычислительное устройство, извлечение квадратного корня, возведение в степень, гигагерцовый диапазон, оптический разветвитель, транспарант, оптический объединитель

В работе рассмотрен новый подход к конструированию средств высокопроизводительной вычислительной техники на основе оптоэлектронных технологий. Оптические устройства обработки информации применяются в областях, где требуется максимальное быстродействие и простота вычислений, например, при синтезе встраиваемых систем и специализированных вычислителей для обработки высокоскоростных процессов. В общем случае встраиваемые системы обрабатывают поступающие данные вполне определенного и заранее известного диапазона, при этом вычисления допускается выполнять с заданной точностью, но время вычисления весьма ограничено и критично. В связи с этим, для разрабатываемого оптического вычислителя, в результате проведения анализа существующих форматов представления чисел, был выбран формат представления чисел с фиксированной запятой. В качестве примера реализуемой нелинейной функции была выбрана одна из наиболее сложных функций для вычисления в традиционных схемах - функция извлечения квадратного корня. Проведен анализ существующих алгоритмов извлечения квадратного корня как базовой операции многих математических вычислений, находящихся свое практическое применение в геометрии, архитектуре, строительстве и пр. Изложены теоретические основы построения быстродействующего вычислителя квадратного корня для чисел, представленных в двоичном коде. Предложенный подход позволяет производить вычисление квадратного корня на основе комбинационного устройства, без использования метода проб и ошибок, что радикально увеличивает скорость вычислений. Так как при решении практических задач необходимо за минимальное время производить целый набор вычислений над несколькими аргументами, то с этой целью разработана функциональная схема специализированного оптического программируемого устройства для выполнения набора функциональных преобразований на основе принципов построения и оптимизации комбинационных схем. Основными базовыми элементами оптического программируемого устройства являются оптические разветвители, объединители и транспаранты, технология изготовления которых известна и не вызывает трудностей. Программирование оптического вычислителя для реализации конкретной функции осуществляется разрушением оптических связей, не участвующих в формировании логических функций. Минимальное количество элементов и высокая скорость вычислений делают оптическое программируемое устройство перспективным при построении быстродействующих технических средств вычислительной техники и систем связи.

Информация об авторах:

Каменский Владислав Валерьевич, доцент, к.т.н., ФГБОУ ВО "Ростовский государственный университет путей сообщения", Ростов-на-Дону, Россия, ORCID: 0000-0002-0704-8686

Соколов Сергей Викторович, профессор, д.т.н., ФГБОУ ВО "Московский технический университет связи и информатики", Москва, Россия, ORCID: 0000-0002-5246-841X

Лагунова Елена Олеговна, профессор, д.т.н., ФГБОУ ВО "Ростовский государственный университет путей сообщения", г. Ростов-на-Дону, Россия, ORCID: 0000-0002-2761-8068

Для цитирования:

Каменский В.В., Соколов С.В., Лагунова Е.О. Оптические вычислители нелинейных функций // Т-Сотт: Телекоммуникации и транспорт. 2026. Том 20. №3. С. 35-42.

For citation:

V.V. Kamenskij, S.V. Sokolov, E.O. Lagunova, "Fast optical computers," *T-Comm*, 2026, vol. 20, no.3, pp. 35-42. (in Russian)

Введение

Объемы информационных потоков, генерируемых, передаваемых и потребляемых человечеством, неуклонно растут с каждым годом, что требует все возрастающих вычислительных мощностей. Дальнейшее увеличение вычислительных возможностей при обработке информации традиционными (электронными устройствами) невозможно. Это обусловлено невозможностью преодолеть термодинамический, квантово-механический предел и предел скорости распространения сигналов. Что, в свою очередь, приводит к поиску новых подходов при построении вычислительных устройств, базирующихся на принципиально новых идеях и концепциях. В частности, оптических систем, не обладающих недостатками электронных устройств и имеющих высокое быстродействие при возможности микроминиатюризации и минимальном энергопотреблении.

Вычисление нелинейных функций является сложной задачей, решение которой основывается на аналитических или итерационных методах. Одной из таких задач является извлечение корня.

Наибольшее применение операция извлечения корня получила в геометрии и связанных с ней областях практического применения в архитектуре и строительстве.

Методы вычисления корней (квадратных, кубических и высших) включают итеративные алгоритмы, такие как метод Ньютона-Рафсона, бинарный поиск и специализированные схемы для многочленов. Современные подходы фокусируются на ускорении сходимости, аппаратной реализации и обработке множественных корней.

Метод Ньютона-Рафсона остается основным для приближенного вычисления корней уравнений $f(x)=0$, обеспечивая квадратичную сходимость при хорошем начальном приближении [1].

Вавилонский (Геронский) метод специально для квадратных корней использует итерацию $x_{n+1} = \left(x_n + \frac{a}{x_n} \right) / 2$, показывая быструю сходимость для положительных a . Эти методы модифицированы для численной устойчивости в многомерных задачах [2].

Для корней многочленов применяют схему Горнера для оценки, метод Лобачевского для интервалов и итеративное разложение на множители, эффективно выявляя комплексные и кратные корни. Новые алгоритмы сочетают эти подходы с машинным обучением для предсказания интервалов. Аппаратные реализации на ПЛИС ускоряют вычисления сигмоидов и корней в реальном времени [3, 4].

Авторы [5] исследуют эффективность различных методов извлечения корней и делают сравнительный анализ преимуществ и недостатков каждого метода. В работе [6] представлены новейшие разработки и идеи по извлечению корней, демонстрируется использование графических моделей и компьютеров для оптимизации решений. В [7] изложено развитие методов извлечения корней с демонстрацией современных достижений математики в этой сфере.

В двоичной системе описаны схемы для квадратных и кубических корней, аналогично делению столбиком [8-17].

Операция извлечения квадратного корня в двоичном формате может быть решена путем последовательного подбора

цифр результата, с использованием регистров и сумматоров. Метод реализуется путем применения сдвигов и вычитаний. Вторым способом является табличный способ. Заранее рассчитываются результаты вычислений и заносятся в таблицу. При необходимости вычислить корень, в таблице находится нужное значение «число в степени», а результат основание степени выбирается из таблицы.

Материалы и методы

Методы построения описываемых в данной статье оптических вычислителей основаны на оптических технологиях. Оптические технологии – это направление науки и техники, которое использует оптические сигналы для реализации ключевых информационных процессов. Преимуществами оптических технологий является сверхвысокая скорость передачи информации, минимальная мощность передаваемого сигнала и минимальные потери сигнала, позволяющие разрабатывать быстродействующие оптические устройства со сверхнизким потреблением энергии.

В настоящий момент оптические технологии получили распространение в устройствах телекоммуникации и связи, позволяя передавать информацию на дальние расстояния с малыми задержками. В оптических приемо-передающих системах передача информации осуществляется практически без потерь, обусловленных электромагнитными помехами, обеспечивая высокоскоростной обмен данными. Еще одним важным свойством оптических систем является возможность реализации параллельных вычислений.

Сегодняшние решения – микролинзы, дифракционные решетки, фотонные кристаллы, интегральные оптические схемы и метаматериалы – отличаются микронными габаритами и прецизионной точностью, открывая путь к разработке компактных и высокопроизводительных информационных устройств.

Источниками оптического излучения могут быть лампы, светодиоды и лазеры. Спектральный диапазон таких источников охватывает интервал от 10 нм до 1 мм. Также существуют источники оптического излучения на основе фотонных кристаллов, квантовых точек, суперлюминесцентных диодов и др.

Во всех оптических устройствах традиционно используются три вида базовых оптических элементов, составляющих его основу [18]. Это оптические волноводы, оптические разветвители и оптические объединители.

Перечисленные оптические компоненты – волноводы, разветвители и объединители – выполняют исключительно функции передачи, распределения и объединения оптических сигналов. Они не обладают способностью реализовывать логические операции (И, ИЛИ, НЕ и др.). Такими устройствами могут быть трансфазоры, оптически связанные волноводы. Интенсивность сигналов на выходе этих элементов скачкообразно изменяется при плавном изменении сигнала на входе. Схема и подробное функционирование оптических логических элементов описаны в [19-21].

Выбор формата данных

В электронных вычислительных машинах целые числа X , состоящие из N разрядов, хранятся в виде двоичных чисел:

$$X = 2^{N-1}x_{N-1} + \dots + 2^M x_M + \dots + 2^1 x_1 + 2^0 x_0,$$

что позволяет хранить очень большие числа. Несмотря на это, такой тип хранения данных имеет существенный недостаток: отсутствует возможность хранения дробных значений.

Так как при вычислении квадратного корня могут быть получены дробные значения, то целочисленные типы чисел не подходят для операций вычисления корня.

Для этих целей используют числа с плавающей запятой:

$$X = a \cdot q^N$$

состоящие из а-мантисы и N-порядка.

Операции с целыми числами производятся быстро, а операции с числами с плавающей запятой выполняются значительно дольше и требуют специальных аппаратных устройств. Поэтому во встраиваемых системах, обрабатывающих данные вполне определенного и заранее известного диапазона, используются числа с фиксированной запятой. Числа с фиксированной запятой содержат, как целую I, так и дробную часть F:

$$X = I + F,$$

где

$$I = 2^{n-1}x_{n-1} + 2^i x_i + \dots + 2^1 x_1 + 2^0 x_0,$$

$$F = 2^{1/2}x_{f-1} + 2^{1/4}x_{f-2} \dots + 2^{1/f}x_0$$

В результате число с фиксированной запятой имеет вид

$$X = 2^{n-1}x_{n-1} + \dots + 2^1 x_{m+1} + 2^0 x_m + 2^{-1}x_{m-1} + 2^{-2}x_{m-2} \dots + 2^{-f}x_0$$

Положение запятой определяется количеством бит как целой I, так и дробной частей F. В приведенном примере фиксированная запятая располагается между разрядами x_m и x_{m-1}

Для обозначения положения фиксированной запятой используют количество бит в целой части и количество бит в дробной части $Q_i.f$. Количество бит в целой и дробной части может совпадать (Q4.4, Q8.8, Q16.16), а может и отличаться (Q8.16, Q3.5). Например, десятичному числу 1.4375 в формате числа с фиксированной запятой Q4.4 будет соответствовать двоичное число 0001.0111.

Оптическое комбинационное вычислительное устройство

Рассмотрим далее описание структуры и принципа действия оптического вычислителя.

При небольшом диапазоне значений X вычисление корня может осуществляться табличным методом.

Устройство, реализующее табличный метод, состоит из дешифратора, ячеек памяти и схемы коммутации выходных сигналов. Дешифратор осуществляет выбор одной ячейки памяти со значением результата. В ячейку памяти заранее записывается значение вычисленного корня. Схема коммутации выходных сигналов позволяет передать на выход устройства сигналы только одной ячейки памяти.

Другим, более простым, способом получения результата операции извлечения квадратного корня является построение схемы вычислительного устройства как комбинационной схемы.

Это возможно благодаря тому, что состояние каждого выхода вычислительного устройства однозначно определяется входными сигналами:

$$Y_i = f(x_{n-1}, \dots, x_i, \dots, x_1, x_0)$$

Запишем, например, для входных сигналов x_2, x_1, x_0 значения выходных сигналов $y_7, y_6, y_5, y_4, y_3, y_2, y_1, y_0$ в формате с фиксированной запятой Q4.4 и в двоичном виде $M = \{0000.0000, 0001.0000, 0001.0111, 0001.1100, 0010.0000, 0010.0100, 0010.0111, 0010.1010\}$.

Логические функции для выходных сигналов $y_7, y_6, y_4, y_3, y_2, y_1, y_0$, определяющих целую часть, в зависимости от входных сигналов x_2, x_1, x_0 представлены ниже

$$y_7 = 0$$

$$y_6 = 0$$

$$y_5 = x_2 \bar{x}_1 \bar{x}_0 \vee x_2 \bar{x}_1 x \vee x_2 x_1 \bar{x} \quad x_2 \bar{x}_1 x = x_2$$

$$y_4 = \bar{x}_2 \bar{x}_1 x \vee \bar{x}_2 x_1 \bar{x}_0 \vee \bar{x}_2 x_1 x$$

Дробную часть результата определяют выходные сигналы y_3, y_2, y_1, y_0 . Логические функции для выходных сигналов y_3, y_2, y_1, y_0 равны

$$y_3 = \bar{x}_2 x_1 x \vee x_2 x_1 x_0$$

$$y_2 = \bar{x}_2 x_1 \bar{x}_0 \vee \bar{x}_2 x_1 x \vee x_2 \bar{x}_1 x \vee x_2 x_1 \bar{x}_0$$

$$y_1 = \bar{x}_2 x_1 \bar{x}_0 \vee x_2 x_1 \bar{x}_0 \vee x_2 x_1 x$$

$$y_0 = \bar{x}_2 x_1 \bar{x} \vee x_2 x_1 \bar{x}_0$$

Запишем логические функции для выходных сигналов $y_5, y_4, y_3, y_2, y_1, y_0$ в виде чисел на входе, при котором функция принимает состояние единицы:

$$y_5 = 4, 5, 6, 7$$

$$y_4 = 1, 2, 3$$

$$y_3 = 3, 7$$

$$y_2 = 2, 3, 5, 6$$

$$y_1 = 2, 6, 7$$

$$y_0 = 2, 6$$

Отметим, что значения логической функции y_5 полностью соответствует входному сигналу x_2 , и результат упрощения это подтверждает:

$$\begin{aligned} y_5 &= x_2 \bar{x}_1 \bar{x}_0 \vee x_2 \bar{x}_1 x_0 \vee x_2 x_1 \bar{x}_0 \vee x_2 x_1 x_0 = \\ &= x_2 (\bar{x}_1 \bar{x}_0 \vee \bar{x}_1 x_0 \vee x_1 \bar{x}_0 \vee x_1 x_0) = x_2 \cdot 1 = x_2 \end{aligned}$$

В общем виде оптическое устройство извлечения квадратного корня строится на основе оптических инверторов $I_0 \dots I_i$, оптических конъюнкторов $F_0 \dots F_{i-1}$ и оптических дизъюнкторов $V_0 \dots V_{i-1}$ и работает следующим образом.

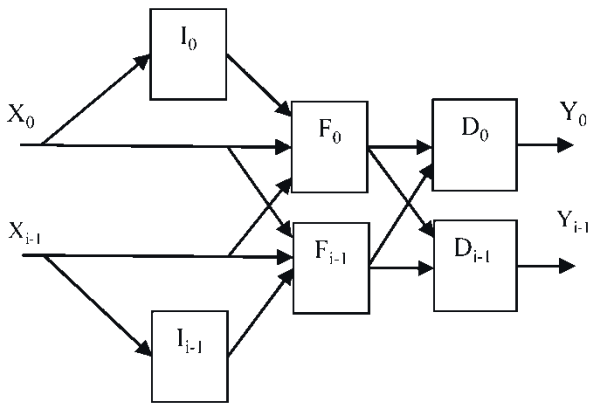


Рис. 1. Оптическое комбинационное вычислительное устройство

Входные оптические сигналы $x_{i-1} \dots x_0$ подаются на входы оптических конъюнкторов $F_0 \dots F_{i-1}$ и при необходимости – на входы оптических инверторов I_i . Оптические конъюнкторы $F_0 \dots F_{i-1}$ осуществляют операции логического умножения «И» в соответствии с логическими функциями для выходных сигналов $y_7, y_6, y_5, y_4, y_3, y_2, y_1, y_0$.

Количество входов каждого оптического конъюнктора $F_0 \dots F_{i-1}$ может быть от 1 до m . Для формирования логической функции одного выхода y_i понадобится от 1 до 2^m оптических конъюнкторов $F_0 \dots F_{i-1}$. С выходов оптических конъюнкторов $F_0 \dots F_{i-1}$ оптические сигналы $p_m \dots p_0$ подаются на входы оптических дизъюнкторов $B_0 \dots B_{i-1}$. Если результат конъюнкции P_i используется при формировании выходной функции Y_i , то выход оптического конъюнктора F_i подключается ко входу соответствующего оптического дизъюнктора B_i .

Таким образом, при подаче на вход устройства оптических сигналов $x_{i-1} \dots x_0$, на выходе устройства будут появляться оптические сигналы $y_7, y_6, y_5, y_4, y_3, y_2, y_1, y_0$, соответствующие результату вычисления квадратного корня числа, поданного на вход устройства.

Оптическое программируемое вычислительное устройство

При решении практических задач необходимо за минимальное время вычислять не одну единственную функцию, а произвести целый набор вычислений над несколькими аргументами.

Например, сложная функция может иметь вид:

$$y = f(h(g(a), g(b)))$$

В первую очередь вычисляются значения функций $ta=g(a)$ и $tb=g(b)$.

Результаты вычисления ta и tb становятся аргументами для функции $q=h(ta, tb)$ и в последнюю очередь вычисляется функция $y=f(q)$.

Допустим, требуется выполнить несколько действий над двумя числами a и b .

$$c = \sqrt{a^2 + b^2}$$

Традиционный подход заключается в следующем:

1. Разрабатывается устройство, обеспечивающее возведение числа в квадрат.

2. Разрабатывается устройство суммирования результата возведения первого и второго числа в квадрат (в традиционных способах суммирование выполняется с использованием переносов, что приводит к значительным временным затратам).

3. Разрабатывается устройство извлечения квадратного корня.

4. Разрабатывается общее устройство на основе ранее разработанных элементов.

Таким образом, результат вычисляется в виде нескольких последовательных операций. Причем, каждое следующее действие осуществляется только после полного завершения расчета предыдущего значения. Устройство, реализующее каждое действие, разрабатывается и оптимизируется независимо от остальных. В результате, у таких устройств наблюдается большое время вычисления результата за счет последовательного вычисления и переизбытка логических элементов по причине отсутствия оптимизации между ними.

Рассмотрим другой подход вычисления сложной функции. Подход основан, во-первых, на проектировании устройства вычисления конкретной сложной функции, а не взаимосвязанного набора различных специализированных устройств, а во-вторых, на использовании результата, а не вычислении значения путем пробных значений.

Введем некоторые ограничения. Числа, подаваемые на вход, a и b – целые. Требуется получить результат в формате с фиксированной запятой (точкой) Q4.4.

Например, для входных сигналов a_1, a_0, b_1, b_0 , которые могут принимать значения $\{0000, 0001, \dots, 1111\}$ вычислим значения выходных сигналов в десятичном виде $D = \{0.0000, 1.0000, 2.0000, 3.0000, 1.0000, 1.4142, 2.2361, 3.1623, 2.0000, 2.2361, 2.8284, 3.6056, 3.0000, 3.1623, 3.6056, 4.2426\}$. Затем переведем их в двоичный формат с фиксированной запятой Q4.4 $M[\] = \{0000.0000, 0001.0000, 0010.0000, 0011.0000, 0001.0000, 0001.0111, 0010.0100, 0011.0011, 0010.0000, 0010.0100, 0010.1110, 0011.1010, 0011.0000, 0011.0011, 0011.1010, 0100.0100\}$.

Для каждого разряда целой части запишем логические функции для выходных сигналов c_7, c_6, c_5, c_4 , определяющих целую часть в зависимости от входных сигналов a_1, a_0, b_1, b_0

$$c_7 = 0$$

$$c_6 = a_1 a_0 b_1 b_0$$

$$c_5 = \bar{a}_1 b_1 \vee a_1 \bar{b}_1 \vee \bar{a}_0 b_1 \vee b_1 \bar{b}_0$$

$$c_4 = \bar{a}_1 b_0 \vee a_0 \bar{b}_1 \vee \bar{a}_0 b_1 b_0 \vee a_1 a_0 \bar{b}_0$$

Логические функции для выходных сигналов дробной части результата c_3, c_2, c_1, c_0 представлены ниже

$$c_3 = a_1 \bar{a}_0 b_1 \vee a_1 b_1 \bar{b}_0$$

$$c_2 = \bar{a}_1 a_0 \bar{b}_1 b_0 \vee \bar{a}_1 a_0 b_1 \bar{b}_0 \vee a_1 \bar{a}_0 \bar{b}_1 b_0 \vee a_1 \bar{a}_0 b_1 \bar{b}_0 \vee a_1 a_0 b_1 b_0$$

$$c_1 = \bar{a}_1 a_0 b_0 \vee a_1 \bar{a}_0 \bar{b}_1 \vee a_0 \bar{b}_1 b_0 \vee a_1 b_1 \bar{b}_0$$

$$c_0 = \bar{a}_1 a_0 b_0 \vee a_0 \bar{b}_1 b_0$$

Функциональная схема оптического вычислителя

Так, функциональная схема оптического вычислителя, осуществляющего вычисление трех арифметических функций, состоит из:

- 1) оптических инверторов I_i , $i=0,1,2,\dots,m$;
- 2) оптических конъюнкторов F_i , $i=0,1,2,\dots,m$;
- 3) оптических транспарантов B_i , $i=0,1,2,\dots,m$;
- 4) оптических разветвителей S_i , $i=0,1,2,\dots,m$;
- 5) оптических дизъюнкторов D_i , $i=0,1,2,\dots,m$;

Входные сигналы $X_0 \dots X_m$ поступают на входы оптических инверторов I_i и оптических конъюнкторов F_i . Выходы оптических конъюнкторов F_i подключены к входам оптических транспарантов B_i . Выходы оптических транспарантов B_i через оптические разветвители подключены ко входам оптических дизъюнкторов D_i .

Особенности оптических конъюнкторов B_i в том, что единичный сигнал на выходе составляет K усл. ед.

Оптические транспаранты имеют такой коэффициент поглощения, что оптические сигналы на выходах каждого оптического разветвителя S_i равны 1 усл. ед.

Особенности оптических дизъюнкторов D_i состоят в том, что единичный сигнал на его входе интенсивностью больше 0,5 усл. ед. приводит к появлению сигнала на выходе 1 усл. ед.

Оптические транспаранты B_i , оптические разветвители S_i и оптические дизъюнктеры D_i составляют оптическую логическую матрицу.

Для реализации требуемых логических зависимостей между входами и выходами устройства осуществляется разрушение оптических связей в области O_1 между входами устройства или выходами оптических инверторов и входами оптических конъюнктеров F_i , а также разрушение связей в области O_2 между выходами оптических разветвителей S_i и входами оптических дизъюнктеров D_i .

Принцип действия оптического вычислителя

На входы оптического устройства, осуществляющего вычисление трех арифметических функций, подаются числа A и B . Из чисел A и B формируется число двойной разрядности X , так, что старшая часть числа XH соответствует числу A , а младшая часть числа XL соответствует числу B .

Полученное число X подается на вход устройства в двоичном позиционном коде " $x_m x_{m-1} \dots x_0$ " в виде набора из m бит x_i :

$$X = \sum_{i=0}^{m-1} x_i 2^i$$

Входные оптические сигналы с интенсивностью 1 усл. ед. (овных) ед. (иниц) поступают на входы оптических конъюнктеров F_i .

Рассмотрим формирование выходных сигналов для каждого выхода c_7, c_6, c_5, c , определяющего целую часть.

Сигнал c_7 при указанном выше наборе входных сигналов всегда равен нулю. Поэтому разрушаются все связи между выходами оптических разветвителей S_i и входами оптического дизъюнктера D_7 .

Для формирования сигнала c_6 используется оптический конъюнктер F_0 . На входы конъюнктера F_0 подаются четыре сигнала a_1, a_0, b_1, b_0 . Сигнал c_6 принимает значение единицы, если на всех входах устройства присутствуют единицы. При

наличии на входах оптического конъюнктера F_0 комбинации сигналов

$$a_1, a_0, b_1, b_0 = 1, 1, 1, 1$$

оптический конъюнктер F_0 выдает оптический сигнал в K условных единиц на свой выход. Пройдя оптический разветвитель S_i , оптический сигнал уменьшится и с интенсивностью 1 усл. ед. поступит на вход оптического дизъюнктера D_6 . В данном случае кроме конъюнктера F_0 больше ни один конъюнктер F_i не выдает сигнал высокого уровня интенсивности K усл. ед. на свой выход, остальные сигналы на выходах конъюнктеров F_i равны 0. Так как интенсивность сигнала на входе дизъюнктера D_6 больше 0,5, то интенсивность сигнала на его выходе будет равна 1 усл. ед. В результате интенсивность сигнала на входе дизъюнктера D_6 будет равна 1.

$$c_6 = D_6(F_0(a_1 a_0 b_1 b_0))$$

Для формирования сигнала c_5 используется 4 конъюнктера $F_1 \dots F_4$. На входы конъюнктеров $F_1 \dots F_4$ входные сигналы подаются парами. На входы конъюнктера F_1 подаются два сигнала: инвертированный сигнал a_1 и сигнал b_1 . На входы конъюнктера F_2 подаются сигналы a_1 и b_1 , но в этом случае сигнал a_1 не инвертированный, а сигнал b_1 инвертированный. Также по два сигнала подаются на конъюнктеры F_3 и F_4 .

Неиспользуемые сигналы на входы конъюнктеров $F_1 \dots F_4$ не подаются. Для этого разрушаются оптические связи между входами устройства, выходами инверторов и входами конъюнктеров $F_1 \dots F_4$:

$$c_5 = D_5(F_1(\bar{a}_1 b_1), F_2(a_1 \bar{b}_1), F_3(\bar{a}_0 b_0), F_4(b_0 \bar{b}_0))$$

Оптический конъюнктер F_1 выдает оптический сигнал в K условных единиц на выход, если у него на входе комбинация сигналов равна 01, а оптические конъюнктеры F_2, F_3, F_4 выдают оптический сигнал в K условных единиц на свой выход, если у них на входе присутствует соответственно комбинация сигналов 10, 01 и 10. Пройдя оптический разветвитель S_i , оптический сигнал уменьшится и с интенсивностью 1 усл. ед. поступит на вход оптического дизъюнктера D_5 . В данном случае один из конъюнктеров $F_1 \dots F_3$ выдает сигнал высокого уровня интенсивности K усл. ед. на свой выход. Так как интенсивность сигнала на входе дизъюнктера D_5 больше 0,5, то интенсивность сигнала на его выходе будет равна 1 усл. ед. В результате интенсивность сигнала на входе дизъюнктера D_5 будет равна 1.

Для формирования сигнала c_4 используется 4 конъюнктера $F_5 \dots F_8$. На входы конъюнктеров F_5 и F_6 входные сигналы подаются парами. На входы конъюнктера F_5 подаются сигналы: инвертированный a_1 и b_0 , а на входы конъюнктера F_6 подаются сигналы a_0 и инвертированный сигнал b_0 .

На входы конъюнктеров F_7 и F_8 подается по три входных сигнала. На входы конъюнктера F_7 подается инвертированный сигнал a_0 и сигналы b_1 и b_0 , а на входы конъюнктера F_8 подаются сигналы a_1, a_0 и инвертированный сигнал b_0 :

$$c_4 = D_4(F_5(\bar{a}_1 b_0), F_6(a_0 \bar{b}_0), F_7(\bar{a}_0 b_1 b_0), F_8(a_1 a_0 \bar{b}_0))$$

Если на входе устройства присутствуют необходимые оптические сигналы, то оптический конъюнктер F_i выдает оптический сигнал в K условных единиц на свой выход.

Схема формирования сигналов дробной части c_3, c_2, c_1, c_0 работает аналогично.

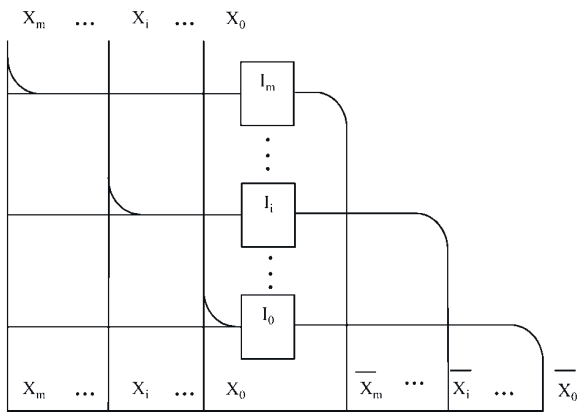


Рис. 2. Оптические инверторы

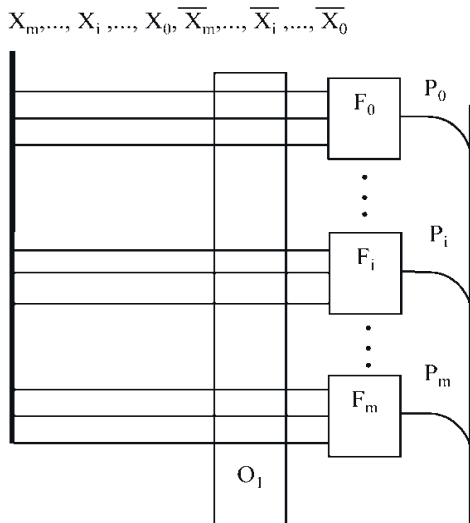


Рис. 3. Оптические конъюнкторы

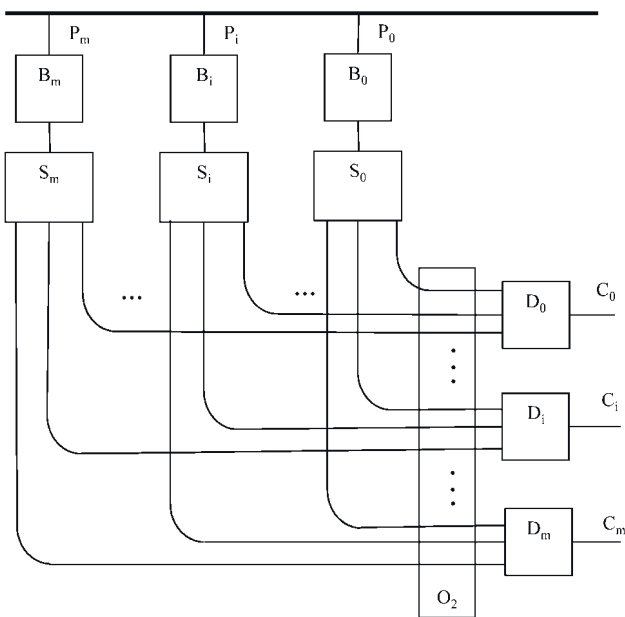


Рис. 4. Оптическая логическая матрица

Программирование оптического вычислителя

В исходном состоянии все оптические связи со входами оптических конъюнкторов F_i не разрушены и на входы оптических конъюнкторов F_i подаются все входные сигналы $X_0 \dots X_m$ и все сигналы с выходов оптических инверторов I_i . При программировании оптического вычислителя для реализации конкретной функции, во-первых, производится разрушение оптических связей, обеспечивающих прохождение оптических сигналов на входы оптических конъюнкторов F_i для сигналов, не являющихся аргументами логических функций, а во-вторых, производится разрушение оптических связей, обеспечивающих прохождение оптических сигналов с выходов оптических разветвителей S_i на входы оптических дизъюнкторов D_i .

Например, для формирования сигнала C_6 используется оптический конъюнктор F_0 , на который до программирования подаются сигналы $a_1, a_0, b_1, b_0, \bar{a}_1, \bar{a}_0, \bar{b}_1, \bar{b}_0$. Сигналы $\bar{a}_1, \bar{a}_0, \bar{b}_1, \bar{b}_0$ при формировании функции данного разряда не используются, поэтому оптические связи между выходами инверторов и входами оптического конъюнктора F_0 разрушаются. Кроме этого, для формирования сигнала C_6 используется оптический дизъюнктор D_0 , на который до программирования подаются сигналы с выходов оптических разветвителей $S_0 \dots S_m$. Оптические связи между выходами оптических разветвителей $S_j \dots S_m$ и входами оптического дизъюнктора D_0 не используются при формировании функции данного разряда, поэтому эти оптические связи также разрушаются.

Таким образом, метод разработки оптического программируемого вычислителя состоит из следующих этапов:

- исходя из параметров технической задачи, определяется сложная функция, для которой требуется производить вычисления;
- определяется диапазон входных аргументов;
- определяется требуемый диапазон выходных значений;
- определяется необходимая точность вычислений;
- определяется требуемое количество бит в целой и дробной части;
- производится вычисление значений функции;
- производится перевод значений из десятичной системы счисления в двоичную с фиксированной запятой;
- для каждого выходного бита определяется зависимость выходного значения от входных значений;
- производится оптимизация функций на основе законов алгебры логики, методами Карно и др;
- осуществляется проверка полученных зависимостей;
- осуществляется построение оптической схемы путем удаления неиспользуемых оптических связей.

Пример. Для проверки правильности логических функций была написана программа на языке Си++. Программа осуществляет перебор всех входных комбинаций. Для каждой входной комбинации сигналов a_1, a_0, b_1, b_0 осуществляется вычисление логических функций для каждого выхода устройства $c_6, c_5, c_4, c_3, c_2, c_1, c_0$.

```
for(i=0; i<=15; i++)
{
    a1 = (i&0x08) >> 3; a0 = (i&0x04) >> 2;
    b1 = (i&0x02) >> 1; b0 = (i&0x01);
```

```

c6 = a1&a0&b1&b0;
c5 = (~a1&b1) | (a1&~b1) | (~a0&b1) |
(b1&~b0);
c4 = (~a1&b0) | (a0&~b1) | (~a0&b1&b0) |
(a1&a0&~b0);
c3 = (a1&~a0&b1) | (a1&b1&~b0);
c2 = (~a1&a0&~b1&b0) | (~a1&a0&b1&~b0) |
(a1&~a0&~b1&b0) | (a1&~a0&b1&~b0)
| (a1&a0&b1&b0);
c1 = (~a1&a0&b0) | (a1&~a0&b1) |
(a0&~b1&b0) | (a1&b1&~b0);
c0 = (~a1&a0&b0) | (a0&~b1&b0);

std::cout << a1 << " " << a0 << " " <<
b1 << " " << b0 << " ";
std::cout << c5 << " " << c4 << " " <<
c3 << " " << c2 << " " << c1 << " " <<
<< c0 << " \n";
}

```

Реализация данной программы подтвердила эффективность предложенного подхода.

Заключение

Проведен обзор существующих алгоритмов извлечения квадратного корня как базовой операции многих математических вычислений.

В ходе исследования изложены теоретические основы построения быстродействующего вычислителя квадратного корня для чисел, представленных в двоичном коде. Предложенный способ позволяет производить вычисление квадратного корня на основе комбинационного устройства, без использования метода проб и ошибок, что значительно увеличивает скорость вычислений при получении результата. Устройство может быть легко модифицировано для вычисления других функций или набора функций.

Разработана функциональная схема оптического программируемого устройства для выполнения функциональных преобразований. Рассмотрена работа составных элементов и всего устройства в целом.

Изложены принципы программирования оптического вычислителя для реализации конкретной функции.

Минимальное количество элементов и высокая скорость вычислений делают предложенное оптическое программируемое устройство перспективным при построении технических средств вычислительной техники и систем связи.

Литература

1. Стаценко И.В. Исследование скорости сходимости одного обобщенного ньютоновского метода и классического метода Ньютона в процедуре уточнения корней многочлена // Точная наука. 2020. №78. С. 2-9.

2. Мачарадзе Г.Т., Морозова Е.А. Алгоритм быстрого вычисления квадратного корня и его актуальность // Вестник Российского нового университета. 2018. С. 106-109 DOI: 10.25586/RNU.V9187.18.04.P.106.

3. Ушенина И.В. Реализация на ПЛИС и сравнительный анализ вычислителей сигмоида, работающих с полным диапазоном аргумента с учетом симметрии // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2025. №71. С. 120-129. doi: 10.17223/19988605/71/1.

4. Гаишков С.Б., Фролов А.Б., Попова Е.П. Об оценке сложности алгоритмов извлечения квадратных корней в конечных полях и кольцах вычетов // Вестник МЭИ. №5. 2018. С. 79-88.

5. Фёдоров А., Васнецова Т. Анализ эффективности методов извлечения корней // Вопросы математики, выпуск 1, 2023. С. 15-25.

6. Соколов Д. Новые подходы к извлечению корней // Труды Международной конференции по математике, 2024. С. 35-45.

7. Краснова А. Извлечение корней: эволюция методов // Электронный журнал «Математика XXI века», 2025. С. 12-20.

8. Златопольский Д.М. Методика извлечения квадратного и кубического корней в двоичной системе исчисления // Информатика в школе. 2021. № 1(1). С. 42-45 [https:// doi.org/10.32517/2221-1993-2021-20-1-42-45](https://doi.org/10.32517/2221-1993-2021-20-1-42-45).

9. Николаев С. Классические и современные методы извлечения корней. М.: Издательство: Мир науки, 2024, 320 с.

10. Сергеева О. Простота и сложность извлечения корней. М.: Издательство: Альфа-Бук, 2025, 200 с.

11. Andraos K., Ghiasi N.M. A High-Throughput FPGA Implementation of a Digit-Recurrence Square Root Unit with BFloat16 Support // IEEE Transactions on Computers. 2024. Vol. 73, No. 5, pp. 1245-1258. DOI: 10.1109/TC.2024.3355678

12. de Aledo M.R.G., de Guzmán P.L. A Formal Verification of the Fast Inverse Square Root Algorithm in the Coq Proof Assistant // Journal of Automated Reasoning. 2023. Vol. 67, Issue 4, Article 28. DOI: 10.1007/s10817-023-09670-6

13. Gupta S., Martinez A., Roberts M. Numerical Stability Analysis of Iterative Methods for Square Root in Mixed-Precision ML Training // Proceedings of the Machine Learning and Systems Conference (MLSys). 2024. Vol. 6, pp. 45-59.

14. Kaur J., Nguyen T.H., Bruguera J.D. Efficient Square Root and Inverse Square Root for the Posit Number System // IEEE Transactions on Emerging Topics in Computing. 2023. Vol. 11, No. 2, pp. 456-469 DOI: 10.1109/TETC.2023.3245671

15. Petrov D., Takahashi Y., Mosca M. A Scalable Quantum Algorithm for Square Root and Division with Arbitrary Precision // Quantum. 2024. Vol. 8, p. 1245. DOI: 10.22331/q-2024-01-08-1245

16. Wang C., Kim Y., Fahmy S.A. Hardware Optimizations of the Newton-Raphson Square Root for 3D Graphics Pipelines // ACM Transactions on Graphics (TOG). 2024. Vol. 43, No. 4, Article 78 DOI: 10.1145/3654662

17. Zhang F., Li H., McCluske E.J. Error Analysis and Correction for Digit-by-Digit Square Root in Safety-Critical Systems // IEEE Transactions on Computers. 2023. Vol. 72, No. 9, pp. 2541-2554. DOI: 10.1109/TC.2023.3267890

18. Pshko A. Mohammed, Integration of self-standing X- and Y-shaped polymer coupler and splitter with single mode optical fiber – Optical Materials <https://doi.org/10.1016/j.optmat.2020.110685>

19. Каменский В.В., Соколов С.В. Синтез преобразователей кодов на основе оптических технологий // Радиотехника. 2023. Т. 87. №11. С. 202-209.

20. Каменский В.В., Соколов С.В., Лагунова Е.О. Оптический цифро-аналоговый вычислитель // Т-Comm: Телекоммуникации и транспорт. 2025. Т. 19. № 4. С. 37-44.

21. Каменский В.В., Соколов С.В., Лагунова Е.О. Метод построения оптических вычислителей // Т-Comm: Телекоммуникации и транспорт. 2025. Т. 19. № 6. С. 33-42.

FAST OPTICAL COMPUTERS

Vladislav V. Kamenskij, Rostov State Transport University (RSTU), Rostov-on-Don, Russia, kam-vladislav@yandex.ru
Sergei V. Sokolov, Moscow Technical University of Communications and Informatics, Moscow, Russia, s.v.s.888@yandex.ru
Elena O. Lagunova, Rostov State Transport University (RSTU), Rostov-on-Don, Russia, lagunova@rambler.ru

Abstract

The paper discusses a new approach to the design of high-performance computing systems based on optoelectronic technologies. Optical information processing devices are used in areas where maximum speed and simplicity of calculations are required, such as the synthesis of embedded systems and specialized processors for high-speed processing. In general, embedded systems process incoming data within a specific and predefined range, and calculations can be performed with a specified accuracy, but the calculation time is highly limited and critical. In this regard, for the developed optical calculator, as a result of conducting an analysis of existing formats of number representation, the format of number representation with a fixed point was chosen. As an example of the implemented nonlinear function, one of the most difficult functions for calculation in traditional schemes – the function of extraction of the square root was chosen. The theoretical foundations of building a fast-acting square root calculator for numbers represented in binary code are presented. The proposed approach allows for the calculation of the square root based on a combinational device, without the use of trial and error, which radically increases the speed of calculations. Since practical problems require a minimum amount of time to perform a set of calculations on several arguments, a functional diagram of a specialized optical programmable device has been developed to perform a set of functional transformations based on the principles of building and optimizing combinational circuits. The main basic elements of an optical programmable device are optical splitters, combiners, and transparencies, the technology of which is known and does not cause difficulties. Programming an optical computer to implement a specific function is done by destroying the optical. The minimum number of elements and high calculation speed make the optical programmable device promising for building fast-acting technical means of computer engineering and communication systems.

Keywords: optical computing device, square root extraction, exponentiation, gigahertz range, optical splitter, transparent, optical combiner.

References

- [1] I.V. Statsenko, "Research of the convergence rate of one generalized Newton method and the classical Newton method in the procedure of refining the roots of a polynomial," *Precise Science*. 2020. No. 78, pp. 2-9.
- [2] G.T. Macharadze, E.A. Morozova, "The algorithm of fast calculation of the square root and its relevance," *Bulletin of the Russian New University*. 2018, pp. 106-109 DOI: 10.25586/RNU.V9I87.18.04.P.106.
- [3] I.V. Ushenina, "Implementation on FPGA and comparative analysis of the sigmoid calculators working with the full range of the argument taking into account the symmetry," *Bulletin of Tomsk State University. Management, computer science and computer engineering*. 2025. No. 71, pp. 120-129. doi: 10.17223/19988605/71/1.
- [4] S.B. Gashkov, A.B. Frolov, E.P. Popova, "On the complexity of algorithms for extracting square roots in finite fields and residue rings," *Vestnik MEI*. No. 5. 2018, pp. 79-88.
- [5] A. Fedorov, T. Vasnetsova, "Analysis of the Effectiveness of Root Extraction Methods," *Journal "Issues of Mathematics*, Issue 1, 2023, pp. 15-25.
- [6] D. Sokolov, "New Approaches to Root Extraction," *Proceedings of the International Conference on Mathematics*, 2024, pp. 35-45.
- [7] A. Krasnova, "Extraction of Roots: Evolution of Methods," *Electronic Journal "Mathematics of the 21st Century*, 2025, pp. 12-20.
- [8] D.M. Zlatopolsky, "Methodology for Extracting Square and Cube Roots in the Binary Number System," *Informatics at School*. 2021. No. 1(1), pp. 42-45, <https://doi.org/10.32517/2221-1993-2021-20-1-42-45>.
- [9] C. Nikolaev, "Classical and Modern Methods of Root Extraction," Publisher: World of Science, 2024, 320 p.
- [10] O. Sergeeva, "Simplicity and Complexity of Root Extraction," Publisher: Alfa-Book, 2025, 200 p.
- [11] K. Andraos, N. M. Ghiasi, "A High-Throughput FPGA Implementation of a Digit-Recurrence Square Root Unit with BFloat16 Support," *IEEE Transactions on Computers*. 2024. Vol. 73, No. 5, pp. 1245-1258. DOI: 10.1109/TC.2024.3355678
- [12] M. R. G. de Aledo, P. L. de Guzman, "A Formal Verification of the Fast Inverse Square Root Algorithm in the Coq Proof Assistant," *Journal of Automated Reasoning*. 2023. Vol. 67, Issue 4, Article 28. DOI: 10.1007/s10817-023-09670-6
- [13] S. Gupta, A. Martinez, M. Roberts, "Numerical Stability Analysis of Iterative Methods for Square Root in Mixed-Precision ML Training," *Proceedings of the Machine Learning and Systems Conference (MLSys)*. 2024. Vol. 6, pp. 45-59.
- [14] J. Kaur, T. H. Nguyen, J. D. Bruguera, "Efficient Square Root and Inverse Square Root for the Posit Number System," *IEEE Transactions on Emerging Topics in Computing*. 2023. Vol. 11, No. 2, pp. 456-469 DOI: 10.1109/TETC.2023.3245671
- [15] D. Petrov, Y. Takahashi, M. Mosca, "A Scalable Quantum Algorithm for Square Root and Division with Arbitrary Precision," *Quantum*. 2024. Vol. 8, p. 1245. DOI: 10.22331/q-2024-01-08-1245
- [16] C. Wang, Y. Kim, S. A. Fahmy, "Hardware Optimizations of the Newton-Raphson Square Root for 3D Graphics Pipelines," *ACM Transactions on Graphics (TOG)*. 2024. Vol. 43, No. 4, Article 78 DOI: 10.1145/3654662
- [17] F. Zhang, H. Li, E. J. McCluske, "Error Analysis and Correction for Digit-by-Digit Square Root in Safety-Critical Systems," *IEEE Transactions on Computers*. 2023. Vol. 72, No. 9, pp. 2541-2554. DOI: 10.1109/TC.2023.3267890
- [18] Pshko A. Mohammed, Integration of self-standing X- and Y- shaped polymer coupler and splitter with single mode optical fiber - *Optical Materials* <https://doi.org/10.1016/j.optmat.2020.110685>
- [19] V.V. Kamensky, S.V. Sokolov, "Synthesis of code converters based on optical technologies," *Radio engineering*. 2023. Vol. 87. No. 11, pp. 202-209.
- [20] V.V. Kamensky, S.V. Sokolov, E.O. Lagunova, "Optical digital-analog computer," *T-Comm*. 2025. Vol. 19. No. 4, pp. 37-44.
- [21] V.V. Kamensky, S.V. Sokolov, E.O. Lagunova, "Method of constructing optical computers," *T-Comm*. 2025. Vol. 19. No. 6, pp. 33-42.

РАЗРАБОТКА МЕТОДОВ И АЛГОРИТМОВ РЕШЕНИЯ ЗАДАЧ УПРАВЛЕНИЯ В ОРГАНИЗАЦИОННЫХ СИСТЕМАХ ТОРГОВЫХ КОМПАНИЙ

DOI: 10.36724/2072-8735-2026-20-3-43-52

Леохин Юрий Львович,

Московский технический университет связи и информатики, Москва, Россия, y.l.leokhin@mtuci.ru

Дымкова Светлана Сергеевна,

Московский технический университет связи и информатики, Москва, Россия, s.s.dymkova@mtuci.ru

Фатхулин Тимур Джалилевич,

Московский технический университет связи и информатики, Москва, Россия, t.d.fatkhulin@mtuci.ru

Зозуля Ирина Сергеевна,

Московский технический университет связи и информатики, Москва, Россия

Manuscript received 15 December 2025;

Accepted 25 February 2026

Ключевые слова: метод, алгоритм, задача управления, объемы продаж, организационная система торговых компаний, прогнозирование, показатели, эффективность

В статье показана разработка и экспериментальная проверка модифицированного метода на основе алгоритма XGBoost, применяемого для прогнозирования объемов продаж. Цель работы – повышение точности методов и алгоритмов прогнозирования объемов продаж, используемых в организационных системах торговых компаний. Актуальность работы обусловлена тем, что одной из основных проблем предметной области продаж является наличие шумов и аномалий в данных. При прогнозировании это приводит к снижению точности результатов и возникновению эффекта переобучения. Настройка существующих параметров в алгоритмах прогнозирования при управлении в организационных системах торговых компаний не всегда является решением данных проблем и приводит к необходимости дополнительных модификаций для уточнения прогнозов. В работе разработан метод модификации алгоритма градиентного усиления, использующий корректировки расчета градиента с помощью оценок аномальности образцов набора данных методом леса изоляции. Также использованы дополнительные гиперпараметры и метод остановки обучения при получении информации о начале переобучения. В ходе исследования были рассмотрены подходы к регуляризации и снижению влияния аномалий в данных на примере алгоритма XGBoost. Внесены дополнительные модификации в предлагаемый метод с применением кросс-валидации. Все модификации протестированы в разных комбинациях на разных наборах данных. Таким образом выявлена закономерность повышения точности при обучении на кросс-валидации с применением разработанного метода взвешивания аномалий. Практическая значимость работы заключается в разработке универсального подхода к модификации алгоритма XGBoost, который может быть применен для повышения точности прогнозирования продаж различных категорий товаров при управлении в организационных системах торговых компаний.

Информация об авторах:

Леохин Юрий Львович, Московский технический университет связи и информатики, профессор, д.т.н., orcid.org/0000-0003-3321-4497

Дымкова Светлана Сергеевна, Московский технический университет связи и информатики, к.т.н., orcid.org/0000-0003-1945-9850

Фатхулин Тимур Джалилевич, Московский технический университет связи и информатики, кафедра "Интеллектуальный анализ данных", доцент, к.т.н., orcid.org/0000-0003-0998-1055

Зозуля Ирина Сергеевна, Московский технический университет связи и информатики, кафедра "Математическая кибернетика и информационные технологии", магистрант, orcid.org/0009-0006-3445-8070

Для цитирования:

Леохин Ю.Л., Дымкова С.С., Фатхулин Т.Д., Зозуля И.С. Разработка методов и алгоритмов решения задач управления в организационных системах торговых компаний // Т-Comm: Телекоммуникации и транспорт. 2026. Том 20. №3. С. 43-52.

For citation:

Yu. L. Leokhin, S. S. Dymkova, T. D. Fatkhulin, I. S. Zozulya, "Development of Methods and Algorithms for Solving Management Problems in Organizational Systems of Trading Companies," *T-Comm*, 2026, vol. 20, no. 3, pp. 43-52. (in Russian)

Введение

Среди ключевых задач управления в организационных системах торговых компаний выделяют своевременное прогнозирование объемов продаж товаров и услуг [5, 11]. В настоящее время задача управления объемами продаж различных товаров, в том числе телекоммуникационного или компьютерного оборудования, имеет первостепенное значение [3, 22-24]. Это связано как с резко возросшим спросом на такое оборудование из-за развития систем искусственного интеллекта, так и с санкционным давлением отдельных стран [7, 23-25]. В результате управление в организационных системах торговых компаний, которые подразумевают под собой формальную структуру компаний, происходящие в них процессы, реализуемые политики и процедуры, регулирующие деятельность торговых организаций, становится значительно затруднено. Оптимизация управления даст возможность повысить эффективность предприятий [26].

Существующие методы и алгоритмы прогнозирования различных показателей [1, 2, 14, 27] не всегда дают максимально эффективный результат [8, 14]. Это относится и к прогнозированию таких динамически изменяющихся величин, как пропускная способность сетей, построенных на определенном типе оборудования [6, 7, 15-21], так и экономических показателей торговых компаний, напрямую влияющих на результативность управления в этих организационных системах [19, 20]. В работах [8, 12-14] показано, что разработка модифицированного метода и алгоритма решения задач управления спросом в организационных системах телекоммуникационных компаний позволило значительно улучшить показатели таких систем.

Еще одним примером необходимости разработки новых методов и алгоритмов решения задач управления в разнообразных организационных системах является необходимость прогнозирования динамики изменения криптовалют [6]. В работе [7] показано, что разработка современных методов получения данных и прогнозирования валютных значений курсов криптовалют дает возможность эффективнее управлять организационными системами криптобирж на основе ретроспективной, текущей и экспертной информации.

Все это указывает на актуальность разработки методов и алгоритмов решения задач управления объемами продаж в организационных системах торговых компаний. Настоящее исследование строится на уже проведенном анализе существующих методов и алгоритмов прогнозирования объемов продаж, который позволил выявить наиболее перспективные из них для решения поставленной задачи [5, 11]. В данной работе будет предложен модифицированный вариант метода и алгоритма прогнозирования.

1 Постановка задачи

В современном обществе предоставляют много товаров и услуг. Для прогнозирования разных показателей требуется крайне много сил и затрат, как временных, так и человеческих ресурсов. На помощь приходят методы машинного обучения, которые систематизируют, находят закономерности и прогнозируют показатели, например, объемы продаж.

Среди множества алгоритмов машинного обучения алгоритм градиентного бустинга XGBoost (eXtreme Gradient

Boosting) зарекомендовал себя как один из наиболее эффективных инструментов для решения задач регрессии, в том числе и прогнозирования продаж [4, 5, 12, 13].

Однако, несмотря на свои преимущества, XGBoost обладает и определенными недостатками, главным из которых является склонность к переобучению, особенно при работе с определенными типами данных, имеющими высокий уровень шума, аномалий или недостаточный объем. Это ограничивает его прямое применение и создает предпосылки для дальнейшей модификации и адаптации алгоритма под специфику конкретных задач и наборов данных.

Таким образом, возникает проблема, заключающаяся в необходимости повышения устойчивости и обобщающей способности метода XGBoost для задач прогнозирования продаж за счет разработки и применения комплекса модификаций, направленных на борьбу с переобучением и адаптацией к различным особенностям исходных данных.

2 Определение вариантов модификаций для повышения точности метода XGBoost

По результатам экспериментов [5] оптимальным методом для прогнозирования объемов продаж товаров признан алгоритм градиентного бустинга XGBoost. Этот алгоритм характеризуется высокой точностью, быстротой вычислений и возможностью тонкой настройки параметров [1-2, 13-16], хотя и обладает недостатками такими, как склонность к переобучению и снижению точности на определенных типах данных, что открывает перспективы дальнейшего совершенствования.

Основным критерием при проведении модификации метода градиентного усиления является критерий точности. При этом стоит учитывать факт того, что модификация метода и алгоритма предполагает не только увеличение точности, но и возможное увеличение скорости и сложности построенной модели. Поэтому при выборе модификации стоит минимизировать указанные риски.

При проведении экспериментов и анализе полученных данных были выявлены проблемы переобучения и снижения точности, связанные с недостаточностью обрабатываемых данных, наличием шумов и аномалий. Все это приводит к искажению обучения [11]. Под шумами понимают случайные колебания продаж, которые не несут полезной информации для обучающего алгоритма.

Модифицировать алгоритм можно несколькими способами. Рассмотрим некоторые из них.

1. Оптимизация гиперпараметров

Гиперпараметры определяют поведение модели. Грамотная настройка может заметно улучшить качество и производительность модели [9, 10]. Рассмотрим ключевые параметры:

- `max_depth`: ограничивает глубину деревьев, предотвращая переусложнение модели. Чем меньше глубина, тем ниже риск переобучения. Обычно оптимальное значение находится в диапазоне от 3 до 10.
- `min_child_weight`: определяет минимальный суммарный вес (вес сумм всех объектов в узле) для формирования узла дерева. Увеличение `min_child_weight` препятствует созданию небольших узлов, снижая переобучение [17].

- `subsample`: регулирует долю объектов, выбираемых случайно для каждого отдельного дерева. Установив `subsample` менее 1, можно уменьшить эффект переобучения и повысить устойчивость модели.

- `colsample_bytree`: контролирует долю признаков, используемых при построении каждого дерева. Установка `colsample_bytree` менее 1 уменьшает размерность пространства признаков, ускоряя обучение и уменьшая риск переобучения.

- `learning_rate`: шаг обучения влияет на скорость и точность модели. Маленькие `learning_rate` замедляют обучение, но снижают вероятность переобучения. Большие шаги увеличивают скорость, но повышают риск скачков в процессе обучения.

2. Регуляризация

Регуляризация помогает бороться с переобучением, добавляя штраф в случае сложной модели. В XGBoost поддерживаются два основных вида регуляризации [8]:

- `alpha` (L1 regularization): L1-регуляризация поощряет редкие признаки, приводя некоторые веса к нулю, тем самым устраняя лишние признаки.

- `lambda` (L2 regularization): L2-регуляризация равномерно распределяет штрафы, уменьшая влияние экстремально высоких весов.

3. Ранняя остановка обучения (Early Stopping)

Если качество модели перестаёт улучшаться, то ранняя остановка останавливает обучение модели. Данный подход позволяет эффективно предотвращать переобучение и экономить время. Для включения ранней остановки необходимо задать аргумент `early_stopping_rounds`.

4. Параллельное обучение

XGBoost изначально поддерживает параллельную обработку, что позволяет существенно ускорить выполнение алгоритма. Параметр `n_jobs` управляет количеством потоков.

Также можно использовать графические ускорители (GPU), установив соответствующий флажок:

```
xgb_model = xgb.XGBRegressor(tree_method='gpu_hist').
```

Этот режим полезен для больших наборов данных, так как позволяет существенно ускорить процесс обучения.

5. Выбор правильного формата данных

Чтобы ускорить работу XGBoost, важно правильно хранить данные. Лучший вариант – конвертировать их во внутренний формат библиотеки, `DMatrix`, который заточен под высокую скорость обработки. Также стоит сохранять исходные данные в эффективных форматах вроде `Parquet` или `ORC` – это уменьшит их объем и ускорит чтение с диска, что положительно скажется на общей производительности.

6. Кросс-валидация

Кросс-валидация – это важный метод оценки качества модели, помогающий определить оптимальные параметры и выявить потенциальные проблемы переобучения. Выполнять кросс-валидацию стоит на небольшом количестве данных [3-9].



Рис. 1. Блок-схема алгоритма модифицированного метода XGBoost

7. Балансировка классов

Несбалансированные классы приводят к смещению модели в сторону преобладающего класса. Используются следующие техники для борьбы с дисбалансом:

- Class Weights: установка весов для классов позволяет сбалансировать важность редких классов.
- Oversampling: искусственное увеличение числа представителей редкого класса.
- Undersampling: сокращение числа представителей преобладающего класса.

Балансировка классов поможет избежать смещения модели и обеспечить лучшее качество прогнозов.

8. Выявление и устранение лишней корреляции

Коррелированные признаки создают излишнюю избыточность и ухудшают интерпретируемость модели. Для диагностики корреляции строится матрица корреляций, и исключаются признаки с высоким уровнем взаимозависимости.

Уменьшение размерности пространства признаков упрощает модель и ускорит её обучение.

При корректном использовании нескольких методов, возможно улучшить как точность прогнозирования, так и не увеличить скорость обработки информации.

Таблица 1

Значения гиперпараметров для применяемых алгоритмов МО

Алгоритм	Наименование гиперпараметра	Тип значения	Диапазон поиска гиперпараметра
UMAP	n_neighbors	int	2 - 10
	n_components	int	2 - 10
	min_dist	float	0,0 – 0,5
KMeans	n_clusters	int	2 - 11
XGBoost	n_estimators	int	200 – 1000, шаг 200
	max_depth	int	1 - 10
	learning_rate	float	0,001 - 1
	subsample	float	0,05 - 1
	colsample_bytree	float	0,05 - 1
XGBoost с регуляризаторами L1, L2, gamma	min_child_weight	int	1 - 20
	reg_alpha	float	0 - 10
	reg_lambda	float	0 - 10
Isolation Forest	gamma	float	0 - 10
	n_estimators	int	200 – 1000, шаг 200
	max_samples	float	0,05 - 1
	max_features	float	0,05 - 1
	bootstrap	bool	True, False

3 Алгоритм модифицированного метода XGBoost

В результате выбора оптимального сочетания методов улучшения алгоритма был получен модифицированный алгоритм, который будет протестирован с применением датасетов, использованных в экспериментах 1-3 [5] (рис. 1). С целью оценки качества прогнозов необходимо вычислить следующие метрики: RMSE, MAE и R² [5].

4 Экспериментальная апробация разработанных методов и алгоритмов

При обучении рассматриваемых в исследовании моделей были взяты данные из датасетов открытой платформы «Kaggle» (открытый ресурс: <https://www.kaggle.com/datasets>) [5]. Данные выбирались релевантно рассматриваемой предметной области и согласно решаемой практической задаче. В таблице 1 приведены варианты гиперпараметров, используемых при настройке исследуемых методов (алгоритмов) машинного обучения.

В дальнейшем будут показаны расчеты значения метрик в каждом эксперименте для всех моделей.

Эксперимент №1. Рассмотрим датасет продаж (информация из открытых ресурсов платформы «Kaggle»: <https://www.kaggle.com/datasets/ihelon/coffee-sales>).

Данные представляют собой продажи кофейных напитков (капучино, горячий шоколад, какао, латте, американо и др.) в вендинговых аппаратах. Данные были предобработаны и приведены к единому числовому формату.

Всего в датасете указаны 262 факта продажи в разные временные интервалы. Задача спрогнозировать объёмы продаж в ближайшем будущем.

Таблица 2 демонстрирует работу алгоритмов по различным метрикам. Результаты приведены как на тестовой, так и на обучающей выборках (рис. 2). Ранее при рассмотрении алгоритмов машинного обучения на этом датасете оптимальные результаты показал алгоритм XGBoost.

Далее в таблицах 2-4 рассмотрены оценки изменения точности и показатели результатов на метриках.

Как видно из таблицы 2, модели с CV демонстрируют более сбалансированные результаты между тренировочной и тестовой выборками. Также необходимо отметить позитивное влияние кросс-валидации, которая приводит к лучшим результатам метрик модификаций. При этом традиционный вариант с регуляризацией L1, L2, gamma увеличивает ошибки модели на тестовых данных. Таким образом, этот подход является неэффективным.

В моделях с UMAP по отдельности отмечается катастрофическое ухудшение метрик (R² снижается до 0.47), однако в комбинации с Isolation Forest и кросс-валидацией наблюдается некоторое улучшение показателей. Модели с Isolation Forest демонстрируют стабильно хорошие результаты, что свидетельствует о наличии выбросов в данных и эффективности их фильтрации.

Когда в комбинированные подходы добавляем кросс-валидацию, наблюдается повышение показателей точности в сравнении с базовыми модификациями. Следовательно, правильная комбинация предлагаемых модификаций приводит к тому, что они дополняют друг друга.

Таблица 2

Полученные значения работы алгоритмов в моделях в Эксперименте №1

Модель	Метрика	Значения на выборке	
		train	test
XGBoost	MAE	0,0080	0,1375
XGBoost + L1, L2, gamma		0,2120	0,3162
XGBoost+IsolationForest		0,0010	0,1599
XGBoost+UMAP		0,0008	1,5044
XGBoost+IsolationForest+UMAP		0,0006	1,9260
XGBoost+CV		0,0204	0,1335
XGBoost+IsolationForest+CV		0,1390	0,2692
XGBoost+UMAP+CV		0,0349	1,5203
XGBoost+IsolationForest+UMAP+CV		0,0234	2,3413
XGBoost		R²	1,0000
XGBoost + L1, L2, gamma	0,9923		0,9852
XGBoost+IsolationForest	1,0000		0,9899
XGBoost+UMAP	1,0000		0,4718
XGBoost+IsolationForest+UMAP	1,0000		0,3121
XGBoost+CV	0,9999		0,9898
XGBoost+IsolationForest+CV	0,9955		0,9861
XGBoost+UMAP+CV	0,9999		0,5498
XGBoost+IsolationForest+UMAP+CV	0,9999		-0,0157
XGBoost	RMSE		0,0126
XGBoost + L1, L2, gamma		0,3705	0,5176
XGBoost+IsolationForest		0,0016	0,4282
XGBoost+UMAP		0,0011	3,0915
XGBoost+IsolationForest+UMAP		0,0010	3,5281
XGBoost+CV		0,0384	0,4303
XGBoost+IsolationForest+CV		0,2845	0,5203
XGBoost+UMAP+CV		0,0502	2,8540
XGBoost+IsolationForest+UMAP+CV		0,0320	4,2869

Проанализируем графики метрик для Эксперимента №1. Самый лучший результат получен при использовании XGBoost. Таким образом, эта модель подходит для поставленной задачи. Подходы с UMAP продемонстрировали очень низкие показатели ошибок на тренировочных выборках. Можно сделать вывод о явной тенденции к переобучению, так как на тестовой выборке показатели ошибок экстремально высоки. При этом более сбалансированные результаты выдаются моделями с кросс-валидацией, которая уменьшает переобучение и эффективна при применении сложных модификаций.

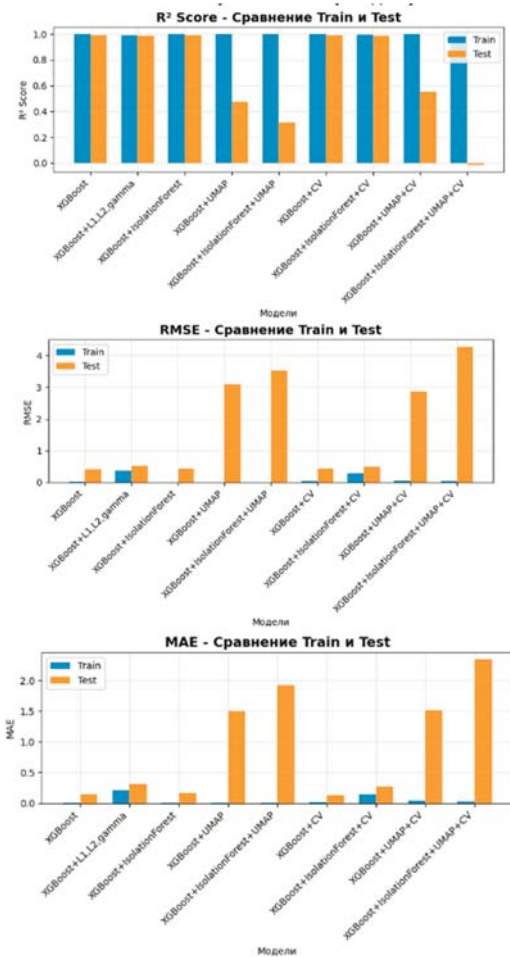


Рис. 2. Значения метрик для Эксперимента №1

Эксперимент №2. Эксперимент проводится на следующем датасете продаж (информация из открытых ресурсов платформы «Kaggle»): <https://www.kaggle.com/datasets/xontology/data-penjualan-zara>).

Данные представляют собой продажи одежды, аксессуаров и иных вещей из магазина Zara. Данные были предобработаны и приведены к единому числовому формату. Всего в датасете указаны 252 факта продажи в разные временные рамки. Задача спрогнозировать объёмы продаж в ближайшем будущем.

В таблице 3 показаны увеличение и уменьшение точности дальнейших модификаций. Оценка производится в сравнении с показателями предыдущего эксперимента, где на этом же датасете оптимальные результаты продемонстрировала модель с алгоритмом XGBoost (рис. 3).

Аналогично предыдущему эксперименту, регуляризация с помощью штрафования L1 и L2 ухудшила показатели точности, однако в данном случае проявился иной эффект - метрики на тренировочных данных значительно ухудшились (R^2 снизился с 0,9997 до 0,9932), что свидетельствует о недостаточной гибкости модели при сильной регуляризации. Следовательно, настраивать параметры регуляризации надо с применением более тонкого подхода.

Таблица 3

Полученные значения работы алгоритмов в моделях Эксперимента №2

Модель	Метрика	Значения на выборке		
		train	test	
<i>XGBoost</i>	MAE	9,4271	674,4052	
<i>XGBoost + L1, L2, gamma</i>		42,4052	667,6066	
<i>XGBoost+IsolationForest</i>		0,2391	771,7242	
<i>XGBoost+UMAP</i>		1,2992	728,8567	
<i>XGBoost+IsolationForest+UMAP</i>		1,7943	635,1774	
<i>XGBoost+CV</i>		535,9774	568,9708	
<i>XGBoost+IsolationForest+CV</i>		518,4127	575,4599	
<i>XGBoost+UMAP+CV</i>		554,0367	583,7050	
<i>XGBoost+IsolationForest+UMAP+CV</i>		530,1759	573,8359	
<i>XGBoost</i>		R ²	0,9997	-0,5753
<i>XGBoost + L1, L2, gamma</i>			0,9932	-0,5012
<i>XGBoost+IsolationForest</i>			1,00	-0,8270
<i>XGBoost+UMAP</i>	1,00		-0,7925	
<i>XGBoost+IsolationForest+UMAP</i>	1,00		-0,4902	
<i>XGBoost+CV</i>	0,1763		-0,0399	
<i>XGBoost+IsolationForest+CV</i>	0,1903		-0,0508	
<i>XGBoost+UMAP+CV</i>	0,1433		-0,0185	
<i>XGBoost+IsolationForest+UMAP+CV</i>	0,1919		-0,0552	
<i>XGBoost</i>	RMSE		12,7994	854,7986
<i>XGBoost + L1, L2, gamma</i>			57,6514	834,4577
<i>XGBoost+IsolationForest</i>			0,3392	920,5702
<i>XGBoost+UMAP</i>		1,7389	911,8384	
<i>XGBoost+IsolationForest+UMAP</i>		2,3661	831,3951	
<i>XGBoost+CV</i>		634,9137	694,5244	
<i>XGBoost+IsolationForest+CV</i>		621,2661	698,1325	
<i>XGBoost+UMAP+CV</i>		647,5251	687,3219	
<i>XGBoost+IsolationForest+UMAP+CV</i>		620,6437	699,6118	

Каждая по отдельности модели с Isolation Forest и UMAP показали идеальные результаты на тренировочных данных ($R^2 = 1.0000$), однако на тестовой выборке продемонстрировали катастрофическое ухудшение метрик (R^2 от -0.79 до -0.83), что указывает на сильное переобучение. Их совместное применение несколько улучшило ситуацию ($R^2 = -0.49$), но не позволило достичь положительных значений, что может свидетельствовать о неадекватной обработке аномалий при текущем разделении данных.

Применение кросс-валидации с XGBoost существенно повысило точность прогнозов на тестовой выборке – модель XGBoost+CV показала $R^2 = -0.0399$ против $R^2 = -0.5753$ у базового XGBoost, что подтверждает эффективность кросс-валидации для снижения переобучения.

Однако добавление Isolation Forest к кросс-валидации незначительно ухудшило результат ($R^2 = -0.0508$), что может

проявляться в необходимости более тонкой настройки параметров обнаружения аномалий. Дальнейшее добавление UMAP к кросс-валидации улучшило результат (XGBoost+UMAP+CV с $R^2 = -0.0185$), а модель XGBoost+IsolationForest+UMAP+CV показала $R^2 = -0.0552$.

Наилучший результат среди всех моделей продемонстрировала XGBoost+UMAP+CV, что свидетельствует о положительном влиянии снижения размерности в комбинации с кросс-валидацией. Одновременное снижение ошибок на тренировочной и тестовой выборках у моделей с кросс-валидацией по сравнению с базовыми подходами подтверждает эффективность методов борьбы с переобучением, хотя ни одной модели не удалось достичь положительного значения R^2 на тестовой выборке, что указывает на необходимость дополнительного анализа данных и feature engineering (рис. 3).

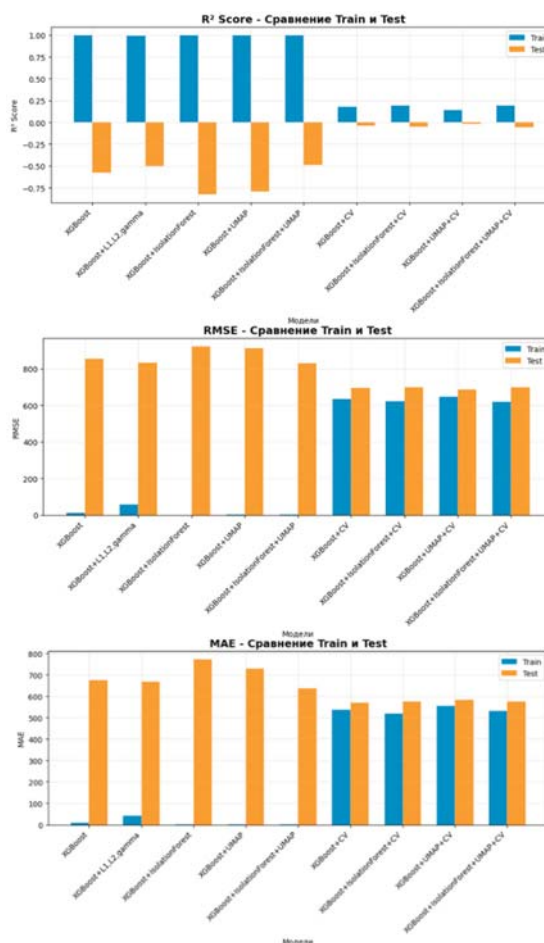


Рис. 3. Значения метрик для Эксперимента №2

Эксперимент №3. Эксперимент проводится на следующем датасете (информация из открытых ресурсов платформы «Kaggle»): <https://www.kaggle.com/datasets/atharvasoundankar/chocolate-sales>. Данные представляют собой продажи шоколадных изделий (белого, молочного и темного шоколада). Данные были предобработаны и приведены к единому числовому формату. Всего в датасете указаны 1094 факта продажи в разные временные рамки. Задача спрогнозировать объёмы продаж в ближайшем будущем. Полученные значения по результатам работы алгоритмов как на тестовых, так и на обучающих выборках отражены в таблице 4.

Таблица 4

Полученные значения работы алгоритмов в моделях для Эксперимента №3

Модель	Метрика	Значения на выборке	
		train	test
XGBoost	MAE	681,002	4003,608
XGBoost + L1, L2, gamma		1132,416	3819,624
XGBoost+IsolationForest		181,073	4052,540
XGBoost+UMAP		60,437	3629,583
XGBoost+IsolationForest+UMAP		49,958	3832,292
XGBoost+CV		3223,656	3221,571
XGBoost+IsolationForest+CV		3194,351	3222,103
XGBoost+UMAP+CV		3217,339	3253,444
XGBoost+IsolationForest+UMAP+CV		3180,561	3227,209
XGBoost		R ²	0,952
XGBoost + L1, L2, gamma	0,872		-0,488
XGBoost+IsolationForest	0,996		-0,680
XGBoost+UMAP	0,999		-0,316
XGBoost+IsolationForest+UMAP	0,999		-0,446
XGBoost+CV	0,041		-0,005
XGBoost+IsolationForest+CV	0,047		-0,002
XGBoost+UMAP+CV	0,054		-0,018
XGBoost+IsolationForest+UMAP+CV	0,070		0,002
XGBoost	RMSE		899,253
XGBoost + L1, L2, gamma		1471,762	4914,831
XGBoost+IsolationForest		237,797	5223,173
XGBoost+UMAP		81,637	4622,154
XGBoost+IsolationForest+UMAP		69,812	4845,172
XGBoost+CV		4031,295	4039,579
XGBoost+IsolationForest+CV		3987,638	4032,989
XGBoost+UMAP+CV		4003,164	4066,136
XGBoost+IsolationForest+UMAP+CV		3941,093	4024,302

Оптимальным для построения точной модели по результатам эксперимента оказался алгоритм XGBoost (рис. 4).

Таким образом, все основные моменты Эксперимента №2 остались неизменными. Как и в предыдущий раз, отсутствует положительный эффект на обучение при регуляризации посредством L1, L2 и gamma. Модель XGBoost+L1,L2,gamma показывает R² = -0.4883 на тестовой выборке против R² = -0.5942 у базового XGBoost, что свидетельствует об ухудшении обобщающей способности. Если отдельно дополнить модель XGBoost алгоритмом Isolation Forest, то R² = -0.6809. Дополнение модели XGBoost алгоритмом

UMAP выдает R² = -0.3163. Таким образом, метрики качества ухудшаются.

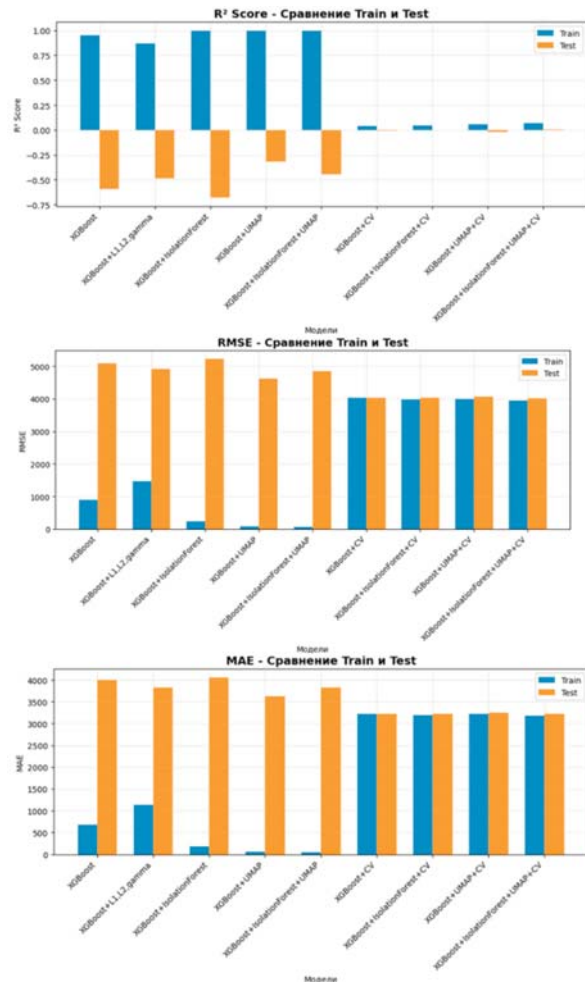


Рис. 4. Значения метрик для Эксперимента №3

Точность итоговой модели значительно улучшается при кросс-валидации: все модели с ней показывают существенное улучшение тестовых метрик.

Наилучший результат демонстрирует модель XGBoost+IsolationForest+UMAP+CV с R² = 0.0022, что является единственным положительным значением среди всех подходов. Также заметное улучшение наблюдается в модели XGBoost+IsolationForest+CV с R² = -0.0021, что подтверждает эффективность комбинации методов борьбы с переобучением.

Особенности датасета могут оказать значительное влияние на результаты эксперимента. Так, неэффективность методов снижения размерности может стать результатом того, что в датасете есть определенное количество схожих по характеристикам товарных групп. В этом случае кластеризация с применением алгоритма UMAP практически не повысила точность при использовании без CV. Это отдельный случай определенного датасета. В целом же применение кластеризации весьма эффективно для обеспечения универсальности модели на разнообразных данных.

Следует отметить, что у модифицированных подходов с кросс-валидацией наблюдается уменьшение переобучения по сравнению с другими моделями. Это выражается в выравнивании ошибок как на тестовых, так и на тренировочных данных, что показано на рисунке 4.

Модели без CV демонстрируют экстремальное переобучение с $R^2 \approx 1.0$ на тренировочных данных и отрицательными значениями на тестовых, в то время как подходы с CV показывают более сбалансированные результаты, что подтверждает их эффективность для улучшения обобщающей способности моделей.

Эксперименты 1-3 позволяют определить основные тенденции и закономерности, на основе чего можно сделать ряд выводов.

Эффект переобучения возможен при обучении моделей и их модификации на недостаточном количестве данных. Эксперимент №2 ярко продемонстрировал данный эффект переобучения, когда на тестовой выборке все значения R^2 оказались отрицательными у всех моделей.

Эксперимент №3 доказывает, что комбинация различных алгоритмов в моделях приводит к улучшению метрик, при последовательном добавлении методов растет точность.

Таким образом, наиболее эффективной стратегией оказалась комбинирование методов обработки аномалий, снижения размерности и кросс-валидации, что позволило достичь сбалансированных результатов и улучшить обобщающую способность моделей.

5 Выводы

Таким образом, в настоящей работе осуществлена разработка методов и алгоритмов решения задач прогнозирования объемов продаж в организационных системах торговых компаний. В исследовании изучена проблема прогнозирования объемов продаж различных (в том числе телекоммуникационных) товаров с применением метода градиентного усиления.

Был проведен ряд модификаций этого алгоритма для решения задач управления в организационных системах торговых компаний: взвешивание образцов выборки с применением IsolationForest, кластеризация данных с использованием UMAP, кросс-валидация. Комплекс перечисленных модификаций позволил существенно улучшить точность прогнозирования. Перечисленные модификации, использующиеся в ходе разработки методов и алгоритмов решения задач управления в организационных системах торговых компаний, позволили улучшить обучение алгоритма XGBoost за счет фильтрации выбросов и снижения их негативного влияния на процесс обучения, учесть специфики данных, улучшить обработку разнонаправленных закономерностей в распределениях шумов, обеспечить стабильность результатов обучения, повысить обобщающую способность моделей.

В результате проведения настоящего исследования можно сделать выводы о том, что разработанные методы и алгоритмы могут быть применены для повышения эффективности решения задач прогнозирования продаж, что используется при управлении в организационных системах торговых компаний.

Литература

1. John D. Kelleher, Deep Learning. The MIT Press Essential Knowledge series, MIT Press, 2019.
2. Simon J.D. Prince, Understanding Deep Learning. MIT Press, 2023.

3. Леохин Ю.Л., Фатхулин Т.Д., Маслов К.В. Разработка методов системного анализа бизнес-процессов в банковской сфере для принятия решений о кредитовании различных организаций // Научные технологии в космических исследованиях Земли. 2025. Т. 17. № 5. С. 59-71. DOI: 10.36724/2409-5419-2025-17-5-59-71.

4. Леохин Ю.Л., Дымкова С.С., Фатхулин Т.Д., Мяlicheva А.А. Методы прогнозирования дефектов в программных продуктах на основе ретроспективной и текущей информации // Т-Comm: Телекоммуникации и транспорт. 2026. Том __. № 1. С. XX-УУ.

5. Леохин Ю.Л., Дымкова С.С., Фатхулин Т.Д., Зозуля И.С. Методы и алгоритмы интеллектуальной поддержки принятия управленческих решений в организационных системах торговых компаний // Т-Comm: Телекоммуникации и транспорт. 2025. Том 19. №12. С. 44-50.

6. Leokhin Y., Fatkhulin T., Dymkova S. Research of Methods for Forecasting Dynamically Changing Data in Solving Applied Problems // 2025 Intelligent Technologies and Electronic Devices in Vehicle and Road Transport Complex (TIRVED), Moscow, Russian Federation, 2025, pp. 1-12.

7. Леохин Ю.Л., Дымкова С.С., Фатхулин Т.Д. Методы машинного обучения в прикладных задачах прогнозирования динамично изменяющихся данных // Т-Comm: Телекоммуникации и транспорт. 2025. Т. 19. №8. С. 49-63.

8. Леохин Ю.Л., Фатхулин Т.Д., Занегин А.В. Модификация метода градиентного усиления для прогнозирования спроса на отдельные виды товаров // Научные технологии в космических исследованиях Земли. 2025. Т. 17. № 2. С. 32-41. DOI: 10.36724/2409-5419-2025-17-2-32-41.

9. Leokhin Y., Fatkhulin T., Boitsov K. Computer Vision Methods in Applied Problems of Classifying Objects in Images // 2025 Wave Electronics and its Application in Information and Telecommunication Systems (WECONF), St. Petersburg, Russian Federation, 2025, pp. 1-10, DOI: 10.1109/WECONF65186.2025.11017109.

10. Леохин Ю.Л., Дымкова С.С., Фатхулин Т.Д. Исследование и разработка инструментальных средств повышения качества изображений // Т-Comm: Телекоммуникации и транспорт. 2025. Т. 19. №4. С. 45-56. (in English).

11. Фатхулин Т.Д., Зозуля И.С. Исследование методов прогнозирования показателей бизнес-процессов // REDS: Телекоммуникационные устройства и системы. 2025. Т. 15, № 2. С. 12-17. EDN ZXTFFD.

12. Fatkhulin T., Leokhin Y., Zanegin A., Rakhmatova A. Development and Research of a Modified Gradient Boosting Method Effectiveness to Solve Applied Problems of Time-Series Forecasting // 2025 Systems of Signals Generating and Processing in the Field of on Board Communications, Moscow, Russian Federation, 2025, pp. 1-10, DOI: 10.1109/IEEECONF64229.2025.10948023.

13. Leokhin Y., Fatkhulin T., Zanegin A., Rakhmatova A. Researching the Efficiency of Machine Learning Methods Used in Forecasting Demand for Certain Types of Goods // 2025 Systems of Signals Generating and Processing in the Field of on Board Communications, Moscow, Russian Federation, 2025, pp. 1-8, DOI: 10.1109/IEEECONF64229.2025.10948113.

14. Вишневецкий В.М., Леохин Ю.Л., Фатхулин Т.Д., Занегин А.В. Методы машинного обучения в решении задачи прогнозирования спроса на отдельные виды товаров // Т-Comm: Телекоммуникации и транспорт. 2024. Т. 18. №10. С. 34-43.

15. Leokhin Y.L., Fatkhulin T.D. Evaluation of Service Availability in Software-Defined Optical Network // 2021 Systems of Signals Generating and Processing in the Field of on Board Communications, Moscow, Russia, 2021, pp. 1-6, DOI: 10.1109/IEEECONF51389.2021.9416122.

16. Leokhin Y.L., Fatkhulin T.D. Approach to Estimating the Probability of Providing "Cloud" Services in the SDN // 2020 Systems of Signals Generating and Processing in the Field of on Board Communications, Moscow, Russia, 2020, pp. 1-9, DOI: 10.1109/IEEECONF48371.2020.9078593.

17. *Деарт В.Ю., Фатхулин Т.Д.* Анализ процесса создания суперканала с необходимой пропускной способностью в сети, построенной по технологии транспортных программно-конфигурируемых сетей (T-SDN) // Т-Comm: телекоммуникации и транспорт. 2018. Т. 12. №10. С. 23-30.

18. *Деарт В.Ю., Фатхулин Т.Д.* Анализ транспортных программно-конфигурируемых сетей (T-SDN) с управляемым оптическим уровнем с целью получения модели, позволяющей оценить возможность предоставления сервиса Bandwidth on Demand // Т-Comm: Телекоммуникации и транспорт. 2018. Т.12. №4. С.35-42.

19. *Deart V., Fatkhulin T.* Analysis of the functioning of a multi-domain transport software-defined network with controlled optical layer // 2017 21st Conference of Open Innovations Association (FRUCT), Helsinki, Finland, 2017, pp. 79-87, DOI: 10.23919/FRUCT.2017.8250168.

20. *Деарт В.Ю., Фатхулин Т.Д.* Анализ современного состояния транспортных сетей с целью внедрения технологии программно-конфигурируемых сетей (SDN) // Т-Comm: Телекоммуникации и транспорт. 2017. Т.11. №6. С. 4-9.

21. *Портнов Э.Л., Фатхулин Т.Д.* Технологии достижения высоких скоростей передачи в современных когерентных DWDM-системах связи // Т-Comm: Телекоммуникации и транспорт. 2015. Т.9. №8. С. 34-37.

22. *Ринас Н.А., Золкин А.Л., Каберова А.Р., Скибин Ю.В.* Влияние

автоматизации и искусственного интеллекта на социальное неравенство // Экономика и управление: проблемы, решения. 2025. Т. 7, № 1(154). С. 116-125. DOI 10.36871/ek.up.p.r.2025.01.07.015. EDN BGBKZZ.

23. *Беспалова В.В., Каберова А.Р., Белинская Д.Б.* и др. Методический подход к управлению устойчивостью развития региона // Экономика и управление: проблемы, решения. 2024. Т. 11, № 11(152). С. 88-93. DOI 10.36871/ek.up.p.r.2024.11.11.011. EDN GDEJYC.

24. *Драгуленко В.В., Золкин А.Л., Есина О.И., Каберова А.Р.* Влияние численности населения на экономический рост и развитие стран // Экономика и управление: проблемы, решения. 2024. Т. 11, № 9(150). С. 67-75. DOI 10.36871/ek.up.p.r.2024.09.11.009. EDN ANEUBG.

25. *Pavlov S.V., Dokuchaev V.A., Maklachkova V.V., Mytenkov S.S.* Features of supporting decision making in modern enterprise infocommunication systems // Т-Comm: Телекоммуникации и транспорт. 2019. Т. 13. № 3. С. 71-74.

26. *Бурлов В.Г., Грачев М.И.* Аналитическо-динамическая модель управленческого решения в социально-экономических системах на примере руководителя учебного заведения высшего образования // Т-Comm: Телекоммуникации и транспорт. 2019. Т. 13. № 10. С. 27-34.

27. *Шелухин О.И., Раковский Д.И.* Бинарная классификация многоатрибутных размеченных аномальных событий компьютерных систем с помощью алгоритма SVDD // Научные исследования в космических исследованиях Земли. 2021. Т. 13. № 2. С. 74-84.

DEVELOPMENT OF METHODS AND ALGORITHMS FOR SOLVING MANAGEMENT PROBLEMS IN ORGANIZATIONAL SYSTEMS OF TRADING COMPANIES

Yuri L. Leokhin, Moscow Technical University of Communications and Informatics, Moscow, Russia, y.l.leokhin@mtuci.ru

Svetlana S. Dymkova, Moscow Technical University of Communications and Informatics, Moscow, Russia, s.s.dymkova@mtuci.ru

Timur D. Fatkhulin, Moscow Technical University of Communications and Informatics, Moscow, Russia, t.d.fatkhulin@mtuci.ru

Irina S. Zozulya, Moscow Technical University of Communications and Informatics, Moscow, Russia

Abstract

This paper presents the development and experimental validation of a modified method based on the XGBoost algorithm for sales forecasting. The objective of the study is to improve the accuracy of sales forecasting methods and algorithms employed in organizational systems of trading companies. The relevance of this study lies in the fact that one of the key challenges in the sales domain is the presence of noise and anomalies in data. This leads to a decrease in the accuracy of forecasting results and the occurrence of overfitting. Tuning existing parameters in forecasting algorithms for management in organizational systems of trading companies does not always solve these problems and leads to the need for additional modifications to refine forecasts. This paper develops a method for modifying the gradient boosting algorithm using adjustments to the gradient calculation using anomaly estimates for dataset samples using the isolation forest method. Additional hyperparameters and a method for stopping training upon receiving information about the onset of overfitting are also used. The study examined approaches to regularization and mitigating the impact of data anomalies using the XGBoost algorithm as an example. Additional modifications to the proposed method are made using cross-validation. All modifications were tested in various combinations on different datasets. Thus, a pattern of increased accuracy was identified when training on cross-validation using the developed anomaly weighting method. The practical significance of this work lies in the development of a universal approach to modifying the XGBoost algorithm, which can be applied to improve the accuracy of sales forecasting for various product categories in organizational management systems of retail companies.

Keywords: method, algorithm, management problem, sales volumes, organizational system of trading companies, forecasting, indicators, efficiency

References

- [1] John D. Kelleher, *Deep Learning*. The MIT Press Essential Knowledge series, MIT Press, 2019.
- [2] Simon J.D. Prince, *Understanding Deep Learning*. MIT Press, 2023.
- [3] Yu.L. Leokhin, T.D. Fatkhulin, K.V. Maslov, "Development of business processes system analysis methods in the banking sphere for making decisions on lending various organizations," *H&ES Reserch*. 2025. Vol. 17. No. 5, pp. 59-71 DOI: 10.36724/2409-5419-2025-17-5-59-71. (In Rus)
- [4] Yu. L. Leokhin, S. S. Dymkova, T. D. Fatkhulin, A. A. Myalicheva, "Methods of predicting defects in software products based on retrospective and current information," *T-Comm*, 2026, vol. 20, no.1. pp. 52-59. (in Russian)
- [5] Yu.L. Leokhin, S.S. Dymkova, T.D. Fatkhulin, I.S. Zozulya, "Methods and algorithms of intellectual support for making management decisions in organizational systems of trading companies," *T-Comm*, 2025, vol. 19, no.12, pp. 44-50. (in Russian)
- [6] Y. Leokhin, T. Fatkhulin and S. Dymkova, "Research of Methods for Forecasting Dynamically Changing Data in Solving Applied Problems," *2025 Intelligent Technologies and Electronic Devices in Vehicle and Road Transport Complex (TIRVED)*, Moscow, Russian Federation, 2025, pp. 1-12, in press.
- [7] Yu.L. Leokhin, S.S. Dymkova, T.D. Fatkhulin, "Machine learning methods in applied problems of forecasting dynamically changing data," *T-Comm*, 2025, vol. 19, no.8, pp. 49-63. (in Russian)
- [8] Yu.L. Leokhin, T.D. Fatkhulin, A.V. Zanein, "The gradient boosting method modification to forecast demand for individual types of goods," *H&ES Reserch*. 2025. Vol. 17. No. 2, pp. 32-41. DOI: 10.36724/2409-5419-2025-17-2-32-41. (In Rus)
- [9] Y. Leokhin, T. Fatkhulin and K. Boitsov, "Computer Vision Methods in Applied Problems of Classifying Objects in Images," *2025 Wave Electronics and its Application in Information and Telecommunication Systems (WECONF)*, St. Petersburg, Russian Federation, 2025, pp. 1-10, DOI: 10.1109/WECONF65186.2025.11017109.
- [10] Yu.L. Leokhin, S.S. Dymkova, T.D. Fatkhulin, "Research and development of image improvement tools," *T-Comm*, 2025, vol. 19, no. 4, pp. 45-56.
- [11] Fatkhulin T.D., Zozulya I.S., "Research of methods for forecasting business process indicators," *REDS: Telecommunication devices and systems*. - 2025. Vol. 15, No. 2, pp. 12-17. (in Russian)
- [12] T. Fatkhulin, Y. Leokhin, A. Zanein and A. Rakhmatova, "Development and Research of a Modified Gradient Boosting Method Effectiveness to Solve Applied Problems of Time-Series Forecasting," *2025 Systems of Signals Generating and Processing in the Field of on Board Communications*, Moscow, Russian Federation, 2025, pp. 1-10, DOI: 10.1109/IEEECONF64229.2025.10948023.
- [13] Y. Leokhin, T. Fatkhulin, A. Zanein and A. Rakhmatova, "Researching the Efficiency of Machine Learning Methods Used in Forecasting Demand for Certain Types of Goods," *2025 Systems of Signals Generating and Processing in the Field of on Board Communications*, Moscow, Russian Federation, 2025, pp. 1-8, DOI: 10.1109/IEEECONF64229.2025.10948113.
- [14] V.M. Vishnevsky, Yu.L. Leokhin, T.D. Fatkhulin, A.V. Zanein, "Machine learning methods in solving the problem of forecasting demand for specific types of goods," *T-Comm*, vol. 18, no. 10, pp. 34-43. (in Russian)
- [15] Y. L. Leokhin and T. D. Fatkhulin, "Evaluation of Service Availability in Software-Defined Optical Network," *2021 Systems of Signals Generating and Processing in the Field of on Board Communications*, Moscow, Russia, 2021, pp. 1-6, DOI: 10.1109/IEEECONF51389.2021.9416122.
- [16] Y. L. Leokhin and T. D. Fatkhulin, "Approach to Estimating the Probability of Providing "Cloud" Services in the SDN," *2020 Systems of Signals Generating and Processing in the Field of on Board Communications*, Moscow, Russia, 2020, pp. 1-9, DOI: 10.1109/IEEECONF48371.2020.9078593.
- [17] V.Yu. Deart, T.D. Fatkhulin, "Analysis of the process of creating a superchannel with the necessary capacity in the network organized according to transport software-defined networks (T-SDN) technology," *T-Comm*, 2018, vol. 12, no.10, pp. 23-30. (in Russian).
- [18] Yu. Deart, T.D. Fatkhulin, "Analysis of transport software-defined networks (T-SDN) with controlled optical layer to obtain a model providing assessment of the possibility of bandwidth on demand service," *T-Comm*, 2018, vol. 12, no.4, pp. 35-42. (in Russian).
- [19] V. Deart and T. Fatkhulin, "Analysis of the functioning of a multi-domain transport software-defined network with controlled optical layer," *2017 21st Conference of Open Innovations Association (FRUCT)*, Helsinki, Finland, 2017, pp. 79-87, DOI: 10.23919/FRUCT.2017.8250168.
- [20] Yu. Deart, T.D. Fatkhulin, "Analysis of current state of transport networks with the purpose of introducing software defined networks (SDN) technology," *T-Comm*, 2017, vol. 11, no.6, pp. 4-9. (in Russian)
- [21] E.L. Portnov, T.D. Fatkhulin, "Technologies aimed at achieving high speed transmission in modern coherent DWDM communication systems," *T-Comm*. 2015. Vol 9. No.8, pp. 34-37. (in Russian)
- [22] N. A. Rinas, A. L. Zolkin, A. R. Kaberova and Yu. V. Skibin, "The Impact of Automation and Artificial Intelligence on Social Inequality," *Economy and Management: Problems, Solutions*, 2025, Vol. 7, No. 1(154), pp. 116-125. DOI 10.36871/ek.up.p.r.2025.01.07.015. (in Russian)
- [23] V. V. Bepalova, A. R. Kaberova, D. B. Belinskaya [et al.], "Methodological approach to managing the sustainability of regional development," *Economy and Management: Problems, Solutions*, 2024, Vol. 11, No. 11(152), pp. 88-93. DOI 10.36871/ek.up.p.r.2024.11.11.011. (in Russian)
- [24] V. V. Dragulenko, A. L. Zolkin, O. I. Esina and A. R. Kaberova, "The Impact of Population on Economic Growth and Development of Countries," *Economy and Management: Problems, Solutions*, 2024, Vol. 11, No. 9(150), pp. 67-75. DOI 10.36871/ek.up.p.r.2024.09.11.009. (in Russian)
- [25] S.V. Pavlov, V.A. Dokuchaev, V.V. Maklachkova, S.S. Mytenkov, "Features of supporting decision making in modern enterprise infocommunication systems," *T-Comm*. 2019, vol. 13, no. 3, pp. 71-74.
- [26] V.G. Burlov, M.I. Grachev, "Analytical-dynamic model of management decision in socio-economic systems on the example of the head of a educational institution of higher education," *T-Comm*, 2019, vol. 13, no.10, pp. 27-34.
- [27] O.I. Sheluhin, D.I. Rakovskiy, "Binary classification of multi-attribute tagged data about anomalous events in computer systems using the SVDD algorithm," *H&ES Research*. 2021, vol. 13, no. 2, pp. 74-84. Doi: 10.36724/2409-5419-2021-13-2-74-84.

Information about authors:

Yuri L. Leokhin, Moscow Technical University of Communications and Informatics, Rectorate, Vice-Rector for Scientific Work, full professor, Dr. Sc. (Tech.), orcid.org/0000-0003-3321-4497

Svetlana S. Dymkova, Moscow Technical University of Communications and Informatics, Candidate Sc. (Tech.), orcid.org/0000-0003-1945-9850

Timur D. Fatkhulin, Moscow Technical University of Communications and Informatics, Dpt. of IDA, Docent, Candidate Sc. (Tech.), orcid.org/0000-0003-0998-1055

Irina S. Zozulya, Moscow Technical University of Communications and Informatics, Dpt. of MC and IT, master's student, orcid.org/0009-0006-3445-8070

FOCUSED OVERVIEW OF HIGH-THROUGHPUT SATELLITE SYSTEMS: FROM STANDARDIZATION TO QOS MANAGEMENT

Aleksandr Yu. Troegubov,
Moscow Institute of Physics and Technology (MIPT),
Moscow, Russia, troegubov.aiu@phystech.edu

Alexander A. Maslov,
Moscow Institute of Physics and Technology (MIPT),
Moscow, Russia, maslov.aa@mipt.ru

Gennady V. Sebekin,
Moscow Institute of Physics and Technology (MIPT),
Moscow, Russia, sebekin.gv@mipt.ru

Sergey N. Stepanov,
Moscow Technical University of Communications and
Informatics, Moscow, Russia, s.n.stepanov@mtuci.r

Alexey O. Shchurkov,
Moscow Institute of Physics and Technology (MIPT),
Moscow, Russia, shchurkov.ao@mipt.ru

DOI: 10.36724/2072-8735-2026-20-3-53-64

Manuscript received 18 December 2025;
Accepted 02 March 2026

The work of the second, third, and fifth authors was supported by the NTI Foundation under Contract No. 70-2025-000804, May 26, 2025, and was carried out at MIPT (Phystech).

Keywords: High-throughput satellite, satellite communications, multibeam systems, resource management, quality of service

Ubiquitous communication networks are crucial for many spheres of human life, yet many regions remain underserved because terrestrial infrastructure is costly or difficult to deploy. Satellite communication systems address this limitation by providing global coverage and supporting broadcast and multicast services, making them an important part of modern communication networks that complement terrestrial systems. The paper provides a focused overview of modern high-throughput satellite (HTS) systems, emphasizing architectural evolution, standardization, and quality-of-service (QoS) resource management. The development of HTS systems involves numerous technical challenges. The paper examines current standardization efforts and spectrum regulation frameworks, along with common applications and use cases. Moreover, it overviews HTS network architecture covering geostationary and non-geostationary orbits, multibeam payloads, and multi-layer deployments, and analyzes physical- and medium-access-layer design principles. The paper further discusses service differentiation and QoS support in HTS networks, outlining traffic characteristics and relevant performance metrics at different system time scales. Overall, the paper shows that queuing theory and teletraffic models can provide a suitable analytical foundation for evaluating HTS performance and guiding the design of efficient satellite networks in future communication systems.

Information about authors:

Aleksandr Yu. Troegubov, Moscow Institute of Physics and Technology (MIPT), Moscow, Russia, <https://orcid.org/0000-0001-9105-9486>

Alexander A. Maslov, Moscow Institute of Physics and Technology (MIPT), Moscow, Russia

Gennady V. Sebekin, Moscow Institute of Physics and Technology (MIPT), Moscow, Russia

Sergey N. Stepanov, Moscow Technical University of Communications and Informatics, Moscow, Russia, <https://orcid.org/0000-0002-9815-9532>

Alexey O. Shchurkov, Moscow Institute of Physics and Technology (MIPT), Moscow, Russia

Для цитирования:

Троегубов А. Ю., Маслов А. А., Себекин Г. В., Степанов С. Н., Щурков А. О. Ключевые аспекты спутниковых систем с высокой пропускной способностью: от стандартизации к управлению качеством обслуживания // Т-Comm: Телекоммуникации и транспорт. 2026. Том 20. №3. С. 53-64.

For citation:

A. Yu. Troegubov, A. A. Maslov, G. V. Sebekin, S. N. Stepanov, A. O. Shchurkov, "Focused overview of high-throughput satellite systems: from standardization to QoS management," *T-Comm*, 2026, vol. 20, no. 3, pp. 53-64.

1 Introduction

Nowadays ubiquitous communication networks are essential everywhere: from economic to educational and social development [1]. At the same time, vast regions of the planet remain underserved by terrestrial infrastructure due to geography or cost reasons [2]. Satellite communications (SatCom) can fill this gap, providing complete coverage from space. In particular, satellite systems complement ground networks to provide broadcast and multicast capabilities over entire Earth's surface. Recent studies [3-6] highlight the key role of SatCom networks in forming future communication landscape and creating global digital space.

Satellites offer unique advantages over ground-based systems [7, 8]. For example, a single satellite in geostationary orbit can cover an entire continent, and large satellite constellations in low Earth orbit (LEO) can provide global broadband services. Unlike terrestrial base stations, satellites can broadcast the same content simultaneously to all users located in a wide area, making them ideal for radio and TV broadcasting, data distribution, or software updates. Moreover, satellite links are also resilient to natural disasters, while ground infrastructure can fail. All these features make SatCom an important complement to fiber and cellular networks.

Traditionally, communication satellites operated with one or a few wide beams covering large regions. Such an approach is simple, but spectrally inefficient, as a single beam cannot reuse frequency, so capacity is limited by the available spectrum. In contrast, High Throughput Satellites (HTS) employ a multi-beam architecture, in which the service area is divided into many narrow spot beams with aggressive frequency reuse as depicted in Fig. 1. Modern HTS systems employ tens and even hundreds of beams, each focused on a small region. By reusing the same frequency bands across non-adjacent beams, HTS achieve much higher spectral efficiency, which can multiply the total throughput compared to a conventional wide-beam satellite.

In addition, multi-beam HTS systems employ advanced digital payloads and onboard processing to manage complexity. Digital beamforming and signal processing enable flexible beam shapes and power allocation. In conventional geostationary Earth orbit (GEO) multi-beam HTS systems, all the satellite beams are constantly illuminated, even if there is no demand to be satisfied. Techniques like beam hopping (BH) allow the satellite to dynamically allocate capacity across beams in time, matching varying traffic demands. Furthermore, narrow spot beams have

higher antenna gain, allowing smaller user terminals and higher-order modulation for further increase in data rates. Note that in LEO constellations, the dynamic demand is handled with inter-satellite links and dynamic routing [10]. As a result, HTS platforms can deliver aggregate throughputs on the order of hundreds of gigabits up to beyond one terabit per second per satellite.

Altogether, modern HTS networks offer significant capacity growth and cost reduction over legacy systems. Novel SatCom systems enable entirely new services and applications. For instance, HTS and LEO constellations are considered key components of the future 6G non-terrestrial networks framework [2, 8]. Another example is the integration of SatCom networks with the Internet of Things (IoT) [11, 12]. Moreover, the research community is working on applying AI (artificial intelligence) methods to improve the performance of SatCom networks [2, 13].

As we can see, the HTS system development involves plenty of technical challenges ranging from high-frequency radio hardware to air interfaces and communication protocols. Many of these aspects are already addressed by the research community and industry, but the specific implementation details are part of proprietary technologies and therefore are not in open access. Despite diverse challenges, the overall system performance is fundamentally constrained by resource management. Resource planning is a persistent challenge – satisfying non-uniform and time-varying traffic demands are further complicated by various system constraints.

In this sense, there is a need for analytical tools capable of describing stochastic traffic behavior and service differentiation. We believe that queuing theory and teletraffic models can be a suitable framework for analyzing HTS performance. In this paper, we offer a targeted analysis of HTS systems intended for readers working in operational research, satellite network resource management, and performance modeling of HTS networks.

The rest of the paper is organized as follows. We discuss standardization and spectrum regulation aspects of SatCom networks in Section 2. Section 3 is devoted to HTS use cases and applications. The overview of HTS networks architecture is in Section 4. Sections 5 and 6 cover PHY- and MAC-level design principles. In Section 7, we consider supported service types and QoS in HTS networks. Section 8 concludes the paper and outlines future research directions in satellite networking. The list of the accepted abbreviations is in Table 1.

2 Standardization & Spectrum

In this section, we consider two fundamental components of all wireless communication systems – standardization aspects and spectrum regulation.

2.1 Standardization bodies

Unlike well-known terrestrial systems with a single standardization body (e.g., IEEE for Wi-Fi or 3GPP for cellular), SatCom involves several overlapping organizations due to the variety of applications and stakeholders. Moreover, some operators (e.g., Inmarsat, Iridium) developed proprietary protocols. Nevertheless, we observe a clear trend toward open standards to promote multi-vendor interoperability [4, 14].

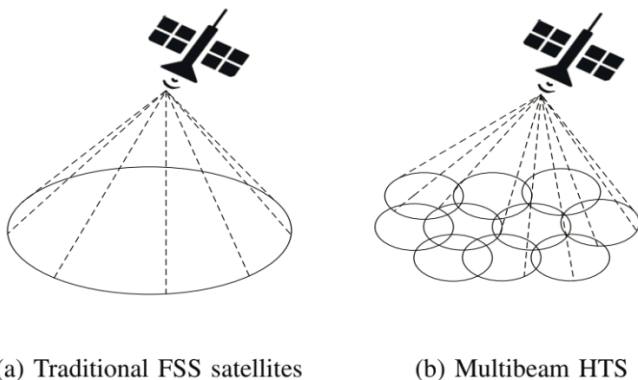


Fig. 1. Comparison of single-beam and multibeam architectures [9].

Table 1

List of abbreviations

Abbreviation	Meaning
ACM	Adaptive Coding and Modulation
ADS-B	Automatic Dependent Surveillance-Broadcast
AFR	Array-Fed Reflector
AIS	Automatic Identification System
BCH	Bose-Chaudhuri-Hocquenghem code
BH	Beam Hopping
CCSDS	Consultative Committee for Space Data Systems
CNR	Carrier-to-Noise Ratio
DRA	Direct Radiating Array
DVB	Digital Video Broadcasting
eMBB	Enhanced Mobile Broadband
FCC	Federal Communications Commission
GEO	Geostationary Earth Orbit
GPS	Global Positioning System
GW	Gateway
HAP	High-Altitude Platform
HEO	Highly Elliptical Orbit
HTS	High-Throughput Satellite
IoT	Internet of Things
ISL	Inter-Satellite Link
LAP	Low-Altitude Platform
LDPC	Low-Density Parity-Check code
LEO	Low Earth Orbit
MAC	Medium Access Control Layer
MCS	Modulation and Coding Scheme
MEO	Medium Earth Orbit
MF-TDMA	Multi-Frequency Time-Division Multiple Access
MFPB	Multiple-Feed-Per-Beam
mMTC	Massive Machine-Type Communications
NR	New Radio
NTN	Non-Terrestrial Networks
OFDMA	Orthogonal Frequency-Division Multiple Access
PHY	Physical Layer
QoS	Quality of Service
RF	Radio Frequency
SC-FDMA	Single-Carrier Frequency-Division Multiple Access
SFPB	Single Feed Per Beam
SatCom	Satellite Communication
TDM	Time-Division Multiplexin
TT&C	Telemetry, Tracking, and Control
UT	User Terminal
URLLC	Ultra-Reliable Low-Latency Communications
VLEO	Very Low Earth Orbit
VSAT	Very Small Aperture Terminal

ETSI and DVB Project. The European Telecommunications Standards Institute (ETSI) publishes and maintains the Digital Video Broadcasting (DVB) family of open standards for satellite broadcast services (e.g., TV and radio) and broadband systems [4, 14]. ETSI specifies standards for both forward link (gateway-to-user) and the return link (user-to-gateway). The latest amendment for the forward link, DVB-S2X [15], introduces wide support for HTS systems, while the corresponding amendment for the return link is DVB-RCS2 [16].

3GPP NTN. The 3rd Generation Partnership Project (3GPP) has gradually integrated satellites into the cellular framework [17]. After initial studies in Release 14, non-terrestrial networks (NTN) became a formal part of 5G architecture from Release 15. By the

latest 3GPP Release 17 [18], NTN specifications describe transparent and regenerative payloads, inter-satellite links, and air-interface adaptations for uncommon channel conditions. Moreover, NTN includes support for satellite IoT (NB-IoT and eMTC). The primary goal of 3GPP NTN is seamless operation with terrestrial 5G/6G to extend coverage and enable disaster-resilient services [4, 8].

CCSDS. The Consultative Committee for Space Data Systems (CCSDS) is a consortium of major space agencies. It recommends standards for telemetry, tracking, command, and space data handling in space missions. However, the CCSDS standards are mainly relevant to scientific and space missions than commercial HTS broadband [14].

2.2 Spectrum regulation

Generally, spectrum is regulated globally by the International Telecommunication Union (ITU). In addition, national agencies, e.g. FCC in the U.S, can impose additional restrictions for local spectrum assignments [4, 14].

Growing demand from both terrestrial and satellite operators makes some bands crowded and leads to congestion. Thus, there is a trade-off between the robustness of lower bands and the capacity of higher bands. Specifically, modern HTS systems increasingly exploit Ku- and Ka-band with advanced frequency reuse techniques. Another challenge is the coexistence between satellite and terrestrial 5G/6G systems, especially in the C- and Ka-bands, where frequency sharing is under active study [8]. A brief description of key frequency bands (see Fig. 2) is presented below [4, 19].

- **L-band** (1-2 GHz). Offers long wavelengths that are resilient to weather conditions (clouds, rain, foliage), making it suitable for reliable services: mobile satellite services (Inmarsat, Iridium), TT&C services, and navigation systems (GPS). Although limited bandwidth constrains data rates, L-band enables compact user terminals.
- **S-band** (2-4 GHz). Offers moderate capacity with robustness against weather. It is employed for TT&C services, digital audio radio, space research and operations, and weather radars.
- **C-band** (4-8 GHz). Offers robust communication against rain fading. It is traditionally used by fixed satellite services, supporting TV broadcasting and corporate networks. However, C-band has become crowded and is a subject to coexistence conflicts with 5G cellular systems, particularly in the 3.4-3.8 GHz range.
- **X-band** (8-12 GHz). Reserved for governmental and military purposes.
- **Ku-band** (12-18 GHz). Offers higher capacity than C-band with smaller antennas. It became a preferable band for broadcast and broadband VSAT systems: direct-to-home television, enterprise VSAT networks, and aviation broadband. However, it is more sensitive to rain fade and, as C-band, suffers from spectrum scarcity.
- **Ka-band** (26-40 GHz). Offers wide bandwidth at the cost of significant attenuation in adverse weather. It forms the basis for modern HTS with spot-beam architecture. Although it supports multi-Gbps broadband

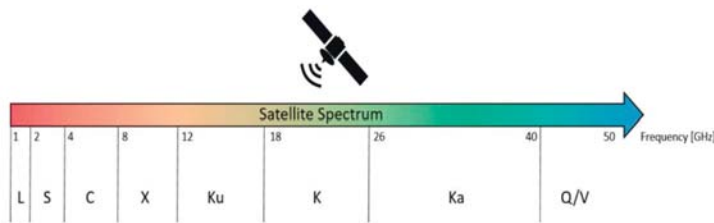


Fig. 2. SatCom frequency bands [14].

services, it requires advanced PHY-level methods.

- **Q/V-band** (40-75 GHz). Promising band for solving congestion issues in lower bands. Its feasibility for feeder links (from gateway to satellite) is still under investigation.

3 Use cases & Applications

HTS systems support a wide spectrum of applications across both commercial and governmental domains. In this section, we highlight the most relevant use cases.

3.1 5G/6G Non-Terrestrial Networks

The ITU-R IMT-2020 requirements defined three main directions for 5G: enhanced mobile broadband (eMBB), massive machine-type communications (mMTC), and ultra-reliable low latency communications (URLLC) [20]. Moreover, another ITU document [21] defines novel scenarios and guidelines for 2030 networks: ubiquitous connectivity, artificial intelligence integration and sensing integration. As part of the NTN framework, HTS systems are expected to play a key role in satisfying user requirements for beyond-5G networks [4, 8, 14].

First, HTS systems can extend broadband coverage to rural and remote regions (such as oceans and forests), where common terrestrial networks are impractical or impossible. Second, SatCom systems can provide reliable service in high-mobility scenarios (in-flight and maritime cases). Third, compared to terrestrial networks, HTS systems have extremely wide service area, which can be useful in mMTC applications: HTS constellations can handle data flows from billions of IoT devices, supplementing terrestrial networks in smart agriculture or remote monitoring [22, 23]. Although SatCom systems cannot meet URLLC latency requirements due to propagation delay, they can significantly contribute by offloading terrestrial resources in hybrid terrestrial-satellite architectures.

3.2 Positioning and navigation

As discussed above, mobility support remains one of the key features of HTS networks. SatCom is suitable for safety-critical applications such as tracking and communication in aeronautical and maritime systems [8, 24].

In aviation, SatCom can become a part of Automatic Dependent Surveillance-Broadcast (ADS-B) system. Originally, the navigation is performed with terrestrial infrastructure: an Air Traffic Management (ATM) network and an Air Traffic Controller (ATC). However, ground-based facilities are insufficient in polar regions or in the middle of the oceans, and a satellite-assisted ADS-B system can be considered as an alternative [25].

In maritime, the Automatic Identification System (AIS) is used for tracking ships [26]. The system became especially widespread when it was supported by commercial SatCom systems due to extended coverage and reliability.

This demonstrates that vehicle, aircraft and ship tracking are promising applications for further development because advanced positioning and coordination are key for unmanned worldwide transportation system [8].

3.3 Space observation, exploration, and communications

Beyond communications, satellite networks play an important role in Earth observation [8, 27]. Modern constellations generate large volumes of imagery and sensor data for agriculture, logistics, and disaster management. In traditional Earth-observation systems, the downlink of payload data is typically performed only when the satellite passes over a dedicated ground station, which limits the data delivery delay to processing services on the ground. HTS systems can provide high-capacity backhaul to transfer collected data and time-sensitive messages worldwide, thereby improving data availability and operational responsiveness.

Deep-space missions depend on robust SatCom links, with unique challenges such as long delays and extreme distances. Advances at the PHY-level, such as beamforming and novel coding schemes, directly contribute to the evolution of deep-space communications. Some concepts imply hybrid infrastructures where HTS networks provide broadband-like services to support human presence on the Moon.

3.4 Multi-layer and multi-orbit architectures

A notable trend in SatCom is the development of multi-layer and multi-orbit network architectures to combine both extensive coverage and low latency [9, 28, 29]. HTS networks can be integrated with aerial networks, creating multi-layer aerial-satellite architectures [8]. Examples of such networks are outlined below.

- Very Low Earth Orbit (VLEO) satellites at 100-450 km: offer reduced latency and smaller link budgets but require frequent fleet replenishment. Several operators are exploring VLEO for high-performance mobile services.
- High Altitude Platforms (HAPs) at 15-25 km: represented by balloons or solar-powered UAVs, act as relay nodes, providing regional coverage and complementing HTS backhaul. They offer lower latency and rapid deployment but face operational and endurance constraints.
- Low Altitude Platforms (LAPs) at 0-4 km: represented by tethered balloons or drones, extend HTS services for disaster recovery and temporary events, ensuring reliable coverage in critical situations.

Another trend is to combine GEO, MEO, and LEO systems to form a multi-orbit HTS network [9]. While GEO systems offer wide coverage and long service lifetimes, LEO provides low-latency links. Hybrid architecture demonstrates that different orbital layers can complement each other to support extended number of services. At the same time, such solutions require more advanced user terminals and introduce new challenges in dynamic resource allocation and inter-orbit handover management.

4 Orbits & Network aspects

In this section, we consider the common network topology of a SatCom system. Section 4.1 provides a general description and Sections 4.2-4.4 discuss individual network segments.

4.1 Network architecture

The network architecture of a typical HTS system has several parts as shown in Fig. 3 [14, 30].

The space segment is represented by a satellite constellation. The term “satellite” typically refers to a satellite platform equipped with complex payloads, including antennas and transponders. To reduce latency, direct communication within the space segment can be performed via inter-satellite links (ISLs).

The ground segment is represented by facilities managed by network and satellite operators. Some facilities serve as interfaces between the satellite systems and terrestrial networks, while others are responsible for tracking and monitoring the satellite's health.

The user segment is represented by various user terminals (UTs) ranging from very small aperture terminals (VSATs) to compact mobile devices, which access the system via user links. User links are typically divided into the forward link (gateway-to-user) and the return link (user-to-gateway).

Altogether, these segments form the architecture of a SatCom network.

4.2 Space segment

A satellite constellation consists of satellites themselves and their orbital configurations [14]. A satellite platform is responsible for power supply and attitude (orientation) control, while its payload performs communication functions. The key aspect of constellation design is the orbit type. There are four primary orbit classes: geostationary Earth orbit (GEO), highly elliptical orbit (HEO), medium Earth orbit (MEO), and low Earth orbit (LEO) [9].

GEO systems (~36 000 km) operate at a fixed altitude where

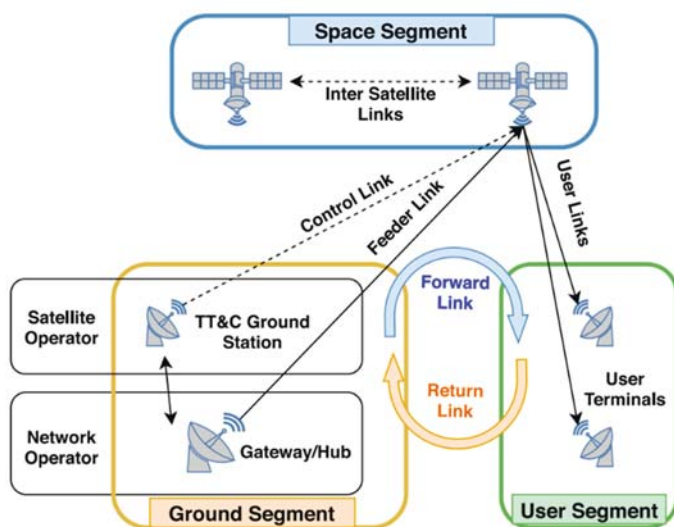


Fig. 3. Overview of a common SatCom network topology comprising space, ground, and user segments [14].

satellite position relative to the Earth's surface remains constant or near constant in the case of close-to-geostationary orbits. The orbit provides consistent, wide-area coverage, and is used for common satellite services (television and radio broadcasting, broadband Internet).

GEO systems have several advantages. First, user terminals for GEO systems do not require tracking antennas (except mobile terminals), which reduces complexity and cost. Second, they have low per-satellite deployment cost due to high coverage (one-third of the Earth's surface) and long satellite lifetime (15+ years) [9]. Third, the complex satellite payload can employ multi-beam techniques to provide high throughput. On the other hand, there are limitations: high signal attenuation, which affects capacity, and long propagation delays.

HEO systems differ from circular orbits by their high eccentricity. The key characteristic is long line-of-sight time near apogee, making them suitable for communications in high-latitude regions [14]. For example, several HEO satellite systems, e.g. Molniya and Tundra, are notably employed in Russia.

MEO systems (2 000-36 000 km) offer a compromise between GEO and LEO, with lower latency than GEO and longer visibility windows than LEO, which simplifies service continuity but requires more satellites to ensure global coverage. Currently, the orbit is used by navigation satellite systems (GLONASS, Galileo) and a few operators (e.g., SES's O3b mPOWER constellation) [9].

LEO systems (500-2 000 km) offer low-latency communication (30-50 ms), suitable for real-time and interactive applications [5, 30]. The proximity to Earth allows the use of high-frequency bands such as Ku, Ka, Q, and V, enabling high throughput and large system capacity, making such systems attractive for global broadband coverage. However, rapid satellite motion entails frequent handovers and complex constellation management to maintain continuous connections. A typical LEO constellation has hundreds to thousands of satellites, and adding new planes or satellites is straightforward, which makes the system scalable.

Inter-satellite links in LEO systems reduce dependence on ground stations, thus decreasing delays and improving connectivity.

4.3 Ground segment

The ground segment has two main responsibilities: controlling and maintaining the space segment through telemetry, tracking, and control (TT&C) operations, and enabling network access and managing traffic via gateway stations (GWs) [8, 9, 14].

GWs are responsible for user data transmission and reception, frequency conversion, baseband signal processing, and feeder link establishment with visible satellites. Specifically, a GW consists of radio frequency (RF), intermediate frequency (IF), and baseband subsystems. The RF chain integrates amplifiers, duplexers, and directive antennas. The IF stage performs frequency translation. The baseband subsystem manages modulation, multiplexing, and some advanced functions such as precoding for multi-beam operations.

TT&C stations monitor and control satellite operations. They are responsible for antenna control and tracking, updating satellite configurations, mitigating orbital drift, and monitoring overall satellite system health. Generally, TT&C stations use a special control link to communicate with satellites.

The ground segment can also incorporate elements of the terrestrial layer, e.g., cellular or Wi-Fi connectivity, that extend service to mobile and fixed users and enable network backhauling through the core network.

4.4 User segment

The user segment consists of user terminals (UTs) that provide end-users with access to HTS networks [9, 14]. UTs primarily operate in the Ka and Ku bands and can be broadly classified into fixed and mobile types.

Fixed terminals are typically implemented as very small aperture terminals (VSATs), which are two-way satellite earth stations. A VSAT has an outdoor unit (antennas, amplifiers, frequency converters) and an indoor unit (modem and customer interface). These units enable bidirectional satellite communication.

Mobile terminals encompass handheld devices, vehicle-mounted systems, and aeronautical terminals. Compared to fixed UTs, they require compact and lightweight antennas to ensure portability and optimized power management.

Future UTs are expected to evolve toward multi-radio devices integrating satellite connectivity with terrestrial technologies.

5 PHY-level design

In this section, we consider several aspects of PHY-level design in SatCom networks: channel characteristics (Section 5.1), antenna design (Section 5.2), waveform design (Section 5.3), and payload types that can be carried by a satellite (Section 5.4).

5.1 Channel characteristics

The PHY-level design of satellite systems is notably affected by a unique propagation condition determined by orbit altitude and operating frequency [3]. Channels in GEO and LEO systems differ from those in terrestrial networks due to the long transmission paths, minimal scattering, and predominance of line-of-sight propagation [14, 19, 30]. Fixed satellite links, especially in the Ku- and Ka-bands, can often be modeled as additive white Gaussian noise channels, but they are strongly affected by atmospheric phenomena. These long- and short-term fading mechanisms degrade signal quality, especially at higher frequencies. In contrast, mobile satellite links experience additional diffuse multipath components caused by reflections from terrain and buildings. Typically, such channels are modeled using multi-state Markov channel models, which capture transitions between deep-shadowed, intermediate, and clear-sky states.

From a link-budget perspective, LEO systems generally achieve higher carrier-to-noise ratios due to a lower path loss component [30]. For handheld terminals, GEO links often achieve poor CNR, whereas LEO systems can provide performance gains of up to 15-16 dB in uplink for comparable configurations. However, the mobility of LEO satellites introduces challenges such as Doppler shifts, time-varying received power, handovers - which require advanced channel estimation and tracking techniques.

5.2 Antenna design

As satellite services move from broadcast to broadband, the antenna systems move from single-beam to multi-beam antenna architectures that generate numerous high-gain spot beams to maximize spectral efficiency [4]. Two main categories of satellite antennas are passive and active arrays [9, 31, 32].

Passive antennas rely on reflectors and feed systems to form beams through mechanical means. There are two common designs: single-feed-per-beam (SFPB) and multiple-feed-per-beam (MFPB) configurations. An SFPB system is simpler to implement but requires multiple reflectors to achieve continuous coverage, which increases payload mass. In contrast, an MFPB system generates each beam from a cluster of feeds, allowing contiguous coverage with fewer reflectors and improved efficiency, but requires a more complex beamforming network. In addition, passive antennas employ mechanical steering mechanisms, which increase payload weight and provide slower beam tracking compared to digital counterparts.

Active antennas integrate amplifiers directly with the radiating elements, providing distributed amplification. These systems can be implemented in either direct-radiating array or array-fed reflector configurations. Active arrays support electronic beam steering, enabling rapid reconfiguration and adaptive coverage without mechanical movement. Combined with digital beamforming, active arrays can dynamically adjust beam direction and side-lobe levels, thereby enhancing spectral reuse and mitigating inter-beam interference. However, active antennas consume more power and require more complex signal processing.

Recent developments in array technologies for LEO systems, such as electronically steerable antennas, enable fast handovers, simultaneous tracking of multiple satellites, and multi-orbit interoperability [30]. Furthermore, the integration of active antennas with massive MIMO and precoding techniques opens opportunities for spatial multiplexing and adaptive beam management, significantly improving system capacity.

5.3 Waveform design

Modern broadband SatCom systems mainly adopt the DVB family of standards, specifically DVB-S2 and its extension DVB-S2X standards, for both GEO and LEO systems [30]. The DVB-S2 standard supports so-called variable and adaptive coding and modulation (VCM and ACM) to cope with challenging propagation conditions. Another important feature is support for advanced channel codes such as LDPC and BCH codes. Its successor, the DVB-S2X standard, notably extends the modulation granularity and supports multi-stream transmission on a single carrier.

With the evolution toward 5G NTN, waveform design encounters new challenges such as large propagation delays, Doppler shifts, and so on. The research community is evaluating candidate waveforms capable of supporting integrated satellite-terrestrial operation, including adaptations of 5G New Radio air interfaces [3]. Related studies also explore the feasibility for NTN of other 5G NR features such as timing-advance or HARQ procedures.

5.4 Payload types

Satellite payloads can be broadly categorized into transparent and regenerative, with the former referred to as ‘bent-pipe’ transponders and the latter as on-board processing systems [4, 9].

Transparent payloads apply RF processing (filtering, frequency conversion, amplification) to received signals without regeneration. Bent-pipe transponders can be analog or digital. Using digital transponders enables advanced beamforming techniques, thus improving frequency reuse and spectral efficiency.

Regenerative payloads, in addition to RF processing, apply on-board digital processing, including demodulation and decoding, of received signals. This enables packet switching, adaptive coding and modulation, beam hopping and improved interference mitigation. Another key advantage is the support of ISLs and single-hop mesh connectivity between terminals, thereby reducing latency. On the other hand, regenerative payloads are more complex and require additional on-board computational resources.

In practice, transparent payloads remain widely deployed due to their simplicity and low cost, while regenerative payloads are increasingly integrated in next-generation HTS systems.

6 MAC-level design

The section overviews common channel access schemes on the forward (Section 6.1) and return links (Section 6.2) in SatCom systems and discusses resource management challenges (Section 6.3).

6.1 Forward link

The forward link carries data from the GW to UTs [19, 30]. In HTS systems, wideband carriers are typically assigned per spot beam. User data is multiplexed in the time domain, i.e., the satellite transmits a continuous stream of frames in a broadcast manner, and each UT decodes its own data. For example, the DVB-S2X standard employs time-division multiplexing (TDM) with ACM applied on a frame-by-frame basis.

A scheduler determines which UTs will receive data in the current frame and defines the data transmission parameters. The scheduler selects a modulation and coding scheme (MCS) based on channel state feedback from the UT. Moreover, when generating a frame, it considers many aspects: channel conditions, packet or flow priority, QoS class, and buffer status at a GW side. These factors affect both intra-beam packet multiplexing and inter-beam resource allocation across active beams. Modern systems often employ multicast scheduling within a frame, hence the selected MCS is constrained by the user with the poorest channel conditions. Additional features, such as beam hopping and carrier aggregation, are expected to further enhance throughput.

GEO HTS systems, such as Viasat, use the DVB-S2X standard described above: wideband carriers (hundreds of MHz per spot beam) with TDM multiplexing of user packets. In contrast, LEO systems, such as OneWeb and Starlink, employ proprietary protocols with waveform and channel access designs more closely aligned with cellular networks.

6.2 Return link

The return link carries data from UTs to the GW [19, 30]. The DVB-RCS2 standard specifies a Multi-Frequency Time-Division Multiple Access (MF-TDMA) scheme, in which time-frequency resources are divided into frames (time domain) and frequency channels (frequency domain). The network management system assigns specific resources to UTs based on their demand reports by broadcasting a Terminal Burst Time Plan (TBTP) on the forward link. This approach provides fine granularity and flexibility in resource allocation, supporting diverse traffic profiles. Part of each frame is reserved for user capacity requests, typically occurred via random access mechanisms.

While GEO systems follow the DVB-RCS2 standard, LEO systems adopt similar uplink principles but employ waveform and access methods closer to those used in cellular networks, such as SC-FDMA or OFDMA-based allocation across time and frequency domains.

Table 2 illustrates nominal downlink and uplink throughputs for several GEO and LEO HTS systems. The GEO satellites offer high aggregate downlink capacity ranging from hundreds of Gbps to beyond one Tbps per satellite [8, 33, 34]. The capacity is distributed across many spot beams, resulting in up to 100-150 Mbps per user in downlink. Uplink data rates on GEO consumer plans are much lower, e.g., around 5 Mbps.

LEO satellites have lower capacity per satellite, e.g., 10-100 Gbps per satellite [10, 35], but compensate it with constellation scale. Individual users can achieve nominal rates of hundreds of Mbps in downlink and tens of Mbps in uplink depending on the consumer plan. Note that nominal rate assumes ideal conditions, actual network throughput is generally lower.

6.3 Resource management

Resource management in multibeam HTS networks encompasses dynamic power and frequency resource allocation. The scheduler is responsible for optimizing resource assignment to meet QoS requirements for different traffic flows.

Physically limited transmit power should be sensibly distributed across beams to consider traffic loads and propagation conditions. Flexible per-beam power assignment enhances system performance, but finding the optimal power allocation is a challenging task due to interference and payload constraints [9, 36]. This optimization problem often requires iterative or heuristic algorithms to approximate solutions under the satellite's total power budget. Co-channel interference among adjacent beams further complicates power allocation decisions.

A notable dynamic strategy is beam hopping, where only a subset of beams is illuminated at any given time according to a predefined pattern, i.e., all available RF power is concentrated on

Table 2

Nominal throughput metrics for HTS systems

System	DL Throughput	UL throughput
ViaSat-3 (GEO)	~1 Tbps per satellite	~3-5 Mbps
Hughes JUPITER-3 (GEO)	> 500 Gbps per satellite	up to 5 Mbps
SpaceX Starlink (LEO)	20-250 Mbps per user	10-20 Mbps
OneWeb (LEO)	~150 Mbps per user	~30 Mbps

a smaller set of active beams [37, 38]. Beam hopping can track spacial and temporal demand fluctuations by periodically switching active beams, thus improving amplifier utilization and reducing simultaneous inter-beam interference. At the same time, it raises scheduling complexity. Using ML-based methods is a promising approach for solving complex resource allocation problems in BH systems [39, 40].

Effective spectrum allocation is also important. Various frequency reuse schemes can be employed. Two extreme cases are no frequency reuse (so-called orthogonal scheme) and full frequency reuse among beams [14]. In practice, many multibeam systems adopt a carefully chosen reuse factor (e.g., 3 or 4-color reuse) to balance capacity and interference, and future systems aim for higher reuse by implementing advanced interference mitigation schemes [41]. In addition to static patterns, bandwidth can be dynamically distributed among beams according to the spatial traffic demands. Digital payloads support this flexibility and can adapt the bandwidth per beam on the fly. Furthermore, the carrier aggregation feature is introduced to aggregate multiple non-contiguous carriers for a single user or beam, providing greater transmission flexibility.

With aggressive frequency reuse, managing co-channel interference is a central aspect that requires joint PHY/MAC-layer solutions [19, 42, 43]. One approach is multi-user precoding on the forward link, which involves pre-processing the signals transmitted on each beam to mitigate inter-beam interference. Precoding enables full frequency reuse by significantly reducing interference, but its performance depends on accurate channel state information and user scheduling. Therefore, precoding is typically integrated with scheduling strategies at the MAC layer: the network scheduler selects groups of users or beams to serve simultaneously in a frame that are suitable for precoding, thereby maximizing the interference cancellation benefits. In addition to precoding, advanced resource scheduling can exploit spatial separation by allowing resource reuse only among beams that are sufficiently far apart to cause negligible interference to each other [44]. Such scheduling strategies, involving spacial or time clustering of beams, effectively balance scheduling complexity to keep low co-channel interference.

7 HTS services and QoS

7.1 System timescale

SatCom networks can be viewed and analyzed at multiple time scales, which determine the detail level and relevant performance metrics in modeling [4]. In general, we can define four levels for call-driven networks: customer service time (call-level), session/application activity, flow or burst communications, and packet-level transmission. Call and session durations are measured in seconds and minutes, whereas packet-level events occur on the order of milliseconds. Different levels capture different aspects of system behavior. Queueing theory can be used to model system performance at both packet and call levels [45, 46].

Call-level models treat an entire connection as a single arrival and focus on new call admission control, mobility management, and resource planning [47, 48]. The key performance metrics include the probability of blocking a new call, the probability of dropping a handoff call, and average call service time. The aim of

call-level models is to estimate the number of channels/carriers/terminals required to maintain the desired metrics at a predefined level. Such models typically use Poisson arrival processes and exponential call durations, leading to closed-form expressions.

Packet-level models [45, 49] consider each packet within the system. They can capture link-layer protocols, adaptive coding and modulation, and cross-layer scheduling. The key performance metrics include throughput, packet loss probability, and latency. Detailed discrete-event simulators or stochastic queueing models are used to compute these metrics, often at the cost of higher computational complexity.

It is important to note that call-level and packet-level dynamics are interdependent, e.g., accepting a handoff call increases the load and may cause additional losses. Therefore, it is possible to combine both approaches, using effective bandwidth to translate variable-rate traffic into equivalent traffic demands and jointly optimize admission control and packet-level scheduling [46].

7.2 Service types in HTS networks

HTS networks can support services with different QoS requirements [4, 7, 50]. Typical HTS systems include: broadband Internet access (web browsing, file downloads), real-time communication (VoIP, video conference), media streaming (live broadcasting), corporate or cellular backhaul, and IoT sensors (bursty low-rate traffic). Table 3 illustrates the diversity of traffic types in HTS networks along with their typical traffic volume, session duration, and QoS requirements [51, 52].

Each service has a distinct traffic profile: web browsing and file downloads are tolerant to variable rate and latency (so called elastic traffic), whereas VoIP/telephony requires low latency; video conferencing and streaming require consistent bandwidth, whereas video-on-demand platforms use buffers to smooth the delay. Elastic traffic can utilize any spare capacity in a network. If some beams or time slots are underutilized by real-time services, bandwidth can be dynamically reassigned to elastic data flows, allowing temporal increase of their rate. This maximizes overall throughput without compromising the QoS of delay-sensitive traffic.

Table 3

Representative HTS service characteristics

Service type	Typical volume	Session duration	QoS needs
Broadband Internet (web, downloads)	50-100 Mb per session	10-30 min	Delay tolerant; elastic throughput
Real-time communication (VoIP, video call)	1-5 Mb per session	5-10 min	Low latency; jitter sensitive
Video streaming (HD)	0.7-1.5 Gb per session	30-60 min	Stable rate; buffering is acceptable
Backhaul (corporate / cellular)	3-20 Gb per hour	continuous link	High reliability; moderate latency tolerance
IoT sensors (telemetry)	< 0.1 Mb per session	< 1 min	Delay-tolerant; minimal bandwidth

In practice, network traffic can be modeled as a mix of real-time (delay-sensitive) and elastic (delay-tolerant) flows [47, 48, 53]. Modern HTS payloads with multi-beam architectures allow dynamic power allocation among beams, allowing reconfiguring capacity on-the-fly to meet current QoS requirements. Moreover, queuing theory methods are applicable to both LEO and GEO SatCom systems [48, 54, 55] providing analytical tools to describe both data and control channels [56, 57].

8 Conclusion

High-Throughput Satellite (HTS) systems represent a major trend in satellite communications. Their design innovations enable capacities ranging from hundreds of gigabits per second up to terabits per satellite, supporting modern broadband applications over wide geographical areas.

The development of HTS systems has many technical challenges. One of the most important issues is resource planning. The paper provides a focused overview of HTS architecture, discussing design principles at PHY and MAC layers and their impact on system performance. We believe that a suitable analytical framework can be built on queuing theory and teletraffic models. Such a framework can guide the development of resource allocation strategies and support future system design.

The complexity of HTS systems will continue to grow. Future satellite networks are likely to become multi-orbit and multi-layered to optimize coverage and latency. Moreover, dynamic and AI-driven resource management will become important for handling highly variable traffic. Advanced concepts, such as Non-Terrestrial Networks for 6G, suggest new ways to exploit HTS capacity. Addressing these challenges will require further development of both system design techniques and analytical tools.

References

[1] M. Giordani and M. Zorzi, “Non-terrestrial networks in the 6G era: Challenges and opportunities,” *IEEE Network*, 2021, vol. 35, no. 2, pp. 244–251.

[2] S. Mahboob and L. Liu, “Revolutionizing future connectivity: A contemporary survey on AI-empowered satellite-based non-terrestrial networks in 6G,” *IEEE Communications Surveys & Tutorials*, 2024, vol. 26, no. 2, pp. 1279–1321.

[3] J. Heo, S. Sung, H. Lee, I. Hwang, and D. Hong, “MIMO satellite communication systems: A survey from the PHY layer perspective,” *IEEE Communications Surveys & Tutorials*, 2023, vol. 25, no. 3, pp. 1543–1570.

[4] O. B. Yahia, Z. Garroussi, O. Bélanger, B. Sansò, J.-F. Frigon, S. Martel, A. Lesage-Landry, and G. K. Kurt, “Evolution of high-throughput satellite systems: A vision of programmable regenerative payload,” *IEEE Communications Surveys & Tutorials*, 2025, vol. 27, no. 3, pp. 1565–1597.

[5] J. Zheng, T. H. Luan, G. Li, J. Zhao, Z. Yin, N. Cheng, and J. Pan, “Low Earth Orbit satellite networks: Architecture, key technologies, measurement, and open issues,” *IEEE Network*, 2025.

[6] W. Shao, S. Zhai, T. Hui, X. G. Zhang, and J. Li, “Optimization of a two-color reuse scheme in multi-beam high throughput satellite systems,” *Electronics Letters*, 2025, vol. 61, no. 1, p. e70131.

[7] P. Wang, J. Zhang, X. Zhang, Z. Yan, B. G. Evans, and W. Wang, “Convergence of satellite and terrestrial networks: A comprehensive survey,” *IEEE Access*, 2019, vol. 8, pp. 5550–5588.

[8] M. Höyhty, S. Boumard, A. Yastrebova, P. Järvensivu, M.

Kiviranta, and A. Anttonen, “Sustainable satellite communications in the 6G era: A European view for multilayer systems and space safety,” *IEEE Access*, 2022, vol. 10, pp. 99973–100005.

[9] R. Chen, W.-X. Long, B.-Q. Wang, Y. He, R.-J. Sun, N. Cheng, G. Zheng, and D. Niyato, “Multibeam high throughput satellite: Hardware foundation, resource allocation, and precoding,” *arXiv preprint*, 2025, arXiv:2508.00800.

[10] M. Handley, “Delay is not an option: Low latency routing in space,” in *Proc. 17th ACM Workshop on Hot Topics in Networks*, 2018, pp. 85–91.

[11] Z. Lin, Y. Mao, B. Clerckx, and G. K. Kurt, “Supporting IoT with rate-splitting multiple access in satellite and aerial-integrated networks,” *IEEE Internet of Things Journal*, 2021, vol. 8, no. 14, pp. 11123–11134.

[12] M. Centenaro, C. E. Costa, F. Granelli, C. Sacchi, and L. Vangelista, “A survey on technologies, standards and open challenges in satellite IoT,” *IEEE Communications Surveys & Tutorials*, 2021, vol. 23, no. 3, pp. 1693–1720.

[13] D. Zhao, H. Qin, N. Xin, and B. Song, “Flexible resource management in high-throughput satellite communication systems: A two-stage machine learning framework,” *IEEE Transactions on Communications*, 2023, vol. 71, no. 5, pp. 2724–2739.

[14] O. Kodheli, E. Lagunas, N. Maturo, S. K. Sharma, B. Shankar, J. F. M. Montoya, J. C. M. Duncan, D. Spano, S. Chatzinotas, S. Kisseleff, et al., “Satellite communications in the new space era: A survey and future challenges,” *IEEE Communications Surveys & Tutorials*, 2020, vol. 23, no. 1, pp. 70–109.

[15] ETSI, “Digital Video Broadcasting (DVB); Second generation framing structure, channel coding and modulation systems for broadcasting, interactive services, news gathering and other broadband satellite applications—Part 2: DVB-S2 Extensions (DVB-S2X),” ETSI EN 302 307-2 V1.3.1, 2021.

[16] ETSI, “Digital Video Broadcasting (DVB); Second generation DVB interactive satellite system (DVB-RCS2)—Part 3: Higher layers satellite specification,” ETSI TS 101 545-3 V1.3.1, 2020.

[17] X. Lin, S. Rommer, S. Euler, E. A. Yavuz, and R. S. Karlsson, “5G from space: An overview of 3GPP non-terrestrial networks,” *IEEE Communications Standards Magazine*, 2021, vol. 5, no. 4, pp. 147–153.

[18] 3GPP, “System architecture for the 5G system (5GS),” 3GPP TS 23.501, Version 17.9.0, Release 17, 2023.

[19] M. Khammassi, A. Kammoun, and M.-S. Alouini, “Precoding for high-throughput satellite communication systems: A survey,” *IEEE Communications Surveys & Tutorials*, 2023, vol. 26, no. 1, pp. 80–118.

[20] ITU-R, “IMT vision—Framework and overall objectives of the future development of IMT for 2020 and beyond,” Recommendation ITU-R M.2083-0, 2015.

[21] ITU-R, “Framework and overall objectives of the future development of IMT for 2030 and beyond,” Recommendation ITU-R M.2160-0, 2023.

[22] A. A. Maslov, G. V. Sebekin, M. S. Stepanov, S. N. Stepanov, and A. O. Shchurkov, “Model' seti peredachi dannykh IoT na baze kosmicheskikh apparatov na nizkikh krugovykh orbitakh. Chast' 1. Rezhim odnostoronnego sluchainogo mnozhestvennogo dostupa,” *Informatsionnye protsessy*, 2025, vol. 25, no. 3, pp. 456–471.

[23] A. A. Maslov, G. V. Sebekin, M. S. Stepanov, S. N. Stepanov, and A. O. Shchurkov, “Model' seti peredachi dannykh IoT na baze kosmicheskikh apparatov na nizkikh krugovykh orbitakh. Chast' 2. Rezhim sluchainogo mnozhestvennogo dostupa s podtverzheniem priema paketov,” *Informatsionnye protsessy*, 2025, vol. 25, no. 3, pp. 472–489.

[24] F. S. Prol, J. A. Del Peral-Rosado, R. A. Rosa, J. A. Lopez-Salcedo, and G. Seco-Granados, “Position, navigation, and timing (PNT) through low earth orbit (LEO) satellites: A survey on current status, challenges, and opportunities,” *IEEE Access*, 2022, vol. 10, pp. 83971–84002.

[25] M. Strohmeier, M. Schäfer, V. Lenders, and I. Martinovic, “Realities and challenges of nextgen air traffic management: The case of

ADS-B," *IEEE Communications Magazine*, 2014, vol. 52, no. 5, pp. 111–118.

[26] D. Yang, L. Wu, S. Wang, H. Jia, and K. X. Li, "How big data enriches maritime research—A critical review of automatic identification system (AIS) data applications," *Transport Reviews*, 2019, vol. 39, no. 6, pp. 755–773.

[27] P. Wang, H. Li, B. Chen, and S. Zhang, "Enhancing earth observation throughput using inter-satellite communication," *IEEE Transactions on Wireless Communications*, 2022, vol. 21, no. 10, pp. 7990–8006.

[28] H. Al-Hraishawi, M. Minardi, H. Chougrani, O. Kodheli, J. F. M. Montoya, and S. Chatzinotas, "Multi-layer space information networks: Access design and softwarization," *IEEE Access*, 2021, vol. 9, pp. 158587–158598.

[29] M. Höyhtyä, A. Anttonen, M. Majanen, A. Yastrebova-Castillo, M. Varga, L. Lodigiani, et al., "Multi-layered satellite communications systems for ultra-high availability and resilience," *Electronics*, 2024, vol. 13, no. 7, p. 1269.

[30] H. Al-Hraishawi, H. Chougrani, S. Kisseleff, E. Lagunas, and S. Chatzinotas, "A survey on nongeostationary satellite systems: The communication perspective," *IEEE Communications Surveys & Tutorials*, 2022, vol. 25, no. 1, pp. 101–132.

[31] L. Lei, J. Wan, and Q. Gong, "Multibeam antennas with reflector for high throughput satellite applications," *IEEE Aerospace and Electronic Systems Magazine*, 2021, vol. 37, no. 2, pp. 34–46.

[32] G. Bacci, R. De Gaudenzi, M. Luise, L. Sanguinetti, and E. Sebastiani, "Formation-of-arrays antenna technology for high-throughput mobile nonterrestrial networks," *IEEE Transactions on Aerospace and Electronic Systems*, 2023, vol. 59, no. 5, pp. 4919–4935.

[33] X. Tao, W. Ding, S. Wang, X. Zhang, W. Cui, and Q. Cui, "An active multi-beam antenna design method and its application for the future 6G satellite network," *Space: Science and Technology*, 2024, vol. 4, p. 0149.

[34] Y. Antia, S. Morrar, D. Roos, and the Hughes Network Systems Team, "JUPITER high throughput satellite system—500 Gbps from space," *Microwave Journal*, 2019, vol. 62, no. 10, pp. 106–109.

[35] J. Garrity and A. Husar, *Digital connectivity and low Earth orbit satellite constellations: Opportunities for Asia and the Pacific*, Asian Development Bank, 2021.

[36] S. Dimitrov, S. Erl, B. Barth, S. Jaekel, A. Kyrgiazos, and B. G. Evans, "Radio resource management techniques for high throughput satellite communication systems," in *Proc. European Conference on Networks and Communications (EuCNC)*, 2015, pp. 175–179.

[37] L. Lei, E. Lagunas, Y. Yuan, M. G. Kibria, S. Chatzinotas, and B. Ottersten, "Beam illumination pattern design in satellite networks: Learning and optimization for efficient beam hopping," *IEEE Access*, 2020, vol. 8, pp. 136655–136667.

[38] Y. Feng, Y. Sun, and M. Peng, "Performance analysis in satellite communication with beam hopping using discrete-time queuing theory," *IEEE Internet of Things Journal*, 2024, vol. 11, no. 7, pp. 11679–11692.

[39] X. Hu, Y. Zhang, X. Liao, Z. Liu, W. Wang, and F. M. Ghannouchi, "Dynamic beam hopping method based on multi-objective deep reinforcement learning for next generation satellite broadband systems," *IEEE Transactions on Broadcasting*, 2020, vol. 66, no. 3, pp. 630–646.

[40] Z. Lin, Z. Ni, L. Kuang, C. Jiang, and Z. Huang, "Dynamic beam pattern and bandwidth allocation based on multi-agent deep reinforcement learning for beam hopping satellite systems," *IEEE Transactions on Vehicular Technology*, 2022, vol. 71, no. 4, pp. 3917–3930.

[41] P. J. Honnaiah, E. Lagunas, D. Spano, N. Maturo, and S. Chatzinotas, "Demand-based scheduling for precoded multibeam high-throughput satellite systems," in *Proc. IEEE Wireless Communications and Networking Conference (WCNC)*, 2021, pp. 1–6.

[42] P. D. Arapoglou, A. Ginesi, S. Cioni, S. Erl, F. Clazzer, S. Andrenacci, and A. Vannelli-Coralli, "DVB-S2X-enabled precoding for high throughput satellite systems," *International Journal of Satellite Communications and Networking*, 2016, vol. 34, no. 3, pp. 439–455.

[43] C. Zhang, X. Zhao, and G. Zhang, "Joint precoding schemes for flexible resource allocation in high throughput satellite systems based on beam hopping," *China Communications*, 2021, vol. 18, no. 9, pp. 48–61.

[44] T. Van Chien, E. Lagunas, T. H. Ta, S. Chatzinotas, and B. Ottersten, "User scheduling and power allocation for precoded multi-beam high throughput satellite systems with individual quality of service constraints," *IEEE Transactions on Vehicular Technology*, 2022, vol. 72, no. 1, pp. 907–923.

[45] M. Ghaderi, R. Boutaba, and G. W. Kenward, "Joint call and packet QoS in cellular packet networks," *Computer Networks*, 2007, vol. 51, no. 4, pp. 1060–1071.

[46] D. Niyato and E. Hossain, "A novel analytical framework for integrated cross-layer study of call-level and packet-level QoS in wireless mobile multimedia networks," *IEEE Transactions on Mobile Computing*, 2007, vol. 6, no. 3, pp. 322–335.

[47] G. Sebekin, A. Maslov, and A. Shchurkov, "Simulation of joint maintenance of real-time multiservice traffic and elastic data traffic in networks based on high-throughput satellite," *Journal Information-Measuring and Control Systems*, 2024, vol. 22, no. 1, pp. 11–22.

[48] A. A. Maslov, G. V. Sebekin, S. N. Stepanov, A. O. Shchurkov, and A. P. Vasilyev, "Model of processes for joint maintenance of real-time multiservice traffic and elastic data traffic in a network of low-power mobile subscriber terminals based on high-throughput satellites," *T-Comm*, 2024, vol. 18, no. 3, pp. 41–49.

[49] J. Puttonen, S. Rantanen, F. Laakso, J. Kurjenniemi, K. Aho, and G. Acar, "A packet level simulator for future satellite communications research," in *Proc. AIAA International Communications Satellite Systems Conference*, 2014, p. 4437.

[50] A. Troegubov, S. Stepanov, E. Malikova, and D. X. Thu, "Modeling of multiservice traffic service process in a high-throughput satellite network," in *Proc. Systems of Signals Generating and Processing in the Field of On-Board Communications*, 2025, pp. 1–7.

[51] Cisco Systems, Inc., *Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update*, 2015.

[52] A. S. G. Andrae, "New perspectives on internet electricity use in 2030," *Engineering and Applied Science Letters*, 2020, vol. 3, no. 2, pp. 19–31.

[53] S. N. Stepanov and M. S. Stepanov, "Approximate method for estimating characteristics of joint service of real-time traffic and elastic data traffic in multiservice access nodes," *Automation and Remote Control*, 2023, vol. 84, no. 11, pp. 1191–1207.

[54] A. A. Maslov, G. V. Sebekin, M. S. Stepanov, S. N. Stepanov, and A. O. Shchurkov, "Modeling the processes of subscriber servicing in a data transmission network based on spacecraft in a low circular orbit. I," *Journal of Communications Technology and Electronics*, 2024, vol. 69, no. 10, pp. 376–385.

[55] A. A. Maslov, G. V. Sebekin, M. S. Stepanov, S. N. Stepanov, and A. O. Shchurkov, "Modeling of subscriber service processes in a data transmission network based on high-throughput satellites. II," *Journal of Communications Technology and Electronics*, 2025, vol. 70, no. 4, pp. 168–181.

[56] A. A. Maslov, G. V. Sebekin, M. S. Stepanov, S. N. Stepanov, and A. O. Shchurkov, "Model' obsluzhivaniya mult'iservisnogo trafika v uzle dostupa sputnikovoi seti svyazi s dinamicheski izmenyaemoi skorost'yu predostavleniya servisov," *Avtomatika i Telemekhanika*, 2025, no. 11, pp. 75–91.

[57] A. A. Maslov, G. V. Sebekin, M. S. Stepanov, S. N. Stepanov, and A. O. Shchurkov, "Modelirovanie raboty sluzhebnogo kanala sputnikovykh setei shirokopolosnogo dostupa," *Informatsionnye Protsesty*, 2025, vol. 25, no. 4, pp. 765–786.

КЛЮЧЕВЫЕ АСПЕКТЫ СПУТНИКОВЫХ СИСТЕМ С ВЫСОКОЙ ПРОПУСКНОЙ СПОСОБНОСТЬЮ: ОТ СТАНДАРТИЗАЦИИ ДО К УПРАВЛЕНИЮ КАЧЕСТВОМ ОБСЛУЖИВАНИЯ

Троегубов Александр Юрьевич, Московский физико-технический институт (национальный исследовательский университет), Москва, Россия, trogubov.aiu@phystech.edu

Маслов Александр Александрович, Московский физико-технический институт (национальный исследовательский университет), Москва, Россия, maslov.aa@mipt.ru

Себекин Геннадий Валерьевич, Московский физико-технический институт (национальный исследовательский университет), Москва, Россия, sebekin.gv@mipt.ru

Степанов Сергей Николаевич, Московский технический университет связи и информатики, Москва, Россия, s.n.stepanov@mtuci.ru

Щурков Алексей Олегович, Московский физико-технический институт (национальный исследовательский университет), Москва, Россия, shchurkov.ao@mipt.ru

Работа второго, третьего и пятого авторов выполнена в МФТИ, Физтех при финансовой поддержке Фонда НТИ в рамках Договора №70-2025-000804 от 26.05.2025

Аннотация

Повсеместное распространение сетей связи является ключевым для многих сфер человеческой деятельности, однако значительная часть регионов имеет низкое покрытие из-за высокой стоимости или сложности развертывания наземной инфраструктуры. Спутниковые системы связи позволяют преодолеть данное ограничение за счет глобального покрытия и поддержки широкоэшелонных и многоадресных сервисов, что делает их важной составляющей современных сетей связи, дополняющей наземные системы. В работе представлен обзор современных спутниковых систем с высокой пропускной способностью (HTS) с акцентом на эволюцию архитектур, вопросы стандартизации и управление ресурсами для выполнения требований к качеству обслуживания. Развитие HTS-систем сопровождается рядом технических вызовов. Рассматриваются текущие инициативы в области стандартизации и регулирования использования радиочастотного спектра, а также типовые приложения и сценарии применения. Кроме того, анализируется архитектура HTS-сетей, охватывающая геостационарные и негеостационарные орбиты, многолучевые полезные нагрузки и многоуровневые развертывания, а также принципы проектирования на физическом уровне и уровне управления доступом к среде. В работе также обсуждаются вопросы дифференциации сервисов и поддержки требований к качеству обслуживания в HTS-сетях, приводятся характеристики трафика и соответствующих показателей производительности на различных временных масштабах функционирования системы. Показано, что методы теории массового обслуживания и телетрафика могут служить подходящей аналитической основой для оценки производительности HTS-систем и обоснования проектных решений при создании эффективных спутниковых сетей в будущих системах связи.

Ключевые слова: Спутниковые системы с высокой пропускной способностью, спутниковая связь, многолучевые системы, управление ресурсами, качество обслуживания.

Литература

1. Giordani M., Zorzi M. Non-terrestrial networks in the 6G era: Challenges and opportunities // IEEE Network. 2021. Vol. 35, № 2. P. 244-251.
2. Mahboob S., Liu L. Revolutionizing future connectivity: A contemporary survey on AI-empowered satellite-based non-terrestrial networks in 6G // IEEE Communications Surveys & Tutorials. 2024. Vol. 26, № 2. P. 1279-1321.
3. Heo J., Sung S., Lee H., Hwang I., Hong D. MIMO satellite communication systems: A survey from the PHY layer perspective // IEEE Communications Surveys & Tutorials. 2023. Vol. 25, № 3. P. 1543-1570.
4. Yahia O.B., Garroussi Z., B?langer O., Sanso B., Frigon J.-F., Martel S., Lesage-Landry A., Kurt G.K. Evolution of high-throughput satellite systems: A vision of programmable regenerative payload // IEEE Communications Surveys & Tutorials. 2025. Vol. 27, № 3. P. 1565-1597.
5. Zheng J., Luan T. H., Li G., Zhao J., Yin Z., Cheng N., Pan J. Low Earth Orbit satellite networks: Architecture, key technologies, measurement, and open issues // IEEE Network. 2025.
6. Shao W., Zhai S., Hui T., Zhang X. G., Li J. Optimization of a two-color reuse scheme in multi-beam high throughput satellite systems // Electronics Letters. 2025. Vol. 61, № 1. P. e70131.
7. Wang P., Zhang J., Zhang X., Yan Z., Evans B. G., Wang W. Convergence of satellite and terrestrial networks: A comprehensive survey // IEEE Access. 2019. Vol. 8. P. 5550-5588.
8. Hoyhtya M., Boumard S., Yastrebova A., Jarvensivu P., Kiviranta M., Anttonen A. Sustainable satellite communications in the 6G era: A European view for multilayer systems and space safety // IEEE Access. 2022. Vol. 10. P. 99973-100005.
9. Chen R., Long W.-X., Wang B.-Q., He Y., Sun R.-J., Cheng N., Zheng G., Niyato D. Multibeam high throughput satellite: Hardware foundation, resource allocation, and precoding // arXiv preprint. 2025. arXiv:2508.00800.
10. Handley M. Delay is not an option: Low latency routing in space // Proc. 17th ACM Workshop on Hot Topics in Networks. 2018. P. 85-91.
11. Lin Z., Mao Y., Clerckx B., Kurt G. K. Supporting IoT with rate-splitting multiple access in satellite and aerial-integrated networks // IEEE Internet of Things Journal. 2021. Vol. 8, № 14. P. 11123-11134.
12. Centenaro M., Costa C. E., Granelli F., Sacchi C., Vangelista L. A survey on technologies, standards and open challenges in satellite IoT // IEEE Communications Surveys & Tutorials. 2021. Vol. 23, № 3. P. 1693-1720.
13. Zhao D., Qin H., Xin N., Song B. Flexible resource management in high-throughput satellite communication systems: A two-stage machine learning framework // IEEE Transactions on Communications. 2023. Vol. 71, № 5. P. 2724-2739.
14. Kodheli O., Lagunas E., Maturro N., Sharma S. K., Shankar B., Montoya J. F. M., Duncan J. C. M., Spano D., Chatzinotas S., Kisseleff S. et al. Satellite communications in the new space era: A survey and future challenges // IEEE Communications Surveys & Tutorials. 2020. Vol. 23, № 1. P. 70-109.
15. ETSI. Digital Video Broadcasting (DVB); Second generation framing structure, channel coding and modulation systems for broadcasting, interactive services, news gathering and other broadband satellite applications-Part 2: DVB-S2 Extensions (DVB-S2X). ETSI EN 302 307-2 V1.3.1. 2021.
16. ETSI. Digital Video Broadcasting (DVB); Second generation DVB interactive satellite system (DVB-RCS2)-Part 3: Higher layers satellite specification. ETSI TS 101 545-3 V1.3.1. 2020.
17. Lin X., Rommer S., Euler S., Yavuz E. A., Karlsson R. S. 5G from space: An overview of 3GPP non-terrestrial networks // IEEE Communications Standards Magazine. 2021. Vol. 5, № 4. P. 147-153.
18. 3GPP. System architecture for the 5G system (5GS). 3GPP TS 23.501. Version 17.9.0. Release 17. 2023.
19. Khammassi M., Kammoun A., Alouini M.-S. Precoding for high-throughput satellite communication systems: A survey // IEEE Communications Surveys & Tutorials. 2023. Vol. 26, № 1. P. 80-118.
20. ITU-R. IMT vision-Framework and overall objectives of the future development of IMT for 2020 and beyond. Recommendation ITU-R M.2083-0. 2015.

21. ITU-R. Framework and overall objectives of the future development of IMT for 2030 and beyond. Recommendation ITU-R M.2160-0. 2023.
22. Маслов А.А., Себекин Г.В., Степанов М.С., Степанов С.Н., Щурков А.О. Модель сети передачи данных IoT на базе космических аппаратов на низких круговых орбитах. Часть 1. Режим одностороннего случайного множественного доступа // Информационные процессы. 2025. Т. 25, № 3. С. 456-471.
23. Маслов А.А., Себекин Г.В., Степанов М.С., Степанов С.Н., Щурков А.О. Модель сети передачи данных IoT на базе космических аппаратов на низких круговых орбитах. Часть 2. Режим случайного множественного доступа с подтверждением приема пакетов // Информационные процессы. 2025. Т. 25, № 3. С. 472-489.
24. Prol F.S., Del Peral-Rosado J.A., Rosa R.A., Lopez-Salcedo J.A., Seco-Granados G. Position, navigation, and timing (PNT) through low earth orbit (LEO) satellites: A survey on current status, challenges, and opportunities // IEEE Access. 2022. Vol. 10. P. 83971-84002.
25. Strohmeier M., Sch?fer M., Lenders V., Martinovic I. Realities and challenges of nextgen air traffic management: The case of ADS-B // IEEE Communications Magazine. 2014. Vol. 52, № 5. P. 111-118.
26. Yang D., Wu L., Wang S., Jia H., Li K.X. How big data enriches maritime research-A critical review of automatic identification system (AIS) data applications // Transport Reviews. 2019. Vol. 39, № 6. P. 755-773.
27. Wang P., Li H., Chen B., Zhang S. Enhancing earth observation throughput using inter-satellite communication // IEEE Transactions on Wireless Communications. 2022. Vol. 21, № 10. P. 7990-8006.
28. Al-Hraishawi H., Minardi M., Chougrani H., Kodheli O., Montoya J.F.M., Chatzinotas S. Multi-layer space information networks: Access design and softwarization // IEEE Access. 2021. Vol. 9. P. 158587-158598.
29. Hoyhtya M., Anttonen A., Majanen M., Yastrebova-Castillo A., Varga M., Lodigiani L. и др. Multi-layered satellite communications systems for ultra-high availability and resilience // Electronics. 2024. Vol. 13, № 7. P. 1269.
30. Al-Hraishawi H., Chougrani H., Kisseleff S., Lagunas E., Chatzinotas S. A survey on nongeostationary satellite systems: The communication perspective // IEEE Communications Surveys & Tutorials. 2022. Vol. 25, № 1. P. 101-132.
31. Lei L., Wan J., Gong Q. Multibeam antennas with reflector for high throughput satellite applications // IEEE Aerospace and Electronic Systems Magazine. 2021. Vol. 37, № 2. P. 34-46.
32. Bacci G., De Gaudenzi R., Luise M., Sanguinetti L., Sebastiani E. Formation-of-arrays antenna technology for high-throughput mobile nonterrestrial networks // IEEE Transactions on Aerospace and Electronic Systems. 2023. Vol. 59, № 5. P. 4919-4935.
33. Tao X., Ding W., Wang S., Zhang X., Cui W., Cui Q. An active multi-beam antenna design method and its application for the future 6G satellite network // Space: Science and Technology. 2024. Vol. 4. P. 0149.
34. Antia Y., Marrar S., Roos D., Hughes Network Systems Team. JUPITER high throughput satellite system-500 Gbps from space // Microwave Journal. 2019. Vol. 62, № 10. P. 106-109.
35. Garrity J., Husar A. Digital connectivity and low Earth orbit satellite constellations: Opportunities for Asia and the Pacific. Asian Development Bank. 2021.
36. Dimitrov S., Erl S., Barth B., Jaeckel S., Kyrgiazos A., Evans B.G. Radio resource management techniques for high throughput satellite communication systems // Proc. European Conference on Networks and Communications (EuCNC). 2015. P. 175-179.
37. Lei L., Lagunas E., Yuan Y., Kibria M.G., Chatzinotas S., Ottersten B. Beam illumination pattern design in satellite networks: Learning and optimization for efficient beam hopping // IEEE Access. 2020. Vol. 8. P. 136655-136667.
38. Feng Y., Sun Y., Peng M. Performance analysis in satellite communication with beam hopping using discrete-time queueing theory // IEEE Internet of Things Journal. 2024. Vol. 11, № 7. P. 11679-11692.
39. Hu X., Zhang Y., Liao X., Liu Z., Wang W., Ghannouchi F.M. Dynamic beam hopping method based on multi-objective deep reinforcement learning for next generation satellite broadband systems // IEEE Transactions on Broadcasting. 2020. Vol. 66, № 3. P. 630-646.
40. Lin Z., Ni Z., Kuang L., Jiang C., Huang Z. Dynamic beam pattern and bandwidth allocation based on multi-agent deep reinforcement learning for beam hopping satellite systems // IEEE Transactions on Vehicular Technology. 2022. Vol. 71, № 4. P. 3917-3930.
41. Honnaiah P.J., Lagunas E., Spano D., Maturo N., Chatzinotas S. Demand-based scheduling for precoded multibeam high-throughput satellite systems // Proc. IEEE Wireless Communications and Networking Conference (WCNC). 2021. P. 1-6.
42. Arapoglou P.D., Ginesi A., Cioni S., Erl S., Clazzer F., Andrenacci S., Vanelli-Coralli A. DVB-S2X-enabled precoding for high throughput satellite systems // International Journal of Satellite Communications and Networking. 2016. Vol. 34, № 3. P. 439-455.
43. Zhang C., Zhao X., Zhang G. Joint precoding schemes for flexible resource allocation in high throughput satellite systems based on beam hopping // China Communications. 2021. Vol. 18, № 9. P. 48-61.
44. Van Chien T., Lagunas E., Ta T.H., Chatzinotas S., Ottersten B. User scheduling and power allocation for precoded multi-beam high throughput satellite systems with individual quality of service constraints // IEEE Transactions on Vehicular Technology. 2022. Vol. 72, № 1. P. 907-923.
45. Ghaderi M., Boutaba R., Kenward G.W. Joint call and packet QoS in cellular packet networks // Computer Networks. 2007. Vol. 51, № 4. P. 1060-1071.
46. Niyato D., Hossain E. A novel analytical framework for integrated cross-layer study of call-level and packet-level QoS in wireless mobile multimedia networks // IEEE Transactions on Mobile Computing. 2007. Vol. 6, № 3. P. 322-335.
47. Sebekin G., Maslov A., Shchurkov A. Simulation of joint maintenance of real-time multiservice traffic and elastic data traffic in networks based on high-throughput satellite // Journal of Information-Measuring and Control Systems. 2024. Vol. 22, № 1. P. 11-22.
48. Maslov A.A., Sebekin G.V., Stepanov S.N., Shchurkov A.O., Vasilyev A.P. Model of processes for joint maintenance of real-time multiservice traffic and elastic data traffic in a network of low-power mobile subscriber terminals based on high-throughput satellites // T-Comm. 2024. Vol. 18, № 3. P. 41-49.
49. Puttonen J., Rantanen S., Laakso F., Kurjenniemi J., Aho K., Acar G. A packet level simulator for future satellite communications research // Proc. AIAA International Communications Satellite Systems Conference. 2014. P. 4437.
50. Troegubov A., Stepanov S., Malikova E., Thu D.X. Modeling of multiservice traffic service process in a high-throughput satellite network // Proc. Systems of Signals Generating and Processing in the Field of On-Board Communications. 2025. P. 1-7.
51. Cisco Systems, Inc. Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update. 2015.
52. Andrae A.S.G. New perspectives on internet electricity use in 2030 // Engineering and Applied Science Letters. 2020. Vol. 3, № 2. P. 19-31.
53. Stepanov S.N., Stepanov M.S. Approximate method for estimating characteristics of joint service of real-time traffic and elastic data traffic in multiservice access nodes // Automation and Remote Control. 2023. Vol. 84, № 11. P. 1191-1207.
54. Maslov A.A., Sebekin G.V., Stepanov M.S., Stepanov S.N., Shchurkov A.O. Modeling the processes of subscriber servicing in a data transmission network based on spacecraft in a low circular orbit. I // Journal of Communications Technology and Electronics. 2024. Vol. 69, № 10. P. 376-385.
55. Maslov A.A., Sebekin G.V., Stepanov M.S., Stepanov S.N., Shchurkov A.O. Modeling of subscriber service processes in a data transmission network based on high-throughput satellites. II // Journal of Communications Technology and Electronics. 2025. Vol. 70, № 4. P. 168-181.
56. Маслов А.А., Себекин Г.В., Степанов М.С., Степанов С.Н., Щурков А.О. Модель обслуживания мультисервисного трафика в узле доступа спутниковой сети связи с динамически изменяемой скоростью предоставления сервисов // Автоматика и телемеханика. 2025. № 11. С. 75-91.
57. Маслов А.А., Себекин Г.В., Степанов М.С., Степанов С.Н., Щурков А.О. Моделирование работы служебного канала спутниковых сетей широкополосного доступа // Информационные процессы. 2025. Т. 25, № 4. С. 765-786.

DATA MODEL DEVELOPMENT FOR LTE ENODEB MANAGING IN THE OPENRAN RADIO ACCESS NETWORK

Sergey V. Terentiev,
Lomonosov Moscow State University;
Softaym LLC, Moscow, Russia, s.ter@mail.ru

Valery O. Tikhvinsky,
Bauman Moscow State Technical University, Moscow, Russia;
and Sevastopol State University, Sevastopol, Russia;
M.I. Krivosheev National Research Centre for Telecommunication
(NRCT), Moscow, Russia, vtniir@mail.ru

Evgeny E. Devyatkin,
M.I. Krivosheev National Research Centre for Telecommunication
(NRCT), Moscow, Russia, deugene@list.ru

Grigoriy A. Fokin,
Bonch-Bruевич Saint-Petersburg State University of
Telecommunications, Saint-Petersburg, Russia,
grihafokin@gmail.com

DOI: 10.36724/2072-8735-2026-20-3-65-72

Manuscript received 02 December 2025;
Accepted 05 February 2026

Keywords: O-RAN, YANG, NETCONF, LTE, JSON, XML

The trends in the open eNodeB/gNB base station interfaces utilization for both 4G/LTE and 5G/NR technologies are relevant not only for network interfaces of the signaling (Control Plane) and user data (User Plane) planes, but also for management plane (Management Plane) interfaces. WG4 working group of OpenRAN Alliance has done important work on standardizing these management interfaces and developed the O-RAN.WG4.TS.MP.0 "Management Plane Specification" technical specification for managing Radio units (O-RU) of distributed eNodeB/gNB base stations. However, the issue of further standardization of configuring the distributed O-DU unit by the eNB/gNB base station in accordance with the operator's frequency-territorial plan during deployment on the site is currently open. A solution to this issue would allow for the unification of the EMS-RAN management system for OpenRAN network modules when using multivendor radio access networks for building 4G/LTE and 5G/NR networks. The 3GPP defines possibility of managing and orchestrating for 5G radio access network (RAN) using the NETCONF protocol and the YANG data modeling language as one of technical solutions for 5G Service-Based Management Architecture (SBMA). The article is devoted to the YANG model development of configuration data for eNodeB base station of OpenRAN network of LTE technology using NETCONF network protocol, as well as the development of sysrepo applications for applying configurations.

Information about authors:

Sergey V. Terentiev, Ph.D. (Tech.), Associate Professor at Lomonosov Moscow State University, Deputy Director of the Radio Subsystem Development Department (R&D), Softaym LLC, Moscow, Russia

Valery O. Tikhvinsky, Professor in Bauman Moscow State Technical University Moscow, Russia; Professor in Sevastopol State University, Sevastopol, Russia; Doctor of Economics, Chief Researcher of the M.I. Krivosheev National Research Centre for Telecommunication (NRCT), Moscow, Russia

Evgeny E. Devyatkin, Ph.D. (Econ.), Director of the Center for Research of Advanced Wireless Communication Technologies (CRWCT) of the M.I. Krivosheev National Research Centre for Telecommunication (NRCT), Moscow, Russia

Grigoriy A. Fokin, Doctor of Technical Sciences, Associate Professor, Head of the Department of Wireless Technologies and Systems (WTS) of the Bonch-Bruевич Saint-Petersburg State University of Telecommunications, Saint-Petersburg, Russia

Для цитирования:

Терентьев С.В., Тихвинский В.О., Десяткин Е.Е., Фокин Г.А. Разработка модели данных для управления базовой станцией eNodeB технологии LTE в сети радиодоступа OpenRAN // Т-Comm: Телекоммуникации и транспорт. 2026. Том 20. №3. С. 65-72.

For citation:

S.V. Terentiev, V.O. Tikhvinsky, E.E. Devyatkin, G.A. Fokin, "Data model development for LTE eNodeB managing in the OpenRAN radio access network," T-Comm, 2026, vol. 20, no. 3, pp. 65-72.

Introduction

The trends in the open eNodeB/gNB base station interfaces utilization for both 4G/LTE and 5G/NR technologies are relevant not only for network interfaces of the signaling (Control Plane) and user data (User Plane) planes [1], but also for management plane (Management Plane) interfaces.

WG4 working group of OpenRAN Alliance [2] has done important work on standardizing these management interfaces and developed the O-RAN.WG4.TS.MP.0 “Management Plane Specification” [3] technical specification for managing Radio units (O-RU) of distributed eNodeB/gNB base stations. However, the issue of further standardization of configuring the distributed O-DU unit by the eNB/gNB base station in accordance with the operator's frequency-territorial plan during deployment on the site is currently open. A solution to this issue would allow for the unification of the EMS-RAN management system for OpenRAN network modules when using multivendor radio access networks for building 4G/LTE and 5G/NR networks.

The 3GPP defines possibility of managing and orchestrating for 5G radio access network (RAN) using the NETCONF protocol and the YANG data modeling language as one of technical solutions for 5G Service-Based Management Architecture (SBMA) [22]. Key technical specifications in this area are (see Table 1):

- TS 28.533 [4] defines a service-oriented SBMA architecture in which management services (MnS) are represented by various protocols, including NETCONF/YANG;
- TS 28.532 [5] defines general CRUD (Create, Read, Update and Delete) management operations and options for their execution by standard NETCONF protocol operations, for example, <edit-config>;
- TS 28.541 [6] contains definitions of information Network Resource models for 5G RAN (New Radio interfaces), including 5G base stations gNB, Centralized Units (CU) and Distributed Units (DU), presented as YANG modules.

Table 1

Review of control tasks and protocols used [4]

Management Feature	Management Capability	MnS definition	Solution Sets
Network and network slicing management	NR Provisioning	CRUD operations/ notifications (3GPP TS 28.532 [5]) + NR NRM fragment (3GPP TS 28.541 [6])	RESTFUL NETCONF/ YANG
	5GC Provisioning	CRUD operations/ notifications (TS 28.532 [5]) + 5GC NRM fragment (3GPP TS 28.541 [6])	RESTFUL NETCONF/ YANG
	Network Slicing Provisioning	CRUD operations/ notifications (3GPP TS 28.532 [5]) + Network Slicing NRM fragment (3GPP TS 28.541 [6]) Network slicing provisioning service (3GPP TS 28.531 [7])	RESTFUL NETCONF/ YANG RESTFUL

Analysis of the execution options for general CRUD management operations for various solution sets (Table 1) will be determined by standard operations of the RESTFUL and NETCONF/YANG protocols.

The advantages of a management approach using a combination of NETCONF/YANG are:

- Interoperability – operators can manage equipment from different vendors within a single management environment;
- Automation – programmability of data models reduces the risk of manual configuration errors;
- Security – using SSH/TLS to transmit NETCONF messages ensures high security of management data.

Authors presented a developed model based on the YANG data modeling language, which allows for the practical configuration of eNodeB in to LTE network in accordance with the operator's frequency-territorial plan at all stages of the network life cycle.

1. Management of eNodeB base stations at different stages of the network life cycle

LTE mobile network and, correspondingly, the base station eNodeB of Radio access network (RAN) must be managed throughout its entire lifecycle – from deployment to dismantling (Fig. 1). The eNodeB lifecycle includes a structured process of planning, deploying, operating, and updating RAN parameters by configuring it to ensure efficient LTE network operation. The final stage of the eNodeB lifecycle involves removing the eNodeB from the LTE network and dismantling its configurations as part of a network resource upgrade.

During the development of the proposed configuration model, the following stages were considered as main stages of RAN LTE base station lifecycle, where eNodeB configuration is required in accordance with the operator's frequency – territorial plan:

- Deployment and commissioning of eNodeB base stations;
- Monitoring the status and configuration parameters of eNodeBs for subsequent optimization during network operation.
- Network optimization during operation to improve key performance indicators and service quality.

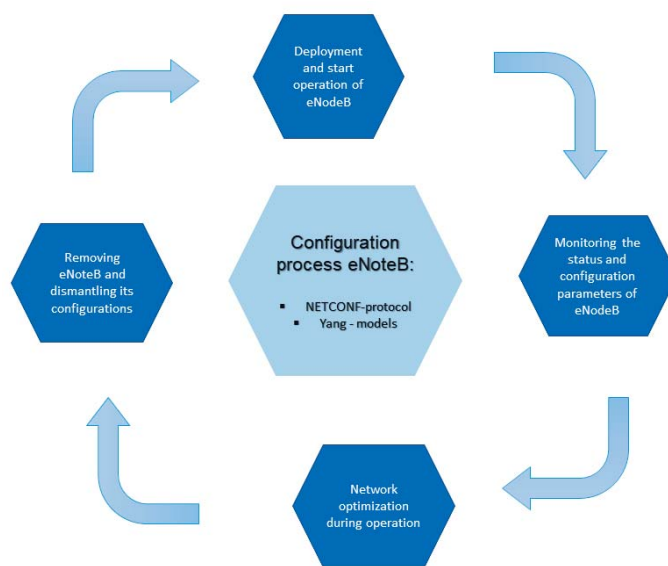


Fig. 1. Configuration process eNodeB from deployment to dismantling

The developed YANG model and applications enable management of eNodeB base stations at the following stages of the network lifecycle:

- during commissioning of the network deployment during the commissioning of the base station;
- during network optimization to improve key performance indicators for both the base station itself and the radio access network as a whole;
- during monitoring of the base station status and configuration parameters during operation.

At the stage of commissioning work for network deployment, when the Base station is put into operation, the operability of the eNodeB base station equipment is checked, its operating parameters are configured and its integration with the operator's network is ensured.

During the network optimization stage, when improving key performance indicators of the eNodeB base station, it is configured to improve coverage quality, increase cell capacity, connection reliability, and data speed.

During the monitoring stage of status and configuration parameters for eNodeB base station, in operation continuous remote monitoring of equipment's performances, collection of telemetry data, and analysis of key performance indicators are carried out to identify traffic anomalies for subsequent configuration.

Thus, the configuration of eNodeB base stations is a critical process that requires the development of a protocol for managing the configuration of LTE base station utilization, one of configuration data models of eNodeB base stations applying specialized software that defines the Network protocols stack of Radio access networks.

2. Capabilities analysis of NETCONF OpenRAN control protocol and problem statement

In accordance with the technical specification O-RAN.WG4.TS.MP.0 [3], O-RU units are controlled using the NETCONF network configuration protocol (RFC 6241) [8]. The NETCONF protocol is a multi-layer basic protocol for managing network units, replacing the outdated SNMP protocol (RFC 3411) [9].

As shown in Fig. 2, the NETCONF protocol [9] is based on either the SSH (RFC 4253) [10] or TLS (RFC 8446) [11] transport protocol and contains four layers:

1. **“Content” level** – management content containing configuration data or status data in XML format;

2. **“Operations” level** is a service part that designates the type of operation being performed, for example, `<get>`, `<get-config>`, `<edit-config>`, `<copy-config>`, `<delete-config>`, `<lock>`, `<unlock>`, `<close-session>` or `<kill-session>` for a given datastore:

- "Running" datastore – storage of configuration data previously successfully applied by the network unit and currently running;

- "Startup" datastore – storage of configuration data loaded by the network unit after it is turned on;

- "Candidate" datastore – storage of valid configuration data that does not affect the operation of the network unit and can be overwritten in other data stores: running or startup.

3. **“Messages” level** is a service part that forms the final message transmitted between the NETCONF client and the NETCONF server within the framework of the operation being

performed and corresponds to the standardized model of the remote procedure call (RPC) (RFC 5717) [12], for example, requests `<rpc>` or responses `<rpc-reply>`, `<rpc-error>`, `<ok>` or notifications `<notification>`;

4. **“Secure Transport” level** – SSH/TLS transport layer protocol, ensuring secure data exchange over the network.

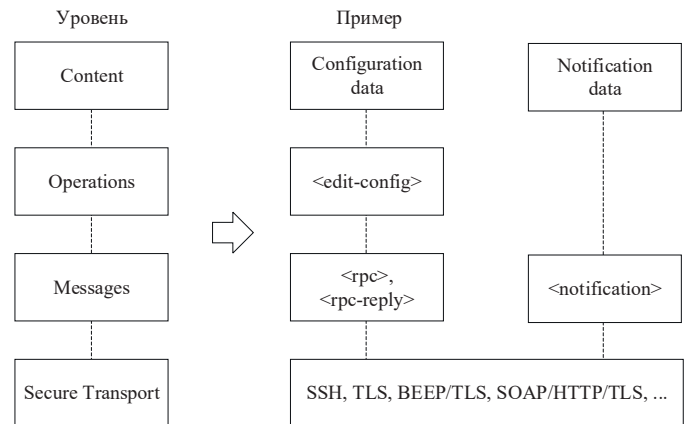


Fig. 2. Structure of the multi-layer architecture of the NETCONF protocol

The main advantages of NETCONF protocol using are the following:

- protocol openness, standardized by the IETF in the form of RFC technical specifications;
- separation of configuration data and operational state data;
- separation of configuration and configuration application operations;
- configuration of a network element in one or more transactions;
- selective data retrieval with filtering;
- streaming of notification data
- validation of configuration data before use;
- support for multiple data storage types, such as running, candidate, and startup;
- simultaneous opening of multiple sessions by one or more users;
- full or selective configuration of multiple parameters with a single `<edit-config>` operation;
- querying configuration data and operational state data with a single `<get>` operation;
- configuration application and rollback: commit and rollback, respectively;
- provide a high level of security and a wide range of authentication and cryptographic methods based on the properties of the SSH and TLS protocols [21].

A comparative analysis of existing control protocols for network units used by radio access networks 4G/5G is shown in Table 2.

An important aspect of the management standardization development in OpenRAN network is not only a messaging protocol development, but also management content unification itself based on the development of a configuration unified model and operational state data. For this purpose, NETCONF protocol uses the modern data modeling language YANG (Yet Another Next Generation, RFC 7950) [13].

Table 2

Network unit management protocols

Protocol type	SNMP	NETCONF	SOAP	REST
Developer of Standard	IETF	IETF	W3C	
Identification of the data object	OIDs	Paths		URLs
Data model	MIBs	YANG	WSDL	WSDL, WADL, text
Configuration operations	SNMP	NETCONF	XML schem	HTTP transactions
Data presentation	BER	XML	XML	XML, JSON
Transport protocol	UDP	SSH, TCP	SSL, HTTP, TCP	SSL, HTTP, TCP

To unify the management of O-RU units, OpenRAN Alliance has developed approximately forty of its own YANG data models of various types [14], and has also adopted YANG data models from IANA, IEEE, and IETF. As an example, some of main YANG models of OpenRAN Alliance are:

- Transceiver path configuration and status data models: o-ran-uplane-conf and o-ran-uplane-conf-option8;
- Performance measurement data model: o-ran-performance-management;
- Software management data model: o-ran-software-management;
- CPRI interface control and status data model: o-ran-processing-element;
- Synchronization control and status data model: o-ran-sync;
- RET servo drive control and status data models: o-ran-ald and o-ran-ald-port;
- File management data model: o-ran-file-management;
- Hardware data model: o-ran-hardware;
- Performance and characteristics data model: o-ran-module-cap.

The analysis revealed that the main advantages of YANG data models are:

- simple and understandable data modeling logic;
- hierarchical data structure;
- modularity and nesting of data;
- reuse of data types;
- use of structured data groups;
- increase in data volume through augmentation;
- formulation of data requirements and constraints for data validation purposes;
- management of data model versioning.

The models discussed above provide basic functionality management for O-RU units of eNodeB. However, in practice, radio modules from different manufacturers differ in so-called vendor-specific functionality. For this purpose, each manufacturer can use its own YANG – models as an add-on, such as vendor-specific-performance-management or vendor-specific-config-data.

3 Realization of LTE base station configuration management protocol

Using the above approach, the authors developed and implemented a structural diagram of the control software and a YANG model of the LTE eNodeB base station **softime-nodeb@2025-12-16.yang**, presented below on Fig. 3. The eNodeB base station, complete with one of type specialized software [15], was used as a prototype.

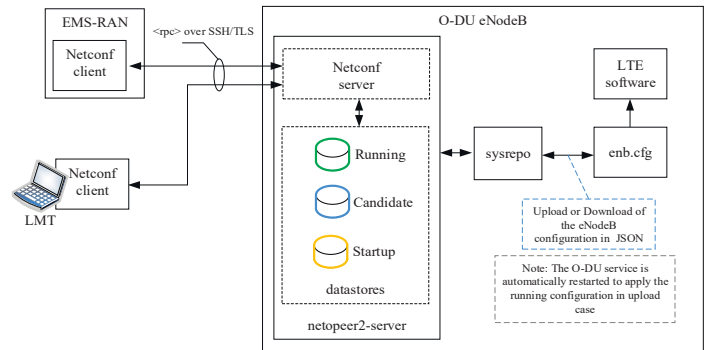


Fig. 3. Structural scheme of the control software for the LTE eNodeB base station

NETCONF server software includes three eNodeB base station configuration data stores: "candidate," "running," and "startup." The directions for copying and committing configuration data between these data stores are illustrated in Fig. 4.



Fig. 4. Directions for copying and committing configuration data between datastores

Before making changes or copying data, it is possible to lock datastores to prevent other sessions from making data changes by parallel. To implement LTE eNodeB base station management via the NETCONF protocol, the following open-source software packages were used:

- netopeer2-2.4.1 (netopeer2-server application);
- sysrepo-3.6.11 (sysrepoctl and sysrepoctl applications – interfaces with the netopeer2-server application).

In the Linux operating system (OS), the NETCONF server is implemented as a systemctl service as netopeer2.service as shown in Listing 1.

```
# systemctl cat netopeer2.service
# /etc/systemd/system/netopeer2.service
[Unit]
Description=netopeer2 NETCONF server
After=network.target

[Service]
#Type=notify
ExecStart=/usr/local/sbin/netopeer2-server -d -v2
Restart=always
RestartSec=10

[Install]
WantedBy=multi-user.target
```

Listing 1. Configuration of systemctl service by netopeer2.service

After installation and initial launch of `netopeer2.service`, the developed YANG model of the base station is loaded using the command `"sysrepoctl --install softime-nodeb\@2025-12-16.yang -v3"` in the Linux OS console, and the initial configuration of the eNodeB base station is imported from the `enb.cfg` file. To import the initial configuration, the authors developed `sysrepo_reload` application in the C programming language (designated as `"sysrepo app"` in Fig. 2), which performs the following actions:

1. Reads the original `enb.cfg` configuration file and converts it to JSON (JavaScript Object Notation) (RFC 8259) [16, 17] by removing macros and conditional statements (`./json_util dump enb.cfg`);

2. Converts configuration data from JSON to eXtensible Markup Language (XML) [18, 19, 20];

3. Adds a header to the XML data with the yang module name and the namespace: `"<enodeb xmlns='urn:softime:nodeb:1.1'>...</enodeb>"`;

4. Validates the XML configuration data against the model, forming a `lyd_node` data node in the `libyang` context based on it;

5. Imports the configuration data as `lyd_node` into the "running" data store and applies the imported configuration to `netopeer2-server` application.

Restoring the NETCONF server configuration after restarting `netopeer2.service` or rebooting the hardware platform of eNodeB distributed unit (O-DU) is ensured by copying the configuration data from the "running" data store to the "startup" data store, for example, using the command in the Linux OS console `"sysrepoctl --copy-from=running --datastore startup"`.

A sample example of the hierarchical structure of the developed YANG model of the configuration data of the eNodeB base station with specialized software is shown in Listing 2:

```
# pyang -f tree softime-nodeb@2025-12-16.yang
module: softime-nodeb
  +--rw enodeb
    +--ro enb_type?   enb_type
    +--rw enb_name?  string
    +--rw enb_id?    union
    +--rw com_addr?  string
    +--rw log_options? string
    +--rw log_filename? string
    +--rw mme_list* [mme_addr]
      | +--rw mme_addr string
      | +--rw gtp_ext_addr? string
      | +--rw slap_bind_addr? string
      | +--rw qci_dscp_mapping* [qci]
      |   +--rw qci uint16
      |   +--rw dscp? uint16
    +--rw gtp_addr? string
    +--rw cell_list* [cell_id]
      | +--rw cell_id union
      | +--rw tac? union
      | +--rw rf_port? uint16
      | +--rw n_id_cell? uint16
      | +--rw root_sequence_index? uint16
      | +--rw dl_earfcn? uint16
      | +--rw ul_earfcn? uint16
      | +--rw ncell_list* [cell_id]
      | | +--rw n_id_cell? -> /enodeb/cell_list/n_id_cell
      | | +--rw cell_id union
      | | +--rw dl_earfcn? -> /enodeb/cell_list/dl_earfcn
      | | +--rw rat? rat_type
      | | +--rw tac? -> /enodeb/cell_list/tac
    ...
```

```
+--rw cell_default
...
+--rw drb_config* [qci]
  +--rw qci uint8
  +--rw ims_dedicated_bearer? boolean
  +--rw rlc_config
...
+--rw pdcp_config
  | +--rw discardTimer? uint16
  | +--rw pdcp_SN_Size? uint8
  | +--rw pdcp_SN_Size_v1130? boolean
  | +--rw pdcp_SN_Size_v1310? boolean
  | +--rw statusReportRequired? boolean
  | +--rw headerCompression
  |   +--rw maxCID? uint16
  |   +--rw profile0x0001? boolean
  |   +--rw profile0x0002? boolean
  |   +--rw profile0x0004? boolean
+--rw logical_channel_config
  +--rw priority? uint8
  +--rw prioritisedBitRate? int16
  +--rw bucketSizeDuration? uint16
  +--rw logicalChannelGroup? uint8
  +--rw logicalChannelSR_Mask? boolean
  +--rw logicalChannelSR_Prohibit? Boolean
```

Listing 2. Example of the developed YANG model of eNodeB base station configuration data

The conformity assessment of the developed `softime-nodeb@2025-12-16.yang` model and the configuration file of the specialized software `"enb_full.xml"` for eNB in XML format was performed and confirmed using the `yanglint` utility.

The conformity assessment of the developed `softime-nodeb@2025-12-16.yang` model and `"enb_full.xml"` configuration file of the specialized software for eNodeB in XML format was performed and confirmed using by `yanglint` utility.

```
# yanglint -t data enb_full.xml softime-nodeb\@2025-12-31.yang -f json | jq '.softime-nodeb:enodeb'
```

To development of the YANG model were imposed following requirements:

- Full compliance of data structures with `enb.cfg` configuration file;
- Optimal model description by grouping data contained in multiple objects, such as `cell_list` and `cell_default` objects;
- Model parameters description of in description fields;
- Units measurement indication for model parameters and their permissible values/ranges of change in "range" fields;
- Indication of permissible values for string text parameters in the enumeration fields;
- Separation of parameters by operation type: read and write or read-only ("rw" and "ro", respectively);
- Separation of data into mandatory and optional (mandatory true or false, respectively);
- Key model parameters (key) contained in "list" node must be unique in `enb.cfg` configuration file across all base station configuration scenarios;
- specifying multiple key fields in the "list" node when one key field is not unique;
- establishing dependencies between different data, if necessary, when the value or mandatory presence of one data depends on other data in the model; for example, different sets of data for cell configurations in frequency FDD and time duplex TDD [23].

Command diagram explaining procedure of the eNodeB base station configuration change is shown in Fig. 3. To monitor changes in the "running" data storage, the `sysrepo_subscribe_and_upload` application was developed in the C programming language (designated as "sysrepo app" in Fig. 3), which performs the following actions:

- The `sysrepo_app` application subscribes to notifications about changes in the running datastore;
- Receives notifications about changes in the running datastore;
- Reads the current configuration from the running datastore;
- Converts data to JSON notation (RFC 8259) [17] and writes it to the `enb.cfg` configuration file, while saving (backing up) the previous configuration;
- Automatically applies the generated configuration file by restarting the O-DU unit service.

The `sysrepo_subscribe_and_upload` application is designed for the x86 CPU architecture and requires a minimal number of libraries, as shown in Listing 3:

```
# ldd sysrepo_subscribe_and_upload
linux-vdso.so.1 (0x00007ffc11ea000)
libsysrepo.so.7 => /usr/local/lib/libsysrepo.so.7 (0x00007f1b191ce000)
libyang.so.3 => /usr/local/lib/libyang.so.3 (0x00007f1b19026000)
libjson-c.so.5 => /lib/x86_64-linux-gnu/libjson-c.so.5 (0x00007f1b19013000)
libc.so.6 => /lib/x86_64-linux-gnu/libc.so.6 (0x00007f1b18dea000)
```

```
libm.so.6 => /lib/x86_64-linux-gnu/libm.so.6 (0x00007f1b18d03000)
libpcre2-8.so.0 => /usr/local/lib/libpcre2-8.so.0 (0x00007f1b18ca1000)
/lib64/ld-linux-x86-64.so.2 (0x00007f1b192a5000)
```

Listing 3. The `sysrepo_subscribe_and_upload` application

An example of changing the configuration by sending a NETCONF `rpc <edit-config>` message, which provides for the following actions with the data in the datastore, is shown in Fig. 5:

- The "merge" action (`nc:operation="merge"`, the default action) merges the provided data with the existing data in the corresponding yang model path. If the data does not exist, it is created;
- The "replace" action (`nc:operation="replace"`) replaces the existing data in the corresponding yang model path with the provided data. If the data does not exist, it is created;
- The "create" action (`nc:operation="create"`) adds data according to the yang model path only if it does not already exist. If it does exist, the NETCONF server returns an error message;
- The "delete" action (`nc:operation="delete"`) deletes the specified data according to the yang model path. If the data does not exist, the NETCONF server returns an error message;
- The "erase" action (`nc:operation="remove"`) – deletes the specified data according to the path of the yang model; unlike the "delete" action, if the data does not exist, the NETCONF server does not return an error, but simply reports successful completion.

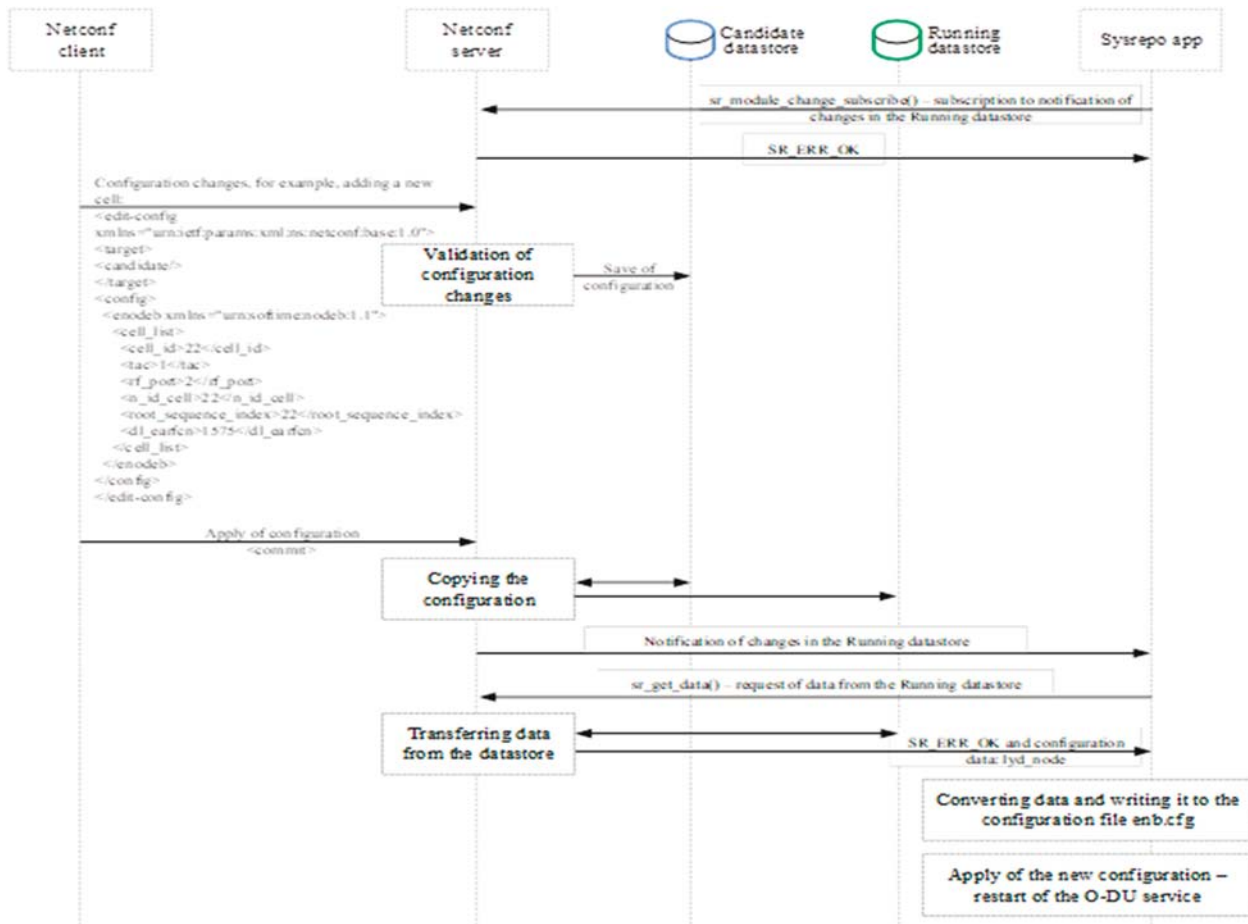


Fig. 5. An example of changing the configuration by sending a NETCONF `rpc <edit-config>` message

An analysis of the command diagram of the eNodeB base station configuration (Fig. 5) shows that as a result of transmitting the `rpc <edit-config>` control command and sequentially executing the steps by the `sysrepo` app application, the data is converted into the JSON notation format and the generated configuration file is applied by restarting the O-DU module service.

Conclusions

An analysis of the OpenRAN Alliance technical specifications revealed opportunities for implementing management of O-RU and O-DU units of the eNodeB base station in the management plane when using multivendor radio access networks for building 4G/LTE and 5G/NR networks.

The developed YANG data model for configuring the eNodeB base station with LTE technology, based on specialized Amarisoft software, enables eNodeB management and configuration.

The NETCONF server deployed on the eNodeB base station in the LTE network enables the use of the YANG data model for base station management, and the developed datastore status monitoring applications in the C programming language can be used to configure the LTE base station in accordance with the operator's frequency and territorial plan. The developed YANG model's compliance with the standard LTE network eNodeB configuration file using the Amarisoft protocol stack was verified in practice.

The open nature of the developed YANG model and the NETCONF network management protocol allows any developer to create new EMS-RAN management systems for Open RAN base stations.

References

- [1] V.O. Tikhvinsky, S.V. Terentyev, V.A. Koval, E.E. Devyatkin, "Development of mobile communication networks from 5G Advanced to 6G: Projects, technologies, architecture," Moscow: Technosfera Publishing House, 2023. 528 p.
- [2] Governance of O-RAN ALLIANCE e.V. in Compliance with WTO Principles. White Paper, July 2023.
- [3] O-RAN ALLIANCE O-RAN.WG4.MP.0: "Management Plane Specification".
- [4] 3GPP TS 28.533 Management and orchestration; Architecture framework.
- [5] 3GPP TS 28.532 Management and orchestration; Generic management services.
- [6] 3GPP TS 28.541 Management and orchestration; Network Resource Model.
- [7] 3GPP TS 28.531 Management and orchestration; Provisioning.
- [8] J. Schönwälder, M. Björklund, P. Shafer, "Network configuration management using NETCONF and YANG," *IEEE Communications Magazine*, no. 48(9): October 2010, pp. 166-173, DOI: 10.1109/MCOM.2010.5560601

Magazine, no. 48(9): October 2010, pp. 166-173, DOI: 10.1109/MCOM.2010.5560601

[9] J. Schönwälder, et al., "SNMP Traffic Analysis: Approaches, Tools, and First Results," *Conference: Integrated Network Management*, May 2007, DOI: 10.1109/INM.2007.374797

[10] Satish Kumar Jangid et al., "Evaluating the Role of Secure Shell (SSH) in Securing Wireless Networks," *3rd International Conference on Smart Generation Computing, Communication and Networking (SMART GENCON)*, December 2023 DOI: 10.1109/SMARTGENCON60755.2023.10441924

[11] H. Krawczyk, K.G. Paterson, H. Wee, "On the Security of the TLS Protocol: A Systematic Analysis," *Lecture Notes in Computer Science*, January 2013, DOI: 10.1007/978-3-642-40041-4_24

[12] G. Hollósi, D. Ficzer, Pal Varga, "Generative AI for Low-Level NETCONF Configuration in Network Management Based on YANG Models," *20th International Conference on Network and Service Management (CNSM)*, October 2024, DOI: 10.23919/CNSM62983.2024.10814410

[13] E. Nataf, O. Festor, "End-to-end YANG-based Configuration Management," *IEEE/IFIP Network Operations and Management Symposium - NOMS 2010, IEEE/IFIP*, Apr 2010, Osaka, Japan, pp. 674-684.

[14] Abderrahim Amlou et al., "Automated Network Programmability Using OpenConfig YANG Models and NETCONF Protocol," *20th ACS/IEEE International Conference on Computer Systems and Applications*, December 2023, DOI: 10.1109/AICCSA59173.2023.10479244.

[15] Fabbryccio A. C. M. Cardoso et al., "A case study on protocol stack integration for 3GPP LTE evolved node B," *2014 IEEE Latin-America Conference on Communications (LATINCOM)*, DOI: 10.1109/LATINCOM.2014.7041860

[16] Changxia Sun et al., "Research and Application of Data Exchange based on JSON," *2020 Asia-Pacific Conference on Image Processing, Electronics and Computers (IPEC)*, DOI: 10.1109/IPEC49694.2020.9115155

[17] Dunlu PENG, Lidong CAO, Wenjie XU, "Using JSON for Data Exchanging in Web Service Applications," *Journal of Computational Information Systems*, no. 7: 16 (2011), pp. 5883-5890.

[18] Mohsen Shekarriz et al., "Query processing optimization in broadcasting XML data in mobile communications," *Journal of Supercomputing*, 2021, 77(2014):1-27, DOI:10.1007/s11227-020-03479-5

[19] M. Javani, M. Mirabi, "An efficient index and data distribution scheme for XML data broad-cast in mobile wireless networks," *J Inf Sci Eng*, 2017, 33, pp. 159-182.

[20] 3GPP TS 28.550 5G; Management and orchestration; Performance assurance.

[21] G.A. Fokin, Abdulwahab Hussein Al-odhari, "Positioning of the moving radiation source using time difference of arrival method," *T-Comm*, vol. 11, no.4, pp. 41-46.

[22] M.G. Bakulin, T.B.C. Ben Rejeb, V.B. Kreyndelin, D.Y. Pankratov, A.E. Smirnov, "Code domain NOMA in 3GPP specifications: 5G or 6G?" *T-Comm*. 2022, vol 16, no. 1, pp. 4-14.

[23] V. Tikhvinskiy, A. Pastukh, S. Dymkova, O. Varlamov, "Compatibility analysis between redcap non-public networks and 5G NR in TDD FR1 and FR2 bands," *Inventions*. 2025, vol. 10, no 1. 12 p.

РАЗРАБОТКА МОДЕЛИ ДАННЫХ ДЛЯ УПРАВЛЕНИЯ БАЗОВОЙ СТАНЦИЕЙ eNodeB ТЕХНОЛОГИИ LTE В СЕТИ РАДИОДОСТУПА OpenRAN

Терентьев Сергей Васильевич, МГУ им. М.В. Ломоносова; ООО "Софттайм", Москва, Россия, s.ter@mail.ru

Тихвинский Валерий Олегович, МГТУ им. Н.Э. Баумана, Москва, Россия;

Севастопольский государственный университет, г. Севастополь, Россия; НИЦ Телеком, Москва, Россия, vtiir@mail.ru

Десяткин Евгений Евгеньевич, НИЦ Телеком, Москва, Россия, deugene@list.ru

Фокин Григорий Алексеевич, Санкт-Петербургский государственный университет телекоммуникаций им. Бонча-Бруевича, г. Санкт-Петербург, Россия, grihafokin@gmail.com

Аннотация

Статья посвящена разработке YANG – модели данных с использованием протокола NETCONF для конфигурирования базовой станции eNodeB в комплектации со специализированным программным обеспечением Amarisoft в сети радиодоступа OpenRAN технологии LTE, а также разработке приложений < sysgero_reload > на языке программирования C в целях применения предложенных конфигураций. Конфигурирование базовой станции eNodeB сети радиодоступа OpenRAN требуется на различных этапах жизненного цикла сети LTE, которые рассмотрены в статье. В основу разработки YANG – модели данных используемой для конфигурирования базовой станции eNodeB положены возможности управления и оркестрирования сетью радиодоступа 5G RAN (Management and Orchestration), стандартизованные Партнерским проектом 3GPP и реализуемые посредством протокола NETCONF применяемого для внедрения технических решений на сервисно-ориентированной архитектуре SBMA (Service-Based Management Architecture). Представленные авторами листинги программ демонстрируют выборочный пример иерархической структуры разработанной YANG-модели конфигурационных данных базовой станции eNodeB со специализированным программным обеспечением Amarisoft. Разработанная YANG-модель и сетевой протокол управления NETCONF позволит любому разработчику архитектуры сети создавать новые системы управления EMS-RAN базовых станций Open RAN.

Ключевые слова: O-RAN, YANG, NETCONF, LTE, JSON, XML.

Литература

1. Тихвинский В.О., Терентьев С.В., Коваль В.А., Десяткин Е.Е. Развитие сетей мобильной связи от 5G Advanced к 6G: Проекты, технологии, архитектура. Издание 2-е доп. М.: Изд-во Техносфера, 2024. 528 с.
2. Governance of O-RAN ALLIANCE e.V. in Compliance with WTO Principles. White Paper, July 2023.
3. O-RAN ALLIANCE O-RAN.WG4.MP.0: "Management Plane Specification".
4. 3GPP TS 28.533 Management and orchestration; Architecture framework.
5. 3GPP TS 28.532 Management and orchestration; Generic management services.
6. 3GPP TS 28.541 Management and orchestration; Network Resource Model
7. 3GPP TS 28.531 Management and orchestration; Provisioning.
8. Schonwalder J., Bjorklund M., Shafer P. Network configuration management using NETCONF and YANG // IEEE Communications Magazine, 48(9): October 2010, pp. 166-173, DOI: 10.1109/MCOM.2010.5560601
9. Schonwalder J. et al. SNMP Traffic Analysis: Approaches, Tools, and First Results, Conference: Integrated Network Management, May 2007, DOI: 10.1109/INM.2007.374797
10. Satish Kumar Jangid and al. Evaluating the Role of Secure Shell (SSH) in Securing Wireless Networks // 3rd International Conference on Smart Generation Computing, Communication and Networking (SMART GENCON), December 2023 DOI: 10.1109/SMARTGENCON60755.2023.10441924
11. Krawczyk H., Paterson K.G., Wee H. On the Security of the TLS Protocol: A Systematic Analysis, January 2013 // Lecture Notes in Computer Science, DOI: 10.1007/978-3-642-40041-4_24
12. Hollosi G., Ficzer D., Pal Varga, Generative AI for Low-Level NETCONF Configuration in Network Management Based on YANG Models // 20th International Conference on Network and Service Management (CNSM), October 2024, DOI: 10.23919/CNSM62983.2024.10814410
13. Nataf E., Festor O. End-to-end YANG-based Configuration Management // IEEE/IFIP Network Operations and Management Symposium - NOMS 2010, IEEE/IFIP, Apr 2010, Osaka, Japan, pp. 674-684.
14. Abderrahim Amlou and al. Automated Network Programmability Using OpenConfig YANG Models and NETCONF Protocol // 20th ACS/IEEE International Conference on Computer Systems and Applications, December 2023, DOI: 10.1109/AICCSA59173.2023.10479244.
15. Fabbryccio A.C.M. Cardoso et al. A case study on protocol stack integration for 3GPP LTE evolved node B // 2014 IEEE Latin-America Conference on Communications (LAT-INCOM), DOI: 10.1109/LATINCOM.2014.7041860
16. Changxia Sun et al. Research and Application of Data Exchange based on JSON // 2020 Asia-Pacific Conference on Image Processing, Electronics and Computers (IPEC), DOI: 10.1109/IPEC49694.2020.9115155
17. Dunlu PENG, Lidong CAO, Wenjie XU. Using JSON for Data Exchanging in Web Service Applications // Journal of Computational Information Systems 7: 16 (2011) 5883-5890
18. Mohsen Shekarriz et al. Query processing optimization in broadcasting XML data in mobile communications // Journal of Supercomputing (2021), 77(2014), pp. 1-27, DOI:10.1007/s11227-020-03479-5
19. Javani M., Mirabi M. An efficient index and data distribution scheme for XML data broad-cast in mobile wireless networks // J Inf Sci Eng 2017, 33, pp. 159-182
20. 3GPP TS 28.550 5G; Management and orchestration; Performance assurance.
21. Фокин Г.А., Аль-Одхари А.Х. Позиционирование подвижных источников радиоионлучения разностно-дальномерным методом // Т-Comm: Телекоммуникации и транспорт. 2017. Т. 11. № 4. С. 41-46.
22. Бакулин М.Г., Бен Режеб Т.Б.К., Крейнделин В.Б., Панкратов Д.Ю., Смирнов А.Э. Технология NOMA с кодовым разделением в 3GPP: 5G или 6G? // Т-Comm: Телекоммуникации и транспорт. 2022. Т. 16. № 1. С. 4-14.

Информация об авторах:

Терентьев Сергей Васильевич, к.т.н., доцент МГУ им. М.В. Ломоносова, Заместитель директора департамента разработки радиоподсистемы (R&D) ООО "Софттайм", Москва, Россия

Тихвинский Валерий Олегович, профессор МГТУ им. Н.Э. Баумана и Севастопольского государственного университета, д.э.н., Главный научный сотрудник ФГАУ НИЦ Телеком, Москва, Россия

Десяткин Евгений Евгеньевич, к.э.н., Директор Центра исследований перспективных беспроводных технологий связи (ЦИ ПБТС) ФГАУ НИЦ Телеком, Москва, Россия
Фокин Григорий Алексеевич, д.т.н., доцент, заведующий кафедрой беспроводных технологий и систем (БТС), Санкт-Петербургский государственный университет телекоммуникаций им. Бонча-Бруевича, г. Санкт-Петербург, Россия