



Том 7. № 2
2021

ТРУДЫ УЧЕБНЫХ ЗАВЕДЕНИЙ СВЯЗИ



Темы номера:

- ✓ Расчет неготовности направления связи по дождевому фактору
- ✓ Теория хаоса для задач динамического управления контакт-центрами
- ✓ Защита от атаки клонирования бумажных сертификатов

Vol. 7. Iss. 2
2021

PROCEEDINGS
OF TELECOMMUNICATION UNIVERSITIES

Федеральное государственное бюджетное образовательное
учреждение высшего образования
«Санкт-Петербургский государственный университет телекоммуникаций
им. проф. М.А. Бонч-Бруевича» (СПбГУТ)

Научный журнал

**ТРУДЫ
УЧЕБНЫХ ЗАВЕДЕНИЙ СВЯЗИ**

Том 7. № 2

Proceedings of Telecommunication Universities

Vol. 7. Iss. 2

Санкт-Петербург

2021

Описание журнала

Научный журнал. Включен в Перечень рецензируемых научных изданий, в которых должны быть опубликованы основные научные результаты диссертаций на соискание ученой степени кандидата наук, на соискание ученой степени доктора наук (распоряжение Минобрнауки России № 21-р от 12.02.2019), по специальностям (приказ № 118 от 24.02.2021):

1.2.2. Математическое моделирование, численные методы и комплексы программ

2.2.6. Оптические и оптико-электронные приборы и комплексы

2.2.13. Радиотехника, в том числе системы и устройства телевидения

2.2.14. Антенны, СВЧ-устройства и их технологии

2.2.15. Системы, сети и устройства телекоммуникаций

2.2.16. Радиолокация и радионавигация

2.3.1. Системный анализ, управление и обработка информации

2.3.6. Методы и системы защиты информации, информационная безопасность

Выпускается с 1960 года. Выходит 4 раза в год (ежеквартально). Издается на русском и английском языках.

Редакционный совет

Дукельский К.В. к.т.н., доцент, АО «Государственный оптический институт имени С.И. Вавилова» (ГОИ),
Главный редактор г. Санкт-Петербург, Россия

Буйневич М.В. д.т.н., проф., Санкт-Петербургский государственный университет телекоммуникаций
Зам. Главного редактора им. проф. М.А. Бонч-Бруевича (СПбГУТ), г. Санкт-Петербург, Россия

Розанов Н.Н. д.ф.-м.н., проф., чл.-корр. РАН, АО «Государственный оптический институт
имени С.И. Вавилова» (ГОИ), г. Санкт-Петербург, Россия

Кучерявый Е. PhD, Технологический университет Тампере, г. Тампере, Финляндия

Гошек И. PhD, Технологический университет Брно, г. Брно, Чешская республика

Тиамийу О.А. PhD, Университет Илорина, г. Илорин, Нигерия

Козин И.Д. д.ф.-м.н., проф., Алматинский университет энергетики и связи, г. Алма-Аты, Казахстан

Самуйлов К.Е. д.т.н., проф., Российский университет дружбы народов (РУДН), г. Москва, Россия

Степанов С.Н. д.т.н., проф., Московский технический университет связи и информатики (МТУСИ),
г. Москва, Россия

Росляков А.В. д.т.н., проф., Поволжский государственный университет телекоммуникаций
и информатики (ПГУТИ), г. Самара, Россия

Кучерявый А.Е. д.т.н., проф., Санкт-Петербургский государственный университет телекоммуникаций
им. проф. М.А. Бонч-Бруевича (СПбГУТ), г. Санкт-Петербург, Россия

Канаев А.К. д.т.н., проф., Петербургский университет путей сообщения имени Александра I (ПГУПС),
г. Санкт-Петербург, Россия

Новиков С.Н. д.т.н., проф., Сибирский государственный университет телекоммуникаций и информатики
(СибГУТИ), г. Новосибирск, Россия

Дворников С.В. д.т.н., проф., Военная академия связи им. Маршала Советского Союза С.М. Буденного (ВАС),
г. Санкт-Петербург, Россия

Коржик В.И. д.т.н., проф., Санкт-Петербургский государственный университет телекоммуникаций
им. проф. М.А. Бонч-Бруевича (СПбГУТ), г. Санкт-Петербург, Россия

Ковалгин Ю.А. д.т.н., проф., Санкт-Петербургский государственный университет телекоммуникаций
им. проф. М.А. Бонч-Бруевича (СПбГУТ), г. Санкт-Петербург, Россия

Владыко А.Г. к.т.н., Санкт-Петербургский государственный университет телекоммуникаций
им. проф. М.А. Бонч-Бруевича (СПбГУТ), г. Санкт-Петербург, Россия

Регистрационная информация

Журнал зарегистрирован Федеральной службой по надзору в сфере связи, информационных технологий и массовых коммуникаций: ПИ № 77-77501 от 17.01.2020 г. (пред. рег. № 77-17986 от 07.04.2004 г.)

Подписной индекс в объединенном каталоге «ПРЕССА РОССИИ»: 59983

Размещение в РИНЦ (elibrary.ru) по договору: № 59-02/2013R от 20.02.2013

Контактная информация

Учредитель и издатель: Федеральное государственное бюджетное образовательное учреждение высшего образования «Санкт-Петербургский государственный университет телекоммуникаций им. проф. М.А. Бонч-Бруевича» (СПбГУТ)

Адрес: 191186, Санкт-Петербург, набережная реки Мойки,
учредителя: д. 61, литера А

Адрес: 193232, Санкт-Петербург,
редакции: пр. Большевиков, 22/1, к. 334/2
Тел.: +7 (812) 326-31-63, м. т. 2022,
+79643759970
E-mail: tuzs@spbgut.ru
Web: <http://tuzs.sut.ru>
ВК: <http://vk.com/spbtuzs>

Description

Scientific journal. The journal is included in the List of reviewed scientific publications, in which the main scientific results of dissertations for the degree of candidate of science and for the degree of doctor of science should be published (order of the Ministry of Education and Science of Russia No 21-r of 12 February 2019) in the field of (order of the Ministry of Education and Science of Russia No 118 of 24 February 2021):

1.2.2. Mathematical modeling, numerical methods and complexes of programs

2.2.6. Optical and optoelectronic devices and complexes

2.2.13. Radio engineering, including television systems and devices

2.2.14. Antennas, microwave devices and its technologies

2.2.15. Systems, networks and telecommunication devices

2.2.16. Radiolocation and radio navigation

2.3.1. System analysis, management and information processing

2.3.6. Methods and systems of information security, cybersecurity

Since 1960. Published 4 times per year. Published in Russian and English.

Editorial Board

K.V. Dukel'skii PhD, associate prof., executive Director of Open Joint Stock Company «S.I. Vavilov State Optical Institute» (SOI), Saint-Petersburg, Russia
Editor-in-chief

M.V. Buinevich DSc, prof., The Bonch-Bruевич Saint-Petersburg State University of Telecommunications (SPbSUT), Saint-Petersburg, Russia
Deputy editor-in-chief

N.N. Rozanov DSc, prof., member-corr. RAS, Open Joint Stock Company «S.I. Vavilov State Optical Institute» (SOI), Saint-Petersburg, Russia

Y. Koucheryav PhD, Tampere University of Technology, Tampere, Finland

I. Hošek PhD, Brno University of Technology, Brno, Czech Republic

O.A. Tiamiyu PhD, University of Ilorin, Ilorin, Nigeria

I.D. Kozin DSc, prof., Almaty University of Power Engineering and Telecommunications, Almaty, Kazakhstan

K.E. Samuilov DSc, prof., Peoples' Friendship University (RUDN), Moscow, Russia

S.N. Stepanov DSc, prof., Moscow Technical University of Communication and Informatics (MTUCI), Moscow, Russia

A.V. Roslyakov DSc, prof., Povolzhskiy State University of Telecommunications and Informatics (PSUTI), Samara, Russia

A.E. Koucheryav DSc, prof., The Bonch-Bruевич Saint-Petersburg State University of Telecommunication (SPbSUT), Saint-Petersburg, Russia

A.K. Kanaev DSc, prof., Emperor Alexander I-st Petersburg State Transport University (PSTU), Saint-Petersburg, Russia

S.N. Novikov DSc, prof., Siberian State University of Telecommunications and Information Sciences (SibSUTIS), Novosibirsk, Russia

S.V. Dvornikov DSc, prof., Military Academy of Telecommunications named after Marshal Union S.M. Budyonny, Saint-Petersburg, Russia

V.I. Korzhik DSc, prof., The Bonch-Bruевич Saint-Petersburg State University of Telecommunication (SPbSUT), Saint-Petersburg, Russia

Yu.A. Kovalgin DSc, prof., The Bonch-Bruевич Saint-Petersburg State University of Telecommunication (SPbSUT), Saint-Petersburg, Russia

A.G. Vladyko PhD, The Bonch-Bruевич Saint-Petersburg State University of Telecommunication (SPbSUT), Saint-Petersburg, Russia

Registration Information

Registered by Federal Service for Supervision of Communications, Information Technology and Mass Media on 17.01.2020: PI No. 77-77501 (prev. reg. on 04.07.2004: No. 77-17986)

Subscription index for joint catalog «PRESSA ROSSII»: 59983

Accommodation in RINC (elibrary.ru) by agreement on 20.02.2013: No. 59-02/2013R

Contact Information

Publisher: Federal State Budget-Financed Educational Institution of Higher Education «The Bonch-Bruевич Saint-Petersburg State University of Telecommunications» (SPbSUT)

Post address: 193232, Saint Petersburg, Prospekt Bolshevikov, 22/1
Phone: +7 (812) 326-31-63, local 2022, +79643759970
E-mail: tuzs@spbgut.ru
Web: <http://tuzs.sut.ru>

Publisher address: 191186, Saint Petersburg, Moika river embankment, 61-A

Студенческая весна – научный и технологический ландшафт

Президент России Владимир Путин объявил 2021 г. Годом науки и технологий. Глава государства убежден, пандемия коронавируса отчетливо показала, что наука и технологии имеют колоссальную значимость, поскольку определяют безопасность, качество жизни человека. Всего в Год науки и технологий планируется провести около 70 федеральных мероприятий. Больше половины из них – тематические форумы, выставки и научно-практические конференции. Причем пройдут они как в дистанционном, так и в очном форматах. По прогнозам экспертов, в них будут вовлечены не менее 50 млн. человек.

Университет телекоммуникаций, носящий имя выдающегося ученого в области радиофизики, основателя российской радиоламповой промышленности, профессора Михаила Александровича Бонч-Бруевича, не остается в стороне от столь важного исторического события. В университете созданы все условия для научной работы студентов, аспирантов, молодых ученых и исследователей. За 10 лет ВУЗ подготовил около 17 тыс. выпускников, и практически все они успешно трудоустроились в ведущие отраслевые компании России. Запускаются программы научных исследований, поддерживается конкурентоспособность в области научных разработок. Такие направления, как искусственный интеллект, способы обработки и хранения больших объемов информации, применение экологически чистых и ресурсосберегающих технологий создание и применение новейших материалов, апробируются и на тематических форумах и конференциях, которые проводятся в стенах нашего ВУЗа. Это, конечно, научно-техническая и научно-методическая конференция «Актуальные проблемы инфотелекоммуникаций в науке и образовании» (АПИНО) и Региональная научно-техническая конференция студентов, аспирантов и молодых ученых «Студенческая весна». Так совпало, что именно в Год науки и технологий оба проводимых мероприятия – юбилейные по счету.

Конференция АПИНО, которая является правопреемницей традиционных ежегодных научных форумов профессорско-преподавательского состава с первых лет образования ЛЭИС, была проведена в феврале текущего года в 10-й раз. С 2012 г. конференция проходит в статусе международной. В конференции принимают участие ученые из стран ближнего и дальнего зарубежья, таких как Финляндия, Словения, Латвия, Казахстан и др. Участники – научно-педагогический состав образовательных учреждений, научно-технические работники предприятий, аспиранты. Обсуждаются такие темы, как радиотехнологии связи, инфокоммуникационные сети и системы, информационные системы и технологии, теоретические основы радиоэлектроники и систем связи, цифровая экономика и др. По итогам конференции АПИНО издается сборник научных статей, который размещается в РИНЦ.

«Студенческая весна» – региональная научно-техническая конференция студентов, аспирантов и молодых ученых, традиционно проходящая в мае, собиралась в 75-й раз. Форум является правопреемником Студенческой научно-технической конференции, которая проводилась в ЛЭИС с 1946 г. В 2014 г. форум получил статус регионального. В работе симпозиума участвуют не только студенты и аспиранты, магистранты, но и молодые ученые-специалисты в области связи, возраст которых не превышает 35 лет. Темы конференции тесно перекликаются с АПИНО. На пленарном и секционных заседаниях в течение двух дней заслушано свыше семисот докладов и сообщений. В выступлениях участников обсудили состояние и перспективы развития мирового и отечественного уровня ИТ и телекоммуникаций.

Выступление перед большой аудиторией – это сложный и ответственный шаг. Это, в первую очередь, выражение собственных мыслей и идей в той или иной научной отрасли, того или иного направления. Поэтому любой доклад, выступление, информационное сообщение как на АПИНО, так и на «Студенческой весне» – это прояснение актуальных на сегодняшний день вопросов, подача идей рационального решения тех или иных проблем по различным тематикам. Работа любого публичного мероприятия, как конференция, не ограничивается только пленарными докладами. Обмен мнениями, обсуждение тех или иных злободневных тем и ответ на вопросы участников и слушателей – все это можно увидеть на «круглых столах» и на различных тематических секциях. Вне всякого сомнения, замечания и предложения, которые появляются здесь, выявляют различные неточности в сообщении выступающего, определяются спорные утверждения, даются советы и подсказки по устранению недочетов или ошибок. Для всех специалистов, которые еще проходят обучение, такие конференции, как АПИНО и «Студенческая весна», являются отличной площадкой для продвижения своих работ, проведения дискуссий, отстаивания своих научных взглядов. «Студенческая весна» сегодня по праву является генератором новых идей.

Научная работа в стенах нашего университета многоаспектна. Начиная с 2013 г. в СПбГУТ проводится конкурс студенческих научных работ. За все время проведения в нем приняли участие более 800 учащихся. На это состязание свой проект может представить любой студент СПбГУТ, но и студенты АКТ(ф), СПбКТ и СмКТ. Победители конкурса традиционно объявляются на «СТУДЕНЧЕСКОЙ ВЕСНЕ». В этом году

в Конкурс студенческих научных работ СПбГУТ добавилось отдельное научное направление – «Студенты и пандемия: коммуникации в борьбе с COVID-19». Конкурсанты представили как научные работы, так и видеоролики, презентации. Это различные предложения об организации жизни человека в условиях удаленной работы или дистанционного обучения, способы предотвращения и профилактики заболевания, а также технологии и сервисы в условиях социального дистанцирования. Поступило 25 работ, и жюри, тщательно рассмотрев представленные материалы, приняло решение присудить: Гран-при – за разработку методов дистанционного мониторинга пациентов с Covid-19 с использованием носимой электроники – Ирине Игнатовой и Анне Парфеновой, факультет инфокоммуникационных сетей и систем, Институт магистратуры; за реализованный проект в социальной среде – Александре Перевышко, Институт магистратуры; за реализованный проект в образовательном процессе – Ярославу Плетневу и Кристине Фроловой, факультет информационных систем и технологий; за инновационную технологию – Анастасии Спиркиной, факультет инфокоммуникационных сетей и систем.

Первые места присудить: на факультете радиотехнологий и связи – Владиславу Тимофееву; на факультете инфокоммуникационных сетей и систем – Марии Шарлаевой, на факультете инфокоммуникационных сетей и систем – Софии Сергиенко, на факультете фундаментальной подготовки – Рашиду Алли, на факультете социальных цифровых технологий – Карине Долгополовой. Присудить первое место – Апполинарии Парадеевой, Архангельский колледж телекоммуникаций СПбГУТ; а также Ивану Поддубному и Дарине Григорьевой, Смоленский колледж телекоммуникаций СПбГУТ. Поздравляем наших дипломантов!

Нельзя не сказать еще про одно событие, которое регулярно проходит в стенах нашего ВУЗа. Речь идет о Международном Балтийском коммуникационном форуме – BAFO. Это масштабное научно-практическое мероприятие, ставшее визитной карточкой факультета социальных цифровых технологий (СЦТ). Программа Форума включает в себя: мастер-классы от специалистов в области социальных, маркетинговых и международных коммуникаций; студенческая и преподавательская научные конференции, круглый стол «Цифровая трансформация в государственном управлении». По итогам Форума издаются студенческий сборник тезисов и «Вестник факультета СЦТ» (РИНЦ).

Конечно, список научных мероприятий, который проводит наш ВУЗ в течение года, этим не ограничивается. Каждую осень под эгидой Международного конгресса по ультрасовременным телекоммуникациям и системам управления ICUMT проходит Международная научно-техническая конференция студентов, аспирантов и молодых ученых «Интернет вещей и его приложения. Искусственный интеллект в сетях связи» – INTNITEN. Тематика – Интернет вещей, сети связи 6G, беспроводные сенсорные сети, тактильный интернет, интернет навыков, дополненная реальность, промышленный интернет вещей, программно-конфигурируемые сети, машинное обучение и искусственный интеллект в сетях связи, беспилотный автотранспорт и т. д. В этом году появится новое направление – COVID-19: технологии телекоммуникаций как средство борьбы с пандемией. Конференция собирает российских и иностранных молодых ученых не только из СПбГУТ, но и из РУДН (Москва), Технологического университета (г. Тампере, Финляндия), Технологического университета (г. Брно, Чехия), что подтверждает ее актуальность, востребованность и перспективность.

Под занавес года СПбГУТ проводит региональную научно-методическую конференцию магистрантов и их руководителей «Подготовка профессиональных кадров в магистратуре для цифровой экономики» – ПКМ. Девиз конференции – апробация результатов научно-исследовательской деятельности магистрантов, их наставников, а также руководителей образовательных программ магистратуры. Участников – обучающихся и их руководителей из ВУЗов Северо-Запада России – привлекают такие тематики ПКМ, как радиотехнологии связи, инфокоммуникационные сети и системы, информационные системы и технологии, теоретические основы радиоэлектроники и систем связи, цифровая экономика и управление в инфокоммуникациях. Обсуждаются инновационные решения в рассматриваемых научных направлениях, привлекается внимание к актуальным проблемам подготовки кадров в магистратуре.

Главная цель Года науки, объявленного Президентом РФ В.В. Путиным – привлечение нашей молодежи, которая у нас, безусловно, талантлива, в сферу высоких технологий. Но этого мало. Нужно стремиться, чтобы вовлечь профессиональное научное сообщество к тому, чтобы благодаря внедренным новейшим технологиям активнее развивалась наша страна, а граждане России имели полную информацию о том, как реализуются те или иные инициативы в области науки и технологий. И работа ведущего в области связи университета телекоммуникаций – этому ярчайший пример.

*Проректор по научной работе СПбГУТ,
доктор технических наук,
старший научный сотрудник
Александр Викторович Шестаков*

СОДЕРЖАНИЕ

CONTENTS

2.2 ЭЛЕКТРОНИКА, ФОТОНИКА, ПРИБОРОСТРОЕНИЕ И СВЯЗЬ	
Бугров В.Н. Динамическое квантование коэффициентов цифрового фильтра Гольдштейн А.Б., Кисляков С.В., Феноменов М.А. Методы теории хаоса для задач динамического управления контакт-центрами Ковалгин Ю.А. О применении программы Radio Mobile для расчета зон покрытия передатчиков, работающих в диапазоне ОВЧ Ковальский С.П., Стахеев И.Г., Титова О.В., Фокин Н.И. Метод ограничений в многокритериальной задаче распределения потоков Орлов А.Е., Новосадов К.С., Сальников Д.В., Беженарь В.Н. Алгоритм расчета показателя неготовности направления связи по дождевому фактору с учетом выполнения требований к пропускной способности, достоверности и ограничения частотно-энергетического ресурса Фокин Г.А., Владыко А.Г. Позиционирование транспортных средств с комплексированием дальномерных, угломерных и инерциальных измерений в расширенном фильтре Калмана	8 18 24 37 44 51
Bugrov V. Dynamic quantization of digital filter coefficients Goldstein A., Kislyakov S., Fenomenov M. Chaos theory methods for contact-center dynamic control Kovalgin Y. On the application of the Radio Mobile software for calculating the coverage zones of transmitters broadcasting in the VHF range Kowalski S., Titova O., Stakheev I., Fokin N. Method of restrictions in the multicriteria task of allocation of flows Orlov A., Novosadov K., Salnikov D., Bezhenar V. Evaluating unavailability algorithm of the satellite link based on the rain factor, taking into account the requirements for bandwidth, reliability and limitation of the frequency and energy resource Fokin G., Vladiko A. Vehicles positioning with the fusion of time of arrival, angle of arrival and inertial measurements in the extended Kalman filter	
2.3 ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ И ТЕЛЕКОММУНИКАЦИИ	
Бобровский А.И. Развитие теоретических основ синтеза космических адаптивных видеоинформационных систем Коржик В.И., Старостин В.С., Флакман Д.А. Разработка метода использования цифровых водяных знаков для защиты от атаки клонирования бумажных сертификатов Молдовян Д.Н., Фахрутдинов Р.Ш., Мирин А.Ю., Костина А.А. Схема цифровой подписи со скрытой группой, обладающей двумерной цикличностью	69 79 85
Bobrovskiy A. Theoretical foundations development of space adaptive video information systems synthesis Korzhik V., Starostin V., Flaksman D. Elaboration of digital watermarking method for a protection of cloning attack on paper certificates Moldovyan D., Fahrutdinov R., Mirin A., Kostina A. Digital signature scheme with hidden group possessing two-dimensional cyclicity	
ТРУДЫ МОЛОДЫХ УЧЕНЫХ	
Жук Р.В. Способ определения потенциала нарушителя безопасности информации и реализуемых им уязвимостей программного обеспечения Салман В.Д. Анализ гомоморфных криптосистем Бенало и Пэйе для построения системы электронного голосования	95 102
Zhuk R. Method for determining the potential of an information security intruder and realizable software vulnerabilities Salman W. Analysis of homomorphic cryptosystems of Benaloh and Paillier for the construction of an electronic voting system	

ЭЛЕКТРОНИКА, ФОТОНИКА, ПРИБОРОСТРОЕНИЕ И СВЯЗЬ

**2.2.6 – Оптические
и оптико–электронные приборы
и комплексы**

**2.2.13 – Радиотехника, в том числе системы
и устройства телевидения**

**2.2.14 – Антенны, СВЧ–устройства
и их технологии**

**2.2.15 – Системы, сети и устройства
телекоммуникаций**

2.2.16 – Радиолокация и радионавигация

Динамическое квантование коэффициентов цифрового фильтра

В.Н. Бугров^{1*}

¹Нижегородский государственный университет им. Н.И. Лобачевского,
Нижегород, 603022, Российская Федерация

*Адрес для переписки: bug@rf.unn.ru

Информация о статье

Поступила в редакцию 11.05.2021

Принята к публикации 01.06.2021

Ссылка для цитирования: Бугров В.Н. Динамическое квантование коэффициентов цифрового фильтра // Труды учебных заведений связи. 2021. Т. 7. № 2. С. 8–17. DOI:10.31854/1813-324X-2021-7-2-8-17

Аннотация: Рассматривается возможность квантования коэффициентов цифрового фильтра в концепции динамического математического программирования, как процесса пошагового квантования коэффициентов с их дискретной оптимизацией на каждом шаге по общей для всего процесса квантования целевой функции. Динамическое квантование позволяет существенно уменьшить функциональную ошибку реализации требуемых характеристик малоразрядного цифрового фильтра в сравнении с классическим квантованием. Приводится алгоритм пошагового динамического квантования методами целочисленного нелинейного программирования с учетом заданного масштабирования сигнала и радиуса полюсов передаточной функции фильтра. Иллюстрируется эффективность применения данного подхода на примере динамического квантования коэффициентов каскадного полосно-пропускающего БИХ-фильтра высокого порядка с минимальной разрядностью представления целочисленных коэффициентов. Проведен сравнительный анализ функциональных ошибок квантования, а также проверка работоспособности квантованного фильтра на тестовом и реальном сигналах.

Ключевые слова: цифровой фильтр, квантование, динамическое программирование, дискретный синтез, целевая функция.

Введение

К основным эффектам конечной разрядности представления данных в системах цифровой фильтрации относится погрешности характеристик, обусловленные квантованием коэффициентов цифрового фильтра. Наиболее часто в алгоритме цифровой фильтрации используются вычисления в формате с фиксированной точкой (ФТ), когда все переменные цифрового алгоритма являются дробными вещественными числами, принадлежащими диапазону от -1 до $1-2^{-R}$, где R – число значащих битов, используемых для двоичного представления числа [1–3]. Таким образом, заданным числом двоичных разрядов можно реализовать только дискретный ряд определенных значений коэффициентов фильтра в интервале $|x_i| < 1,0$, где $X(x_1, x_2, \dots, x_i, \dots, x_n)$ – совокупный вектор коэффициентов цифрового фильтра. Однако аналитический расчет коэффициентов, например, преобразованием аналогового прототипа БИХ-фильтра, методом окна или частотной выборки для КИХ-фильтра, осуществляется в непрерывном n -мерном вещественном пространстве E^n с неограниченной

точностью, поэтому на стадии реализации фильтра необходима процедура квантования его коэффициентов, которая в формате ФТ осуществляется следующим образом: для заданной длины битового представления R целочисленный код (квант) вещественного неквантованного коэффициента $x_i \in E^n$ в варианте округления до ближайшего целого (round) равен целой части выражения в скобках:

$$x_i^{\text{int}} = \text{int}(x_i \cdot 2^R + 0,5), \quad (1)$$

а в варианте усечения (fixed) он равен:

$$x_i^{\text{int}} = \text{int}(x_i \cdot 2^R).$$

Тогда вещественный квант данного коэффициента определяется по выражению:

$$x_i^{\text{float}} = \frac{x_i^{\text{int}}}{2^R}, \quad (2)$$

а ошибка квантования коэффициента составляет:

$$\varepsilon_i = x_i - x_i^{\text{float}}. \quad (3)$$

В рамках статистической модели [2] ошибки квантования коэффициентов считаются равномерно распределенными случайными величинами с нулевым средним значением. Такая статистическая модель позволяет рассчитать дисперсию шума квантования коэффициентов и их статистическую разрядность, необходимую для выполнения определенных функциональных требований. При этом каждому типу цифрового фильтра, его порядку и структуре построения свойственны различные статистические характеристики квантования коэффициентов. Ошибки квантования в классическом аналитическом проектировании приводят к искажению частотных характеристик и необходимости масштабирования коэффициентов фильтра. В рекурсивных структурах возможны также появления предельных циклов.

Однако ошибку квантования можно устранить, осуществив дискретизацию параметрического пространства коэффициентов фильтра только теми значениями, при которых ошибки квантования (3) равны нулю. Математическое задание квантованного пространства при этом возможно как вещественным (2), так целочисленным ее кодом (1). Фильтры, использующие вещественную арифметику вычислений и дискретизацию коэффициентов, принято называть квантованными вещественными цифровыми фильтрами (ВЦФ), а фильтры, использующие целочисленную арифметику вычислений и дискретизацию параметров – целочисленными цифровыми фильтрами (ЦЦФ). Вариант целочисленного представления является более универсальным и практически значимым, так как ЦЦФ имеют минимальную вычислительную сложность и могут быть реализованы на любой цифровой платформе (MCU, DSP, FPGA) без наличия FPU-сопроцессора (*от англ.* Floating Point Unit) в структуре вычислителя. Разрядность такого представления определяется интервалом изменения целочисленных коэффициентов фильтра, что практически весьма удобно при реализации алгоритма динамического квантования коэффициентов. При необходимости целочисленное решение легко преобразовать в вещественное квантованное решение формата ФТ, используя для этого их однозначную связь через соотношение (2). Формирование целочисленного кода часто называют нормализацией коэффициентов фильтра, когда в дискретном n -мерном пространстве I_R^n формируется вектор целочисленных коэффициентов $IX(ix_1, ix_2, \dots, ix_i, \dots, ix_n)$, квантованных с заданной разрядностью R их представления.

Ошибки квантования коэффициентов (3) вызывают, как уже сказано выше, соответствующую функциональную ошибку $\Delta H_{(\omega)}^R$, ошибку реализации требуемой характеристики цифрового фильтра. Для дискретного табулированного представления характеристики наиболее часто функцио-

нальная ошибка выполнения требований формируется по критерию среднеквадратичного отклонения (СКО):

$$\Delta H_{(\omega)}^R = \frac{1}{p} \cdot \sum_{k=1}^p [Y_k^R(IX) - Y_k^T]^2, \quad (4)$$

где $Y_k^R(IX)$ – текущее значение характеристики фильтра на k -й дискретной частоте диапазона определения для заданной разрядности R представления коэффициентов, а Y_k^T – требуемое значение частотной характеристики фильтра.

В некоторых случаях функциональную ошибку выгодно формировать и по критерию максимальной ошибки:

$$\Delta H_{(\omega)}^R = \max_k |Y_k^R(IX) - Y_k^T|.$$

Функциональная ошибка (4) будет, естественно, возрастать с понижением разрядности представления коэффициентов и ограничивать тем самым минимальную разрядность коэффициентов допустимым уровнем ошибки реализации требуемой характеристики. В классическом квантовании ошибки (3) и (4) определяют порог разрядного представления коэффициентов при реализации цифрового фильтра на конкретной цифровой платформе.

Приведенный алгоритм описывает классический «статический» механизм квантования коэффициентов фильтра, квантования в один проход, когда дальнейшее уменьшение разрядности вызывает уже недопустимую функциональную ошибку выполнения требований.

Однако существенное уменьшение функциональной ошибки (4) можно обеспечить применением «динамического» механизма квантования коэффициентов на основе динамического программирования [4, 5], как раздела оптимального программирования, в котором процесс решения общей задачи может быть разбит на отдельные этапы (шаги), причем результаты решения на одном этапе процесса квантования существенно влияют на последующие шаги. Принцип оптимальности Р. Беллмана, являющийся основным содержанием динамического программирования, гарантирует, что решение на любом шаге не локально лучше, а лучше с точки зрения процесса квантования коэффициентов в целом.

Алгоритм динамического квантования

Рассмотрим алгоритм динамического квантования на примере квантования коэффициентов нормализованного БИХ-фильтра, состоящего из каскадного соединения m -звеньев второго порядка ($m = N/2$, где N – порядок фильтра), с передаточной функцией [6, 7]:

$$H(z) = \prod_{i=1}^m \frac{b_{0i} + b_{1i}z^{-1} + b_{2i}z^{-2}}{a_{0i} + a_{1i}z^{-1} + a_{2i}z^{-2}}, \quad (5)$$

где комплексная переменная z при переходе к описанию частотной характеристики $K(e^{j\omega T}) = |K(e^{j\omega T})| \cdot e^{j\Phi(\omega T)}$ принимает значение $z = e^{j\omega T}$, где T – период дискретизации.

Квантование коэффициентов комплексного коэффициента передачи (5) при этом осуществляется целочисленным кодом (1). Тогда разностное уравнение для одного звена второго порядка имеет вид:

$$y_n = \frac{(b_0 x_n + b_1 x_{n-1} + b_2 x_{n-2} - a_1 y_{n-1} - a_2 y_{n-2})}{a_0}, \quad (6)$$

где x_n, y_n – целочисленные входная и выходная временные последовательности; a_0 – нормирующий power-of-two коэффициент:

$$a_{0i} \in \{2^q\}, \quad q = \overline{0, R} \quad i = \overline{1, m}. \quad (7)$$

На рисунке 1 приведена типичная структура звена целочисленного БИХ-фильтра, соответствующая разностному уравнению (6).

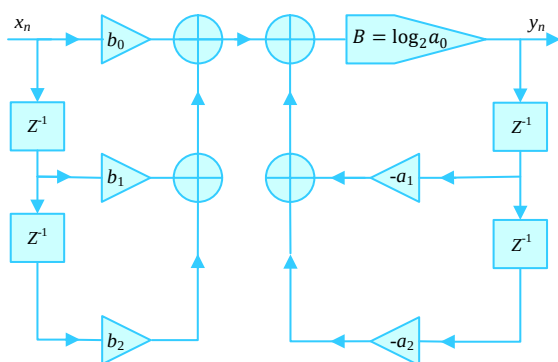


Рис. 1. Структура рекурсивного звена

Fig. 1. Structure of a Recursive Section

Как видно, при вычислении отклика фильтра, кроме традиционных операций сложения, умножения и задержки на такт, присутствует операция сдвига на $B = \log_2 a_0$ бит, с помощью которой реализуется целочисленное деление на нормирующий power-of-two коэффициент.

На рисунке 2 показана общая блок-схема алгоритма динамического пошагового квантования коэффициентов ЦЦФ как последовательности поисковых задач с поэтапным понижением разрядности R представления коэффициентов фильтра вплоть до требуемого малоразрядного их представления $R^T = R_{\min}$.

На первом, стартовом этапе неквантованные вещественные коэффициенты $X \in E^n$ сначала квантуются целочисленным кодом (1) со стартовой высокой разрядностью $R_{\max} = 16$ бит (блоки 1–3), формируя тем самым дискретное целочисленное решение на квантованном множестве коэффициентов $IX \in I_{R_{\max}}^n$ с соответствующей функциональной ошибкой квантования (4).

Затем на этом дискретном множестве ставится задача оптимизации дискретных коэффициентов данной разрядности (блок 4) по общей для всех

этапов динамического квантования целевой функции (4), как задачи целочисленного нелинейного программирования (блок 4), которая для каскадного БИХ-фильтра с передаточной функцией (5) записывается так [6]:

$$F^0(IX_R) = \min F(IX_R), \quad IX_R \in I_R^{5m}, \quad (8)$$

$$\begin{aligned} -2^R < a_{di} < 2^R, \quad d = \overline{1, 2}, \quad i = \overline{1, m}, \\ -2^R < b_{di} < 2^R, \quad d = \overline{1, 2}, \quad i = \overline{1, m}, \end{aligned} \quad (9)$$

$$a_{0i} \in \{2^q\}, \quad q = \overline{0, R} \quad i = \overline{1, m}, \quad (10)$$

$$|Z_{pi}| < r_{\max}, \quad i = \overline{1, m}, \quad (11)$$

$$K_i^{\min} \leq |K_i(e^{j\omega T})| \leq K_i^{\max}, \quad i = \overline{1, m}, \quad (12)$$

где $F(IX_R)$ – целевая функция; d – индекс коэффициента передаточной функции (5); $r_{\max}$ – допустимый максимальный радиус полюсов; $K_i^{\min}, K_i^{\max}$ – допустимые границы изменения коэффициента усиления i -го звена.

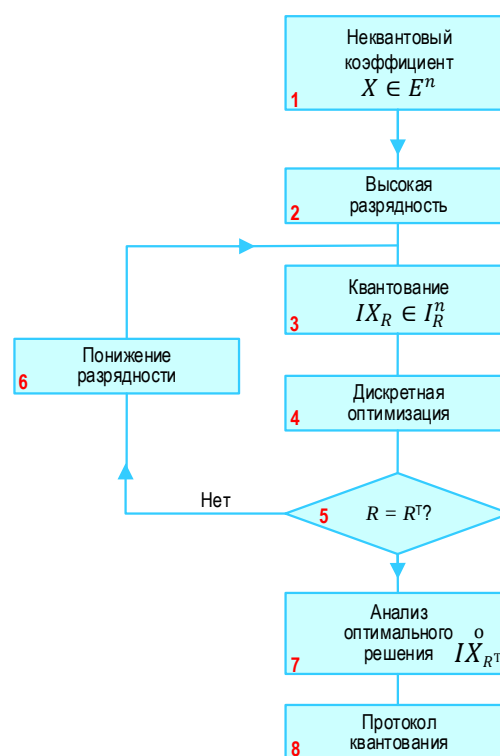


Рис. 2. Структура динамического квантования

Fig. 2. Dynamic Quantization Framework

Задача пошаговой дискретной оптимизации (8) записана относительно целочисленного пространства коэффициентов размерностью $5m$. Решение ее невозможно традиционными аналитическими методами [1, 8–10], а численное решение в дискретном квантованном пространстве коэффициентов формата ФТ возможно только применением сеточных алгоритмов поисковой минимизации, когда дискретность сетки поиска задается числом двоичных разрядов R , которыми в кодовом пространстве отображается каждая i -я переменная. Этим требованиям вполне отвечает поисковый алгоритм гло-

бальной минимизации полимодальных целевых функций на детерминированной сетке [11]. При этом выполнение функциональных ограничений устойчивости (11) по всем полюсам коэффициента передачи с радиусами, не выше $r_{\max}$, а также масштабирования усиления (12) в каскадной структуре реализуется применением штрафных функций [11, 12]. Малые потери на поиск и высокая надежность отделения глобального экстремума в квантованном параметрическом пространстве являются характерными достоинствами данного алгоритма.

При этом необходимо заметить, что квантование коэффициентов на текущем шаге (блок 3) обычно не приводит к потере глобального экстремума общей целевой функции и в зоне притяжения глобального экстремума осуществляется лишь оптимизации, подстройка дискретных параметров с новой разрядностью их представления.

Далее в блоке 5 производится проверка соответствия текущей разрядности квантования R с требуемой минимальной разрядностью R^T и в случае несоответствия осуществляется понижение разрядности представления коэффициентов (блок 6) с последующим их квантованием и оптимизацией на новом дискретном множестве (блоки 3 и 4) также по общей для всех этапов динамического квантования целевой функции ошибки (4). Пересчет целочисленных коэффициентов разрядности L в разрядность представления M при этом осуществляется по следующей простой формуле $x_M^{\text{int}} = x_L^{\text{int}} \times 2^{M-L}$. Так при понижении разрядности на 1 бит мы имеем $x_{R-1}^{\text{int}} = x_R^{\text{int}}/2$. Точно так же пересчитываются и границы нового дискретного множества целочисленных коэффициентов. Таким образом, решение на текущем этапе динамического квантования является стартовой точкой этапа последующего. Цикл повторяется до достижения требуемой минимальной разрядности представления коэффициентов, после чего осуществляется выявление оптимального малоразрядного решения, обеспечивающего допустимый уровень ухудшения функциональных характеристик фильтра, и проводится анализ этого решения во временной и частотной областях. Процесс завершается формированием итогового протокола динамического квантования коэффициентов цифрового фильтра (блоки 7 и 8 общей диаграммы).

Таким образом, процесс динамического квантования состоит из этапов последовательного понижения разрядности коэффициентов с формированием дискретного множества коэффициентов текущей разрядности I_R^n с последующим решением задачи дискретной оптимизации коэффициентов на этом множестве по критерию общей функциональной ошибки реализации требований. Характерно, что дискретные множества, формируемые на этапах динамического квантования, являются

вложенными подмножествами исходного дискретного множества коэффициентов с максимальной разрядностью:

$$I_{R_{\min}}^n \subset I_{R_{\min}+1}^n \subset I_{R_{\min}+2}^n \subset \dots \subset I_{R_{\max}}^n.$$

Это означает, что любое малоразрядное решение всегда может быть реализовано и на цифровой платформе большей разрядности с той же функциональной ошибкой выполнения требований. Обратное же утверждение, естественно, несправедливо.

Динамическое квантование коэффициентов БИХ-фильтра

Рассмотрим пример квантования методом динамического программирования коэффициентов исходного синтезированного полосно-пропускающего БИХ-фильтра при целочисленном описании его квантованного варианта. В итоге необходимо определить минимальную разрядность квантования коэффициентов, при которой выполняются следующие спецификационные требования к функциональным показателям фильтра:

- полоса пропускания: 400–700 Гц;
- коэффициент передачи в полосе пропускания: 0 дБ с допуском ± 1 дБ;
- уровень подавления на частотах ниже 200 и выше 900 Гц: 40 дБ;
- нелинейность фазы в полосе: 10° ;
- порядок рекурсивного фильтра: 20;
- максимально допустимый радиус полюсов: 0,9;
- масштабирование коэффициентов передачи звеньев в интервал $\{1-4\}$;
- частота дискретизации: 2 кГц;
- программная реализация на микроконтроллере с целочисленным RISC-ядром.

График требуемой АЧХ полосно-пропускающего фильтра (ППФ) приведен на рисунке 6а (выделен красным цветом), а исходные неквантованные вещественные его коэффициенты указаны в таблице 1 (где m – звено фильтра; U – усиление звена $K_{\max}$). Приведенные выше требования спецификаций в неквантованном ППФ выполнялись с большим запасом.

Осуществим сначала классическое квантование коэффициентов данного фильтра в варианте округления по соотношению (1) последовательно до минимальной разрядности их представления 3 бита, фиксируя функциональную ошибку (4) реализации требуемой АЧХ (таблица 2, где δ – $\Delta N_{\text{клас}}$; η – $\Delta N_{\text{дин}}$). Из таблицы видно, что разброс квантования по функциональной ошибке превосходит два порядка (рисунок 3а), поэтому график соответствующей зависимости для наглядности представлен и в логарифмическом масштабе (рисунок 3б). Необходимо заметить, что в таблице и графиках разрядность квантования целочисленных коэффициентов указана уже с учетом знакового бита.

ТАБЛИЦА 1. Оптимальные коэффициенты неквантованного фильтра

TABLE 1. Optimal Coefficients of a Non-Quantized Filter

m	Коэффициенты передаточной функции					U
	a_1	a_2	b_0	b_1	b_2	
1	0,414917	0,658203	-0,307983	-0,031860	0,408936	2,1
2	-0,227051	0,469971	-0,072266	-0,815430	-0,235840	1,7
3	0,994995	0,667725	-0,195679	0,709595	0,403198	2,8
4	-0,672852	0,654785	0,170898	-0,992065	0,805298	2,6
5	-0,659668	0,872803	0,979980	-0,740967	0,605591	3,3
6	0,999878	0,675537	0,267944	0,144287	0,552612	1,7
7	0,998901	0,673584	0,729492	0,438477	0,266235	1,6
8	-0,144653	0,528687	0,160278	0,008057	-0,501831	1,4
9	-0,015869	0,409546	0,085083	0,742798	0,532715	1,5
10	-0,130493	0,146484	0,320435	-0,722900	0,250000	1,0

ТАБЛИЦА 2. Ошибки классического и динамического квантования

TABLE 2. Classical and Dynamic Quantization Errors

СКО	Разрядность квантования, бит									
	НК	14	12	10	8	7	6	5	4	3
δ	0,102	0,102	0,103	0,112	0,202	0,507	0,304	3,28	3,82	10,4
η	0,102	0,096	0,103	0,094	0,091	0,097	0,094	0,153	0,40	1,32

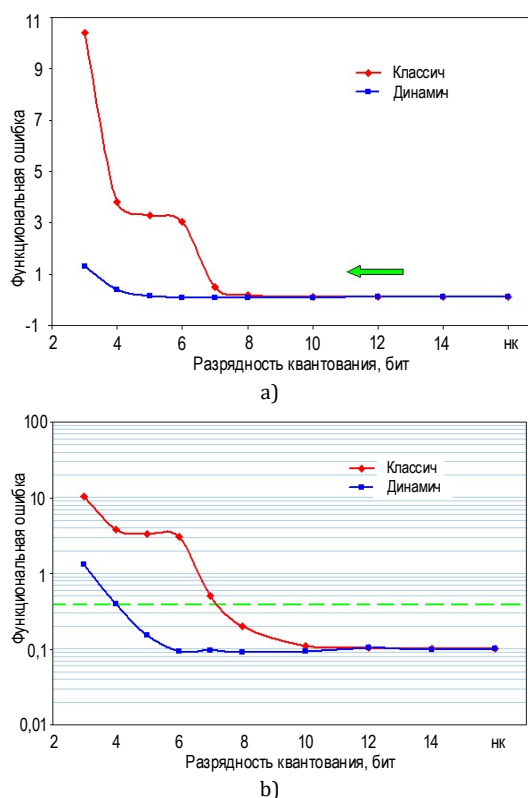


Рис. 3. Ошибки классического и динамического квантования в линейном (а) и логарифмическом (б) масштабе

Fig. 3. Errors of Classical and Dynamic Quantization in Linear (a) and Logarithmic (b) Scales

Можно также отметить, что случая потери устойчивости БИХ-фильтра при квантовании с округлением не наблюдалось, так как исходный неквантованный ППФ был синтезирован с хорошим запасом по устойчивости (радиусы полюсов не превышали 0,9), хотя в противном случае выход полюсов квантованного фильтра за единичную окружность на малой разрядности представления коэффициентов при классическом механизме квантования наблюдается довольно часто.

На рисунке 4 отражена модификация АЧХ квантуемого цифрового фильтра при понижении разрядности его коэффициентов. Осуществим теперь динамическое квантование коэффициентов исходного фильтра (таблица 1) по приведенному выше алгоритму и целевой функции (4) также последовательно от максимальной до минимальной 3-битовой разрядности их представления, фиксируя функциональную ошибку $\Delta H_{\text{дин}}$ реализации требуемой АЧХ в таблице 2. Графики динамической ошибки также представлены на рисунке 3, а соответствующая модификация АЧХ фильтра при квантовании отображена на рисунке 5.

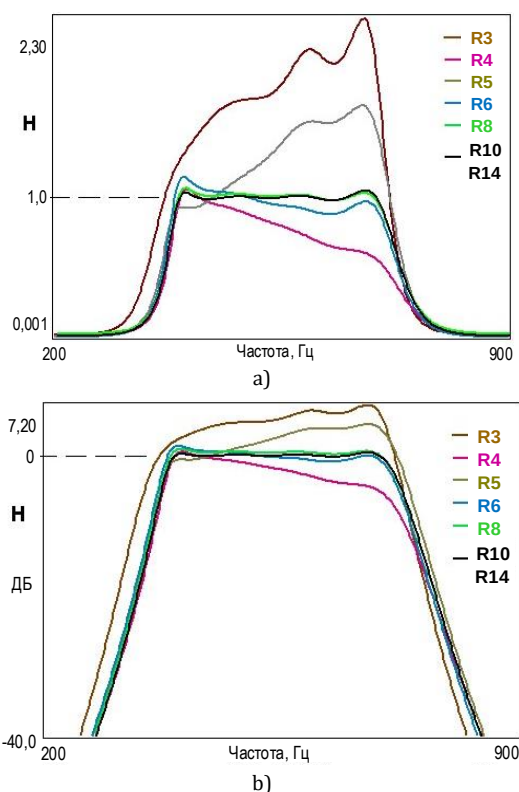


Рис. 4. Модификация АЧХ фильтра при классическом квантовании коэффициентов в линейной (а) и логарифмической (б) шкале

Fig. 4. Modification of the Frequency Response of the Filter with Classical Quantization of Coefficients in Linear (a) and Logarithmic (b) Scales

Из таблицы и приведенных графиков видно, что функциональная ошибка классического квантования резко возрастает, начиная с уровня 10–11 битового представления коэффициентов полосового БИХ-фильтра, в то время как ошибка динамиче-

ского квантования имеет тенденцию к росту, только начиная с уровня 4–5 битового представления коэффициентов. Также можно отметить, что допустимому уровню 0,4 (отображен на рисунке 3b пунктиром) выполнения спецификационных требований по функциональной ошибке соответствует только 8-битовое решение при классическом квантовании, в то время как при динамическом квантовании для представления коэффициентов вполне достаточно 4-х бит. В таблице 3 приведены эти оптимальные значения целочисленных 4-битовых коэффициентов передаточной функции ППФ, а также достигнутый уровень его каскадного масштабирования. Графики частотных характеристик фильтра представлены на рисунке 6.

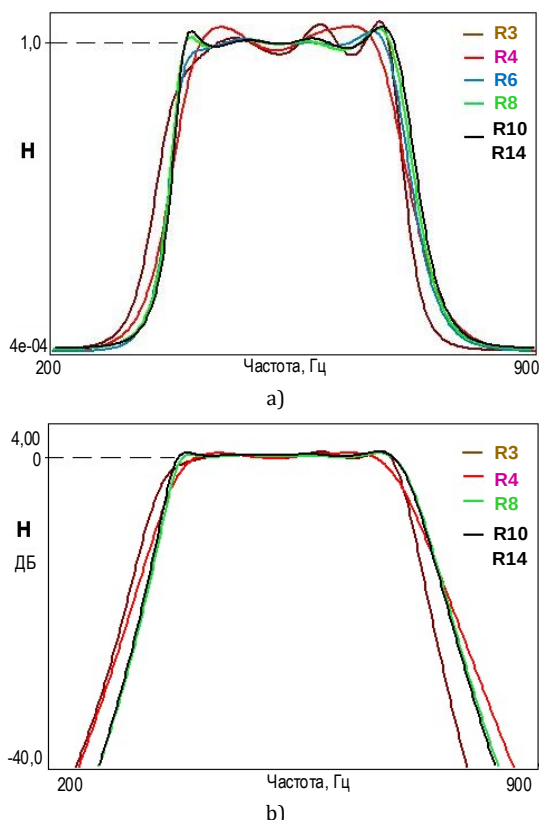


Рис. 5. Модификация АЧХ фильтра при динамическом квантовании коэффициентов в линейной (а) и логарифмической (б) шкале

Fig. 5. Modification of the Frequency Response of the Filter with Dynamic Quantization of Coefficients in Linear (a) and Logarithmic (b) Scales

Как видно, в полосе пропускания неравномерность АЧХ отквантованного ППФ составляла $\Delta K = 0,98$ дБ, а фазовая нелинейность не превышала $\Delta\varphi = 5,5^\circ$, что полностью соответствует требованиям спецификации.

Вещественные 4-битовые коэффициенты формата 4.3 приведены в таблице 4. С помощью *m*-файла данные коэффициенты загружались в пакет MATLAB для контрольного анализа характеристик отквантованного ППФ во временной и частотной области (рисунок 7). Видно, что при динамическом

квантовании полюса передаточной функции не вышли за границы допустимого максимального радиуса 0,9, а инерционность квантованного ППФ по задержке импульсного отклика составила 15 семплов.

ТАБЛИЦА 3. Оптимальные целочисленные коэффициенты фильтра

TABLE 3. Optimal Integer Filter Coefficients

m	Коэффициенты передаточной функции					U
	a_1	a_2	b_0	b_1	b_2	
1	3	5	-2	3	3	1,9
2	-1	3	-7	-1	-1	1,3
3	7	4	-3	5	3	2,1
4	-4	4	0	-6	7	2,0
5	-5	6	7	-7	3	2,7
6	7	5	2	1	4	1,1
7	7	5	5	3	2	1,0
8	-2	4	1	-2	-4	1,4
9	0	1	-6	-5	2	1,4
10	-2	-1	-1	-4	7	1,1

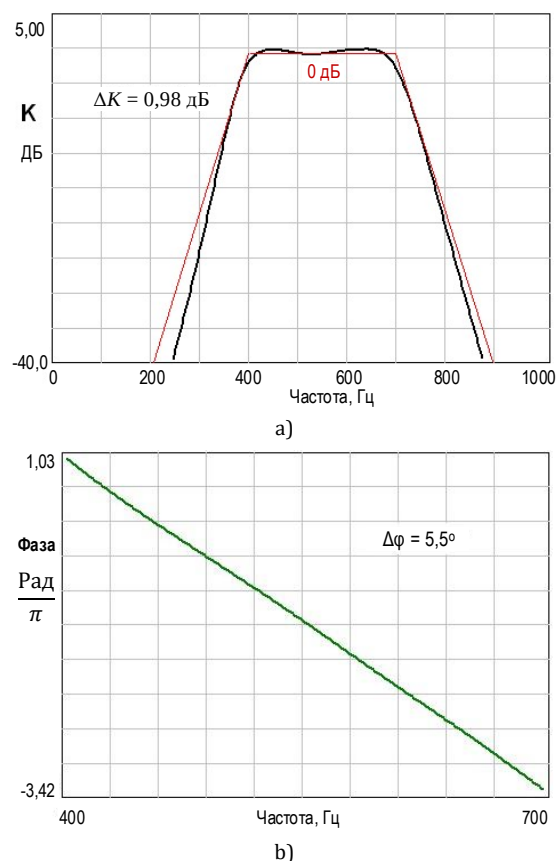


Рис. 6. Характеристики 4-битового ППФ: а) АЧХ; б) ФЧХ в полосе пропускания

Fig. 6. Characteristics of a 4-Bit BPF: a) Frequency Response; b) Phase Response in the Passband

Таким образом, анализ в пакете MATLAB полностью подтверждает квантованные характеристики полосового БИХ-фильтра.

ТАБЛИЦА 4. Квантованные вещественные коэффициенты 4-битового фильтра

TABLE 4. Quantized Real Coefficients of a 4-Bit Filter

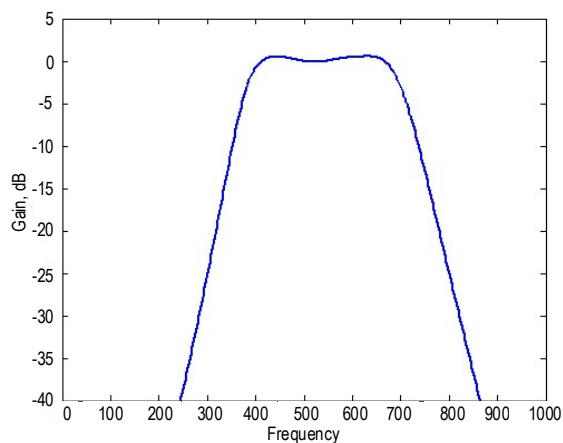
m	Коэффициенты передаточной функции				
	a_1	a_2	b_0	b_1	b_2
1	0,375	0,625	-0,250	0,375	0,375
2	-0,125	0,375	-0,875	-0,125	-0,125
3	0,875	0,500	-0,375	0,625	0,375
4	-0,500	0,500	0,000	-0,750	0,875
5	-0,625	0,750	0,875	-0,875	0,375
6	0,875	0,625	0,250	0,125	0,500
7	0,875	0,625	0,625	0,375	0,250
8	-0,250	0,500	0,125	-0,250	-0,500
9	0,000	0,125	-0,750	-0,625	0,250
10	-0,250	-0,125	-0,125	-0,500	0,875

Проведем модельное тестирование работоспособности данного 4-битового ППФ в пакете MATLAB. Для этого на временном интервале 2 сек конструируем ЛЧМ-сигнал с девиацией вплоть до частоты Найквиста при $f_s = 2$ кГц и пропускаем его через квантованный фильтр.

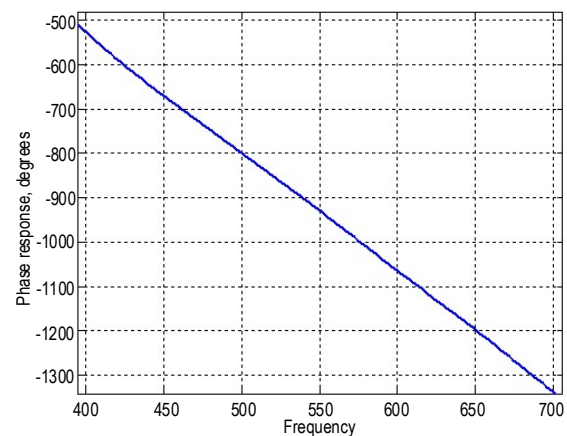
Результаты модельной фильтрации ЛЧМ-сигнала (рисунок 8) подтверждают полную работоспособность квантованного 4-битового ППФ.

Практическая реализация квантованного фильтра осуществлялась на многофункциональном микроконтроллере MSP430F1611 фирмы Texas Instruments [13] с целочисленным RISC-ядром.

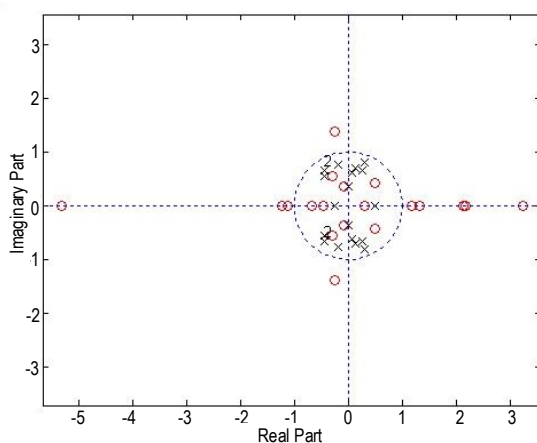
Экспериментальные частотные характеристики фильтра, измеренные на реальном сигнале при частоте дискретизации $f_s = 2$ кГц, приведены на рисунке 9.



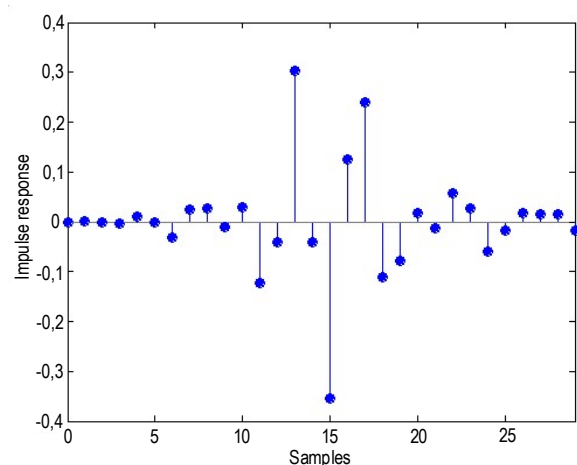
a)



b)



c)



d)

Рис. 7. Характеристики ВЦФ: а) АЧХ; б) ФЧХ в полосе пропускания; в) карта нулей и полюсов в z-плоскости; д) импульсная характеристика фильтра

Fig. 7. Real Filter Characteristics: a) Frequency Response; b) Phase Response in the Passband; c) Map of Zeros and Poles in the z-Plane; d) Impulse Response of the Filter

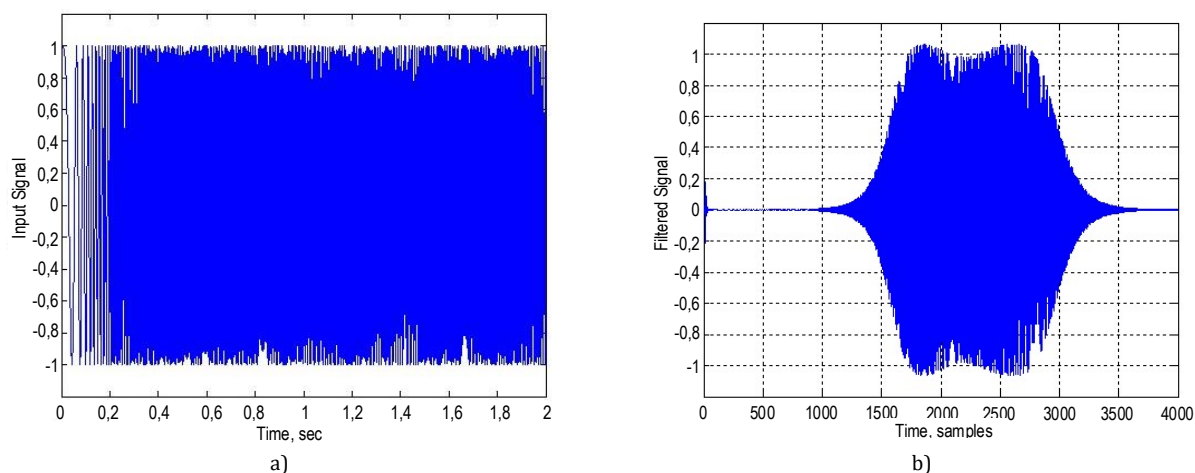


Рис. 8. Входной ЛЧМ-сигнал (а) и его фильтрация квантовым ППФ (б)

Fig. 8. Input Signal (a) and its Filtering (b)

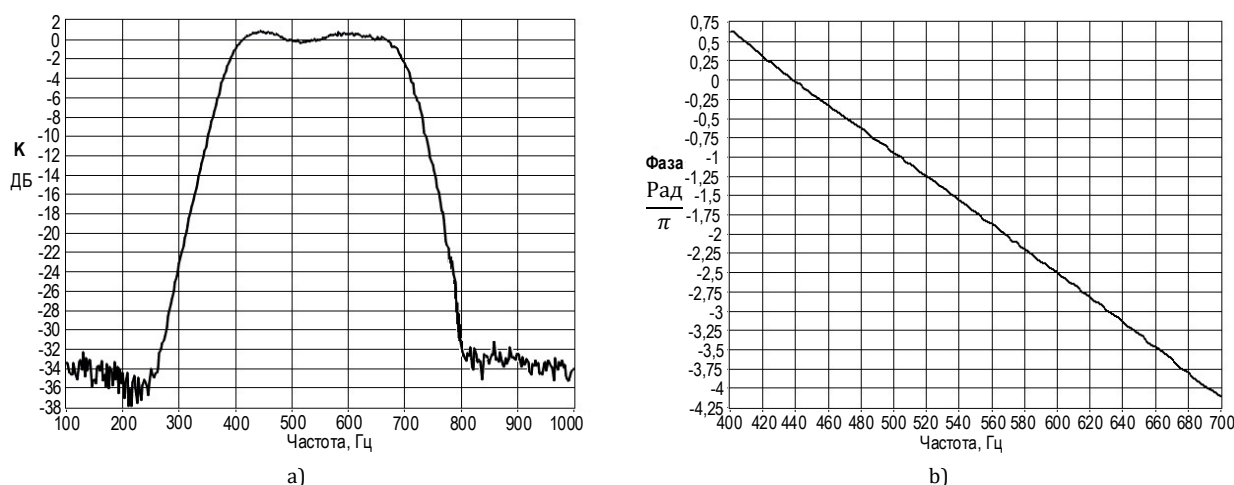


Рис. 9. Экспериментальные измерения АЧХ (а) и ФЧХ (б) в полосе пропускания

Fig. 9. Experimental Measurements of the Frequency Response (a) and Phase Response (b) in the Passband

Исследование устойчивости фильтра по предельным циклам осуществлялось путем анализа выходного сигнала при нулевом входном сигнале. Анализ показал, что при максимальном радиусе полюсов 0,87 в 4-битовом ППФ предельные циклы отсутствуют. Таким образом, все функциональные требования при динамическом квантовании коэффициентов рекурсивного ППФ были выполнены.

Заключение

При проектировании высокоскоростных мало-разрядных цифровых фильтров одной из основных задач является уменьшение ошибки реализации требуемых функциональных характеристик при уменьшении разрядности представления данных в алгоритме цифровой фильтрации. При классическом квантовании коэффициентов фильтра весьма существенные трудности возникают уже при квантовании коэффициентов до 8–10 бит, что подтверждается приведенными в статье данными.

Эти негативные последствия могут во многом быть устранены применением динамического программирования, как процесса пошагового квантования коэффициентов с их дискретной оптимизацией на каждом шаге по общей для всего процесса квантования целевой функции. Для приведенного в статье примера динамическое квантование коэффициентов позволило практически на порядок уменьшить функциональную ошибку реализации требуемых характеристик 4-битового цифрового БИХ-фильтра в сравнении с классическим статистическим квантованием. Современные алгоритмы минимизации на дискретной сетке с заданной разрядностью представления параметров позволяют решать динамическую задачу квантования надежно и эффективно, что позволяет проектировать цифровые фильтры минимальной разрядности представления данных в целочисленном или вещественном квантованном пространстве с полным контролем устойчивости и масштабирования каскадных БИХ-фильтров, существенно повысить качество проектируемых мало-разрядных цифровых фильтров.

Список используемых источников

1. Ifeachor E.C., Jervis B.W. Digital Signal Processing: A Practical Approach. Harlow: Pearson Education, 2002. 640 p.
2. Рабинер Л., Гоулд Б. Теория и применение цифровой обработки сигналов. Пер. с англ. М.: Мир, 1978. 840 с.
3. Мингазин А.Т. Анализ квантованных КИХ-фильтров // Цифровая обработка сигналов. № 4. 2019. С. 3–13.
4. Есипов Б.А. Методы оптимизации и исследования операций. Самара: Изд-во Самар. гос. аэрокосм. ун-та, 2007. 180 с.
4. Вентцель Е.С. Исследование операций. Задачи, принципы, методологии. М.: Наука, 1988. 208 с.
5. Бугров В.Н. Дискретный синтез минимально-фазовых и линейно-фазовых цифровых БИХ-фильтров // Компоненты и технологии. 2019. № 10(219). С. 92–103.
6. Бугров В.Н., Пройдаков В.И., Артемьев В.В. Синтез цифровых фильтров методами целочисленного нелинейного программирования // 17-я международная конференция «Цифровая обработка сигналов и её применение – DSPA-2015»: тезисы докладов. М.: НТО РЭС им. А.С. Попова, 2015. С. 200–204.
7. Мингазин А.Т. Минимум максимальной взвешенной ошибки аппроксимации АЧХ классических аналоговых и цифровых фильтров // Цифровая обработка сигналов. 2018. № 4. С. 18–20.
8. Getu B.N. Digital IIR Filter Design using Bilinear Transformation in MATLAB // International Conference on Communications, Computing, Cybersecurity, and Informatics (CCCI, Sharjah, United Arab Emirates, 3–5 November 2020). IEEE, 2020. DOI:10.1109/CCCI49893.2020.9256625
9. Zhang M., Kwan H.K. FIR filter design using multiobjective teaching-learning-based optimization // Proceedings of the 30th Canadian Conference on Electrical & Computer Engineering (CCECE, Windsor, Canada, 13–16 May 2018). IEEE, 2018. DOI:10.1109/CCECE.2017.7946800
10. Воинов Б.С., Бугров В.Н., Воинов Б.Б. Информационные технологии и системы: поиск оптимальных, оригинальных и рациональных решений. М.: Наука, 2007. 730 с.
11. Zangwill W.I. Non-linear programming via penalty functions // Management Science. 1967. Vol. 13. Iss. 5.
12. Семенов Б.Ю. Микроконтроллеры MSP430. Первое знакомство. М.: Изд-во «Солон-пресс», 2006. 180 с.

* * *

Dynamic Quantization of Digital Filter Coefficients

V. Bugrov¹

¹N.I. Lobachevsky State University of Nizhni Novgorod,
Nizhni Novgorod, 603022, Russian Federation

Article info

DOI:10.31854/1813-324X-2021-7-2-8-17

Received 11th May 2021

Accepted 1st June 2021

For citation: Bugrov V. Dynamic Quantization of Digital Filter Coefficients. *Proc. of Telecom. Universities*. 2021;7(2):8–17. (in Russ.) DOI:10.31854/1813-324X-2021-7-2-8-17

Abstract: *The possibility of quantizing the coefficients of a digital filter in the concept of dynamic mathematical programming, as a dynamic process of step-by-step quantization of coefficients with their discrete optimization at each step according to the objective function, common to the entire quantization process, is considered. Dynamic quantization can significantly reduce the functional error when implementing the required characteristics of a low-bit digital filter in comparison with classical quantization. An algorithm is presented for step-by-step dynamic quantization using integer nonlinear programming methods, taking into account the specified signal scaling and the radius of the poles of the filter transfer function. The effectiveness of this approach is illustrated by dynamically quantizing the coefficients of a cascaded high-order IIR bandpass filter with a minimum bit depth to represent integer coefficients. A comparative analysis of functional quantization errors is carried out, as well as a test of the quantized filter performance on test and real signals.*

Keywords: *the IIR-filter, quantization, dynamic programming, discrete synthesis, criterion function.*

References

1. Ifeachor E.C., Jervis B.W. *Digital Signal Processing: A Practical Approach*. Harlow: Pearson Education; 2002. 640 p.
2. Rabiner R.L., Gold B. *Theory and Application of Digital Signal Processing*. New Jersey: Prentice-Hall, 1975.
3. Mingazin A.T. Analysis of Quantized Fir Filters. *DSPA*.2019;4:3–13. (in Russ.)
4. Esipov B.A. *Optimization Techniques and Operations Research*. Samara: Samara Aerospace University Publ.; 2007. 180 p. (in Russ.)
5. Ventcel E.C. *Operations Research. Objectives, Principles, Methodologies*. Moscow: Nauka Publ.; 1988. 208 p. (in Russ.)
6. Bugrov V.N. Discrete Synthesis of Minimum-Phase and Linear-Phase IIR Digital Filters. *Components and Technologies*.2019;10(219):92–103. (in Russ.)
7. Bugrov V.N., Proidakov V.I., Artemev V.V. Synthesis of Digital Filters by Methods of Integer Nonlinear Programming. *Proceedings of the 17-th International Conference "Digital Signal Processing and its Applications – DSPA-2015"*. Moscow: Russian Scientific and Technical Society of Radio Engineering, Electronics and Communications named after A.S. Popova Publ.; 2015. p.200–204. (in Russ.)
8. Mingazin A.T. Minimum-of-Maximum Weighted Error in Magnitude Response Approximation of Analog and Digital Classical Filters. *DSPA*. 2018;4:18–20. (in Russ.)
9. Getu B.N. Digital IIR Filter Design using Bilinear Transformation in MATLAB. *Proceedings of the International Conference on Communications, Computing, Cybersecurity, and Informatics, CCCI, 3rd–5th November 2020, Sharjah, United Arab Emirates*. IEEE; 2020. DOI:10.1109/CCCI49893.2020.9256625
10. Zhang M., Kwan H.K. FIR filter design using multiobjective teaching-learning-based optimization. *Proceedings of the 30th Canadian Conference on Electrical & Computer Engineering, CCECE, 13–16 May 2018, Windsor, Canada*. IEEE; 2018. DOI:10.1109/CCECE.2017.7946800
11. Voinov B.S., Bugrov V.N., Voinov B.B. *Information Technology and Systems: Search for Optimal, Original and Rational Solutions*. Moscow.: Nauka Publ.; 2007. 730 p. (in Russ.)
12. Zangwill W.I. Non-linear programming via penalty functions. *Management Science*. 1967;13(5).
13. Semenov B.U. *Microcontrollers MSP430. The first acquaintance*. Moscow: Solon-press Publ.; 2006. 180 p.

Сведения об авторе:

БУГРОВ
Владимир Николаевич

кандидат технических наук, доцент, доцент кафедры радиотехники Национального исследовательского Нижегородского университета им. Н.И. Лобачевского – ННГУ, bug@rf.unn.ru

Методы теории хаоса для задач динамического управления контакт-центрами

А.Б. Гольдштейн^{1, 2}, С.В. Кисляков^{1, 2}, М.А. Феноменов^{1, 2}

¹Санкт-Петербургский государственный университет телекоммуникаций им. проф. М.А. Бонч-Бруевича, Санкт-Петербург, 193232, Российская Федерация

²НТЦ АРГУС, Санкт-Петербург, 197198, Российская Федерация

*Адрес для переписки: s.v.kislyakov@gmail.com

Информация о статье

Поступила в редакцию 17.03.2021

Принята к публикации 21.04.2021

Ссылка для цитирования: Гольдштейн А.Б., Кисляков С.В., Феноменов М.А. Методы теории хаоса для задач динамического управления контакт-центрами // Труды учебных заведений связи. 2021. Т. 7. № 2. С. 18–23. DOI:10.31854/1813-324X-2021-7-2-18-23

Аннотация: Работа посвящена поиску оптимальных методов управления контакт-центрами (КЦ), в частности, методов прогнозирования нагрузки для дальнейшего расчета нужного количества операторов КЦ. Если количество операторов будет всегда больше, чем это требуется, то собственники КЦ будут нести дополнительные финансовые затраты на оплату труда операторов. Если сотрудников будет слишком мало, то качество обслуживания упадет. Прогнозирование нагрузки КЦ требуется для того, чтобы заранее выводить на работу оптимальное количество операторов. Для прогнозирования входящей нагрузки КЦ предложено применить теорию хаоса. Расчет показателя Ляпунова – его положительное значение – указывает на хаотическое поведение входного потока нагрузки КЦ. Для прогнозирования нагрузки использованы методы линейного и нелинейного прогнозирования, а также метод глобальной аппроксимации. В работе приведены результаты сравнения этих методов для задачи прогнозирования входящей нагрузки КЦ.

Ключевые слова: контакт-центр, автоматизация, теория хаоса, WFM.

Постановка задачи

Контакт-центры (КЦ) являются мощным инструментом взаимодействия с большим числом клиентов и установления продуктивного контакта с каждым из них. Они предоставляют клиентам информационные услуги с помощью операторов через речевые каналы, а также с использованием чатов, электронной почты, IP-телефонии. Для обеспечения необходимого качества обслуживания клиентов при минимальном числе операторов КЦ, необходимо заранее знать (или прогнозировать с максимально возможной точностью) объем входящей нагрузки, что позволит аналитическим путем рассчитать число операторов КЦ.

Показатели оценки качества услуг, предоставляемых КЦ

Одним из основных показателей, характеризующих обработанные входящие вызовы, является время ожидания клиента в очереди. Этот показатель сильно влияет на общее впечатление от пользования услугами КЦ. Считается, что оптимальным

значением будет формула 80/20, то есть 80 % звонков ожидают обработки менее 20 с.

Еще один важный показатель – это среднее время обработки вызова оператором. Слишком продолжительные обработки вызовов могут говорить о непрофессиональной работе операторов, а слишком короткие – о том, что они реально не предоставляют услуг потребителям. Если вызов не дожидается обработки оператором, значит услуги клиенту не будут предоставлены. Оптимальным значением этого показателя считается 4–8 %.

Оценка удовлетворенности клиентов является наиболее важным показателем и обычно определяется в ходе опросов после вызова, хотя в оценку могут быть включены и другие показатели, такие как, например, Net Promoter Score (пер. с англ. индекс потребительской лояльности).

На все эти показатели так или иначе влияет организация работы самих операторов (работников КЦ), их расписание и количество в смене. Если операторов будет меньше оптимального, то очередь вырастет.

тет и упадет уровень обслуживания. Если же их будет слишком много, то увеличится время простоя и возрастут потери на оплату труда. Поэтому крайне важно оптимизировать расписание операторов, которое напрямую зависит от входящей нагрузки – величины переменной, складывающейся из целого ряда факторов, например, – времени суток, дня недели и даже погоды.

Автоматизация управления работниками КЦ

В документах организации TeleManagement Forum комплекс таких задач определяется как Workforce management (WFM, *пер. с англ.* – управление рабочей силой). Это общее название совокупности процессов планирования, результатом которого является расписание для работников на некоторый будущий период. Анализ производится на основе данных о входящем трафике за предыдущие периоды и производительности операторов. Итогом работы становится расписание для каждого оператора КЦ.

WFM применительно к КЦ можно разделить на несколько задач [1]:

- прогнозирование нагрузки на определенных временных интервалах (обычно 15–30 мин);
- определение количества операторов и, при необходимости, операторов с определенной квалификацией, которые должны находиться в определенный временной интервал на рабочих местах;
- формирование расписания работы сотрудников КЦ.

Добиться минимальной ошибки для прогнозирования довольно сложно, т. к. требуется учесть множество факторов, влияющих на поток поступающих вызовов. Для разных профилей бизнеса это могут быть:

- всплески количества вызовов как результат маркетингового продвижения;
- изменения спроса – например, приобретение новой компании или появление новых продуктов;
- погодные факторы – снег, наводнения и очень жаркая погода могут оказать большое влияние на количество поступающих звонков;
- специальные мероприятия – такие события, как чемпионат мира, могут вызвать большой всплеск звонков;
- отказы оборудования – отключение питания, обрыв телефонных линий и т. д.

Влияние этих факторов необходимо минимизировать для получения максимально точных результатов прогнозирования.

Входные данные

В качестве исходных данных использовались данные о поступивших вызовах анонимного банка (за 1999 г.). Данные находятся в свободном доступе для проведения различных исследований.

Банковский КЦ предоставлял следующие услуги:

- информацию по транзакциям и банковским операциям для клиентов;
- интерактивный речевой ответ (с использованием модулей VRU (*аббр. от англ.* Voice Response Unit));
- поддержку пользователей сайта банка.

Данные организованы в 12 текстовых файлах, каждый из которых содержит данные о вызовах за месяц – от 20 000 до 30 000 вызовов в месяц. Каждый вызов характеризуется 17 полями.

Проверка на хаотичность временного ряда

Для продолжения работы необходимо убедиться в хаотичности исследуемого процесса, воспользовавшись следующими критериями [2–4]:

- 1) неотрицательный показатель Ляпунова, который указывает на хаотическую динамику;
- 2) фрактальная структура траектории в фазовом пространстве (пространство состояний), указывающая на наличие странного аттрактора.

Воспользуемся первым критерием. С помощью показателя Ляпунова проверяется чувствительность системы к вариации начальных условий. Максимальный показатель Ляпунова (Maximum Lyapunov Exponent) характеризует скорость разбегания близких траекторий, положительное значение которого говорит о хаотическом поведении системы.

Расчеты проводились с использованием пакетов программ TISEAN 3.0.1. и MATLAB.

Вычисление максимального показателя Ляпунова производилось с использованием утилиты lyap_k пакета TISEAN.

Результатом работы утилиты lyap_k является набор данных, представляющих собой зависимость логарифма коэффициента разбегания траекторий от времени:

$$S(\epsilon, m, \Delta n) = \frac{1}{N} \sum_{n=0}^N \ln \left(\frac{1}{|U(S_{n0})|} \sum_{S_n \in U(S_{n0})} |S_{n0+\Delta n} - S_{n+\Delta n}| \right),$$

где ϵ – окрестность точки S_{n0} ; m – размерность пространства; Δn – время; $U(S_{n0})$ – окрестность точки S_{n0} диаметра ϵ .

В качестве исходных данных выберем входящие вызовы за январь. Сформируем временные ряды и определим показатель Ляпунова для разных Δt . Полученные значения представлены в таблице 1.

ТАБЛИЦА 1. Расчет показателя Ляпунова

TABLE.1. Calculation of the Lyapunov Index

Δt , ч	0,1	0,5	1	2
λ_1	0,022935	0,051363	0,010193	0,010186

При малом Δt длительность разговора теоретически может превышать шаг временного ряда, приводя к сильным всплескам нагрузки. Это связано с тем, что при формировании временного ряда мы относим тот или иной разговор к соответствующему временному интервалу, опираясь на его начало. То есть, чем больше шаг временного ряда, тем меньше сказывается описанный выше эффект.

Из таблицы 1 видно, что условие наличия хаоса в системе удовлетворено, т. к. полученные показатели Ляпунова больше нуля.

Далее измерения производились при $\Delta t = 0,5$ ч, при максимальном значении показателя Ляпунова.

Применение методов прогнозирования

Сравним методы прогнозирования. В качестве интервала прогнозирования в эксперименте выберем одну неделю.

Приведем общую схему эксперимента (рисунок 1). Сначала будем формировать временной ряд для прогнозирования по описанному выше алгоритму. Первую часть ряда будем подавать в качестве входных данных на блок прогнозирования, вторую же часть используем для проверки вычисленного прогноза. Таким образом, сравнивая спрогнозированные значения и реальные, получим погрешность прогнозирования, которую позже будем использовать в качестве критерия для сравнения методов теории хаоса.

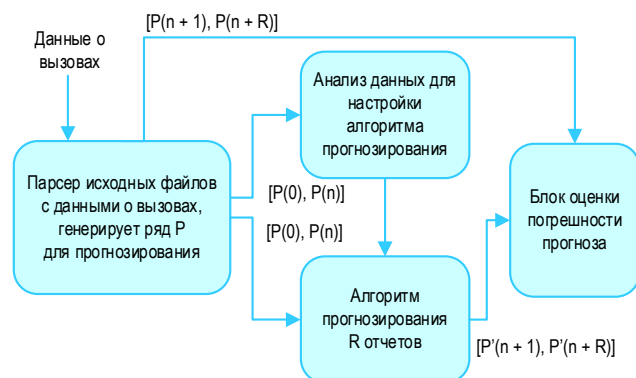


Рис. 1. Общая структура экспериментальной установки

Fig.1. High-Level Algorithm of Experiment

Нелинейное прогнозирование

Первым рассмотренным методом было нелинейное прогнозирование. Суть его заключается в локальной аппроксимации ближайших соседей в фазовом пространстве [5–7].

Приведем общий вид прогноза в случае использования данного метода:

$$x(t+T) = \frac{1}{|U_m|} \sum_{x(t') \in U_m} x(t'+T), \quad (1)$$

где U_m – окрестность $x(t)$, кроме самой $x(t)$.

Данный алгоритм реализован в утилите lzo-gun, lzo-test уже упомянутого пакета TISEAN. Первым шагом необходимо определить минимальную фрактальную размерность фазового пространства m . Метод определения минимальной достаточной размерности вложения m был предложен Кернеллом и назван методом ложных ближайших соседей [8]. Для каждой точки x_i в m -мерном фазовом пространстве ищется ближайший сосед x_j , вычисляется расстояние $\|x_j - x_i\|$.

Далее итерируем систему и вычисляем:

$$R_i = \frac{|x_{i+1} - x_{j+1}|}{|x_i - x_j|}. \quad (2)$$

Если R_i превышает некоторый заданный порог, то данная точка помечается как имеющая ложного соседа. Такой алгоритм реализован в утилите false-nearest пакета TISEAN, результатом которой будет зависимость доли ложных соседей в зависимости от размерности вложения для данной системы.

Критерием того, что фрактальная размерность выбрана достаточно большой, является нулевая или малая доля ложных соседей. В нашем случае после применения утилиты false-nearest получим размерность фазового пространства равной 23 ($m = 23$), исходя из 10-процентного порога.

Для расчета задержки d использовалась утилита lzo-test. Утилита оценивает ошибки одношагового прогнозирования модели на многомерном временном ряду. Приведенные ошибки прогноза нормируются к стандартным отклонениям каждого компонента. С помощью этой утилиты подбиралась пара m, d , при которой погрешность прогноза оказывалась наименьшей. Фрактальная размерность была определена на предыдущем шаге исследования ($m = 23$). Расчет позволил определить задержку, равной 2 ($d = 2$), т. к. она дает наилучший результат на начальных шагах прогнозирования, пока величина погрешности еще остается приемлемой.

Последняя величина, которую следует определить, это размер окрестности, в которой будет происходить поиск ближайших соседей в процессе прогнозирования. Как показывает практика, она не сильно влияет на качество прогноза, тем не менее, в данной работе не будем пренебрегать ей и получим более точное значение. Идея ее вычисления крайне проста. Необходимо получить нормированную погрешность прогноза, как функцию радиуса окрестности. Такой алгоритм реализован утилитой lzo-gm. Для расчетов будем пользоваться парой значений, которые были определены на предыдущем шаге: $m = 23, d = 2$.

Рассчитанный размер окрестности $r = 3,438$ получен для пары параметров $m = 23, d = 2$. Перейдем непосредственно к процессу прогнозирования. В качестве элементарного шага временного ряда был

выбран $\Delta t = 0,5$ ч. Интервал прогнозирования выбран равным одной неделе. Таким образом, количество шагов прогнозирования составит:

$$L = \frac{7 * 24}{0,5} = 336. \quad (3)$$

В качестве исходных данных используем временной ряд, сформированный ранее на основе данных о входящих вызовах КЦ за май. Прогноз будет осуществляться на основе первых трех недель, последняя неделя будет сравниваться с полученным прогнозом и использоваться как эталонная для вычисления погрешности прогноза (рисунок 2).

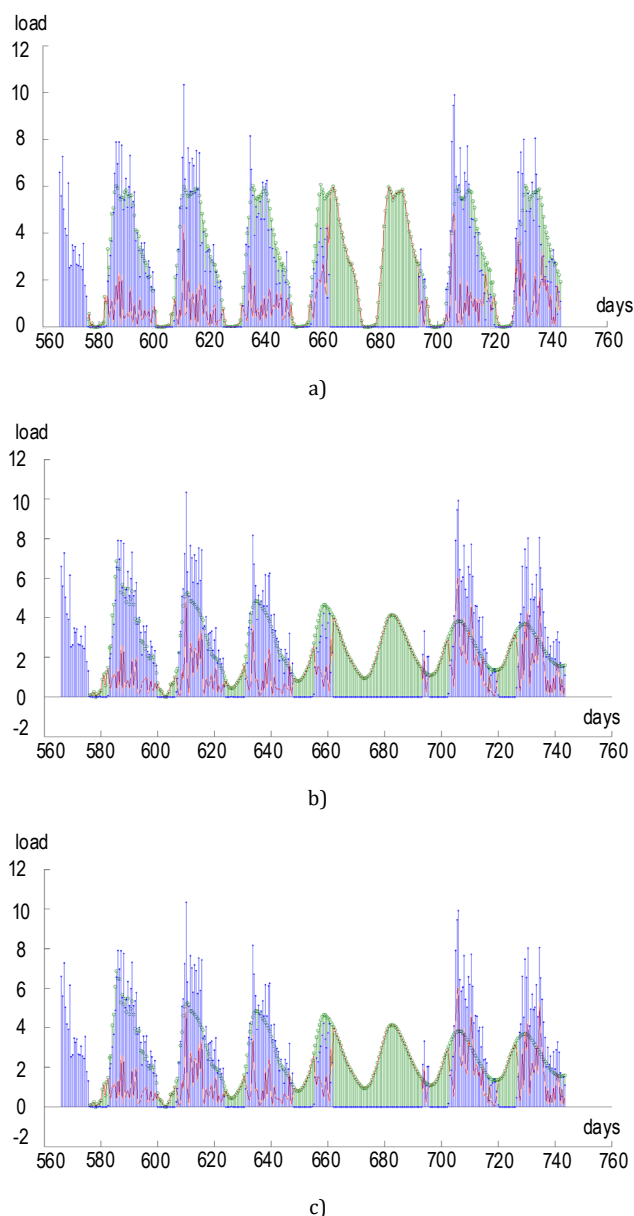


Рис. 2. Результат прогноза с использованием методов: нелинейного прогнозирования (а), линейного прогнозирования (б) и глобальной полиномиальной аппроксимации (с)

Fig. 2. Forecast Using the Nonlinear Forecasting Method (a), the Linear Forecasting Method (b) and the Global Polynomial Approximation Method (c)

На графике показан участок оригинального временного ряда (синий график с маркером точка) – реальные значения последней недели мая. Зеленый график с маркером круг – полученный прогноз. Красная кривая является абсолютной погрешностью прогноза, представляющая собой абсолютное значение разности спрогнозированных значений и реальных значений на протяжении всего недельного интервала прогнозирования. Прогнозирование выполняем методом нелинейного прогнозирования, реализованным в утилите lzo-run (см. рисунок 2а).

Линейное прогнозирование

Отличие метода локального линейного прогнозирования от предыдущего заключается в использовании локальной линейной аппроксимации для получения прогноза [6, 7]. Прогноз в данном случае рассчитывается следующим образом:

$$x(t + T) = a_n x(t) + b_n, \quad (4)$$

где a_n, b_n находятся из выражения:

$$\sum_{x(t') \in U_m} (x(t' + T) - a_n x(t') - b_n)^2 \rightarrow \min. \quad (5)$$

Используем аналогичный алгоритм для значений размерности и задержки, соответственно, 30 и 2 ($m = 30, d = 2$), а размер окрестности r равным 6.

Как и в предыдущем эксперименте, количество шагов прогнозирования равно $L = 336$ (при шаге временного ряда $\Delta t = 0,5$ ч и интервале прогнозирования в 1 неделю). Прогноз (см. рисунок 2б) осуществлен на основе данных первых трех недель мая (как и в прошлом эксперименте). Последнюю неделю используем в качестве реальных значений для вычисления погрешности.

Метод глобальной полиномиальной аппроксимации

Рассмотрим метод глобальной полиномиальной аппроксимации [9]. При его использовании предиктор F_T ищется из условия:

$$\sum_t (x(t' + T) - f_T(x(t')))^2 \rightarrow \min. \quad (6)$$

Для этого метода применим пару параметров размерности и задержки, соответственно, 16 и 3 ($m = 16, d = 3$), дающих наименьшие значения погрешности прогнозирования.

Следует отметить, что несмотря на то, что параметры для методов прогнозирования подбирались с опорой на один и тот же временной ряд, они варьируются в зависимости от методов. Изменения этого параметра обусловлены желанием получить наименьшую погрешность прогноза.

Теперь получим прогноз (см. рисунок 2с), исходя из данных $m = 16, d = 3, L = 336$. Порядок полинома

выберем $p = 2$ (значение подбираем исходя из экспериментальных данных).

Теперь необходимо провести сравнение описанных выше методов. Для этого выберем меру сравнения – погрешность прогнозирования.

Сравнение полученных прогнозов

Для того, чтобы сравнить рассмотренные методы прогнозирования, вычислим нормированную погрешность прогнозирования для каждого метода по выражению:

$$\sigma_n(T, t) = \frac{|x_{pred}(t, T) - x(t + T)|}{\sigma_x}, \quad (7)$$

где $\sigma_x = \sqrt{\langle (x - \langle x \rangle)^2 \rangle}$ – среднеквадратичное отклонение от среднего значения ряда.

Анализ полученных результатов показал, что наиболее эффективно со своей задачей справился метод простого нелинейного прогнозирования, показав лучшие результаты как на краткосрочный прогноз (около суток), так и среднесрочный (неделя). Остальные методы давали приемлемый результат примерно на суточном интервале. Причем

метод локального линейного прогнозирования оказался ожидаемо точнее полиномиальной аппроксимации.

Выводы

Воспользовавшись одним из критериев хаотичности (неотрицательный показатель Ляпунова), на реальных данных о вызовах была доказана хаотичная природа процесса поступления нагрузки на КЦ. Таким образом, с уверенностью можно говорить о возможности использования теории хаоса для прогнозирования данного процесса.

Лучше всего для этой цели по результатам эксперимента подходит метод нелинейного прогнозирования. Метод локального линейного прогнозирования дал неплохие результаты только для краткосрочного предсказания (одни сутки). Метод глобальной аппроксимации позволяет получить удовлетворительный результат только при краткосрочном прогнозе. Таким образом, для прогнозирования нагрузки КЦ целесообразно выбрать метод нелинейного прогнозирования.

Список используемых источников

1. Общее описание системы WFMCC. URL: <https://argus-wfmcc.ru> (дата обращения 25.05.2021)
2. Мун Ф. Хаотические колебания: Вводный курс для научных работников и инженеров. Пер. с англ. М.: Мир, 1990. 312 с.
3. Contopoulos G. Highlights of chaos research. 2018. URL: <https://arxiv.org/pdf/1807.09492.pdf> (дата обращения 07.04.2019)
4. Moon F.C. Chaotic and Fractal Dynamics: Introduction for Applied Scientists and Engineers. New York: Wiley, 1992.
5. Farmer D.J., Sidorowich J.J. Predicting chaotic time series // Physical Review Letters. 1987. Vol. 59. Iss. 8. PP. 845–848.
6. Лоскутов А. Математические основы хаотических динамических систем: курс лекций. М.: МГУ, 2007. 78 с.
7. Лоскутов А.Ю., Михайлов А.С. Основы теории сложных систем. М.: Ижевск: Институт компьютерных исследований, 2007. 620 с.
8. Oestreicher C. A history of chaos theory // Dialogues Clin Neurosci. 2007. Vol. 9. № 3. PP. 279–289. DOI:10.31887/DCNS.2007.9.3/coestreicher
9. Федоров А.Я., Мелентьева Т.А., Мелентьева М.А. Стохастическая динамика систем // Фундаментальные исследования. 2008. № 2. С. 112–113. URL: <http://fundamental-research.ru/ru/article/view?id=2659> (дата обращения 31.05.2019)

* * *

Chaos Theory Methods for Contact-Center Dinamic Control

A. Goldstein^{1, 2} , S. Kislyakov^{1, 2} , M. Fenomenov^{1, 2} 

¹The Bonch-Bruевич Saint-Petersburg State University of Telecommunications,
St. Petersburg, 193232, Russian Federation

²Scientific and Technical Center "ARGUS",
St. Petersburg, 197198, Russian Federation

Article info

DOI:10.31854/1813-324X-2021-7-2-18-23

Received 17th March 2021

Accepted 21st April 2021

For citation: Goldstein A., Kislyakov S., Fenomenov M. Chaos Theory Methods for Contact-Center Dynamic Control. *Proc. of Telecom. Universities*. 2021;7(2):18–23. (in Russ.) DOI:10.31854/1813-324X-2021-7-2-18-23

Abstract: The work is devoted to searching for optimal control methods for contact center, in particular, methods for predicting the load for further calculation of required number of operators. If the number of operators is always more than required, then the owners of the contact center will incur financial losses. If there are too few employees, the quality of service will decline. Predicting the load of the contact center is required in order to bring the optimal number of operators to work in advance. It is proposed to apply chaos theory to predict the incoming load of a contact center. Positive value of the Lyapunov index indicates the chaotic behavior of the input flow of the load. To predict the load, the methods of linear and nonlinear forecasting and the method of global approximation are used. The paper presents the results of comparing these methods for the problem of predicting the incoming load of contact center.

Keywords: contact center, automation, chaos theory, WFM.

References

1. General Description of the System WFMCC. Available from: <https://argus-wfmcc.ru> [Accessed 25th May 2021]
2. Moon F. *Chaotic Oscillations*. Translate from English. Moscow: Mir Publ.; 1990. 312 p. (in Russ.)
3. Contopoulos G. *Highlights of chaos research*. 2018. Available from: <https://arxiv.org/pdf/1807.09492.pdf> [Accessed 7th April 2019]
4. Moon F.C. *Chaotic and Fractal Dynamics: Introduction for Applied Scientists and Engineers*. New York: Wiley; 1992.
5. Farmer D.J., Sidorowich J.J. Predicting chaotic time series. *Physical Review Letters*. 1987;59(8):845–848.
6. Loskutov A. Mathematical Foundations of Chaotic Dynamical Systems: a Course of Lectures. Moscow: Lomonosov Moscow State University Publ.; 2007. 78 p. (in Russ.)
7. Loskutov A.Yu., Mikhailov A.S. Fundamentals of the Theory of Complex Systems. Izhevsk: Institute for Computer Research Publ.; 2007. 620 p. (in Russ.)
8. Oestreicher C. A history of chaos theory. *Dialogues Clin Neurosci*. 2007;9(3):279–289. DOI:10.31887/DCNS.2007.9.3/coestreicher
9. Fedorov A.Ya., Melentjeva T.A., Melentjeva M.A. Stochastic dynamics of systems. *Fundamental research*. 2008;2:112–113. Available from: <http://fundamental-research.ru/ru/article/view?id=2659> [Accessed 31th May 2019]

Сведения об авторах:

ГОЛЬДШТЕЙН
Александр Борисович

кандидат технических наук, доцент кафедры инфокоммуникационных систем Санкт-Петербургского государственного университета телекоммуникаций им. проф. М.А. Бонч-Бруевича, директор НТЦ АРГУС,
agold@argustelecom.ru
 <https://orcid.org/0000-0002-4136-4703>

КИСЛЯКОВ
Сергей Викторович

кандидат технических наук, доцент кафедры инфокоммуникационных систем Санкт-Петербургского государственного университета телекоммуникаций им. проф. М.А. Бонч-Бруевича, бизнес-аналитик НТЦ АРГУС,
s.v.kislyakov@gmail.com
 <https://orcid.org/0000-0001-8842-7903>

ФЕНОМЕНОВ
Михаил Александрович

аспирант кафедры инфокоммуникационных систем Санкт-Петербургского государственного университета телекоммуникаций им. проф. М.А. Бонч-Бруевича, заместитель директора НТЦ АРГУС,
m.fenomenov@argustelecom.ru
 <https://orcid.org/0000-0003-4918-8481>

О применении программы Radio Mobile для расчета зон покрытия передатчиков, работающих в диапазоне ОВЧ

Ю.А. Ковалгин^{1*}

¹Санкт-Петербургский государственный университет телекоммуникаций им. проф. М.А. Бонч-Бруевича, Санкт-Петербург, 193232, Российская Федерация

*Адрес для переписки: kowalgin@sut.ru

Информация о статье

Поступила в редакцию 24.03.2021

Принята к публикации 24.05.2021

Ссылка для цитирования: Ковалгин Ю.А. О применении программы Radio Mobile для расчета зон покрытия передатчиков, работающих в диапазоне ОВЧ // Труды учебных заведений связи. 2021. Т. 7. № 2. С. 24–36. DOI:10.31854/1813-324X-2021-7-2-24-36

Аннотация: Рассмотрена особенность применения программы Radio Mobile для расчета зон покрытия аналоговых ЧМ- и цифровых DRM-передатчиков; показано, при каких условиях вводимых первичных данных обеспечивается получение достоверных результатов; установлено соотношение уровней обоих передатчиков, при котором обеспечивается равенство их зон покрытия.

Ключевые слова: цифровое вещание, зона покрытия ОВЧ-передатчика.

Вводная часть

Используя первичные данные рекомендаций МСЭ-R таких, как рекомендация МСЭ-R BS.1660-8. «Техническая основа для планирования наземного цифрового звукового радиовещания в полосе ОВЧ» [1] и рекомендация МСЭ-R P.1546-6. «Метод прогнозирования для трасс связи "пункта с зоной" для наземных служб в диапазоне частот от 30 МГц до 4000 МГц» [2] и ряда других, мы можем разработать методику расчета зоны покрытия цифровых DRM-передатчиков, работающих в диапазоне ОВЧ. Подход к ее разработке, последовательность процедур, необходимых для расчета, подробно изложены в [1–5]. Общий подход к расчету зон покрытия аналоговых и цифровых передатчиков представлен в [6]. Анализ представленного материала свидетельствует о том, что результаты вычислений зависят от множества факторов, тщательный учет которых достаточно трудоемок, требует выполнения очень большого числа рутинных операций. В частности, необходим учет таких факторов, как выходная мощность передатчика, чувствительность приемника, несущая частота передатчика, усиление и диаграммы направленности передающей и приемной антенн, потери в фидере, затухание радиоволн при распространении в воздушной среде, профили трасс с высотными отметками для выбранных радиальных направлений. Кроме учета сведений, представленных в указанных рекомендациях, расчет предпо-

лагает также и нанесение полученных зон покрытия на карты местности, что неизбежно сопровождается применением дополнительных программных продуктов с детальной проработкой всех особенностей рельефа местности для каждого из выбранных радиальных направлений, иначе говоря точного учета профилей трасс с их высотными отметками. А это, в свою очередь, требует выполнения дополнительного пласта вычислений, необходимых для учета распределения напряженности поля передатчика по покрываемой территории. С этой точки зрения автоматизация данных вычислений полезна для практики, востребована жизнью, что подтверждается наличием соответствующих программных пакетов. Однако до сих пор в публикациях отсутствуют сведения о достоверности получаемых с их помощью результатов.

Целью настоящей работы и является решение данной задачи.

Основная часть

Для расчета зоны покрытия прежде всего необходимо знание минимального уровня напряженности поля передатчика в точке приема. Обычно он выполняется для высоты приемной антенны 10 м над уровнем Земли для $T = 50 \%$ времени и $L = 50 \%$ мест охвата в зоне покрытия и затем корректируется с учетом условий приема.

Рекомендация МСЭ-R BS.1660-8 [1] определяет следующие условия приема:

- FX (аббр. от англ. Fixed Reception) – стационарный прием, высота расположения приемной антенны составляет 10 м над уровнем земли;
- PO (аббр. от англ. Portable Outdoor) – прием вне помещения на переносные устройства;
- PO-H (от англ. Handheld Portable Outdoor) – прием вне помещения на портативные переносные устройства;
- PI (аббр. от англ. Portable Indoor) – прием в помещении на переносные устройства;
- PI-H (от англ. Handheld Portable Indoor) – прием в помещении на портативные переносные устройства.

Условия PO и PI моделируют прием в пригородной зоне с хорошими условиями и приемником с всенаправленной антенной; условия PO-H и PI-H моделируют ситуацию приема в городской среде с плохими условиями приема и с приемником, оборудованным внешней антенной. Прием в помещении предполагает, что приемная антенна расположена на высоте не менее 1,5 м над уровнем пола в помещениях первого этажа, имеющих окно. Прием вне помещения также предполагает, что антенна расположена на высоте 1,5 м над уровнем земли.

Рекомендация МСЭ-R BS.1660-8 [1] расчет минимальной напряженности поля в точке приема требует вести последовательно по приведенным ниже формулам.

Входная мощность шума приемника $P_{ш.пр}$:

$$P_{ш.пр} = F + 10 \log_{10}(k \cdot T_0 \cdot B), \text{ дБВт},$$

где: $F = 7$ дБ – коэффициент шума приемника (дБ); k – постоянная Больцмана ($k = 1,38 \cdot 10^{-23}$ Дж/К); $T_0 = 290$ К – абсолютная температура; B – ширина полосы шума приемника (Гц).

Минимальный уровень мощности полезного сигнала на входе приемника $P_{s,min}$:

$$P_{s,min} = (C/N)_{min} + P_{ш.пр}, \text{ дБ(Вт)},$$

где $(C/N)_{min}$ – минимальное отношение сигнал/шум:

- для DRM-тракта – это значение на входе декодера DRM-приемника (дБ), необходимое для получения битовой ошибки на выходе декодера не более чем $BER = 10^{-4}$, зависит от вида модуляции и скорости кода, т. е. от помехозащищенности цифрового потока;

- для ЧМ-тракта – это значение, необходимое для получения отношения сигнал/шум в канале звука ЧМ-приемника не менее 60 дБ.

Минимальная плотность потока мощности (величина вектора Пойнтинга) в месте приема φ_{min} :

$$\varphi_{min} = P_{s,min} - A_a + L_f, \text{ дБ(Вт/м}^2\text{)},$$

где L_f – потери в фидере приемника (дБ) не зависят от вида модуляции и скорости кода, зависят от выбранного значения несущей частоты f ; A_a – эффективная апертура антенны (дБм²):

$$A_a = 10 \log_{10}[(1,64/4\pi) \cdot (300/f)^2] + G_0, \text{ дБм}^2,$$

где f – центральная (несущая) частота радиоканала (МГц); G_0 – усиление антенны относительно полуволнового вибратора (дБм); оба параметра зависят от частоты, но не зависят от вида модуляции и скорости кода.

Требуемый минимальный уровень напряженности поля в месте расположения приемной антенны E_{min} :

$$E_{min} [\text{дБ(мкВ/м)}] = \varphi_{min} [\text{дБВт/м}^2] + 10 \log_{10} Z_{\rho} [\text{дБОм}] + 20 \log_{10} (1\text{В}/1\text{мкВ}),$$

где $Z_{\rho} \approx 120\pi$ [Ом] – волновое сопротивление в свободном пространстве; в итоге имеем:

$$E_{min} [\text{дБ(мкВ/м)}] = \varphi_{min} [\text{дБВт/м}^2] + 145,8 [\text{дБОм}];$$

после подстановки имеем:

$$E_{min} [\text{дБ(мкВ/м)}] = P_{s,min} - 10 \log_{10}[(1,64/4\pi) \cdot (300/f)^2] - G_0 + L_f + 145,8, \text{ дБ(мкВ/м)}.$$

Для аналогового тракта полоса частот радиоканала $\Delta f_{рк}$ при частотной модуляции определяется, как известно, выражением $\Delta f_{рк} = 2(F_v + \Delta f_d)$, где F_v – верхняя модулирующая частота; Δf_d – девиация несущей частоты. Для системы с пилот-тоном верхняя модулирующая частота комплексного стереосигнала составляет 53 кГц, девиация частоты в нашей стране принята равной ± 50 кГц, что дает нам значение $\Delta f_{рк} = 206$ кГц. С учетом этого значение входной мощности шума ЧМ-приемника $P_{ш.пр}$ соответствует $-143,9$ дБВт. Для системы DRM это значение равно $\Delta f_{рк} = 96$ кГц [4, 5], что дает нам значение $P_{ш.пр} = -146,98$ дБВт. В обоих случаях входная мощность шума приемника не зависит от характеристик передатчика, определяется только коэффициентом шума приемника и полосой частот радиоканала.

Кроме того, для одного и того же значения несущей частоты усиление приемной антенны G_0 и потери в фидере L_f не зависят от вида модуляции, определяются только условиями приема (FX, PI, PI-H, PO, PO-H), одинаковы для аналогового и цифрового трактов приема.

Из опубликованных данных следует, что современные ЧМ-приемники обеспечивают отношение сигнал/шум в канале звука (после частотного детектора и стереодекодера) не менее 60 дБ при отношении сигнал/шум в точке приема не хуже 30 дБ.

Для аналоговых систем ЧМ-радиовещания частотное планирование ведется из условия фиксированного приема, FX: высота расположения приемной антенны принята равной 10 м над уровнем

Земли, число мест охвата в зоне покрытия составляет $L = 50 \%$. При этом в зоне покрытия требуемое минимальное медианное значение напряженности поля передатчика превышает также в $T = 50 \%$ времени. Для этих условий в соответствии с рекомендациями ITU-R BS.412-9 [7] и ITU-R BS.450-3 [8] на границе зоны покрытия независимо от условий приема напряженность поля при работе передатчика в стереорежиме должна быть для сельской местности не менее 54 дБ(мкВ/м). В этом случае все характеристики, заявленные в стандарте для данной системы, будут выполнены (ГОСТ Р. 51107-97 [9] и ГОСТ 51741-2001, [10]).

Для системы DRM (ETSI standard ES 201980 V3.2.1. (2012-6)), носящей пороговый характер,

минимальный уровень требуемой напряженности поля передатчика на границе зоны покрытия кроме перечисленных выше факторов зависит также и от вида модуляции поднесущих частот в OFDM-блоке и скорости кода, установленной для цифрового потока в канале пользователя (МСЭ-R BS.1660-8, [1]).

Результаты вычислений с использованием вышеприведенных формул и сведений рекомендаций ITU-R BS.412-9 [7] и МСЭ-R BS.1660-8 [1] дают нам значения требуемой минимальной напряженности поля на антенне приемника $E_{\min}$, представленные в таблице 1.

ТАБЛИЦА 1. Требуемый минимальный медианный уровень напряженности поля на антенне приемника для разных условий ЧМ- и DRM-приема, несущая частота передатчика 95,7 МГц (ITU-R BS.412-9, [7] и МСЭ-R BS.1660-8, [1])

TABLE 1. Required Minimal Median Level of the Field Strength on the Receiver's Antenna for Different FM- and DRM-Receiving Conditions, Main Transmitter Frequency of 95,7 MHz (ITU-R BS.412-9, [7] and ITU-R BS.1660-8, [1])

№	Наименование параметра		Условия приема					
			FX	PI	PI-H	PO	PO-H	MO
1.	Потери в фидере	L_c (дБ)	1,37	0,00	0,00	0,00	0,00	0,274
2.	Усиление антенны	G_D , дБ	0,00	-2,20	-19,4	-2,20	-19,4	-2,20
3.	Эффективный раскрыв антенны	A_a , (дБм) ²	1,081	-1,12	-18,319	-1,12	-18,319	-1,12
4.	Минимальная плотность потока мощности в месте приема	$\Phi_{\min}$ (дБВт/м ²)	-146,02	-138,92	-118,36	-138,92	-118,36	-140,50
5.	Минимальный уровень напряженности поля на принимающей антенне ЧМ-приемника	$E_{\min}$, дБ(мкВ/м)	54					
6.	Минимальный уровень напряженности поля на принимающей антенне DRM-приемника при модуляции QAM-4 для скорости кода:	$E_{\min}$, дБ(мкВ/м)						
	0,25		2,97	9,799	26,999	9,799	26,999	8,273
	0,33		3,41	10,239	27,439	10,239	27,439	8,713
	0,40		3,63	10,459	27,659	10,459	27,659	8,933
7.	Минимальный уровень напряженности поля на принимающей антенне DRM-приемника при модуляции QAM-16 для скорости кода:	$E_{\min}$, дБ(мкВ/м)						
	0,33		7,409	15,79	32,939	15,79	32,939	13,413
	0,411		8,989	17,37	34,519	17,37	34,519	14,993
	0,50		10,009	18,39	35,539	18,39	35,539	16,013
	0,625		11,169	19,55	36,699	19,55	36,699	17,29

Приведенные в рекомендации МСЭ-R BS.1660-8 [1] значения требуемой минимальной медианной напряженности поля на принимающей антенне DRM-приемника диапазона ОВЧ II (87,5...108 МГц) приведены для каждого вида модуляции поднесущих частот только для одного значения скорости кода $R = 0,33$ (модуляция QAM-4) и $R = 0,50$ (модуляция QAM-16). Напомним, что от скорости кода зависит значение только одного параметра $(C/N)_{\min}$. При этом минимальная медианная напряженность поля при выбранном виде моду-

ляции для другого значения скорости кода может быть найдена как:

$$E_{\min 2} = E_{\min 1} - (SNR_1 - SNR_2), \text{ дБ(мкВ/м)},$$

где $E_{\min 1}$ – минимальная медианная напряженность поля при выбранных значениях вида модуляции и скорости кода при требуемом стандартом отношении сигнал/шум равном SNR_1 , $E_{\min 2}$ – тоже самое, но для другой скорости кода и требуемым стандартом отношении сигнал/шум, обеспечивающие получение вероятности появления битовой ошибки, не превышающей 10^{-4} , как это и требует-

ся стандартом, приведены в таблице 2. Эти данные соответствуют стационарному приему на штыревую антенну в помещении на расстоянии 12 км от телебашни в Санкт-Петербурге, несущая частота DRM-передатчика 95,7 МГц [15]. Они получены экспериментальным путем и позволяют, используя приведенное выше выражение и данные таблицы 2, найти требуемое минимальное значение напряженности поля при DRM-приеме и для других значений скорости кода, не представленных в данной рекомендации.

Итак, приведенные в таблице 1 данные позволяют рассчитать радиус зоны покрытия передатчика при аналоговом ЧМ- и DRM-приеме, считая, что полученные значения E_{min} выполняются на границе зоны покрытия.

ТАБЛИЦА 2. Значения требуемого SNR для разных режимов модуляции и скорости кода в канале MSC

TABLE 2. Values of the Required SNR for Different Modulation Regimes and Coding Rates in MSC Channel

Модуляция поднесущих частот в канале пользователя	Скорость кода, R	Требуемое значение SNR, дБ
4-QAM	0,250 (PL0)	2,27
4-QAM	0,333 (PL1)	2,71
4-QAM	0,400 (PL2)	2,93
4-QAM	0,500 (PL3)	4,09
16-QAM	0,330 (PL0)	7,64
16-QAM	0,411 (PL1)	9,22
16-QAM	0,500 (PL2)	10,24
16-QAM	0,625 (PL3)	11,40

Кроме требуемого при приеме значения E_{min} , на размер зоны покрытия передатчиков существенно влияют: электрические характеристики Земли; значение несущей частоты; условия приема; изменения напряженности поля передатчика, вызванные климатическими факторами, особенностями распространения и условиями приема радиоволн, рельефом местности, наличием различных препятствий на пути распространения радиоволн; уровень промышленных и атмосферных помех; шумовые характеристики и чувствительность приемника; эффективная высота подвеса передающей и приемной антенн; параметры антенно-фидерных устройств.

Учет столь большого множества факторов достаточно трудоемок, важно, что он в значительной степени опирается часто на эмпирические данные рекомендаций международного союза электросвязи ITU-R, требует при расчете зон покрытия проведения очень большого числа вычислений. При этом получаемый на основе этих данных результат всегда требует экспериментальной проверки, ибо при расчетах, как правило, используются усредненные данные, заимствованные из рекомендаций ITU-R P.372 [11], P.527-4 [12], P.528-3 [3],

P. 832-4 [13], P.1546-7 [2], BS.1660-8 [1], BS.412-9 [7] и BS.450-3 [8].

Потери при распространении радиоволн зависят, прежде всего, от диэлектрической проницаемости среды ϵ и проводимости почвы σ , значения которых ко всему прочему являются функцией частоты f , состава и структуры почвы, особенностей климата конкретного региона. Рекомендация ITU-R P.527-4 приводит следующие значения для диапазона ОВЧ: проводимость почвы σ (См/м) лежит в пределах от 0,001 (почва средней сухости) до 0,01 (влажная почва); диэлектрическая проницаемость среды ϵ составляет соответственно 15 (почва средней сухости) и 30 (влажная почва). Рекомендация ITU-R P.832-4 [13] дает нам значение $\epsilon = 17$ для района Санкт-Петербурга, правда, для частоты 1 МГц. Опираясь на эти данные и данные рекомендации ITU-R P.527-4 [12], примем, что для региона Санкт-Петербурга $\epsilon = 20$ и $\sigma = 0,005$ См/м без учета сезонных климатических изменений. Эти данные будем использовать при расчете зоны покрытия.

Особую трудность при расчете зоны покрытия вызывает необходимость учета рельефа местности для профиля трассы каждого радиального направления. Наличие в зоне покрытия различного рода индивидуальных особенностей на пути распространения радиоволн (искусственных и естественных препятствий, имеющейся растительности) оказывают существенное влияние на размер зоны покрытия, особенно на появление внутри нее так называемых малых зон, где прием сигнала вообще становится невозможным.

Применение специализированного программного обеспечения (ПО) при оценке зоны покрытия может не только повысить достоверность получаемых результатов по сравнению с ручным расчетом, но и позволяет существенно экономить время, необходимое для его выполнения и для последующего учета возможных изменений напряженности поля для разных условий приема, с целью выявления малых зон, где прием сигнала оказывается невозможным. Однако, эти преимущества достигаются только в случае, если при разработке ПО с особой тщательностью выполнен учет всего множества факторов, влияющих на особенности распространения радиоволн в диапазоне ОВЧ, что имеет место далеко не всегда.

В настоящее время при оценке зоны покрытия наибольшее распространение в нашей стране получили две программы: Radio Planner v.2.1 (для коммерческого применения) и Radio Mobile v.11.6.6, распространяемая на бесплатной основе. Обе они, как это следует из доступных источников, учитывают все множество факторов, влияющих на особенности распространения радиоволн в диапазоне ОВЧ, а именно: значение несущей частоты радиоканала, высоту подвеса передающей и при-

емной антенн, особенности рельефа местности и растительного покрова для каждого из радиальных направлений, электрические характеристики Земли, влияние второго луча при приеме, индивидуальные особенности аналоговых и цифровых приемопередающих трактов. В основе учета каждого из перечисленных факторов лежат сведения, имеющиеся в соответствующих документах ITU-R. Однако полноценных данных, подтверждающих достоверность получаемых с помощью этих программ результатов, в публикациях нет.

Для детального исследования зоны покрытия передатчиков компанией ООО «Дигитон Системс» был разработан и введен в эксплуатацию экспериментальный аналого-цифровой тракт в формате ЧМ/DRM для обслуживания Санкт-Петербурга и прилегающей к нему территории. Несущая частота ЧМ-передатчика составила 95,9 МГц и мощность 3000 Вт; для цифрового DRM-передатчика эти значения, соответственно, оставили 95,7 МГц и 640 Вт. Передающий тракт аналогового радиовещания включал стереокодер системы с пилот-тоном, ЧМ-модулятор и радиочастотный усилитель передатчика «Полюс ПТ». Передающий тракт цифрового DRM-радиовещания включал контент-сервер компании Fraunhofer, DRM-модулятор компании Rfmondial и радиочастотный усилитель передатчика «Полюс ПТ». Предыскажение в канале ЦРВ системы DRM соответствовало рекомендации 651 МККР, для компрессии цифровых аудиоданных использовался кодек MPEG-4 xHE-AAC. Оба передающих тракта работали на общее антенно-фидерное устройство и использовали общий модернизированный радиочастотный усилитель. Высота подвеса передающей антенны относительно Земли составила 278 м, усиление передающей антенны равно 6 дБм, потери в фидере составили 1,575 дБм.

Все характеристики выбранного оборудования передающих трактов ЧМ- и DRM-радиовещания удовлетворяли требованиям соответствующих стандартов [9, 10, 14], что было подтверждено результатами проведенных измерений.

Для натурных измерений в опытной зоне аналого-цифрового радиовещания использовался модернизированный профессиональный контрольно-измерительный приемник компании Rfmondial (Model RF-SE19) с питанием от бортовой сети автомобиля с четверть волновой калиброванной антенной типа Kathrein K 51 16 4/510 351 с антенным фактором $K = 7,69$ дБ(1/м). Ослабление сигнала в кабеле, соединяющем антенну и высокочастотный вход приемника, было измерено анализатором ZVL Network Analyzer компании «Rohde & Schwarz» и составило 0,49 дБ на частоте 96 МГц.

При проведении измерений были учтены требования рекомендаций ITU-R. Все измерения были выполнены для условия мобильного приема. Вы-

сота приемной антенны составляла 1,5 м над уровнем Земли.

В общей сложности были получены данные изменения напряженности поля и отношения сигнал/шум для каждого из передатчиков с изменением расстояния до телебашни, а также установлены границы зон покрытия с учетом требований рекомендаций МСЭ-R BS.1660-8 [1], ITU-R BS.450-3 [8] и ITU-R BS.412-9 [7]. Карта маршрутов перемещения автомобиля представлена на рисунке 1.

Заметим также, что Санкт-Петербург находится в низине, по мере удаления от места расположения передающей антенны наблюдается подъем уровня суши, возрастающий с удалением от передатчика. Последнее сопровождается увеличением радиуса оптической видимости, а значит и возможного максимального значения радиуса зоны покрытия. Основание телебашни расположено на высоте 10,5 м над уровнем моря, подъем суши в зоне покрытия по мере удаления от телебашни изменяется для разных радиальных направлений в интервале 35...75 м, составляя в среднем в пределах зоны покрытия около 50...60 м. Результаты оценки радиуса зоны покрытия представлены в таблице 3 для мощности ЧМ-передатчика 3000 Вт и DRM-передатчика 640 Вт при высоте подвеса передающей антенны относительно Земли 278 м, усилении антенны 6 дБм и потерях в фидере АФУ 1,575 дБм.

ТАБЛИЦА 3. Радиус зоны покрытия для условий мобильного приема

TABLE 3. Radius of the Coverage Zone for Mobile Receiving Conditions

Тип приемопередающего тракта системы радиовещания	Модуляция: – направление	Радиус зоны покрытия, км
Система с пилот-тоном	Частотная	49...52
Система DRM	QAM-4, $R = 0,25$:	67
	– западное,	64
	– северное,	
	QAM-16, $R = 0,625$:	49
	– северное,	57
	– юго-западное,	54
	– юго-восточное,	64
	– восточное	

Достоверность этих данных подтверждена многочисленными измерениями и может служить основой при расчете зон покрытия передатчиков с помощью ПО Radio Mobile.

Прежде всего, в программе Radio Mobile необходимо выбрать место расположения передающей станции и точек приема, расположенных в разных радиальных направлениях. Для открытия карты местности в программе нужно нажать клавишу F8 (рисунок 2), в открывшемся окне, нажав клавишу «Ввести широту, долготу», ввести координаты телебашни Санкт-Петербурга (широта 59°, 58', 35" и долгота 30°, 19', 14", установить размер извлекаемой карты в пикселях: 1900×1000 пикселей), после чего нажать в данном окне клавишу «Извлечь».

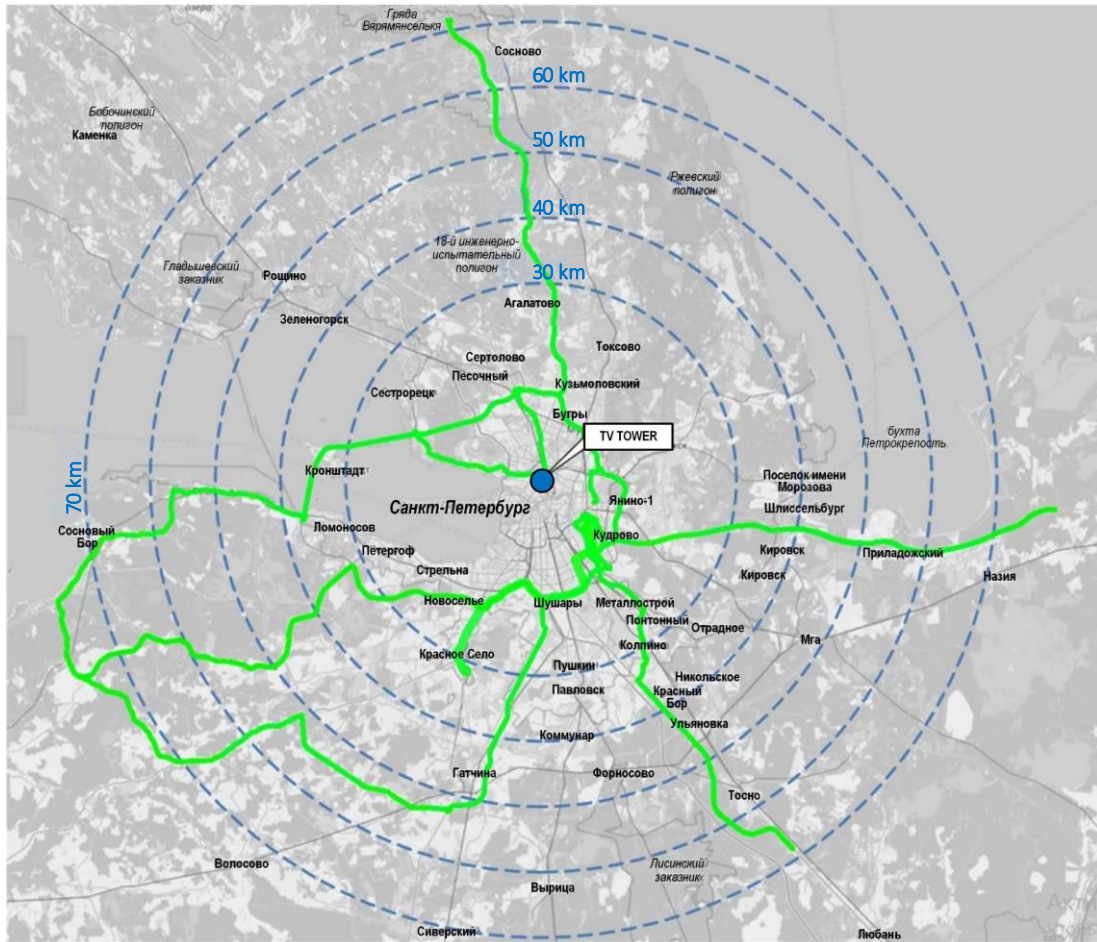


Рис. 1. Карта маршрутов при проведении измерений

Fig. 1. Travelling Maps for Carrying out the Measurements

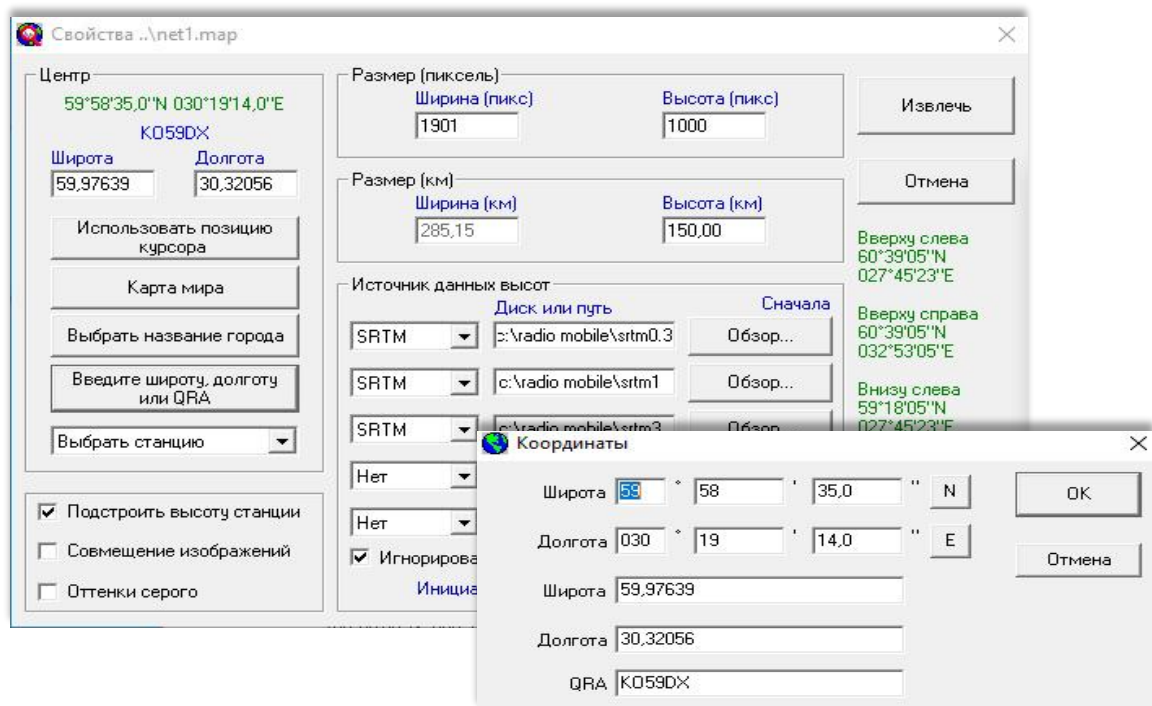


Рис. 2. Окно «Свойства» и вкладка для ввода координат расположения передатчика на карте местности

Fig. 2. Window "Properties" and the Inset for Entering Receiver Coordinates on the Landscape Map

Программа нарисует карту местности Санкт-Петербурга и прилегающей к нему территории. Для выбора радиальных направлений приема на полученной карте местности нужно курсор установить в позицию места приема, нажав на мышке левую кнопку, затем в меню «Файл» открыть окно «Свойства станций», в этом окне набрать имя станции (например, ЧМ-приемник_1), выбрать для нее пиктограмму, затем нажать клавишу «Поместить станцию в позицию курсора», после чего на карте местности появится выбранное нами название станции приема в выбранной нами точке и соответствующее ей радиальное направление. Аналогичные процедуры следует повторить, выбирая радиальные

направления и места расположения других точек приема. В общей сложности число выбранных радиальных направлений должно быть не менее пяти. После выполнения всех этих процедур на карте местности будут показаны все выбранные места приема и радиальные направления для каждой из них так, как это показано на рисунке 3.

Расчеты в программе Radio Mobile выполняются отдельно для каждого из выбранных направлений (например, представленных на рисунке 3), затем полученные данные могут быть обобщены (усреднены) для получения окончательных результатов.



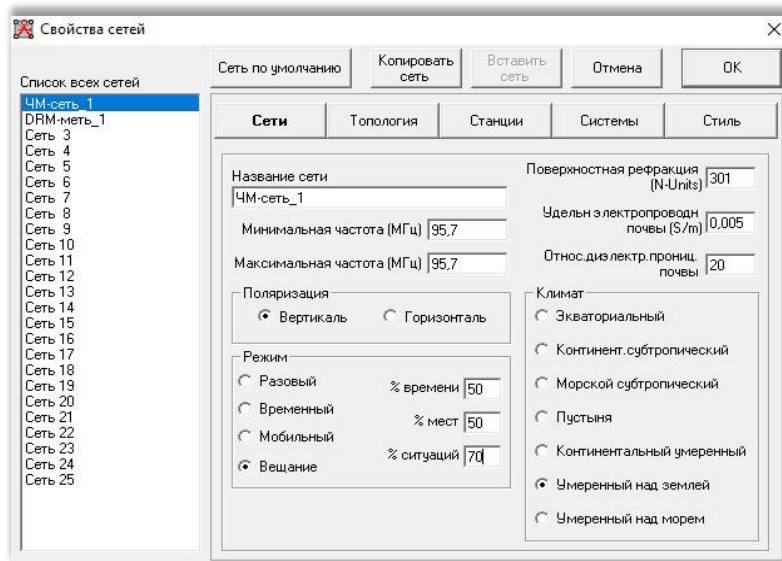
Рис. 3. Карта местности с показом расположения мест передающей и приемной станций для последующего расчета зон покрытия для выбранных радиальных направлений

Fig. 3. A Landscape Map Showing the Positions of the Transmitting and Receiving Stations in Order to Calculate Coverage Zones in the Radial Directions

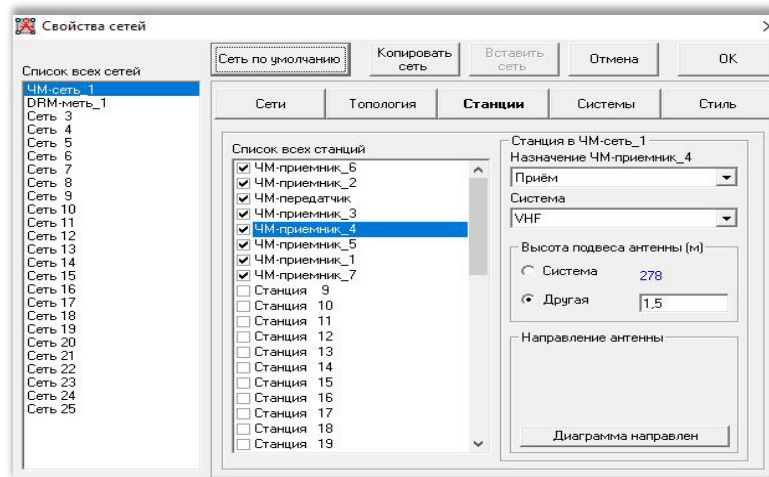
После выбора радиальных направлений необходимо ввести исходные данные для расчета зон покрытия отдельно для аналоговой и цифровой сетей. Для этого необходимо (рисунок 4а) в окне «Свойства сетей», нажать клавишу «Сети», в поле «Название сети» ввести при расчете зоны покрытия аналогового передатчика название данной сети, например, «ЧМ-сеть_1», а для цифровой сети радиовещания ввести соответственно «DRM-сеть_1». Далее в этом же окне нужно указать значение несущей частоты передатчика, например, 95,7 МГц, вид поляризации (в нашем случае – вертикальная), выбрать режим работы «Вещание», указать для исследуемого региона значения удельной проводимости почвы σ и диэлектрической проницаемости среды ϵ (для Санкт-Петербурга $\sigma = 0,005$ См/м, $\epsilon = 20$). Кроме того, в окне «Свойства сетей» следует указать также проценты времени T и мест L , в течение которых в зоне покрытия превышаются требуемые рекомен-

дацией ITU-R минимальные медианные значения напряженности поля передатчика: для аналоговой ЧМ-сети эти значения составляют 50 и 50 % [7], для цифровой DRM-сети [1] соответственно 70 и 99 % для фиксированного приема, FX; 99 и 95 % для условий приема PI, PI-H, PO, PO-H; 99 и 99 % для условий мобильного приема MO.

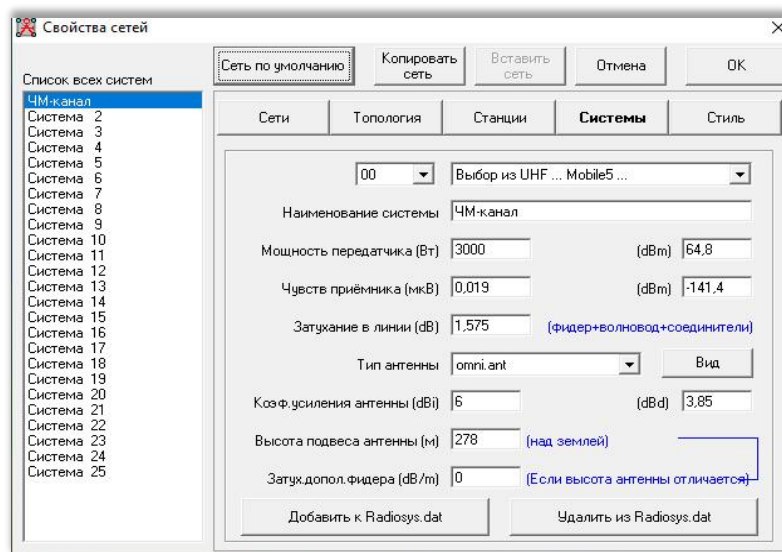
Затем следует поставить галочку в одной из позиций поля «Климат», в нашем случае в районе Санкт-Петербурга «Умеренный над землей». Что же касается такого параметра как «% ситуаций», то в рекомендациях ITU-R по частотному планированию такой показатель отсутствует, нет также соответствующих пояснений к его применению и в информации, поясняющей работу данной программы. Тем не менее изменение его величины существенно влияет на размер зоны покрытия. Ниже при проведении расчетов значение данного параметра будет уточнено.



a)



b)



c)

Рис. 4. Пример введенных исходных данных для расчета зоны покрытия ЧМ-передатчика в окне «Свойства сетей» на вкладках: а) «Сети», б) «Станции», в) «Системы»

Fig. 4. An Example of the Use of Input Data to Calculate Coverage Zones of an FM-Transmitter in Window "Network Properties" by Selecting: а) "Networks", б) "Stations", в) "Systems"

Затем в окне «Свойства сетей» следует нажать кнопку «Топология» и в открывшейся вкладке для аналоговой сети активировать точки «В зоне прямой видимости» и «Аналоговая сеть...». Для цифровой сети нужно в этом окне активировать показатель «Цифровая сеть. Звездная топология». После чего следует нажать кнопку «Станции» (рисунок 4b) и в открывшейся вкладке для аналоговой ЧМ-сети, выделяя галочкой соответствующие позиции, ввести названия станций приема, например, вместо «Станция 1» в этой позиции написать ЧМ-передатчик_1, ниже вместо «Станция 2» ввести «ЧМ-приемник_1», еще ниже «ЧМ-приемник_2» и так далее для всех станций. Для цифровой сети ввести, соответственно, «DRM-передатчик_1», «DRM-приемник_1», «DRM-приемник_2» и т. д. В правой части этой вкладки необходимо указать назначение каждой станции: для передающей станции это «Передача» (Мастер), для приемных станций, соответственно, – «Прием» (Слейв). Ниже в окошечке «Система» нужно ввести название канала, например, «ЧМ-канал» для аналоговой сети или «DRM-канал» для цифровой сети.

После чего необходимо перейти в раздел «Системы» и в открывшейся вкладке (рисунок 4c) для каждого вида исследуемого канала заполнить исходные данные: «Наименование системы» (ЧМ-или DRM-канал), «Мощность передатчика» (Вт), «Чувствительность приемника» (мкВ), «Затухание в линии» передающей антенны с учетом всех соединений (дБм), «Тип антенны» с учетом ее характеристик направленности (дБм), «Коэффициент усиления антенны» (дБм), «Высота подвеса антенны» относительно Земли (м), «Затухание дополнительного фидера» (дБм), если таковой имеется. После ввода этих данных следует нажать клавишу «Добавить...». После заполнения этих данных необходимо открыть вкладку «Стиль», выбрать здесь «Режим распространения», например, «Два луча» при прямой видимости, а также пределы изменения чувствительности приемника в дБ от введенного ранее значения. На этом ввод исходных данных для расчета зоны покрытия завершен. Для ее построения достаточно выбрать любое радиальное направление. Программа при расчете зоны покрытия учитывает особенности рельефа местности, наличие искусственных сооружений, особенности распределения растительного покрова в каждом из радиальных направлений.

Для оценки особенностей профиля трассы следует в основном окне программы Radio Mobile в меню «Инструменты» выбрать вкладку «Радиоканал» (рисунок 5a), далее в нижней ее части слева из числа возможных станций выбрать «ЧМ-передатчик» (или DRM-передатчик, если это цифровой канал), а справа – одну из возможных станций приема, например, ЧМ-приемник_1 (см. рисунок 5a). Затем следует в меню «Инструменты» открыть окно «Зона охвата», при построении зоны

покрытия в полярных координатах перейти в подменю «Зона охвата в полярных координатах» открыть соответствующую вкладку (рисунок 5b).

В окошечке «Радиус» указать минимальное и максимальное расстояния, затем выбрать цвет для отображения зоны покрытия и в левой нижней части этой вкладки указать пороговое значение напряженности поля на границе зоны покрытия, дБ(мкВ/м). Для аналоговой сети это значение равно 54 дБ(мкВ/м), для цифровой сети – значение минимальной напряженности поля, зависящее от вида модуляции поднесущих частот и скорости кода, приведено в таблице 1. Теперь следует нажать клавишу «Рисовать» и программа нарисует зону покрытия в выбранном нами цвете и другими деталями, выделенными в подразделе «График».

В качестве примера на рисунке 6 (сверху) представлена зона покрытия для аналогового ЧМ-передатчика (несущая частота 95,7 МГц, затухание в фидере 1,575 дБм, высота подвеса и коэффициент передачи ненаправленной передающей антенны, соответственно, 278 м и 6 дБм при $\sigma = 0,005$, $\varepsilon = 20$; шаг сетки равен 10 км), а на рисунке 6 (по середине и внизу – тоже самое, но для DRM-передатчика (при $T = 99\%$, $L = 99\%$, $C = 95\%$).

Для получения сетки расстояний от передатчика в меню «Правка» выбрать позицию «Создать кольца», в открывшейся вкладке выбрать позицию станции передачи (например, ЧМ-передатчик), единицу измерения (км), шаг сетки (например, 10 км), и ее цвет, после чего нажать кнопку «Рисовать». Программа нарисует кольцевую сетку с выбранным шагом.

Сравнение результатов расчетов в программе Radio Mobile и натурных измерений при мобильном приеме показало следующее:

- программа Radio Mobile обеспечивает получение достоверных результатов при расчете зон покрытия для аналоговых ЧМ-передатчиков при значении $C = 70\%$, а при расчете зон покрытия DRM-передатчиков при $C = 95\%$;

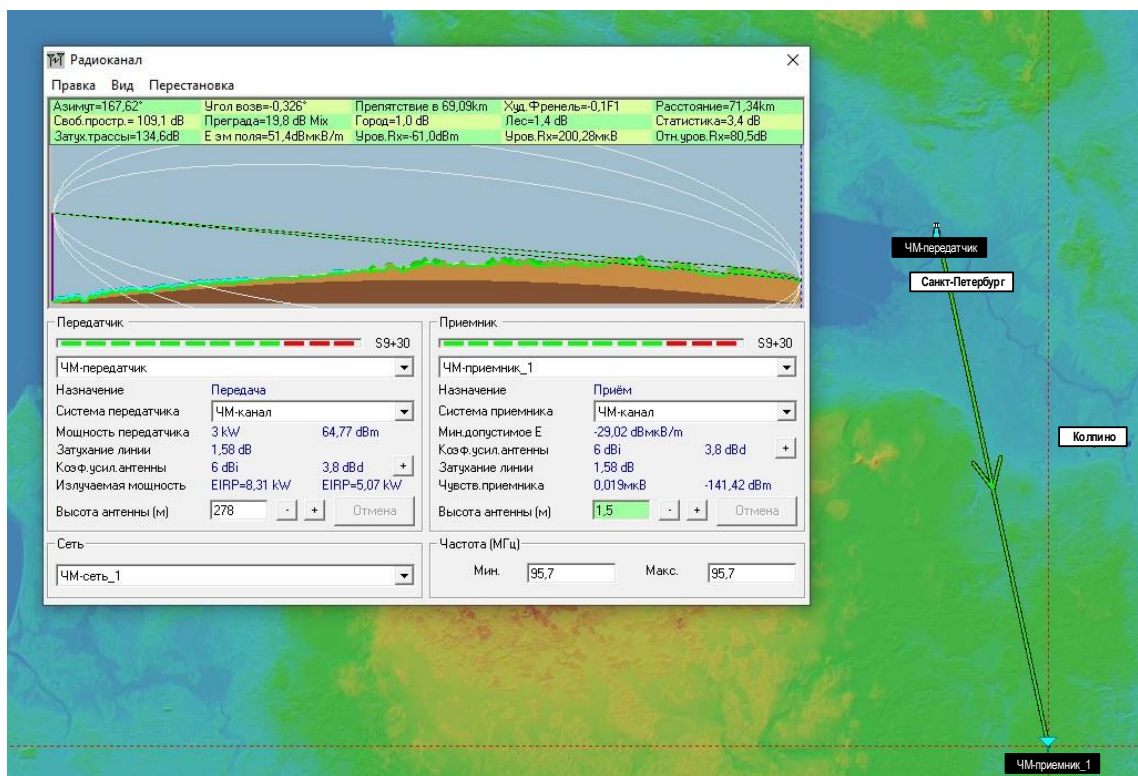
- равенство зон покрытия аналогового (ЧМ) и цифрового (DRM) передатчиков при мобильном приеме наблюдается при соотношении их уровней равном ~ 7 дБм при работе цифрового передатчика в режиме модуляции QAM-16 и скорости кода $R = 0,625$ (что соответствует наименьшей помехозащищенности DRM-сигнала; переход к модуляции QAM-4 и скорости кода $R = 0,25$ (режим наибольшей помехозащищенности) увеличивает радиус зоны покрытия цифрового DRM-передатчика в среднем на 10...15 км, что соответствует результатам натурных измерений (таблица 3); для других условий приема это соотношение будет другим, что следует из таблицы 1;

- программа Radio Mobile позволяет также для каждого радиального направления получить профиль трассы с целым рядом дополнительных све-

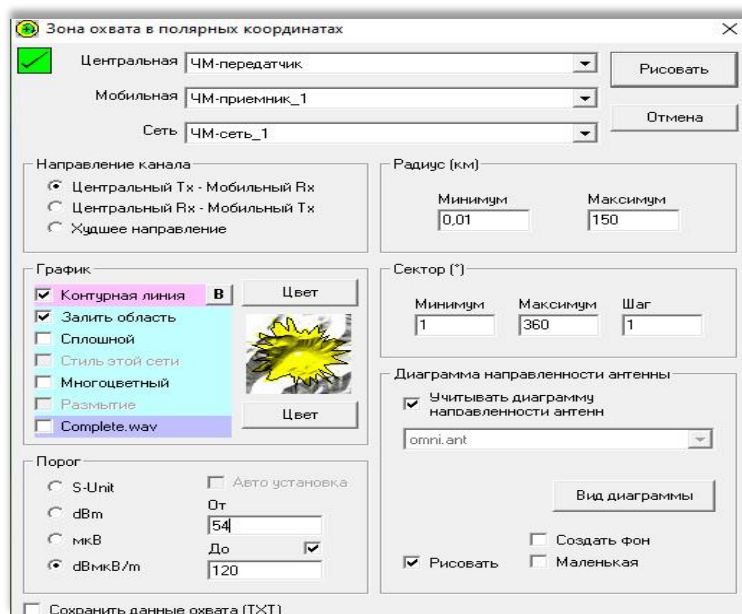
дений, включая и изменение напряженности поля передатчика с расстоянием, что полезно для выявления проблемных мест приема или так называемых проблемных зон приема внутри зоны покрытия.

Наличие порогового эффекта, свойственное любым системам цифрового радиовещания, является, пожалуй, наиболее существенным их недостатком,

требующим очень тщательного частотного планирования. Однако он с лихвой окупается существенным увеличением числа и качества передаваемых программ, возможностью передачи дополнительной информации, и, что самое важное, наличием существенного энергетического выигрыша и более экономным использованием ограниченного радиочастотного ресурса.



a)



b)

Рис. 5. Пример введенных данных для расчета зоны покрытия ЧМ-передатчика в меню «Инструменты» на вкладках: а) «Радиоканал»; б) «Зона охвата в полярных координатах»

Fig. 5. An Example of the Input Data for Calculating Coverage Zones of an FM-Transmitter in "Instruments" Menu by Selecting: а) "Radio Channel", б) "Coverage Zone in Polar Coordinates"

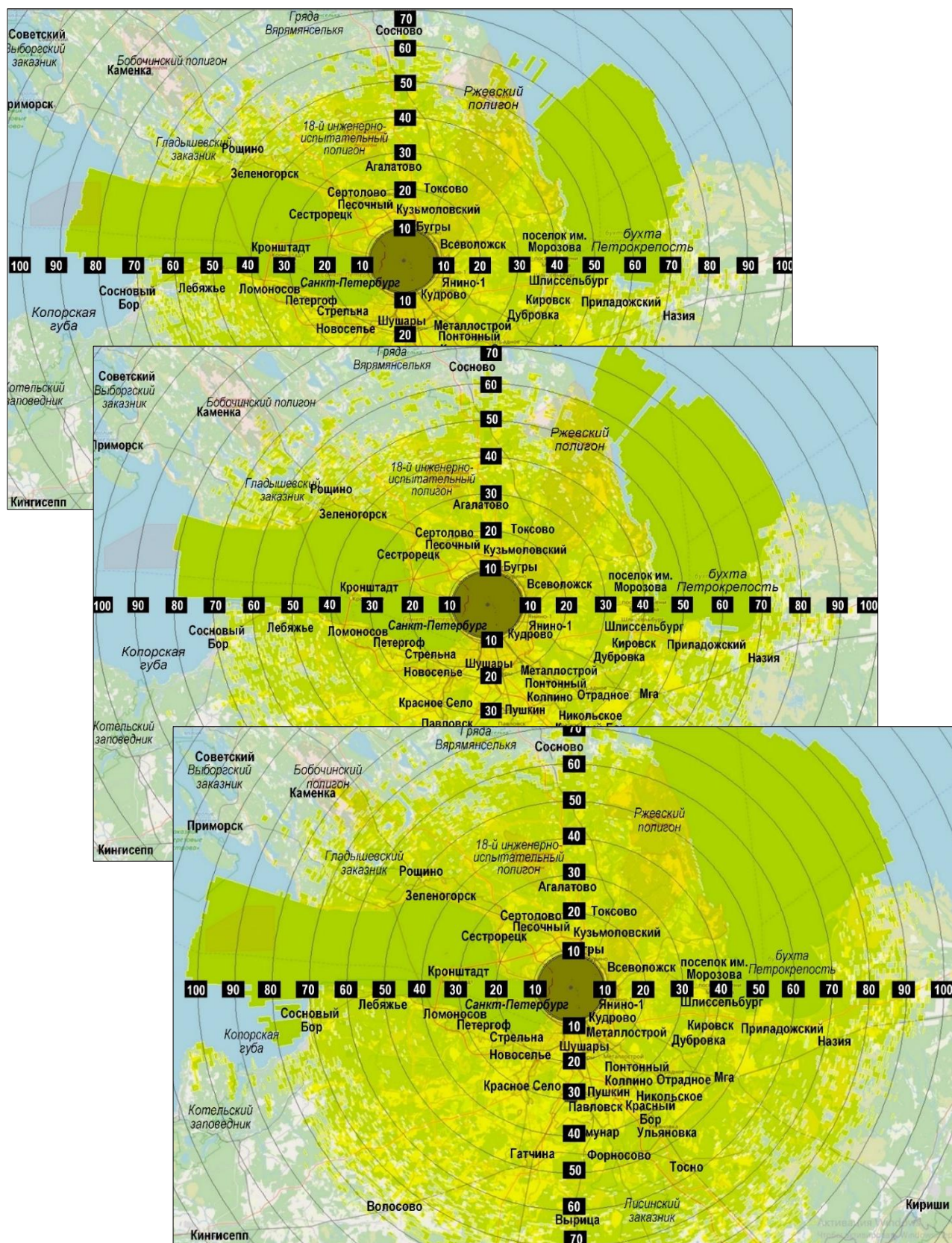


Рис. 6. Зона покрытия аналогового ЧМ-передатчика (сверху) и цифрового DRM-передатчика: модуляция поднесущих частот QAM-16 (посередине) и QAM-4 (внизу)

Fig. 6. Coverage Zone of an FM-Transmitter (Upper Image) and a Digital DRM-Transmitter; the Sub-Frequency Modulation is QAM-16 (Middle Image) and QAM-4 (lower image) for Mobile Receiving

Закключение

1) Исследована возможность применения программы Radio Mobile для расчета зон покрытия аналоговых ЧМ- и цифровых DRM-передатчиков.

2) Найдены путем сравнения расчетных данных с результатами натурных измерений значения параметра «% ситуаций», необходимое для получения достоверных результатов при оценке размера зоны покрытия для разных условий приема.

3) Установлено, что при мобильном приеме обеспечивается равенство зон покрытия аналого-

вого ЧМ- и цифрового DRM-передатчиков при соотношении их уровней мощности около 7 дБм в случае, когда DRM-передатчик работает в режиме наихудшей помехозащищенности при модуляции поднесущих частот QAM-16 и скорости кода, равной $R = 0,625$. Переход в режим наилучшей помехозащищенности (модуляция QAM-4 и скорость кода $R = 0,25$) обеспечивает при сохранении прежней мощности DRM-передатчика увеличение радиуса зоны покрытия в среднем на 10...15 км.

БЛАГОДАРНОСТИ

Автор выражает благодарность руководителю НОЦ СПбГУТ «Технологии информационных образовательных систем» С.В. Мышьянову за любезно предоставленные экспериментальные данные (рисунок 1, таблицы 2 и 3).

Список используемых источников

1. Рекомендация МСЭ R BS.1660-8 (06/2019). Техническая основа для планирования наземного цифрового звукового радиовещания в полосе ОВЧ.
2. Рекомендация МСЭ-R P.1546-6 (08/2019). Метод прогнозирования для трасс связи «пункта с зоной» для наземных служб в диапазоне частот от 30 МГц до 4000 МГц.
3. Рекомендация МСЭ-R P.528-3 (02/2012). Кривые распространения радиоволн для воздушной подвижной и радионавигационной служб, работающих в диапазонах ОВЧ, УВЧ и СВЧ.
4. Ковалгин Ю.А. Цифровое радиовещание: системы и технологии. М.: Горячая линия – Телеком, 2021. 580 с.
5. Planning Parameters for DRM Mode E ('DRM+') concerning the use in VHF bands I, II and III. V 3.0 – 04/05/2011.
6. Быховский М.А. Основы управления использованием радиочастотного спектра: Частотное планирование сетей телерадиовещания и подвижной связи. Автоматизация управления использованием радиочастотного спектра. Том 3. М.: КРАСАНД, 2012. 368 с.
7. Rec. ITU-R BS.412-9 (12/1998). Planning standards for terrestrial FM sound broadcasting at VHF.
8. Rec. ITU-R BS.450-3 (1982-1995-2001). Transmission standards for FM sound broadcasting at VHF.
9. ГОСТ Р. 51107-97. Системы стереофонического радиовещания. Основные параметры. Методы измерений. М.: ИПК Издательство стандартов, 1998.
10. ГОСТ Р. 51741-2001. Передатчики радиовещательные стационарные диапазона ОВЧ. Основные параметры, технические требования и методы измерений. М.: ИПК Издательство стандартов, 2001.
11. Rec. ITU-R P.372-14 (08/2019). Radio noise.
12. Рекомендация МСЭ-R P.527-5 (08/2019). Электрические характеристики земной поверхности.
13. Рекомендация МСЭ-R. P. 832-4 (07/2015). Мировой атлас проводимости почвы.
14. ETSI ES 201 980 v.4.2.1 (2021-01). Digitale Radio Mondiale (DRM). System Specification. EBU-UER, 184 p.

* * *

On the Application of the Radio Mobile Software for Calculating the Coverage Zones of Transmitters Broadcasting in the VHF Range

Y. Kovalgin¹

¹The Bonch-Bruевич Saint-Petersburg State University of Telecommunications, St. Petersburg, 193232, Russian Federation

Article info

DOI:10.31854/1813-324X-2021-7-2-24-36

Received 24th March 2021

Accepted 25th May 2021

For citation:, Kovalgin Y. On the Application of the Radio Mobile Software for Calculating the Coverage Zones of Transmitters Broadcasting in the VHF Range. *Proc. of Telecom. Universities*. 2021;7(2):24–36. (in Russ.) DOI:10.31854/1813-324X-2021-7-2-24-36

Abstract: *This work considers the application of the Radio Mobile software for calculating the coverage zones of analogue FM- and digital DRM-transmitters. The requirements for the initial input data, enabling the reliable results, are demonstrated. It is shown that equality of the coverage zones of both transmitter types is determined by the proper ratio of the transmitter's signal levels.*

Keywords: *digital broadcasting, coverage zone of a VHF-transmitter.*

ACKNOWLEDGMENT

The author expresses his gratitude to S.V. Myshiyarov, head of the Scientific-Educational Centre of St. Petersburg State University of telecommunications, for the kindly supplied experimental data. (figure 1, tables 2 u 3).

References

1. Rec. ITU-R BS.1660-8 (06/2019). *Technical basis for planning of terrestrial digital sound broadcasting in the VHF band.*
2. Rec. ITU-R P.1546-6 (08/2019). *The prediction method for "point-zone" communication of ground services in the frequency range from 30 MHz to 4000 MHz*
3. Rec. ITU-R P.528-3 (02/2012). *Propagation curves for aeronautical mobile and radionavigation services using the VHF, UHF and SHF bands.*
4. Kovalgin Yu.A. *Digital Broadcasting: Systems and Technologies*. Moscow: Goriachaia liniia – Telekom Publ.; 2021. 580 p. (in Russ.)
5. *Planning Parameters for DRM Mode E ('DRM+') concerning the use in VHF bands I, II and III. V 3.0 – 04/05/2011.*
6. Bykhovsky M.A. *Fundamentals of Radio Frequency Spectrum Management: Frequency Planning of Broadcasting and Mobile Networks. RF Spectrum Management Automation. Vol. 3*. Moscow: KRASAND Publ.; 2012. 368 p. (in Russ.)
7. Rec. ITU-R BS.412-9 (12/1998). *Planning standards for terrestrial FM sound broadcasting at VHF.*
8. Rec. ITU-R BS.450-3 (1982-1995-2001). *Transmission standards for FM sound broadcasting at VHF.*
9. GOST P. 51107-97. *Stereophonic broadcasting systems. Main parameters. Methods of measurements*. Moscow: IPK Izdatelstvo standartov Publ.; 1998. (in Russ.)
10. GOST P. 51741-2001. *Broadcasting transmitters, fixed Very High Frequency (VHF). Main parameters, technical requirements and methods of measurement*. Moscow: IPK Izdatelstvo standartov Publ.; 2001.
11. Rec. ITU-R P.372-14 (08/2019). *Radio noise.*
12. Rec. ITU-R P.527-4 (08/2019). *Electrical characteristics of the surface of the Earth.*
13. Rec. ITU-R P. 832-4 (07/2015). *World Atlas of Ground Conductivities.*
14. ETSI ES 201 980 v.4.2.1 (2021-01). *Digitale Radio Mondiale (DRM). System Specification*. EBU-UER, 222 p.

Сведения об авторах:

**КОВАЛГИН
Юрий Алексеевич**

доктор технических наук, профессор, профессор кафедры телевидения и метрологии Санкт-Петербургского государственного университета телекоммуникаций им. проф. М.А. Бонч-Бруевича, kowalgin@sut.ru

Метод ограничений в многокритериальной задаче распределения потоков

С.П. Ковальский¹, О.В. Титова², И.Г. Стахеев², Н.И. Фокин¹

¹Академия ФСО России,

Орел 302015, Российская Федерация

²Санкт-Петербургский государственный университет телекоммуникаций им. проф. М.А. Бонч-Бруевича,

Санкт-Петербург, 193232, Российская Федерация

*Адрес для переписки: ni_fokin@mail.ru

Информация о статье

Поступила в редакцию 25.02.2021

Принята к публикации 11.05.2021

Ссылка для цитирования: Ковальский С.П., Титова О.В., Стахеев И.Г., Фокин Н.И. Метод ограничений в многокритериальной задаче распределения потоков // Труды учебных заведений связи. 2021. Т. 7. № 2. С. 37–43. DOI:10.31854/1813-324X-2021-7-2-37-43

Аннотация: В статье предлагается новый подход к решению многокритериальной задачи формирования ресурсов транспортной сети связи, основанный на сочетании одного из методов многокритериальной оптимизации – методе ограничений и симплекс-методе. Задача декомпозируется и решается поэтапно. На первом этапе формируется множество путей допустимого ранга. На втором производится расчет структурной надежности сформированного на первом этапе множества путей. На заключительном этапе осуществляется выбор оптимальной по целевым функциям, характеризующим стоимость и пропускную способность, совокупности путей транспортирования потоков корреспондирующих пар узлов, который сводится к решению многокритериальной задачи формирования ресурсов транспортной сети связи методом ограничений.

Ключевые слова: многокритериальная оптимизация, многопродуктовый многополюсный потоковый граф, симплекс метод, метод ограничений, корреспондирующая пара узлов, эффективная альтернатива.

1. ВВЕДЕНИЕ

Практика решения задач формирования ресурсов (ЗФР) транспортной сети связи (ТСС) свидетельствует о необходимости учета целого ряда показателей, таких как расход сил и средств на построение и эксплуатацию сети, структурная надежность сети, пропускная способность сети, пропускная способность сечения, пропускная способность квазисечения и т. д. На основании этого можно сделать вывод о том, что задача формирования ресурсов транспортной сети связи является многокритериальной. Следовательно, сам характер задачи приводит к необходимости использования многокритериальных методов решения данной задачи [1–4]. При этом в настоящее время задачи синтеза транспортных сетей решаются, как правило, по одному критерию, а остальные показатели представляются в виде ограничений [5–13]. Следовательно, весьма актуальной является разработка подхода, при котором учитываются сразу несколько показателей в качестве целевых функций. В настоящей работе приведен такой многокритериальный подход к решению задачи формирования ресурсов транспортной сети.

2. МАТЕМАТИЧЕСКАЯ МОДЕЛЬ ТРАНСПОРТНОЙ СЕТИ СВЯЗИ

Для решения ЗФР целесообразно использовать теоретическую модель многополюсной многопродуктовой сети. Совокупность направлений связи (НС) создают в сети потоки различных продуктов, которые не взаимозаменяемы, не смешиваются, существуют одновременно и независимо друг от друга. Единицей потока k -го продукта считается составной типовой цифровой канал, образованный в интересах k -ой корреспондирующей пары узлов (КПУ).

В качестве математической модели структуры ТСС используется многопродуктовый многополюсный потоковый граф [1–7]:

$$G(A, B, U, H),$$

где $A = \{a_j\}, j = \overline{1, N}$, $B = \{b_{ji}\}, j, i = \overline{1, N}, j \neq i, |B| = n$ – множество узлов и линий; $U = \{u_{ji}\}, j, i = \overline{1, N}, j \neq i, H = \{h_{ji}\}, j, i = \overline{1, N}, j \neq i$, – векторы пропускных способностей и надежности линий, соответственно.

НС формируют не взаимозаменяемые, не смешиваемые и существующие одновременно и независимо друг от друга потоки различных продуктов.

3. ПОСТАНОВКА МНОГОКРИТЕРИАЛЬНОЙ ЗАДАЧИ ФОРМИРОВАНИЯ РЕСУРСОВ

Для решения многокритериальной задачи формирования ресурсов транспортной сети связи методом ограничений перейдем от множеств дискретных элементов сети – узлов и ребер к комбинаторным пространствам путей [1–7].

Для постановки многокритериальной задачи формирования ресурсов введем следующие понятия.

Направление связи, организованное на сети в интересах k -ой КПУ, представляет собой совокупность путей $\Pi_k = \{\pi_k^t\}$, $k = \overline{1, m}$, $t = 1, 2, \dots$ из узла a_{sk} в узел a_{gk} , t – номер пути.

Понятия пути π_k^t из узла a_s в узел a_g , ранга пути – $r(\pi_{sg}^t)$ (число ребер, образующих путь) и допустимого пути – $\pi_k^{r \leq r^{\text{доп}}} = \pi_k^{r, t}$ (путь, ранг которого не превышает допустимого значения) введены в [11].

Допустимый ранг пути – $r^{\text{доп}}(\pi_k^t)$ между КПУ на сети ограничен допустимым числом транзитов составного цифрового канала (ЦК), исходя из выполнения самых жестких требований со стороны услуг электросвязи к эксплуатационным нормам на параметры составного ЦК [1–7].

Поэтому при решении ЗФР целесообразно рассматривать не все множество возможных путей – $\Pi_k = \{\pi_k^t\}$, $k = \overline{1, m}$, $t = 1, 2, \dots$, а только допустимые пути, для которых выполняется неравенство:

$$r(\pi_k^{r, t}) \leq r^{\text{доп}}(\pi_k^t), \quad k = \overline{1, m}. \quad (1)$$

Следовательно, для решения ЗФР будем учитывать (1) между КПУ:

$$Z = \{z_k\}, z_k = (a_{sk}, a_{tk}), \quad k = \overline{1, m}.$$

Синтезируемая транспортная сеть связи формируется с использованием узлового R_y и линейного ресурсов R_l [1–7]:

$$\begin{aligned} R_y &= \{r^\psi: \omega_j^\psi, \vartheta_j^\psi, \gamma_j^\psi, U_\psi, \vec{H}_j\}, r = \overline{1, Q_r}, \psi = \overline{1, Q_\psi}, \\ R_l &= \{r^\mu: \omega_{ji}^\mu, \vartheta_{ji}^\mu, \gamma_{ji}^\mu, U_\mu, \vec{L}_\mu, \vec{a}_\mu, \vec{H}_\mu\}, \\ &r = \overline{1, Q_r}, \mu = \overline{1, Q_\mu}. \end{aligned}$$

где r^μ – тип системы передачи; r^ψ – тип сетевого узла; $\omega_{ji}^\mu, \vartheta_{ji}^\mu$ и $\omega_j^\psi, \vartheta_j^\psi$ – нормированные коэффициенты аппроксимации приведенных затрат на развертывание и эксплуатацию μ -го линейного и ψ -го узлового средства, соответственно; γ_{ji}^μ и γ_j^ψ – стоимость одного канала-километра ij -й линии μ -й системы передачи и стоимость одной точки коммутации i -го узла ψ -го узлового средства, со-

ответственно; U_μ, U_ψ – количество каналов и трактов, образуемых r^μ и коммутируемых r^ψ средствами, соответственно; $\vec{H}_\mu, \vec{H}_j$ – векторы эксплуатационных надежностей для r^μ и r^ψ оборудования, соответственно; $\vec{L}_\mu$ – вектор параметров структуры r^μ системы передачи (длины регенерационных и усилительных участков, секций дистанционного питания и т. п.); $\vec{a}_\mu$ – вектор эксплуатационных норм на характеристики каналов и трактов.

В соответствии с целевым использованием ТСС, целевыми критериями ее синтеза могут быть пропускная способность Π и расход сил и средств C на построение и эксплуатацию сети:

$$\Pi = f_1(\overline{\Pi}_Z, r^\mu, r^\psi, V_Z, \vec{H}^{\text{тп}}) = f_1(\beta),$$

$$C = f_2(\overline{\Pi}_Z, r^\mu, r^\psi, V_Z, \vec{H}^{\text{тп}}) = f_2(\beta).$$

Тогда многокритериальная ЗФР состоит в выборе альтернативы $\delta \in \beta$: $\Pi_k = \{\pi_k^t\}$, при которой:

$$\min_{\beta} \Delta f_j(\beta) = \begin{cases} f_1^0 - f_1(\beta) \\ f_2(\beta) - f_2^0 \end{cases}, \quad \mu = \overline{1, Q_\mu}, \quad \psi = \overline{1, Q_\psi}.$$

При решении многокритериальной ЗФР учитывают следующие ограничения:

$$\sum_{t=1}^{h_{\text{св}k}} u_k^t = V_k \quad (2)$$

$$\sum_{k=1}^m \sum_{t=1}^{h_{\text{св}k}} \omega_{j,k}^t v_k^t = v_j,$$

$$\omega_{j,k}^t = 1, \forall b_j \in \pi_k^t, \omega_{j,k}^t = 0, \forall b_j \notin \pi_k^t, \quad (3)$$

$$u_j = \zeta_j OTUL, v_k^t = \zeta_j OTUL, l = 1, \dots, 4,$$

$$k = \overline{1, m}, t = \overline{1, h_{\text{св}k}}, j = \overline{1, n}.$$

Получение содержательного решения данной задачи связано с трудностями принципиального характера, являющимися следствием содержания этой задачи, требующей увязки различных противоречивых требований в рациональном решении. Учитывая эти факты, ЗФР декомпозируется и решается поэтапно. На первом этапе формируется множество путей допустимого ранга. На втором производится расчет структурной надежности сформированного на первом этапе множества путей. На заключительном этапе осуществляется выбор оптимальной по целевым функциям, характеризующим стоимость и пропускную способность, совокупности путей транспортирования потоков корреспонденций пар узлов.

4. РЕШЕНИЕ МНОГОКРИТЕРИАЛЬНОЙ ЗАДАЧИ ФОРМИРОВАНИЯ РЕСУРСОВ

4.1. Нахождения множества допустимых путей

Для нахождения множества допустимых путей разработан ряд алгебраических методов. Наиболее

распространенным является метод, основанный на возведении структурной матрицы в степень [14]. Этот метод практически исключает появления ошибок и позволяет получать результат в довольно удобной для дальнейшего использования форме.

В [14] дано определение структурной матрицы $B = \|\beta_{ji}\|$. В этом же источнике матрица $B = \|\beta_{ji}\|$, определяется как булева матрица, к которой применяется аппарат булевой алгебры логики, и при этом используются следующие преобразования:

$$\begin{aligned} a \wedge a &= a; a \vee a = a; 1 \wedge a = a; a \wedge (a \vee b) = a; \\ a \vee a \wedge b &= a; a \wedge a' = 0; a \vee a' = 1. \end{aligned} \quad (4)$$

Для нахождения множества путей между заданными КПУ (как всех, так и удовлетворяющих заданному свойству (рангу)) следует последовательно возводить структурную матрицу во вторую, третью и т. д. В итоге мы получим характеристическую матрицу. Если требуется найти все пути не более $r^{\text{доп}}$, матрицу следует возводить только до $r^{\text{доп}}$ -ой степени, тогда каждое вхождение данной матрицы будет содержать все пути от узла a_j к узлу a_i ранга не более $r^{\text{доп}}$.

Нахождение множества всех путей π_k^t от a_j к a_i осуществляется путем раскрытия определителя матрицы B_{ij} с вычеркнутыми j -м столбцом и i -й строкой:

$$\pi_k^t = \det B_{ij} = |B_{ij}|. \quad (5)$$

4.2. Расчет структурной надежности множества путей

Исходя из основного предназначения ТСС, наиболее конструктивным показателем структурной надежности является вероятность наличия связи между КПУ – $H_{Z_k}, k = \overline{1, m}$. Наличие связи означает существование хотя бы одного исправного пути между Z_k [1–6]. Тогда показателем структурной надежности ТСС станет вероятность связности H , которая представляет собой вероятность наличия, по крайней мере, одного исправного пути между любой парой узлов сети. Один из возможных подходов преобразования требований к структурной надежности сети в требования к связности заключается в следующем. Направление связи представлено математической вероятностной моделью, состоящей из h путей с надежностями каждого пути – $H_{\pi_k^t}, t = \overline{1, h}, k = \overline{1, m}$. Считается, что отказы элементов путей (ребра графа): независимые события. Тогда надежность ИНС может быть рассчитана с помощью алгоритма ортогонализации.

В [14] *имеется следующее утверждение*: отрицание конъюнкции $tK_j = x_1^{\alpha_1} x_2^{\alpha_2} \dots x_t^{\alpha_t}$ эквивалентно дизъюнкции:

$$K_j' = x_1^{\eta_1} \vee x_1^{\eta_1} x_2^{\eta_2} \vee \dots \vee x_1^{\eta_1} x_2^{\eta_2} \dots x_{t-1}^{\eta_{t-1}} x_t^{\eta_t}. \quad (6)$$

Выражение (6) можно переписать в следующем виде:

$$(x_1 x_2 x_3 \dots x_t)' = \begin{vmatrix} x_1' \\ x_2' \\ \vdots \\ \vdots \\ \vdots \\ x_t' \end{vmatrix} = \begin{vmatrix} x_1' \\ x_1 x_2' \\ \vdots \\ \vdots \\ \vdots \\ x_1 x_2 x_3 \dots x_{t-1} x_t' \end{vmatrix}. \quad (7)$$

Справедливость преобразования (6) доказана в монографии [15]. Доказательство построено на основе теоремы разложения, которая применяется последовательно для элементов $x_1, x_2, \dots, x_{t-1}$ к элементарной дизъюнкции, получаемой в соответствии с правилом де Моргана из элементарной конъюнкции K_j :

$$D_j = x_1^{\eta_1} \vee x_2^{\eta_2} \vee \dots \vee x_{t-1}^{\eta_{t-1}} \vee x_t^{\eta_t}. \quad (8)$$

На основании теоремы разложения [15] функцию алгебры логики можно записать как:

$$\begin{aligned} f(x_1, x_2, x_3, \dots, x_j, x_{j+1}, \dots, x_n) &= \\ &= \vee x_1^{\eta_1} x_2^{\eta_2} \dots x_j^{\eta_j} f(\eta_1, \eta_2, \dots, \eta_j, x_{j+1}, \dots, x_n). \end{aligned}$$

Булева функция $f(x_1, x_2, x_3, \dots, x_n)$, представленная в [15] в виде дизъюнктивной нормальной формы (9), эквивалентна функции (10):

$$f(x_1, x_2, x_3, \dots, x_n) = \bigvee_{j=1}^m K_j, j \leq 2^n, \quad (9)$$

$$\begin{aligned} f(x_1, x_2, x_3, \dots, x_n) &= \\ &= K_1 \vee K_1' K_2 \vee K_1' K_2' K_3 \vee \dots \vee K_1' K_2' K_3' \dots K_{m-1}' K_m. \end{aligned} \quad (10)$$

В матричной форме записи уравнения (9) и (10) принимают вид:

$$f(x_1 x_2 x_3 \dots x_t) = \begin{vmatrix} K_1 \\ K_2 \\ \vdots \\ \vdots \\ \vdots \\ K_m \end{vmatrix} = \begin{vmatrix} K_1 \\ K_1' K_2 \\ \vdots \\ \vdots \\ \vdots \\ K_1' K_2' K_3' \dots K_{m-1}' K_m \end{vmatrix}. \quad (11)$$

Справедливость выражения (11) также доказывается с помощью теоремы разложения. Далее дизъюнкция (11) приводится к дизъюнктивной нормальной форме, в результате получаем ортогональную дизъюнктивную нормальную форму (ОДНФ) булевой функции $f(x_1, x_2, x_3, \dots, x_t)$, т. е. производится преобразование условий работоспособности системы к ОДНФ, после чего производится расчет вероятности безотказной работы системы согласно выражению:

$$H \left\{ f(x_1, x_2, x_3, \dots, x_n) = \bigvee_{j=1}^s \Xi_j = 1 \right\} = \sum_{j=1}^s H \{ \Xi_j = 1 \}, \quad (12)$$

где Ξ_j – ортогональные члены записанной в ОДНФ [15] функции $f(x_1, x_2, x_3, \dots, x_n)$.

4.3. Решение многокритериальной задачи формирования ресурсов ТСС методом ограничений

Основой поиска компромиссного решения ЗФР является система неравенств:

$$\rho_j w_j(\beta^*) \leq \kappa_0(\min) \quad (j \in J), \quad (13)$$

где $\delta^* \in \beta$ – эффективная альтернатива при заданном векторе предпочтений – $\rho \in P^+$ [2, 16].

Решение находится путем построения итерационного процесса с параметром $\kappa_0 \in (0, 1/M)$ с проверкой совместимости системы неравенств (13) для $\delta \in \beta$ и заданного вектора ρ .

Переходим к решению конкретного примера решения ЗФР. Дана сеть связи, представленная на рисунке 1.

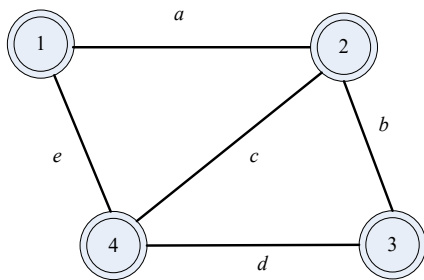


Рис. 1. Сеть связи

Fig. 1. Communication Network

Даны матрицы пропускных способностей, стоимости и надежности линий:

$$P_{ij} = \begin{bmatrix} 0 & 64 & 0 & 16 \\ 64 & 0 & 16 & 16 \\ 0 & 16 & 0 & 256 \\ 16 & 16 & 256 & 0 \end{bmatrix},$$

$$C_{ij} = \begin{bmatrix} 0 & 16 & 0 & 4 \\ 16 & 0 & 4 & 4 \\ 0 & 4 & 0 & 64 \\ 4 & 4 & 64 & 0 \end{bmatrix},$$

$$H_{ij} = \begin{bmatrix} 0 & 0,9 & 0 & 0,9 \\ 0,9 & 0 & 0,9 & 0,9 \\ 0 & 0,9 & 0 & 0,9 \\ 0,9 & 0,9 & 0,9 & 0 \end{bmatrix}.$$

Пусть длина составного ЦК ограничена условием $r^{\text{доп}} \leq 3$. Необходимо организовать два информационных направления связи (ИНС) – $Z_{1,3}$ и $Z_{2,4}$, емкости ИНС – $Z_{1,3} \geq 25$, $Z_{2,4} \geq 15$, надежность ИНС – $\bar{H} \geq \bar{H}^{\text{тр}} \geq 0,95$. Также требуется найти распределение потоков на сети, удовлетворяющее заданным требованиям.

Решение задачи. Определить допустимое распределение потоков $f\langle S \rangle$ по путям $s \in S$, при этом необходимо максимизировать количество каналов по каждому ИНС $\Phi_1\langle S \rangle (j \in J_1)$ и минимизировать функцию стоимости (аренды) каналов $\Phi_2\langle S \rangle (j \in J_2)$. Найти:

$$\Phi_1(f_{hy}) \Rightarrow \max_{f\langle S \rangle, s \in S}, \quad (14)$$

$$\Phi_2(f_{hy}) \Rightarrow \min_{f\langle S \rangle, s \in S}, \quad (15)$$

если

$$\bar{H} \geq \bar{H}^{\text{тр}}, f\langle S \rangle \geq 0, f_{\langle j \rangle i} \leq c_{\langle j \rangle i}. \quad (16)$$

С учетом исходных данных ограничение (16) запишем в виде: $\bar{H} \geq \bar{H}^{\text{тр}} \geq 0,95$, $f_{\langle 1,3 \rangle} \geq 25$, $f_{\langle 2,4 \rangle} \geq 15$. Для нахождения путей для $Z_{1,3}$ и $Z_{2,4}$ (при $r^{\text{доп}} \leq 3$) возводим структурную матрицу $B = \| \beta_{ij} \|$ в третью степень. Вхождения $\chi_{1,3}^3 = ab \vee e\bar{d} \vee e\bar{c}b \vee ac$ и $\chi_{2,4}^3 = c \vee bd \vee \bar{a}e$ дают множество путей для $Z_{1,3}$ и $Z_{2,4}$, соответственно.

Используя алгоритм ортогонализации (6–12), произведем расчет надежности найденных путей (с учетом значений $R_j = R_i = 0,9$, $\forall j, i = 1, \dots, 5$), получим:

$$H_{1,3}(Y_{1,3}) = 0,97848 > H^{\text{тр}} \text{ и } H_{2,4}(Y_{2,4}) = 0,99639 > H^{\text{тр}}.$$

Потоки по информационным направлениям $Z_{1,3} = ab \vee e\bar{d} \vee e\bar{c}b \vee ac\bar{d}$ и $Z_{2,4} = c \vee bd \vee \bar{a}e$ определим следующим образом:

$$f_{\langle 1,3 \rangle} = f_{\langle ab \rangle} + f_{\langle e\bar{d} \rangle} + f_{\langle ac\bar{d} \rangle} + f_{\langle e\bar{c}b \rangle}, \quad (17)$$

$$f_{\langle 2,4 \rangle} = f_{\langle c \rangle} + f_{\langle bd \rangle} + f_{\langle \bar{a}e \rangle}. \quad (18)$$

Запишем условия допустимости потоков:

$$\text{Для ребра (1,2): } f_{\langle ab \rangle} + f_{\langle ac\bar{d} \rangle} + f_{\langle \bar{a}e \rangle} \leq c_{12}, \quad (19)$$

$$\text{Для ребра (1,4): } f_{\langle e\bar{d} \rangle} + f_{\langle e\bar{c}b \rangle} + f_{\langle \bar{a}e \rangle} \leq c_{14}, \quad (20)$$

$$\text{Для ребра (2,3): } f_{\langle ab \rangle} + f_{\langle e\bar{c}b \rangle} + f_{\langle bd \rangle} \leq c_{23}, \quad (21)$$

$$\text{Для ребра (2,4): } f_{\langle c \rangle} + f_{\langle e\bar{c}b \rangle} + f_{\langle ac\bar{d} \rangle} \leq c_{24}, \quad (22)$$

$$\text{Для ребра (3,4): } f_{\langle e\bar{d} \rangle} + f_{\langle ac\bar{d} \rangle} + f_{\langle bd \rangle} \leq c_{34}. \quad (23)$$

Введем следующие обозначения:

$$v_1 = f_{\langle ab \rangle}, v_2 = f_{\langle ac\bar{d} \rangle}, v_3 = f_{\langle \bar{a}e \rangle}, v_4 = f_{\langle e\bar{d} \rangle}, \quad (24)$$

$$v_5 = f_{\langle e\bar{c}b \rangle}, v_6 = f_{\langle bd \rangle}, v_7 = f_{\langle c \rangle}.$$

С учетом выражений (19–24) получим следующую бикритериальную задачу оптимизации:

$$\Phi_1(f): \max_{\vec{v}} \{v_1 + v_2 + v_3 + v_4 + v_5 + v_6 + v_7\}, \quad (25)$$

$$\Phi_2(f): \min_{\vec{v}} \{20v_1 + 84v_2 + 20v_3 + 68v_4 + 12v_5 + 68v_6 + 4v_7\} \quad (26)$$

при ограничениях:

$$g_1(\vec{v}) - g_8(\vec{v}):$$

$$\begin{aligned} v_1 + v_2 + v_3 &\leq 64, \\ v_4 + v_5 + v_3 &\leq 16, \\ v_1 + v_5 + v_6 &\leq 16, \\ v_7 + v_5 + v_2 &\leq 16, \\ v_4 + v_2 + v_6 &\leq 256, \\ v_1 + v_4 + v_5 + v_2 &\geq 25, \\ v_3 + v_6 + v_7 &\geq 15, \\ v_i &\geq 0, \quad i = \overline{1,7}. \end{aligned} \quad (27)$$

Если $\rho_1 = \rho_2 = 1/2$, то $v_1 = 16, v_2 = 16, v_3 = 0, v_4 = 16, v_5 = 0, v_6 = 0, v_7 = 0$ – план, оптимальный по первому критерию; $v_1 = 15, v_2 = 9, v_3 = 0, v_4 = 9, v_5 = 1, v_6 = 0, v_7 = 15$ – план, оптимальный по второму критерию; $\Phi_1^0(f) = 48, \Phi_2^0(f) = 984$ – соответствующие оптимальные значения критериев,

$\Phi_1^{\min}(f) = 40, \Phi_2^{\max}(f) = 3520$ – наихудшие значения критериев $\Phi_1(f)$ и $\Phi_2(f)$ на множестве ограничений.

Функции относительных потерь примут вид:

$$\tilde{w}_1(v^l): \quad \rho_1 w_1(v) = \frac{1}{2} \cdot \frac{48 - v_1 - v_2 - v_3 - v_4 - v_5 - v_6 - v_7}{8}, \quad (28)$$

$$\tilde{w}_2(v^l): \quad \rho_2 w_2(v) = \frac{1}{2} \cdot \frac{20v_1 + 84v_2 + 20v_3 + 68v_4 + 12v_5 + 68v_6 + 4v_7 - 984}{2536}. \quad (29)$$

Эквивалентная задача линейного программирования (25–27) в соответствии с выражениями (28, 29), а также в соответствии с [16] принимает вид:

$$\Psi(f): \quad \min_v \{r_0 = v_8\} \quad (30)$$

при ограничениях:

$$\begin{aligned} v_1 + v_2 + v_3 &\leq 64, v_3 + v_4 + v_5 \leq 16, \\ v_1 + v_5 + v_6 &\leq 16, v_2 + v_5 + v_7 \leq 16, \\ v_1 + v_2 + v_4 + v_5 &\geq 25, v_2 + v_4 + v_6 \leq 256, \\ v_3 + v_6 + v_7 &\geq 15, \\ \frac{1}{2}v_1 + \frac{1}{2}v_2 + \frac{1}{2}v_3 + \frac{1}{2}v_4 + \frac{1}{2}v_5 + \frac{1}{2}v_6 + \\ &+ \frac{1}{2}v_7 + 8v_8 - 24 \geq 0, \\ -\frac{1}{2} \cdot (20v_1 + 84v_2 + 20v_3 + 68v_4 + 12v_5 + \\ &+ 68v_6 + 4v_7) + 3536v_8 + 492 \geq 0, \\ v_i &\geq 0, \quad i = \overline{1, 8}. \end{aligned} \quad (31)$$

Решение:

$$\begin{aligned} \min k_0 &\approx 0,02818991, \\ v_1 &= 16, v_2 = 0, v_3 = 6,5, \\ v_4 &= 9, v_5 = 0, v_6 = 0, v_7 = 16, \\ \Phi_1^0(f) &= 47,5, \Phi_2^0(f) = 1126. \end{aligned}$$

Полученное решение единственное, и оно является точным решением задачи, поскольку для него $\tilde{w}_1(v^l) = \tilde{w}_2(v^l) \approx 0,02818991$, т. е. одинаковы относительные потери от оптимальных значений по обоим критериям.

Список используемых источников

1. Лебедев А.Т., Лебедев И.А., Тумановский В.В. Построение региональных цифровых сетей связи // Научно-технический сборник. Телекоммуникационные технологии. Выпуск 1. СПб.: ГУП НИИ «Рубин», 2000.
2. Фокин Н.И. Многокритериальная задача распределения информационных потоков транспортной сети связи // Вопросы радиоэлектроники. 2019. № 12. С. 64–69. DOI:10.21778/2218-5453-2019-12-64-69
3. Муравцов А.А., Стахеев И.Г., Фокин Н.И., Шинкарев С.А. Синтез потоковой структуры мультипротокольной транспортной сети связи региона с учетом обеспечивающих и взаимодействующих подсистем при заданной надежности передачи их потоков // Научные труды SWORLD. 2011. Т. 6. № 1. С. 32–34.

ЗАКЛЮЧЕНИЕ

Задача формирования ресурсов транспортной сети связи по своей природе является многокритериальной. При этом в настоящее время она решается, как правило, по одному критерию, а остальные показатели представляются в виде ограничений. Поэтому весьма актуальной является разработка подхода, при котором учитываются сразу несколько показателей в качестве целевых функций. При этом следует отметить, что если задачу формирования ресурсов решать на основе введения глобального критерия в виде суммы взвешенных относительных потерь – $\min_{u \in G} \{\rho_1 w_1(v) + \rho_2 w_2(v)\}$, то такое условие не обеспечит равноценность критериев. Следовательно, необходима разработка нового подхода.

В настоящей статье предложен новый подход, основанный на сочетании метода ограничений и симплекс-метода. С учетом высокой сложности задачи решение осуществляется путем ее декомпозиции: формирование множества путей допустимого ранга, расчет структурной надежности сформированного множества путей и выбор оптимальной по целевым функциям, характеризующим расход сил и средств, а также пропускную способность, совокупности путей транспортирования потоков корреспондирующих пар узлов. Конкретный пример, рассмотренный в статье, демонстрирует конструктивность и системность предлагаемого подхода.

При этом следует отметить, что данный подход разработан для задач, допускающих дробное решение. Дальнейшем развитием многокритериального подхода может стать подход, учитывающий комбинаторную природу ряда задач синтеза транспортных сетей.

4. Ковальский С.П., Фокин Н.И., Абдуразаков Р.М., Никитин М.В. Формирование структуры транспортной сети связи // International Journal of Open Information Technologies. 2016. Т. 4. № 6. С. 10–16.
5. Канаев А.К., Лукичев М.М., Муравцов А.А. Синтез потоковой структуры транспортной сети связи с использованием имитационного моделирования // Известия Петербургского университета путей сообщения. 2015. № 2(43). С. 105–111.
6. Трегубов Р.Б., Саитов И.А., Саитов С.И. Элементы теоретико-множественного базиса и системы моделей многопутевых многоадресных инфокоммуникационных систем // Труды СПИИРАН. 2017. Т. 2(51). С. 35–59. DOI:10.15622/sp.51.2
7. Ясинский С.А., Зюзин А.Н. Подход к модернизации топологии транспортных сетей связи на основе кольцевых структур // Информация и космос. 2019. № 1. С. 22–29.
8. Шарифов Ф.А. Задача синтеза надежных сетей // Кибернетика и системный анализ. 2000. Т. 36. № 4. С. 145–157.
9. Шор Н.З., Шарифов Ф.А. Общая задача синтеза надежных сетей // Проблемы информатики и кибернетики. 2006. № 2-3. С. 184–202.
10. Шарифов Ф.А., Гуляцкий Л.Ф. Модели и сложность задач проектирования и реконструкции телекоммуникационных и транспортных систем // Кибернетика и системный анализ. 2014. Т. 50. № 5. С. 49–58.
11. Форд А., Фалкерсон Д. Потоки в сетях. Пер. с англ. М.: Мир, 1966. 276 с.
12. Фрэнк Г., Фриш И. Сети, связь и потоки. Пер. с англ. М.: Связь, 1978. 448 с.
13. Ху Т. Целочисленное программирование и потоки в сетях. Пер. с англ. М.: Мир, 1974. 518 с.
14. Давыдов Г.Б., Рогинский В.Н., Толчан А.Я. Сети электросвязи. М.: Связь, 1977. 360 с.
15. Рябинин И.А., Черкесов Г.Н. Логико-вероятностные методы исследования надежности структурно сложных схем. М.: Радио и связь, 1981. 264 с.
16. Михалевич В.С. Вычислительные методы выбора оптимальных проектных решений. Киев: Наукова думка, 1977. 289 с.

* * *

Method of Restrictions in the Multicriteria Task of Allocation of Flows

S. Kowalski¹ , O. Titova² , I. Stakheev² , N. Fokin¹ 

¹Russian Federation Security Guard Service Federal Academy,
Orel, 302015, Russian Federation

²The Bonch-Bruевич Saint-Petersburg State University of Telecommunications,
St. Petersburg, 193232, Russian Federation

Article info

DOI:10.31854/1813-324X-2021-7-2-37-43

Received 25th February 2021

Accepted 11th May 2021

For citation: Kowalski S., Titova O., Stakheev I., Fokin N. Method of Restrictions in the Multicriteria Task of Allocation of Flows. *Proc. of Telecom. Universities*. 2021;7(2):37–43. (in Russ.) DOI:10.31854/1813-324X-2021-7-2-37-43

Abstract: In article the new approach to the decision of the multicriteria task of formation of resources of a transport communication network based on a combination of one of methods of multicriteria optimization – a method of restrictions and a simplex method is offered. The task is decomposed also dares stage by stage. At the first stage the set of ways of an admissible rank is formed. On the second calculation structural reliabilities of the set of ways generated at the first stage is produced. At the final stage the choice optimal on the criterion functions characterizing cost and carrying capacity, set of ways of transportation of flows corresponding pairs of nodes which is reduced to the decision of the multicriteria task of formation of resources of a transport communication network by a method of restrictions is carried out.

Keywords: multicriteria optimization, multiproduct multipole flow graph, simplex method, method of restrictions, corresponding pairs of nodes, efficient alternative.

References

1. Lebedev A.T., Lebedev I.A., Tumanovsky V.V. Construction of Regional Digital Communication Networks. *Scientific and Technical Collection. Telecommunication Technologies*. Iss. 1. St. Petersburg: Research Institute “Rubin” Publ.; 2000.

2. Fokin N. I. Multicriteria Task of Distribution of Flows of a Transport Communication Network. *Issues of Radio Electronics (Voprosy radioelektroniki)*. 2019;12:64–69. DOI:10.21778/2218-5453-2019-12-64-69
3. Muravtsov A.A., Stakheev I.G., Fokin N.I., Shinkarev S.A. Synthesis of the Streaming Structure of a Multi-Protocol Transport Communication Network of the Region, Taking into Account the Supporting and Interacting Subsystems with a Given Reliability of the Transmission of their Streams. *Nauchnyye trudy SWORLD*. 2011;6(1):32–34.
4. Kovalskiy S.P., Fokin N.I., Abdurazakov R.M., Nikitin M.V. Formation of the Transport Network Structure. *International Journal of Open Information Technologies*. 2016;4(6):10–16.
5. Kanaev A.K., Lukichev M.M., Muravtsov A.A. Synthesis of Streaming Structure of Transportation Communication Network, Using Simulation Modeling. *Proceedings of Petersburg Transport University*. 2015;2(43):105–111.
6. Tregubov R.B., Saitov I.A., Saitov S.I. Elements of the Set-Theoretic Base and System of Models of Multipath Multi-Address Infocommunication Systems. *SPIIRAS Proceedings*. 2017;2(51):35–59. DOI:10.15622/sp.51.2
7. Yasinsky S.A., Zyuzin A.N. An Approach to Modernizing the Topology of Transport Communication Networks Based on Ring Structures. *Information and Space*. 2019;1:22–29.
8. Sharifov F.A. The Problem of Synthesizing Reliable Networks. *Cybernetics and Systems Analysis*. 2000;36(4):145–157.
9. Shor N.Z., Sharifov F.A. General Problem of Synthesis of Reliable Networks. *Problemy informatiki i kibernetiki*. 2006;2-3:184–202.
10. Sharifov F.A., Gulyatsky L.F. Models and Complexity of Design and Reconstruction Tasks for Telecommunication and Transport Systems. *Cybernetics and Systems Analysis*. 2014;50(5):49–58.
11. Ford L.R., Fulkerson D.R. *Flows in Networks*. Translated from English. Moscow: Mir Publ.; 1966. 276 p.
12. Frank H., Frisch I.T. *Communication, Transmission and Transportation Networks*. Translated from English Moscow: Sviaz Publ.; 1978. 448 p.
13. Hu T.C. *Integer Programming and Network Flows*. Translated from English. Moscow: Mir Publ.; 1974. 518 p.
14. Davydov G.B., Roginsky V.N., Tolchan A.Ya. *Telecommunication Networks*. Moscow: Sviaz Publ.; 1977. 360 p.
15. Ryabinin I.A., Cherkesov G.N. Logical-Probabilistic Methods for Investigating the Reliability of Structurally Complex Circuits. Moscow: Radio i sviaz Publ.; 1981. 264 p.
16. Mikhalevich V.S. *Computational Methods for Choosing Optimal Design Solutions*. Kiev: Naukova dumka Publ.; 1977. 289 p.

Сведения об авторах:

КОВАЛЬСКИЙ
Сергей Петрович

кандидат технических наук, сотрудник Академии ФСО России,
metal_forever@inbox.ru

 <https://orcid.org/0000-0002-9000-6195>

ТИТОВА
Ольга Викторовна

кандидат технических наук, доцент кафедры специальных средств связи
Санкт-Петербургского государственного университета телекоммуника-
ций им. проф. М.А. Бонч-Бруевича, olga1110.spb@mail.ru

 <https://orcid.org/0000-0002-7107-3253>

СТАХЕЕВ
Иван Геннадиевич

кандидат технических наук, доцент, доцент военного учебного центра
Санкт-Петербургского государственного университета телекоммуника-
ций им. проф. М.А. Бонч-Бруевича, kisasig@yandex.ru

 <https://orcid.org/0000-0002-1311-5603>

ФОКИН
Николай Иванович

кандидат технических наук, сотрудник Академии ФСО России,
sky79@mail.ru

 <https://orcid.org/0000-0001-9658-8731>

Алгоритм расчета показателя неготовности направления связи по дождевому фактору с учетом выполнения требований к пропускной способности, достоверности и ограничения частотно-энергетического ресурса

А.Е. Орлов¹, К.С. Новосадов¹, Д.В. Сальников¹, В.Н. Беженарь¹

¹Военная академия связи им. Маршала Советского Союза С.М. Буденного,
Санкт-Петербург, 194064, Российская Федерация

*Адрес для переписки: x-bruk@mail.ru

Информация о статье

Поступила в редакцию 02.02.2021

Принята к публикации 11.05.2021

Ссылка для цитирования: Орлов А.Е., Новосадов К.С., Сальников Д.В., Беженарь В.Н. Алгоритм расчета показателя неготовности направления связи по дождевому фактору с учетом выполнения требований к пропускной способности, достоверности и ограничения частотно-энергетического ресурса // Труды учебных заведений связи. 2021. Т. 7. № 2. С. 44–50. DOI:10.31854/1813-324X-2021-7-2-44-50

Аннотация: В статье представлен алгоритм расчета показателя неготовности по дождевому фактору участков спутниковой радиолинии с учетом выполнения требований к пропускной способности, достоверности и ограничения частотно-энергетического ресурса. В отличие от существующих подходов, в предлагаемом задаче решается на основе аналитического выражения, что позволяет снизить вычислительную емкость при приемлемых ограничениях.

Ключевые слова: земная станция, спутниковая радиолиния, сигнально-кодовая конструкция, неготовность.

Введение

В настоящее время в России и во всем мире создаются высокоскоростные спутниковые системы связи в Ka/Q -диапазонах частот с использованием космических аппаратов (КА) на геостационарной орбите. Эти системы обслуживают обширные территории, предоставляя абонентам сервисы широкополосного доступа к удаленным инфокоммуникационным ресурсам. Достигается это за счет применения на КА многоручевых антенн, что обеспечивает высокую эквивалентную изотропно-излучаемую мощность и добротность приемных систем КА в зонах обслуживания абонентов, и тем самым создаются условия для использования абонентами относительно простых терминалов. Сети строятся по топологии «звезда» с центральной земной станцией (ЦЗС), имеющей антенну диаметром до 13 м. Доступный частотный ресурс позволяет поддерживать высокую скорость передачи данных в направлении от ЦЗС к абонентским земным станциям (АЗС). В отличие от спутниковых систем C - и X -диапазонов, одной из особенностей систем миллимет-

рового диапазона являются большие потери в гидрометеорах, которые в пределах России могут составлять от 3 до 20 дБ в зависимости от радиоклиматических параметров районов размещения земных станций.

Необходимость обслуживания больших территорий с разными физико-географическими и климатическими условиями и с разной плотностью пользователей, чьи потребности в пропускной способности могут варьироваться в широких пределах, предопределяет различные требования к частотно-энергетическому ресурсу (ЧЭР). Поэтому одной из основных задач является планирование спутниковой связи. Важнейшей задачей при планировании является определение готовности (неготовности), исходя из заданных требований к пропускной способности и достоверности с учетом ограничений ЧЭР.

В статье предлагается алгоритм, который позволяет упростить вычислительную процедуру определения показателя неготовности по дождевому фактору с учетом заданных требований и ограничений ЧЭР.

Постановка задачи

Известно множество работ, посвященных рассмотрению функционирования и особенностей спутниковых систем связи миллиметрового диапазона. Большинство из них посвящены рассмотрению энергетического расчета. Потери сигнала в гидрометеорах задаются как детерминированный фактор и в качестве дополнительных потерь.

В [1] представлен типовой подход к анализу бюджета спутникового канала – расчету энергетического запаса с учетом требований к отношению энергии сигнала на бит к спектральной плотности мощности шума. На основе полученных результатов возможно определение значения неготовности участка или спутниковой радиолонии. Однако данный расчет предполагает поиск приемлемого соотношения между всеми параметрами, в том числе посредством выбора сигнально-кодовой конструкции (СКК), что приводит к увеличению вычислительной сложности. Вторая проблема заключается в том, что при расчетах не учитывается ограничение ЧЭР.

Наиболее близким к решению поставленной задачи является метод, описанный в [2]. Оценка значения неготовности проводится с использованием зависимостей затухания сигнала от значения процента превышения времени среднего года. В данном методе не учитываются требования к пропускной способности, достоверности и используется итерационный подход, что значительно увеличивает вычислительную емкость.

Основываясь на вышеизложенном, необходима разработка алгоритма для решения задачи определения неготовности участков спутниковой радиолонии, который позволяет снизить вычислительную сложность и при этом позволяет учитывать требования к пропускной способности, достоверности и ограничение ЧЭР.

Неготовность участка спутниковой радиолонии по дождевому фактору

К неготовности спутниковой радиолонии могут приводить события превышения интенсивности выпадения осадков определенных величин в районе ЦЗС или АЗС, а также одновременно. Вероятность последнего события невелика, поэтому данное событие обычно не рассматривается [3, 4].

В [5] приведено выражение для расчета величины затухания радиосигнала в зависимости от процента превышения времени среднего года (процентной вероятности неготовности:

$$A(p) = A_{0,01} \left(\frac{p}{0,01} \right)^l, \quad (1)$$

$$l = -(0,655 + 0,0331 \ln(p) - 0,045 \ln(A_{0,01}) \times \beta(1 - p) \sin \theta),$$

где p – значение неготовности (можно рассматривать как процентную неготовность участка спутниковой радиолонии); $A_{0,01}$ – прогнозируемое значение затухания, превышаемое в течение 0,01 % времени среднего года; β – вспомогательный параметр, зависящий от угла места земной станции θ .

На основе выражения (1) проведем оценку значения неготовности путем перебора значений варьируемых параметров – мощности сигнала на выходе передатчика земной станции и пропускной способности с учетом выполнения требований к отношению энергии сигнала на бит к спектральной плотности шума при заданной вероятности ошибочного приема: $(E_b/N_0) = (E_b/N_0)^*$, где (E_b/N_0) , $(E_b/N_0)^*$ – получаемое и требуемое значение, соответственно. При шаге изменения мощности сигнала $P_i = 1$ Вт, пропускной способности $R_{bj} = 2$ Мбит/с (исходя из максимальной мощности передатчика ЦЗС при работе в многосигнальном режиме 50 Вт), использовании QPSK 1/4, полосе абонентского транспондера 72 МГц, вычислительная сложность может достигать: $\prod_{i=1, j=1}^{n, m} P_i R_{bj}$, т. е. порядка 750 итераций и увеличивается при варьировании СКК из разрешенного множества, соответствующего применяемому стандарту.

В связи с этим для решения поставленной задачи необходимо использование выражения, которое позволяет упростить решение задачи.

Введем следующие обозначения:

$$a = (0,655 - 0,045 \ln(A_{0,01}) \beta \sin \theta);$$

$$b = 0,0331; c = -\beta \sin \theta.$$

Разделив обе части выражения (1) на $A_{0,01}$ и, логарифмировав, получим:

$$\ln \left(\frac{A(p)}{A_{0,01}} \right) = - \left(a + b \ln(p - cp) \ln \left(\frac{p}{0,01} \right) \right).$$

В результате преобразований и замены $x = \ln(p)$ получаем квадратное уравнение:

$$bx^2 + (a + 4,605b - cp)x + \left(4,605(a - cp) + \ln \left(\frac{A(p)}{A_{0,01}} \right) \right) = 0. \quad (2)$$

Рассматривая размещение АЗС на территории Российской Федерации, в соответствии с [3], коэффициент c (через параметр β) примет значение 0. Тогда одним из корней (2) является выражение:

$$p(A) = e^l.$$

$$l = \frac{-(a + 4,605b) + \sqrt{(a + 4,605b)^2 - 4b \left(4,605a + \ln \left(\frac{A}{A_{0,01}} \right) \right)}}{2b} \quad (3)$$

На рисунке 1 представлен график значения процентной неготовности $p(A)$ от прогнозируемых потерь сигнала в гидрометеорах A .

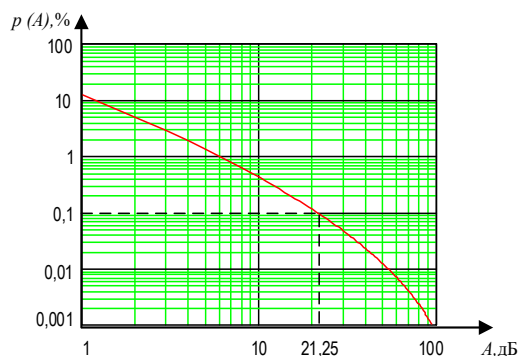


Рис. 1. Зависимость процентной неготовности от прогнозируемых потерь сигнала в гидрометеорах
Fig. 1. Unavailability Percentage in Terms of the Predicted Signal Losses in Hydrometeors

Далее приводятся разработанные алгоритмы оценки неготовности на восходящем и нисходящем участках (рисунки 2, 3).

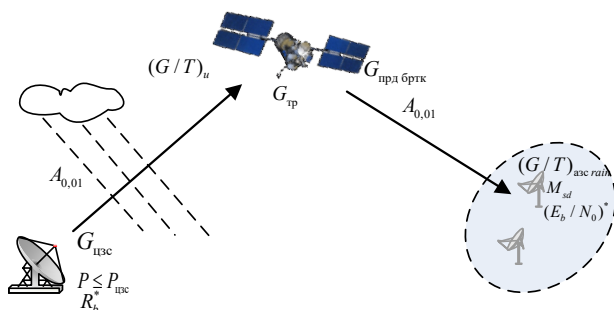


Рис. 2. Модель спутниковой радиосвязи (выпадение осадков в районе ЦЗС)
Fig. 2. Satellite Radio Link Model (the Rain Factor in the Hub Region)

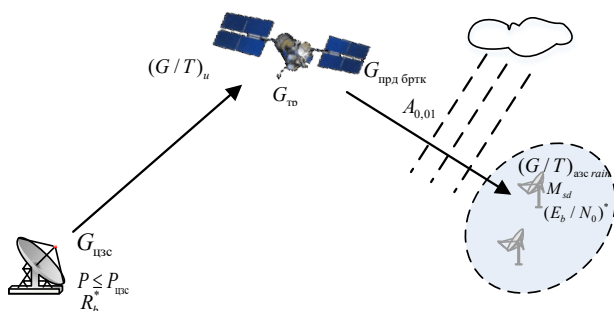


Рис. 3. Модель спутниковой радиосвязи (выпадение осадков в районе АЗС)
Fig. 3. Satellite Radio Link Model (the Rain Factor in the Mobile Terminal Region)

Алгоритм расчета показателя неготовности участков спутниковой радиосвязи

В модели спутниковой радиосвязи, представленной на рисунках 2, 3, рассматривается прямое направление: ЦЗС – АЗС. Энергетический ресурс ЦЗС ограничивается максимальной мощностью передатчика $P \leq P_{\text{цзс}}$, где P – мощность несущей на выходе усилителя мощности ЦЗС; частотным ресурсом выделенной полосы $\Delta f \leq B$, где Δf – полоса, занимаемая сигналом. Выполнение требований к до-

стоверности обеспечивается достижением необходимого значения отношения энергии сигнала на бит к спектральной плотности шума $(E_b/N_0)^*$ при заданной пропускной способности R_b^* , для используемой СКК S_i в соответствии с применяемым стандартом. Значения радиоклиматических параметров районов размещения земных станций в соответствии с [5–15].

Упрощенная структура алгоритма оценки показателей неготовности участков спутниковой радиосвязи представлена на рисунке 4. Расчет показателя неготовности производится по наихудшей точке зоны обслуживания рассматриваемого луча.

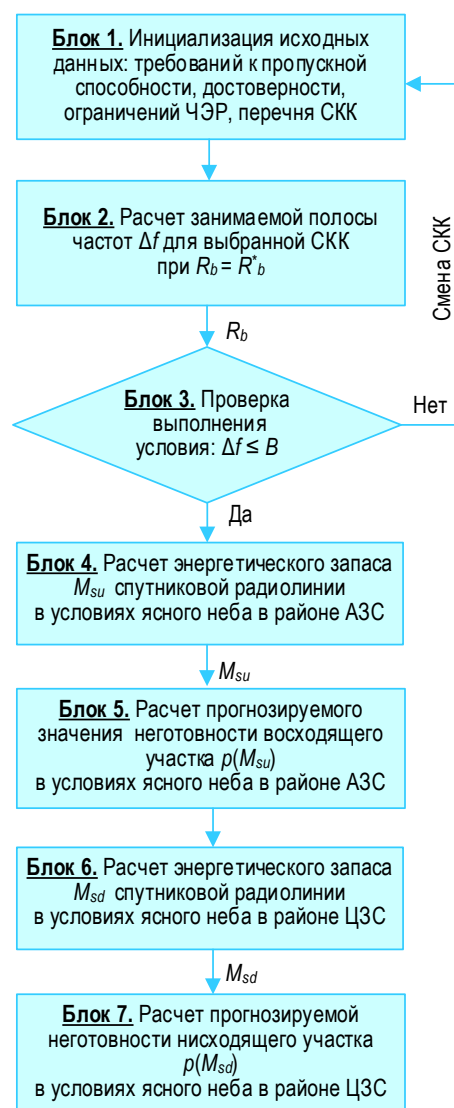


Рис. 4. Упрощенная структура алгоритма оценки показателей неготовности участков спутниковой радиосвязи

Fig. 4. The Simplified Structure of Evaluating Unavailability Algorithm of the Satellite Link Sections

Порядок использования алгоритма заключается в следующем.

В блоке 1 производится инициализация исходных данных: максимальной мощности несущей $P_{\text{цзс}}$ (энергетический ресурс); выделенной полосы ча-

стот B (частотный ресурс); перечня СКК в соответствии с применяемым стандартом S_i ; требований к достоверности: P_n^* или P_b^* – вероятности ошибочного приема информационного пакета или бита, соответственно; требований к пропускной способности R_b^* .

В блоке 2 для выбранной СКК при $R_b = R_b^*$ производится расчет полосы частот, занимаемой сигналом [16] $\Delta f = \frac{R_b^*}{\log_2 M R_c} (1 + \alpha)$, где M – размерность созвездия выбранной СКК, R_c – скорость кодирования, α – ролл-офф фактор (для расчетов принят равным 0,25).

Проверка выполнения критерия $\Delta f \leq B$ производится в блоке 3. При выполнении условия блока осуществляется расчет энергетического запаса M_{su} в условиях ясного неба в районе АЗС (см. рисунок 2) с использованием выражения [16]:

$$M_{su} = \left(\frac{C/N_0}{(E_b/N_0) \cdot R_b} \right) - \left(\frac{\frac{C}{N_0}}{\left(\frac{C}{N_0} \right)_d} \right),$$

где C/N_0 – отношения мощности сигнала к спектральной плотности шума (ОСШ) на входе приемной системы АЗС без учета потерь сигнала в гидрометеорах; $(C/N_0)_d$ – ОСШ на нисходящем участке производится расчет энергетического запаса при заданных ограничениях энергетического ресурса (максимальной мощности несущей).

В противном случае производится смена СКК и повторяются операция в блоках 2, 3.

Расчет прогнозируемого показателя неготовности восходящего участка $p(M_{su})$ с использованием выражения (3) производится в блоке 5.

Операции для расчетов энергетического запаса M_{sd} и показателя неготовности $p(M_{sd})$ нисходящей линии в условиях ясного неба в районе ЦЗС (рисунок 3) осуществляются в блоках 6, 7, соответственно. При расчете данных параметров имеются некоторые особенности, которые рассматриваются в следующем разделе.

Для проведения расчетов используются технические характеристики земных станций и бортового ретрансляционного комплекса современных высокоскоростных спутниковых систем (таблица 1). Основные характеристики некоторых СКК для стандарта DVB-S2 представлены в таблице 2.

Результаты применения алгоритма для расчета показателя неготовности восходящего участка

представлены в графическом виде на рисунках 5а и 5б. При использовании частотно эффективных СКК максимально достижимая пропускная способность на выделенном частотном ресурсе выше (рисунок 5а). Однако необходимо учитывать, что требование к отношению энергии сигнала на бит к спектральной плотности шума ужесточаются при заданном требовании к достоверности, следовательно, и значение показателя неготовности также выше (рисунок 5б).

Требование к пропускной способности выполняется для QPSK 3/4 при использовании выделенного частотного ресурса равного 18,8 МГц, а, например, для 8 PSK 3/4 при $B = 11,1$ МГц (см. рисунок 5а). Однако и показатель неготовности восходящего участка при использовании 8 PSK 3/4 выше и составляет 0,121 % по сравнению с QPSK 3/4 (0,082 %) (см. рисунок 5б). Поэтому при планировании спутниковой связи необходимо учитывать важность предъявляемых требований.

Результаты, представленные на рисунках 5а и 5б, также показывают адекватность работы алгоритма, а его применение в рамках решения задач планирования позволяет упростить процедуру получения необходимых данных.

ТАБЛИЦА 1. Исходные данные

TABLE 1. Initial Data

Технические характеристики, требования и ограничения	Значение
Максимальная мощность несущей (ограничение энергетического ресурса), Вт	12
Выделенная полоса частот (ограничение частотного ресурса), МГц	20
Перечень СКК	согласно стандарта DVB-S2
Требование к пропускной способности, Мбит/с	20
Требование к достоверности, $P_{ош}$	10^{-7}
Требование к отношению энергии сигнала на бит к спектральной плотности шума	технические данные модема CDM-760
Коэффициент недоиспользования мощности по выходу, дБ	ЦЗС – 7; БРТК – 3
Коэффициент усиления антенны, дБ	ЦЗС – 71,4; БРТК – 49,6
Добротность приемной системы, дБ/К	БРТК – 9; АЗС – 23,4
Частота несущей, ГГц	40

ТАБЛИЦА 2. Основные характеристики СКК

TABLE 2. ModCod General Characteristics

СКК	QPSK			8 PSK			16 APSK			32 APSK		
	1/4	2/3	4/5	3/5	3/4	9/10	3/4	5/6	8/9	3/4	5/6	9/10
γ_f , (бит/с)/ Гц	0,48	2,2	1,55	1,74	2,18	2,62	2,9	3,22	3,44	3,62	4,03	4,36
$\gamma_e(E_b/N_0)$, дБ	1,1	1,29	3,0	3,7	4,8	7,1	6,0	6,9	7,8	7,6	8,7	10,1

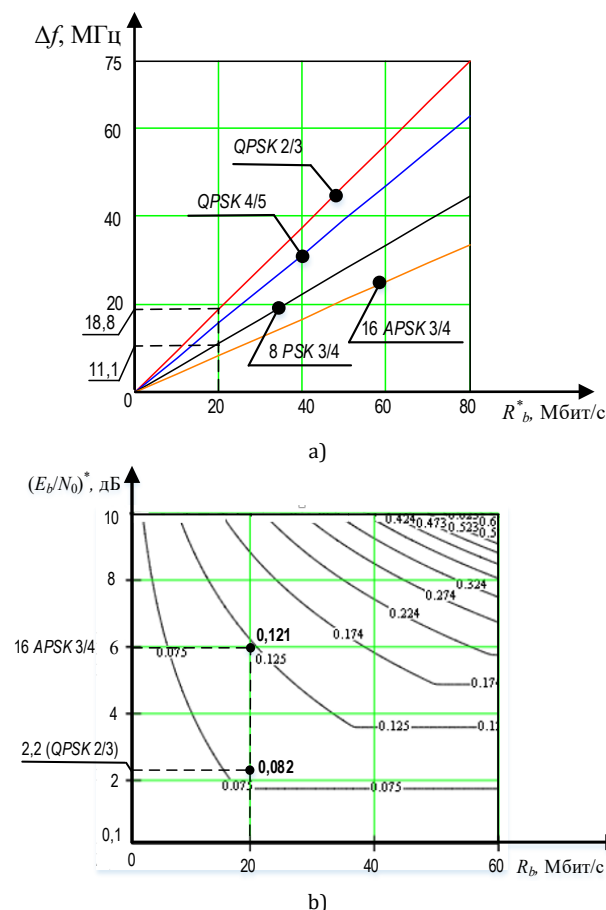


Рис. 5. Зависимость занимаемой полосы от требования к пропускной способности (а) и прогнозируемого значения показателя неготовности от выбранной СКК и требования к пропускной способности (б)

Fig. 5. Occupied Bandwidth in Terms of the Throughput Capacity (a) and Predictable Value of the Unavailability in Terms of the ModCod and Throughput Capacity Requirements (b)

Особенности оценки показателя неготовности нисходящей линии

Отличие процедуры расчета показателя неготовности при выпадении осадков в районе ЦЗС заключается в том, что необходимо учитывать снижение добротности $(G/T)_{\text{азс rain}}$ приемной системы АЗС за счет повышения мощности шумов (шумовой температуры). В некоторых источниках используется термин «мокрая антенна».

Выражение для расчета добротности приемной системы АЗС представляется в следующем виде [16]:

$$(G/T)_{\text{азс rain}} = (G/T)_{\text{азс}} \frac{T_{\text{азс}}}{(T_{\text{азс}} + 273(1 - 1/A(p))), \quad (4)$$

где $(G/T)_{\text{азс}}$ – добротность приемной системы АЗС в условиях ясного неба; $T_{\text{азс}}$ – температура теплового шума приемной системы АЗС.

Выражение для определения энергетического запаса с учетом выражения (4):

$$M_s = \frac{\left(\frac{(C/N_0)_u}{(E_b/N_0)^* R_b} - 1 \right) T_{\text{азс}}}{T_{\text{азс}} + 273}, \quad (5)$$

где $(C/N_0)_u$ – отношение мощности сигнала к спектральной плотности мощности шума на восходящем участке.

Результаты расчета с использованием исходных данных из таблицы 1 представлены в графическом виде на рисунке 6. Анализ расчетов показал, что значение показателя неготовности нисходящей линии выше в 1,75 раза при использовании QPSK 3/4 и в 1,85 раза при использовании 8 PSK 3/4. Следовательно, при планировании спутниковой связи для обеспечения требований к готовности необходимо учитывать энергетический запас для компенсации потерь сигнала при выпадении осадков в районе АЗС.

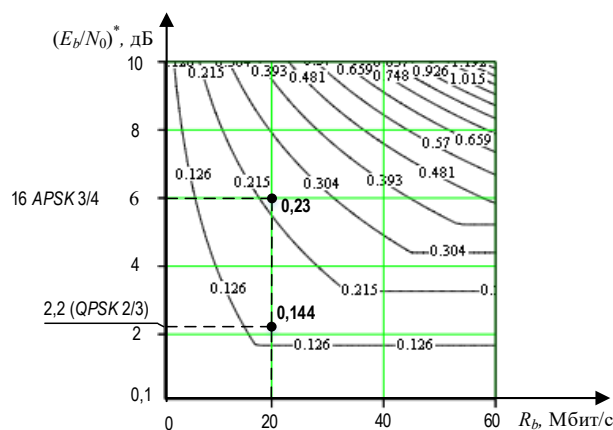


Рис. 6. Зависимость прогнозируемого значения показателя неготовности нисходящего участка от используемой СКК и требования к пропускной способности

Fig. 6. Predictable Value of the Downlink Unavailability in Terms of the ModCod and Throughput Capacity Requirements

Заключение

Разработанный алгоритм может быть использован для оценки показателя неготовности спутниковой радиолнии в целом, а также при рассмотрении дуплексного направления спутниковой связи. Это позволяет осуществлять планирование спутниковой связи, в том числе и при предъявлении требований к готовности как в прямом, так и обратном направлениях, а также при рассмотрении их взаимосвязи. Кроме того, разработанный алгоритм предлагается использовать при проектировании спутниковых систем, в частности, при разработке и обосновании технических требований к наземному сегменту – максимальной мощности передатчика абонентских станций или максимального диаметра антенны с учетом выполнения требования к готовности.

В дальнейшем разработанный алгоритм предлагается использовать при решении задач перераспределения энергетического ресурса – мощности передатчика центральной земной станции между лучами обслуживания с учетом выполнения требований к пропускной способности и достоверности передачи информации.

БЛАГОДАРНОСТИ

Авторы статьи выражают благодарность профессору Санкт-Петербургского государственного университета аэрокосмического приборостроения, доктору технических наук, профессору В.А. Жирову, профессору Военной академии связи, доктору технических наук, профессору С.О. Бурлакову за консультации в процессе обобщения и предоставления материалов для опубликования.

Список используемых источников

1. Скляр Б. Цифровая связь. Теоретические основы и практическое применение. М.: Вильямс, 2003. 1089 с.
2. Sujimol M.R., Rajat A., Gajendra S., Ravikumar G. Rain Attenuation using Ka and Ku-band frequency beacons at Delhi Earth Station // NISCAIR-CSIR. 2015. Т. 44. № 1. С. 45–50. URL: <http://nopr.niscair.res.in/handle/123456789/31326> (дата обращения 02.02.2021)
3. Жиров В.А., Орлов А.Е., Смирнов А.А. Модель радиолинии спутниковой связи в составе высокоскоростной спутниковой системы // Труды учебных заведений связи. 2018. Т. 4. № 3. С. 45–53.
4. Рекомендация МСЭ-R P.1815-1. Дифференциальное ослабление в дожде.
5. Рекомендация МСЭ-R P.618-12. Данные о распространении радиоволн и методы прогнозирования, необходимые для проектирования систем связи Земля – космос.
6. Рекомендация МСЭ-R P.372-11. Радишум.
7. Рекомендация МСЭ-R P.452-12. Процедура прогнозирования для оценки микроволновых помех между станциями, находящимися на поверхности Земли, на частотах выше приблизительно 0,7 ГГц.
8. Рекомендация МСЭ-R P.531-12. Данные об ионосферном распространении радиоволн и методы прогнозирования, необходимые для проектирования спутниковых служб и систем.
9. Рекомендация МСЭ-R P.676-9. Затухание в атмосферных газах.
10. Рекомендация МСЭ-R P.678-3. Оценка изменчивости явлений распространения радиоволн и оценка риска, связанного с запасом на распространение.
11. Рекомендация МСЭ-R P.679-4. Данные о распространении радиоволн, необходимые для проектирования спутниковых радиовещательных систем.
12. Рекомендация МСЭ-R P.681-10. Данные о распространении радиоволн, необходимые для проектирования сухопутных подвижных систем связи Земля-космос.
13. Рекомендация МСЭ-R P.834-8. Влияние тропосферной рефракции на распространение радиоволн.
14. Рекомендация МСЭ-R P.838-3. Модель погонного ослабления в дожде, используемая в методах прогнозирования.
15. Рекомендация МСЭ-R P.840-6. Ослабление из-за облачности и тумана.
16. Tri T.Ha. Digital satellite communication. 2nd ed. Singapore: McGraw-Hill; 1990.

* * *

Evaluating Unavailability Algorithm of the Satellite Link Based on the Rain Factor, Taking into Account the Requirements for Bandwidth, Reliability and Limitation of the Frequency and Energy Resource

A. Orlov¹, K. Novosadov¹, D. Salnikov¹, V. Bezhenar¹

¹Military Signal Academy,
St. Petersburg, 194064, Russian Federation

Article info

DOI:10.31854/1813-324X-2021-7-2-44-50

Received 2nd February 2021

Accepted 11st May 2021

For citation: Orlov A., Novosadov K., Salnikov D., Bezhenar V. Evaluating Unavailability Algorithm of the Satellite Link Based on the Rain Factor, Taking into Account the Requirements for Bandwidth, Reliability and Limitation of the Frequency and Energy Resource. *Proc. of Telecom. Universities*. 2021;7(2):44–50. (in Russ.) DOI:10.31854/1813-324X-2021-7-2-44-50

Abstract: The article presents an algorithm for evaluating the indicator of unavailability for the rain factor of satellite radio link sections, taking into account the requirements for throughput, reliability and limitation of the frequency and energy resource. In contrast to the existing approaches, the proposed problem is based on an analytical expression, which reduces the computational complexity under acceptable constraints.

Keywords: earth station, satellite radio link, signal-code design, unavailability.

ACKNOWLEDGEMENTS

The authors of the article express their gratitude to the professor of the St. Petersburg State University of Aerospace Instrumentation, Doctor of Technical Sciences, Professor V.A. Zhirov, Professor of the Military Academy of Communications, Doctor of Technical Sciences, Professor S.O. Burlakov for consulting in the process of summarizing and providing materials for publication.

References

1. Sklar B. *Digital communications: Fundamentals and Applications*. 2nd ed. Los Angeles: Prentice Hall PTR; 2001.
2. Sujimol M.R., Rajat A., Gajendra S., Ravikumar G. Rain Attenuation using Ka and Ku-band frequency beacons at Delhi Earth Station. *NISCAIR-CSIR*. 2015;44(1):45–50. Available from: <http://nopr.niscair.res.in/handle/123456789/31326> [Accessed 02.02.2021]
3. Zhirov V., Orlov A., Smirnov A. The Model of the Satellite Communication Radio Link in the High-Speed Satellite System. *Proc. of Telecom. Universities*. 2018;4(3):45–53. (in Russ)
4. ITU. Recommendation ITU-R P.1815-1. *Differential rain attenuation*. Geneva: ITU-R; 2009.
5. ITU. Recommendation ITU-R P.618-12. *Propagation data and prediction methods required for the design of Earth-space telecommunication systems*. Geneva: ITU-R; 2015.
6. ITU. Recommendation ITU-R P.372-11. *Radio noise*. Geneva: ITU-R; 2013.
7. ITU. Recommendation ITU-R P.452-12. *Prediction procedure for the evaluation of microwave interference between stations on the surface of the Earth at frequencies above about 0.7 GHz*. Geneva: ITU-R; 2005.
8. ITU. Recommendation ITU-R P.531-12. *Ionospheric propagation data and prediction methods required for the design of satellite services and systems*. Geneva: ITU-R; 2013.
9. ITU. Recommendation ITU-R P.676-9. *Attenuation by atmospheric gases*. Geneva: ITU-R; 2012.
10. ITU. Recommendation ITU-R P.678-3. *Characterization of the variability of propagation phenomena and estimation of the risk associated with propagation margin*. Geneva: ITU-R; 2016.
11. ITU. Recommendation ITU-R P.679-4. *Propagation data required for the design of broadcasting-satellite systems*. Geneva: ITU-R; 2016.
12. ITU. Recommendation ITU-R P.681-10. *Propagation data required for the design of Earth-space land mobile telecommunication systems*. Geneva: ITU-R; 2017.
13. ITU. Recommendation ITU-R P.834-8. *Effects of tropospheric refraction on radiowave propagation*. Geneva: ITU-R; 2016.
14. ITU. Recommendation ITU-R P.838-3. *Specific attenuation model for rain for use in prediction methods*. Geneva: ITU-R; 2005.
15. ITU. Recommendation ITU-R P.840-6. *Attenuation due to clouds and fog*. Geneva: ITU-R; 2013.
16. Tri T.Ha. *Digital satellite communication*. 2nd ed. Singapore: McGraw-Hill; 1990.

Сведения об авторах:

**ОРЛОВ
Алексей Евгеньевич**

кандидат технических наук, преподаватель кафедры военных систем космической, радиорелейной, тропосферной связи и навигации Военной академии связи им. Маршала Советского Союза С.М. Буденного, vorl-2210@mail.ru

**НОВОСАДОВ
Кирилл Сергеевич**

адъюнкт кафедры военных систем космической, радиорелейной, тропосферной связи и навигации Военной академии связи им. Маршала Советского Союза С.М. Буденного, x-bruk@mail.ru

**САЛЬНИКОВ
Денис Владимирович**

кандидат технических наук, доцент, доцент кафедры военных систем космической, радиорелейной, тропосферной связи и навигации Военной академии связи им. Маршала Советского Союза С.М. Буденного, denis_salnikov@mail.ru

**БЕЖЕНАРЬ
Вячеслав Николаевич**

преподаватель кафедры военных систем космической, радиорелейной, тропосферной связи и навигации Военной академии связи им. Маршала Советского Союза С.М. Буденного, 5551541@mail.ru

Позиционирование транспортных средств с комплексированием дальномерных, угломерных и инерциальных измерений в расширенном фильтре Калмана

Г.А. Фокин¹ , А.Г. Владыко¹ 

¹Санкт-Петербургский государственный университет телекоммуникаций им. проф. М.А. Бонч-Бруевича, Санкт-Петербург, 193232, Российская Федерация

*Адрес для переписки: grihafokin@gmail.com

Информация о статье

Поступила в редакцию 08.05.2021

Принята к публикации 31.05.2021

Ссылка для цитирования: Фокин Г.А., Владыко А.Г. Позиционирование транспортных средств с комплексированием дальномерных, угломерных и инерциальных измерений в расширенном фильтре Калмана // Труды учебных заведений связи. 2021. Т. 7. № 2. С. 51–67. DOI:10.31854/1813-324X-2021-7-2-51-67

Аннотация: Настоящая работа посвящена исследованию моделей и методов повышения точности позиционирования в сверхплотных сетях радиодоступа V2X/5G при маневрах транспортных средств за счет комплексирования дальномерных и угломерных первичных измерений с измерениями инерциальных навигационных систем в расширенном фильтре Калмана. Бортовая бесплатформенная инерциальная навигационная система представлена модулями трехосевого акселерометра и гироскопа. Интеграция первичных инерциальных измерений ускорения и угловой скорости с первичными радиотехническими измерениями угла и времени прихода сигнала осуществляются путем преобразования инерциальной системы координат акселерометра и гироскопа в связанную систему координат транспортного средства с использованием кватернионов. Вторичная обработка инерциальных и радиотехнических измерений осуществляется в расширенном фильтре Калмана. Результаты комплексирования показывают повышение точности оценки траектории транспортного средства с нескольких метров до одного метра при повороте на перекрестке.

Ключевые слова: позиционирование, сверхплотные сети 5G, vehicle to everything, дальномерные и угломерные измерения, инерциальная навигационная система, расширенный фильтр Калмана, кватернионы.

1. ВВЕДЕНИЕ

Одними из основных первичных измерений в радиотехнических системах позиционирования являются дальномерные (ДМ, аббр. от англ. Time of Arrival – TOA), разностно-дальномерные (TDOA, аббр. от англ. Time Difference of Arrival и угломерные (УМ, аббр. от англ. Angle of Arrival – AOA) [1, 2]. В сетях мобильной связи четвертого поколения разностно-дальномерные измерения получили распространение для определения местоположения (МП) абонентских станций [3], а УМ измерения – для пеленгации базовых станций [4]. Для позиционирования внутри помещений распространение получили также методы радиокарты, основанные на первичных измерениях индикатора уровня принимаемого сигнала RSSI (аббр. от англ. Received Signal Strength Indication) [5]. Радиокарту внутри помещений можно построить по сигналам RSSI,

принимаемым подвижными устройствами от точек доступа Wi-Fi [6], однако определение МП исключительно методом радиокарты по соотношению трудозатрат и достижимой точности на сегодняшний день оказывается неконкурентоспособным. Для повышения точности позиционирования служит подход, основанный на комплексировании измерений RSSI с измерениями инерциальной навигационной системы (ИНС, аббр. от англ. Inertial Navigation Systems – INS), которой оборудованы современные абонентские устройства [7]. Распространенным инструментом комплексирования разнородных измерений служит подход, основанный на фильтре Калмана (ФК) [8, 9].

Предложенный в 1960 г. [10] Р. Калманом рекурсивный фильтр оценки вектора состояния динамической системы (объекта определения МП) получил широкое распространение в системах позицио-

нирования для вторичной обработки зашумленных первичных измерений; при этом размерность искомого вектора состояния объекта определения МП может превосходить размерность вектора первичных измерений. ФК осуществляет рекурсивную оценку вектора состояния *априорно известной* динамической системы: вектор текущего состояния рассчитывается на основе текущих измерений с использованием предыдущего состояния ФК [10]. Алгоритм ФК включает два этапа: на первом этапе ФК прогнозирует значения переменных вектора состояния; на втором этапе результаты прогноза уточняются [11, 12]. С точки зрения позиционирования подвижных объектов пошаговая работа ФК позволяет прогнозировать траекторию в реальном времени, используя исключительно текущие измерения и вектор предыдущего состояния [13]. Для частного случая гауссовского распределения ошибок первичных измерений ФК дает оптимальную оценку искомым параметрам состояния динамической системы по критерию среднеквадратической ошибки (СКО) [14].

При формализации ФК для конкретного практического приложения используют понятие вектора состояния, включающего параметры, описывающие состояние динамической системы в каждый момент времени. Зная математическую модель наблюдений (измерений) и модель априорного изменения вектора состояния, можно записать уравнение ФК. При нелинейной зависимости оценок координат позиционируемого объекта от первичных измерений для их вторичной обработки используется расширенный фильтр Калмана (РФК, *аббр. от англ.* Extended Kalman Filter – EKF), который осуществляет линеаризацию нелинейных моделей наблюдений с помощью разложения в ряд Тейлора [15].

Использование фильтра Калмана при позиционировании транспортных средств (ТС, *аббр. от англ.* Vehicular User Equipment – VUE) позволяет получать их оптимальные оценки координат (ОК), скорости и/или ускорения в результате вторичной обработки зашумленных первичных измерений в режиме реального времени с непрерывным обновлением оценок по мере движения. В интеллектуальных транспортных системах для сверхплотных сетей радиодоступа V2X/5G решению подлежит задача оценки координат с точностью до 1 м и скорости движения до 150 км/ч [9]. Здесь оценка траектории ТС, включая определение его местоположения (МП), скорости и/или ускорения, осуществляется на основе дальномерных и угломерных первичных измерений, которые собираются опорными пунктами приема (ПП) дорожной инфраструктуры (RSU, *аббр. от англ.* Roadside Unit) с определенной погрешностью измерений времени и угла прихода сигнала. ФК, используя вторичную обработку ДМ/УМ измерений, позволяет прогнозировать траекторию движения ТС с требуемой точностью.

Позиционирование транспортных средств в сверхплотных сетях радиодоступа V2X/5G с использованием комбинирования ДМ и УМ первичных измерений в расширенном фильтре Калмана было исследовано в [9]. Результаты имитационного моделирования (ИМ) показали возможность достижения точности до 1 м для сценария движения ТС по прямой при погрешности в 50 нс для ДМ и 2° для УМ измерений. При маневрировании транспортных средств, в том числе при поворотах на перекрестках, точность позиционирования снижается до нескольких метров (независимо от погрешности первичных измерений), что объясняется фиксированным интервалом их сбора. Проведенный анализ показал, что для повышения точности первичных измерений можно уменьшать периодичность сбора первичных измерений. Однако это не всегда реализуемо на практике. Альтернативным вариантом компенсации ошибок позиционирования на перекрестках является комплексирование радиотехнических измерений с данными бортовых ИНС, которые на интервале поворота позволяют компенсировать неточность ДМ и УМ измерений.

Целью настоящей работы является *разработка и исследование моделей и методов повышения точности позиционирования в сверхплотных сетях радиодоступа V2X/5G при повороте транспортных средств на перекрестках за счет комплексирования ДМ и УМ первичных измерений с измерениями ИНС в РФК*.

Для достижения поставленной цели далее решаются следующие частные задачи. В разделе 2 представлена формализация обобщенной модели ФК. В разделе 3 приводятся особенности использования ИНС применительно к задаче позиционирования ТС, в том числе, преобразование инерциальной системы координат (СК) акселерометра и гироскопа в связанную СК ТС с использованием кватернионов. Раздел 4 содержит математические модели комплексирования радиотехнических ДМ и УМ измерений с измерениями ИНС в РФК. В разделе 5 представлены результаты ИМ позиционирования ТС с использованием ДМ/УМ/ИНС измерений в РФК для сценария поворота на перекрестке. Раздел 6 содержит выводы и направления дальнейших исследований.

2. ОБОБЩЕННАЯ МАТЕМАТИЧЕСКАЯ МОДЕЛЬ ФИЛЬТРА КАЛМАНА

Обобщенная математическая модель ФК формализует работу дискретизированной по времени линейной динамической системы, которая описывается *вектором состояния*. Линейный оператор, воздействуя в каждый дискретный момент времени на *вектор состояния*, переводит его в другой вектор состояния, добавляя *вектор шума* и *вектор управления*. Оценка вектора состояния динамиче-

ской системы производится на основе зашумленных измерений и может быть формализована структурой фильтра, описываемого матричным уравнением (рисунок 1). Переменными уравнения в каждый дискретный момент времени k являются

следующие матрицы: F_k – матрица эволюции системы; H_k – матрица измерений; Q_k – ковариационная матрица системы; R_k – ковариационная матрица шума измерений; B_k – матрица коэффициентов управляющих воздействий.

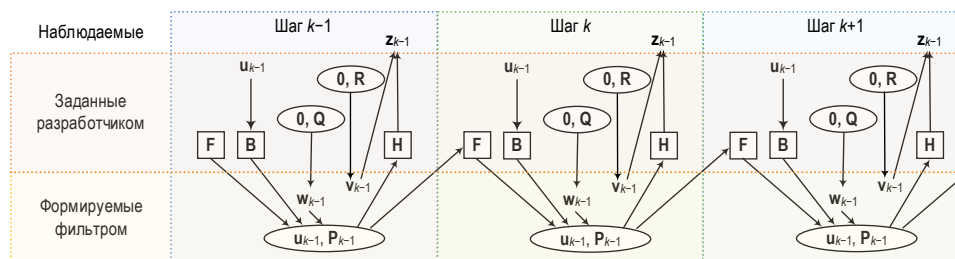


Рис. 1. Обобщенная математическая модель ФК

Fig. 1. Generalized Mathematical Model of the Kalman Filter

Матричное уравнение ФК связывает вектор состояния динамической системы x_k в дискретный момент времени k с ее состоянием x_{k-1} в момент $k-1$:

$$x_k = F_k x_{k-1} + B_k u_k + w_k, \quad (1)$$

где u_k – вектор управляющих воздействий, к которому прикладывается матрица B_k ; w_k – вектор эволюции системы, имеющий нормальное распределение с нулевым математическим ожиданием и ковариационной матрицей Q_k :

$$w_k \sim N(0, Q_k). \quad (2)$$

Вектор состояния динамической системы x_k в дискретный момент времени k связан с вектором первичных измерений z_k соотношением:

$$z_k = H_k x_k + v_k, \quad (3)$$

где матрица измерений H_k связывает вектор состояния x_k с вектором измерений z_k ; v_k – вектор шума измерений с нулевым математическим ожиданием и ковариационной матрицей R_k :

$$v_k \sim N(0, R_k). \quad (4)$$

Рекурсивное свойство ФК проявляется в зависимости оценки вектора состояния системы $\hat{x}_k$ в текущий момент времени k от оценки вектора состояния $\hat{x}_{k-1}$ и погрешности этой оценки в предыдущий момент времени $k-1$, а также от вектора измерения z_k в текущий момент времени k . Введем следующие переменные для формализации состояний ФК: $\hat{x}_k^+$ – апостериорная оценка состояния системы в момент k , полученная по результатам измерений вплоть до момента k ; P_k^+ – апостериорная ковариационная матрица ошибок оценки вектора состояния $\hat{x}_k^+$, включающая дисперсии погрешности вычисленного состояния и ковариации параметров состояния системы. В каждый момент времени k работа ФК включает два этапа: этап прогноза и этап коррекции. На *этапе прогноза* (предсказания) предварительная оценка состояния $\hat{x}_k^-$ в момент k получается по (апостериорной) оценке состояния $\hat{x}_{k-1}^+$ в предыдущий момент вре-

мени $k-1$; оценка $\hat{x}_k^-$ называется *априорной*, так как не использует текущие измерения. На *этапе коррекции* априорная оценка дополняется текущими измерениями; скорректированная оценка $\hat{x}_k^+$ называется *апостериорной* оценкой состояния. С точки зрения позиционирования ТС по радиотехническим и инерциальным измерениям *прогноз* текущего состояния ТС может производиться на основе измерений, собранных в предыдущий момент времени, а коррекция осуществляется по текущим доступным измерениям. При этом, если в какой-то момент времени радиотехнические (апостериорные) измерения становятся недоступными, этап коррекции может быть пропущен, а прогноз выполнен по (априорным) измерениям ИНС.

Этап прогноза (предсказание) априорной оценки вектора состояния $\hat{x}_k^-$ и ковариационной матрицы P_k^- в момент времени k определяются выражениями:

$$\hat{x}_k^- = F_k \hat{x}_{k-1}^+ + B_k u_k, \quad (5)$$

$$P_k^- = F_k P_{k-1}^+ F_k^T + Q_k. \quad (6)$$

Этап коррекции определяется следующими выражениями.

Вектор отклонения (ошибки) полученного в момент k измерения z_k от априорной оценки $\hat{x}_k^-$:

$$y_k = z_k - H_k \hat{x}_k^-. \quad (7)$$

Ковариационная матрица S_k вектора отклонения (ошибки) y_k :

$$S_k = H_k P_k^- H_k^T + R_k. \quad (8)$$

Матрица коэффициентов усиления по Калману получается из ковариационной матрицы этапа предсказания P_k^- и вектора отклонения S_k :

$$K_k = P_k^- H_k^T S_k^{-1}. \quad (9)$$

Коррекция априорной оценки вектора состояния $\hat{x}_k^-$ и ковариационной матрицы P_k^- с получением апостериорной оценки вектора состояния $\hat{x}_k^+$ и ковариационной матрицы P_k^+ в момент времени k определяются выражениями:

$$\hat{\mathbf{x}}_k^+ = \hat{\mathbf{x}}_k^- + \mathbf{K}_k \mathbf{y}_k, \quad (10)$$

$$\mathbf{P}_k^+ = (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \mathbf{P}_k^-, \quad (11)$$

где $\mathbf{I}$ – единичная матрица.

Для позиционирования ТС при его прямолинейном движении с постоянным ускорением a_k (имеющим нормальное распределение с нулевым математическим ожиданием и СКО σ_a) между измерениями k и $k-1$ и периодом их сбора Δt оценка вектора состояния $\mathbf{x}_k$ определяется (значок шляпки далее опускается):

$$\mathbf{x}_k = \mathbf{F} \mathbf{x}_{k-1} + \mathbf{G} a_k, \quad (12)$$

где вектор состояния ФК в $1D$ $\mathbf{x}_k = [x, v_x]^T$ включает координату x и скорость v_x , вектор управления вырождается в скаляр a_k , матрица эволюции системы представлена в виде:

$$\mathbf{F} = \begin{bmatrix} 1 & \Delta t \\ 0 & 1 \end{bmatrix}, \quad (13)$$

матрица управления (коэффициентов управляющих воздействий) – вектор:

$$\mathbf{G} = \begin{bmatrix} \Delta t^2/2 \\ \Delta t \end{bmatrix}. \quad (14)$$

Представленную в настоящем разделе модель ФК далее уточним для РФК при комплексировании ДМ/УМ/ИНС измерений в разделе 4.

3. ИСПОЛЬЗОВАНИЕ ИНС В ЗАДАЧАХ ПОЗИЦИОНИРОВАНИЯ

3.1. Особенности использования ИНС

Рассмотрим особенности использования ИНС в задачах позиционирования [16–23]. Оценка координат ТС осуществляется с использованием РФК на основе ДМ/УМ измерений, а также измерений от модулей ИНС (IMU, аббр. от англ. Inertial Measurement Unit). Первичные ДМ/УМ измерения собираются стационарными ПП дорожной инфраструктуры, положение которых известно; число ПП в каждый момент движения ТС определяется дальностью связи между ПП и ТС.

В качестве модулей IMU используются *микроэлектромеханические системы*, объединяющие в себе микроэлектронные и микромеханические устройства: механическим устройством может быть инерциальный датчик (акселерометр, датчик угловых скоростей, гироскоп, барометрический датчик). В настоящей работе в качестве датчиков рассматривается применение микромеханических гироскопов и микромеханических акселерометров. В РФК ИНС используется на этапе формирования *априорной оценки* на основе первичных измерений *3-осевого акселерометра* и *3-осевого гироскопа*; при этом гироскопы обеспечивают ориентацию акселерометров в пространстве.

Комплексирование радиотехнических ДМ/УМ измерений с измерениями ИНС оправдано в случаях, когда использование глобальной навигационной

спутниковой системы (ГНСС) невозможно или затруднено, а также в сценариях, когда ДМ/УМ измерения оказываются недоступными; примером является движение ТС в тоннелях, а также применение промышленного транспорта на предприятиях, складах и иных объектах, где в явном виде использование радиосигналов проблематично в силу отсутствия условий прямой видимости [24–29].

В отличие от *неавтономных радиотехнических методов* ДМ/УМ измерений ИНС функционирует без использования внешних источников/приемников радиосигналов; оценка траектории объекта с использованием ИНС основана на его инерции и реализуется *бортовыми средствами IMU автономно*. Наиболее распространенными модулями IMU являются датчики линейного ускорения (*акселерометры*) и угловой скорости (*гироскопы*); измерив ускорение и угловую скорость вычислительный модуль IMU путем интегрирования может определить координаты подвижного объекта. ИНС, получив углы ориентации (курс, тангаж и крен) может определить отклонение локальной СК, связанной с корпусом датчика IMU, от глобальной СК, связанной с Землей.

Таким образом, работа ИНС заключается в измерении ускорений и угловых скоростей подвижного объекта относительно трех осей СК для того, чтобы, исходя из этих данных, определить МП и скорость объекта; акселерометры измеряют линейное ускорение, а гироскопы определяют углы тангажа, рысканья и крена подвижного объекта относительно осей СК.

В зависимости от реализации горизонтальной СК, на объекте различают *платформенные ИНС с гиростабилизированной платформой* и *бесплатформенные*. Различие между ними заключается в том, что у платформенных ИНС есть карданов подвес – шарнирная опора, позволяющая закрепленному в ней телу вращаться в нескольких плоскостях одновременно; при этом закрепленное вращающееся тело сохраняет направление оси вращения независимо от ориентации подвеса. Из достоинств платформенных ИНС можно отметить ее высокую точность и отсутствие необходимости пересчета из связанной с корпусом датчика IMU (локальной) СК в необходимую (связанную) систему координат. Недостатком платформенных ИНС являются ограничение в измерении углов поворота, увеличенные массогабаритные показатели, усложнение конструкции, а также большая стоимость. В бесплатформенных ИНС отсутствует карданов подвес, измерительные датчики микромеханических гироскопов и акселерометров жестко связаны с подвижным объектом позиционирования. В настоящее время наибольшее распространение получили бесплатформенные ИНС, обладающие преимуществами неограниченных углов измере-

ний, высокой информативности, надежности и устойчивости к вибрационным и ударным воздействиям, а также уменьшенными массогабаритными параметрами и энергопотреблением [23]. Недостатком бесплатформенных ИНС является рост погрешности даже для неподвижного объекта позиционирования. Далее рассмотрим особенности комплексирования инерциальных с радиотехническими измерениями (РТИ) для компенсации нарастающей погрешности ИНС.

3.2. Особенности комплексирования измерений ИНС

Повышение точности позиционирования с использованием бесплатформенных ИНС реализуется интеграцией автономных ИНС измерений с неавтономными ДМ/УМ измерениями, что позволяет компенсировать свойственные бесплатформенным ИНС накапливающиеся погрешности; комплексирование ДМ/УМ/ИНС измерений при этом осуществляется с использованием расширенного ФК. Достаточно хорошо исследованным направлением комплексирования является компенсация неточностей бесплатформенных ИНС за счет сигналов ГНСС, так как последние не имеют накапливающихся ошибок. Выделяют следующие схемы комплексирования бесплатформенных ИНС с ГНСС: а) раздельная; б) слабо связанная; в) жестко связанная; г) глубоко интегрированная [24]. Приведем особенности данных схем для уточнения рассматриваемого далее способа комплексирования РТИ/ИНС измерений [23, 24].

В раздельной схеме компенсация нарастающей со временем погрешности бесплатформенной ИНС осуществляется ее периодической инициализацией новыми оценками координат и скорости от ГНСС. Недостатком раздельной схемы для высокоподвижных объектов является редкая периодичность сигналов коррекции ГНСС.

В слабо связанной схеме бесплатформенная ИНС и спутниковая навигационная система (СНС) работают независимо. Для коррекции ОК, полученных от ГНСС, используется интегральный ФК, который, работая независимо, позволяет уточнять текущие оценки координат СНС при утрате сигналов ГНСС за счет измерений бесплатформенной ИНС, собранных в предыдущий момент. Недостаток слабо связанной схемы – взаимная корреляция ошибок ОК на выходе приемника СНС.

В жестко связанной схеме бесплатформенной ИНС осуществляет сбор первичных измерений поступательного и вращательного движений, которые вместе с первичными измерениями СНС поступают на фильтр Калмана для вторичной обработки. Данная схема позволяет повысить точность позиционирования по сравнению с раздельной и слабо связанной схемой, исключает взаимную корреляцию ОК, а также снимает проблему син-

хронизации работы бесплатформенной ИНС/СНС. Недостатком жестко связанной схемы является ее низкая надежность, так как при отсутствии первичных измерений ИНС в результате выхода из строя бесплатформенной ИНС система вторичной обработки ФК перестает работать.

В глубоко интегрированной схеме, в отличие от жестко связанной, обработка в ФК, помимо вторичных оценок координат и скорости подвижного объекта, осуществляет вычисление погрешностей бесплатформенной ИНС и первичных ДМ измерений СНС, что позволяет достичь максимальной точности. Недостатком глубоко интегрированной схемы является высокая размерность вектора состояния ФК, что увеличивает вычислительную нагрузку на процедуры вторичной обработки.

Таким образом, рассмотрев известные схемы комплексирования бесплатформенной ИНС с радиотехнической системой позиционирования ГНСС, можно отнести исследуемую далее схему комплексирования ДМ/УМ/ИНС к жестко связанной схеме, в которой ИНС и радиотехнические измерения ДМ/УМ обеспечивают вектор первичных измерений z (3) для вторичной обработки в РФК. Рисунок 2 иллюстрирует обобщенную схему комплексирования первичных измерений в РФК для случая оценки траектории ТС на плоскости в 2D.

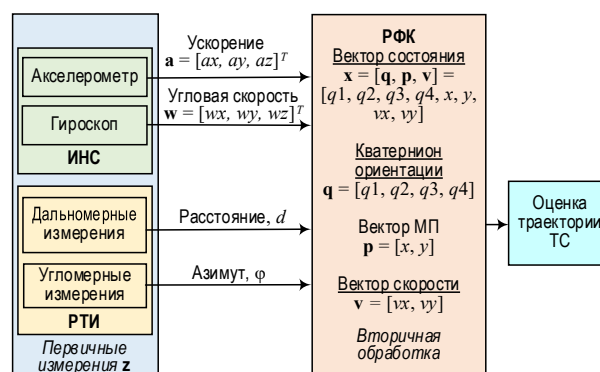


Рис. 2. Обобщенная схема комплексирования ДМ/УМ/ИНС в РФК

Fig. 2. Generalized Scheme for Integrating TOA/AOA/INS in EKF

Для интеграции первичных инерциальных измерений ускорения и угловой скорости с первичными радиотехническими измерениями угла и времени прихода сигнала в единый вектор первичных измерений z (3) необходимо преобразование инерциальной системы координат акселерометра и гироскопа в связанную систему координат транспортного средства.

3.2. Инерциальная и связанная СК

В рассматриваемой далее бесплатформенной ИНС акселерометры и гироскопы жестко связаны с корпусом датчика IMU объекта, поэтому для комплексирования ИНС и ДМ/УМ измерений необходим пересчет из локальной инерциальной СК в связанную СК ТС. Связанная СК традиционно исполь-

зуется для анализа движения летательных аппаратов и может быть применима к прогнозу траектории ТС. Связанная СК в общем случае включает продольную Ox , поперечную Oy и вертикальную Oz оси, проходящие через центр масс объекта – микророзэлектрохимические системы датчика или ТС (рисунок 3а).

Продольной Ox является ось транспортного средства (рисунок 3б); при вращении вокруг нее ТС совершает движение под углом *крена* (*roll*). Вертикальной Oz является ось, лежащая в плоскости симметрии ТС и перпендикулярная его продольной оси Ox ; при вращении вокруг вертикальной оси ТС совершает движение под углом *рыскания* (*yaw*) – поворачивает влево или вправо. Поперечной Oy является ось, перпендикулярная плоскости симметрии ТС, направленная в сторону его правого крыла; при вращении вокруг поперечной оси ТС совершает движение под углом *тангажа* (*pitch*) – опускает и поднимает нос. Рисунок 3 иллюстрирует преобразование бесплатформенной инерциальной СК IMU в связанную СК подвижной платформы – транспортного средства VUE.

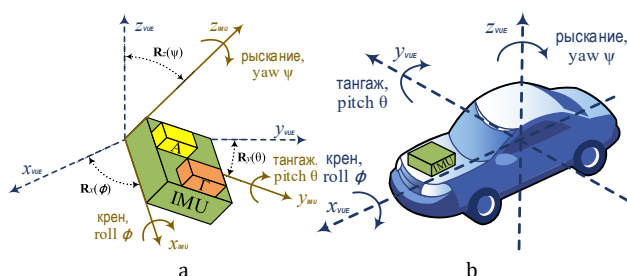


Рис. 3. Преобразование из инерциальной СК IMU (а) в связанную СК VUE (б)

Fig. 3. Conversion from IMU Center-of-Mass System to VUE Frame of Reference

Рассмотрим особенности преобразования СК в специальном программном обеспечении (СПО) Matlab Sensor Fusion and Tracking Toolbox [30–32], которое позволяет отслеживать ориентацию, положение и траекторию позиционируемого подвижного объекта, называемого далее *платформой*.

Ориентация в Matlab Sensor Fusion and Tracking Toolbox определяется угловым смещением/перемещением и может быть описана в терминах поворота отдельной точки или всей СК. В первом

случае СК статична – поворачивается точка. При повороте СК точка статична, поворачивается СК. При заданных осях и углах поворота угловое смещение точки эквивалентно угловому смещению СК с учетом противоположных направлений смещения. В СПО Matlab по умолчанию используется поворот СК: ориентация определяется поворотом исходной/начальной/родительской СК в целевую/конечную/дочернюю СК. Выбор родительской СК зависит от практического приложения. Например, поворот СК независимых датчиков IMU необходим для выравнивания их измерений в единой инерциальной СК. Отслеживание СК платформы используется в задачах их стабилизации. Глобальная СК используется для одновременного сопровождения нескольких платформ.

Для обработки измерений, собранных в разных изначальных ориентациях, используют поворот их СК в терминологию *связанной СК*. Таблица 1 содержит соотношение осей и углов связанной СК.

ТАБЛИЦА 1. Соотношение осей и углов связанной СК

TABLE 1. Inter-Relation of Axes and Angles for Frame of Reference

Ось поворота	Угол поворота (Эйлера)	Обозначение	Диапазон значений
z	Рыскание/Yaw	ψ	$-180^\circ \leq \psi < 180^\circ$
y	Тангаж/Pitch	θ	$-90^\circ \leq \theta \leq 90^\circ$
x	Крен/Roll	ϕ	$-180^\circ \leq \phi < 180^\circ$

Поворот по часовой стрелке вокруг оси в ее положительном направлении образует положительный угол вращения относительно данной оси. Поворот СК в пространстве осуществляется последовательным поворотом относительно каждой из осей СК. Так, например, для поворота родительской СК в последовательности ZYX следует выполнить ее поворот сначала вокруг вертикальной оси Oz на угол ψ с использованием матрицы поворота $R_z(\psi)$, затем вокруг поперечной оси Oy с использованием матрицы поворота $R_y(\theta)$ и в заключении вокруг продольной оси Ox с использованием матрицы поворота $R_x(\phi)$; при этом после каждого поворота оси СК пересчитываются перед очередным поворотом.

Итоговая матрица поворота вектора из родительской в дочернюю СК для заданных углов поворота (рыскание ψ , тангаж θ , крен ϕ) определяется выражением:

$$R(\psi, \theta, \phi) = R_x(\phi)R_y(\theta)R_z(\psi) = \begin{bmatrix} \cos \psi \cos \theta & \sin \psi \cos \theta & -\sin \theta \\ \cos \psi \sin \theta \sin \phi - \sin \psi \cos \phi & \sin \psi \sin \theta \sin \phi + \cos \psi \cos \phi & \cos \theta \sin \phi \\ \cos \psi \sin \theta \cos \phi + \sin \psi \cos \phi & \sin \psi \sin \theta \cos \phi + \cos \psi \cos \phi & \cos \theta \cos \phi \end{bmatrix}. \quad (15)$$

В СПО Matlab Sensor Fusion and Tracking Toolbox координаты обрабатываются массивами размерности $R^{N \times 3}$, где число строк N – число выборок первичных измерений, взятых с интервалом Δt , а число столбцов 3 соответствует переменным СК (x, y, z).

Получение вектора a_{child} в дочерней СК из вектора a_{parent} в родительской СК определяется формулой:

$$a_{\text{child}} = (R(\psi, \theta, \phi) \times a_{\text{parent}}^T)^T. \quad (16)$$

Для исключения использования встроенных функций СПО Matlab Sensor Fusion and Tracking Toolbox и реализации процедур вторичной обработки через примитивы рассмотрим далее особенности использования кватернионов при повороте систем координат [33–37].

3.3. Использование кватернионов при повороте систем координат

Использование кватернионов представляется более удобным для ориентации и поворота СК в пространстве по сравнению с матрицей поворота на углы Эйлера. Кватернионы формализуют алгебраическое представление поворота в 3D и, как система гиперкомплексных чисел, образуют векторное пространство размерностью $\mathbb{R}^4$ над полем вещественных чисел и могут быть формализованы:

$$\mathbf{q} = w + xi + yj + zk, \quad (17)$$

где w, x, y, z – вещественные числа; i, j, k – мнимые единицы со свойством:

$$i^2 = j^2 = k^2 = ijk = -1, \quad (18)$$

при этом результат их попарного произведения зависит от порядка следования (не является коммутативным):

$$ij = k, \quad ji = -k, \quad (19)$$

мнимая часть $xi + yj + zk$ кватерниона $\mathbf{q}$ ведет себя так же, как и вектор $(x, y, z) \in \mathbb{R}^3$, а вещественная часть w ведет себя так же, как скаляр в $\mathbb{R}$.

Определим пространство вращения: пусть точка (w, x, y, z) представляет вращение вокруг осей (x, y, z) на угол:

$$\alpha = 2 \cos^{-1} w = 2 \sin^{-1} \sqrt{x^2 + y^2 + z^2}, \quad (20)$$

тогда кватернион (17) можно представить выражением:

$$\mathbf{q} = w + xi + yj + zk = w + (x, y, z) = \cos \alpha/2 + \mathbf{u} \sin \alpha/2, \quad (21)$$

где $\mathbf{u}$ – единичный вектор.

Поворот вектора $\mathbf{v}$ вокруг оси $\mathbf{u}$ на угол α по часовой стрелке по направлению вектора $\mathbf{u}$ можно представить произведением:

$$\mathbf{v}_\alpha = \mathbf{q} \mathbf{v} \mathbf{q}^{-1}, \quad (22)$$

Обратный кватернион $\mathbf{q}^{-1}$ соответствует вращению вектора $\mathbf{v}_\alpha$ в противоположном направлении:

$$\mathbf{q}^{-1} (\mathbf{q} \mathbf{v} \mathbf{q}^{-1}) \mathbf{q} = \mathbf{v}. \quad (23)$$

Обратный кватернион определяется формулой:

$$\mathbf{q}^{-1} = \mathbf{q}^* / \|\mathbf{q}\|^2, \quad (24)$$

где $\|\mathbf{q}\|$ – модуль кватерниона, определяемый как:

$$\|\mathbf{q}\| = \sqrt{\mathbf{q} \mathbf{q}^*} = \sqrt{w^2 + x^2 + y^2 + z^2}. \quad (25)$$

Для единичного кватерниона присуще: $\|\mathbf{q}\| = 1$.

Кватернион $\mathbf{q}$ также можно определить как вещественную матрицу:

$$\mathbf{q} \rightarrow \mathbf{Q} = \begin{bmatrix} w & -x & -y & -z \\ x & w & -z & y \\ y & z & w & -y \\ z & -y & x & w \end{bmatrix}. \quad (26)$$

При матричном представлении сопряженному кватерниону соответствует транспонированная матрица $\mathbf{q}^* \rightarrow \mathbf{Q}^T$:

$$\mathbf{q}^* \rightarrow \mathbf{Q}^T = \begin{bmatrix} w & x & y & z \\ -x & w & z & -y \\ -y & -z & w & x \\ -z & y & -y & w \end{bmatrix}. \quad (27)$$

Из (22, 24–26) для нормированного кватерниона $\|\mathbf{q}\| = 1$ поворот вектора $\mathbf{v} = [v_1, v_2, v_3, v_4]^T$ можно представить через кватернионы в матричном виде выражением:

$$\mathbf{v}_\alpha = \begin{bmatrix} w & -x & -y & -z \\ x & w & -z & y \\ y & z & w & -y \\ z & -y & x & w \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{bmatrix} = \begin{bmatrix} w & -x & -y & -z \\ x & w & -z & y \\ y & z & w & -y \\ z & -y & x & w \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{bmatrix}, \quad (28)$$

Матричное представление (28) поворота вектора далее будет использовано в преобразовании из инерциальной СК IMU в связанную СК VUE при комплексировании ДМ/УМ/ИНС измерений в РФК.

4. МАТЕМАТИЧЕСКИЕ МОДЕЛИ КОМПЛЕКСИРОВАНИЯ ДМ/УМ/ИНС В РФК

РФК позволяет оценить вектор состояния динамической системы (1), используя модель изменения состояния динамической системы $\mathbf{F}_k$ и ряд измерений (3), косвенно описывающих состояние системы. Алгоритм работы РФК может быть описан в два этапа.

Этап 1. Получение априорной оценки вектора состояния $\mathbf{x}$ (5) и ковариационной матрицы ошибок оценки вектора состояния $\mathbf{P}$ (6) для шага k на основе данных с предыдущего шага $k-1$.

Этап 2. Обновление, на котором формируется апостериорная оценка вектора состояния $\mathbf{x}$ (10) и ковариационной матрицы ошибок $\mathbf{P}$ (11) с учетом новых измерений $\mathbf{z}$ (3).

Априорная оценка вектора состояния $\mathbf{x}$ формируется на основе измерений ИНС:

- измерения *угловой скорости* $\mathbf{w}$ по трем осям, полученные от *гироскопа*, позволяют получить *ориентацию* объекта $\mathbf{q}$ путем интегрирования этих измерений;

- измерения *ускорений* $\mathbf{a}$ по трем осям, полученные от *акселерометра*, позволяют получить *скорость перемещения* $\mathbf{v}$ и само *перемещение* объекта по трем осям путем двойного интегрирования.

Измерения *гироскопа* $\mathbf{w}$ и *акселерометра* $\mathbf{a}$ формируются в локальной инерциальной СК; предпо-

лагается, что начальная ориентация и местоположение ТС в глобальной СК известны с некоторой точностью. Вектор состояния включает в себя следующие величины:

$$\mathbf{x} = [\mathbf{q}, \mathbf{p}, \mathbf{v}]^T = [q_1, q_2, q_3, q_4, x, y, v_x, v_y]^T, \quad (29)$$

где $\mathbf{q} = [q_1, q_2, q_3, q_4]^T$ – кватернион, описывающий ориентацию ТС; $\mathbf{p} = [x, y]^T$ и $\mathbf{v} = [v_x, v_y]^T$ – векторы координат и скоростей ТС, соответственно.

Априорная оценка ориентации ТС формируется путем интегрирования измерений угловой скорости $\mathbf{w}$ в форме кватернионов [26]:

$$\mathbf{q}_k^- = \mathbf{q}_{k-1}^+ + \frac{\Delta t}{2} \mathbf{w}_k \otimes \mathbf{q}_{k-1}^+, \quad (30)$$

$$\mathbf{w}_k = [0, w_x, w_y, w_z]^T, \quad (31)$$

$$\mathbf{q}_{k-1}^+ = [q_1^+, q_2^+, q_3^+, q_4^+]^T. \quad (32)$$

где $\mathbf{w}_k$ – кватернион, составленный из измерений гироскопа w_x, w_y, w_z по трем осям; Δt – период измерений ИНС; $\otimes$ – операция умножения кватернионов; здесь и далее верхние индексы «-» и «+» обозначают априорную и апостериорную оценки, соответственно.

Второе слагаемое выражения (30) является дифференциальным уравнением кватерниона, которое можно выразить через умножение кватернионов в матричном виде (для упрощения индексы не показаны) [26]:

$$\begin{aligned} \dot{\mathbf{q}}_k &= \frac{1}{2} \mathbf{w}_k \otimes \mathbf{q}_{k-1}^+ = \\ &= \frac{1}{2} \begin{bmatrix} 0 & -w_x & -w_y & -w_z \\ w_x & 0 & w_z & -w_y \\ w_y & -w_z & 0 & w_x \\ w_z & w_y & -w_x & 0 \end{bmatrix} \begin{bmatrix} q_1^+ \\ q_2^+ \\ q_3^+ \\ q_4^+ \end{bmatrix}. \end{aligned} \quad (33)$$

Априорная оценка вектора ориентации (30) с учетом (31–33) примет вид:

$$\mathbf{q}_k^+ = \begin{bmatrix} q_1^+ + \frac{\Delta t}{2} (-q_2^+ w_x - q_3^+ w_y - q_4^+ w_z) \\ q_2^+ + \frac{\Delta t}{2} (q_1^+ w_x + q_3^+ w_z - q_4^+ w_y) \\ q_3^+ + \frac{\Delta t}{2} (q_1^+ w_y - q_2^+ w_z + q_4^+ w_x) \\ q_4^+ + \frac{\Delta t}{2} (q_1^+ w_z + q_2^+ w_y - q_3^+ w_x) \end{bmatrix}. \quad (34)$$

В процессе интегрирования амплитуда кватерниона $\mathbf{q}$ может стать отличной от 1, поэтому для корректного выполнения последующей операции вращения необходимо выполнить операцию нормировки кватерниона (25):

$$\|\mathbf{q}\| = \sqrt{q_1^2 + q_2^2 + q_3^2 + q_4^2}. \quad (35)$$

Для интегрирования измерения акселерометра $\mathbf{a}_k$ должны быть представлены в СК VUE – необхо-

димо кватернионом $\mathbf{q}_k^-$ повернуть вектор ускорений $\mathbf{a}_k = [0, a_x, a_y, a_z]$ согласно выражению:

$$\mathbf{a}_k^g = \mathbf{q}_k^{-*} \otimes (\mathbf{q}_k^- \otimes \mathbf{a}_k) = [0, a_x^g, a_y^g, a_z^g]^T, \quad (36)$$

где $\mathbf{a}_k^g$ – вектор ускорений в СК VUE; $\mathbf{q}_k^-$ – кватернион; $\mathbf{q}_k^{-*}$ – комплексно-сопряженный кватернион.

Сначала выполняется умножение кватернионов ($\mathbf{q}_k^- \otimes \mathbf{a}_k$) в матричном виде (28), результатом которого является вектор размерности 4×1 :

$$\begin{aligned} \mathbf{q}_k^- \otimes \mathbf{a}_k &= \begin{bmatrix} 0 & -a_x & -a_y & -a_z \\ a_x & 0 & a_z & -a_y \\ a_y & -a_z & 0 & a_x \\ a_z & a_y & -a_x & 0 \end{bmatrix} \begin{bmatrix} q_1^- \\ q_2^- \\ q_3^- \\ q_4^- \end{bmatrix} = \\ &= \begin{bmatrix} -q_2^- a_x - q_3^- a_y - q_4^- a_z \\ q_1^- a_x + q_3^- a_z - q_4^- a_y \\ q_1^- a_y - q_2^- a_z + q_4^- a_x \\ q_1^- a_z + q_2^- a_y - q_3^- a_x \end{bmatrix}. \end{aligned} \quad (37)$$

Далее выполняется умножение $\mathbf{q}_k^{-*} \otimes (\mathbf{q}_k^- \otimes \mathbf{a}_k)$, в котором сопряженный кватернион $\mathbf{q}_k^{-*}$ записывается в виде матрицы:

$$\begin{aligned} \mathbf{q}_k^{-*} \otimes (\mathbf{q}_k^- \otimes \mathbf{a}_k) &= \\ &= \begin{bmatrix} q_1^- & q_2^- & q_3^- & q_4^- \\ -q_2^- & q_1^- & -q_4^- & q_3^- \\ -q_3^- & q_4^- & q_1^- & -q_2^- \\ -q_4^- & -q_3^- & q_2^- & q_1^- \end{bmatrix} \begin{bmatrix} -q_2^- a_x - q_3^- a_y - q_4^- a_z \\ q_1^- a_x + q_3^- a_z - q_4^- a_y \\ q_1^- a_y - q_2^- a_z + q_4^- a_x \\ q_1^- a_z + q_2^- a_y - q_3^- a_x \end{bmatrix}. \end{aligned} \quad (38)$$

Результатом умножения в (38) является 4-элементный вектор $\mathbf{a}_k^g$, (36) первый элемент которого не используется в дальнейших вычислениях, а элементы a_x^g, a_y^g, a_z^g определяются выражениями:

$$\begin{aligned} a_x^g &= q_1^- (q_1^- a_x + q_3^- a_z - q_4^- a_y) - \\ &- q_2^- (-q_2^- a_x - q_3^- a_y - q_4^- a_z) + \\ &+ q_3^- (q_1^- a_z + q_2^- a_y - q_3^- a_x) - \\ &- q_4^- (q_1^- a_y - q_2^- a_z + q_4^- a_x), \end{aligned} \quad (39)$$

$$\begin{aligned} a_y^g &= q_1^- (q_1^- a_y - q_2^- a_z + q_4^- a_x) - \\ &- q_2^- (q_1^- a_z + q_2^- a_y - q_3^- a_x) - \\ &- q_3^- (-q_2^- a_x - q_3^- a_y - q_4^- a_z) + \\ &+ q_4^- (q_1^- a_x + q_3^- a_z - q_4^- a_y), \end{aligned} \quad (40)$$

$$\begin{aligned} a_z^g &= q_1^- (q_1^- a_z + q_2^- a_y - q_3^- a_x) + \\ &+ q_2^- (q_1^- a_y - q_2^- a_z + q_4^- a_x) - \\ &- q_3^- (q_1^- a_x + q_3^- a_z - q_4^- a_y) - \\ &- q_4^- (-q_2^- a_x - q_3^- a_y - q_4^- a_z). \end{aligned} \quad (41)$$

Априорная оценка местоположения ТС определяется путем интегрирования оценки скорости $\mathbf{v}_k^-$:

$$\mathbf{p}_k^- = \mathbf{p}_{k-1}^+ + \mathbf{v}_k^- \Delta t = \begin{bmatrix} p_x^+ + v_x^- \Delta t \\ p_y^+ + v_y^- \Delta t \end{bmatrix}, \quad (42)$$

где $\mathbf{p}_{k-1}^+ = [p_x^+, p_y^+]^T$, $\mathbf{v}_k^- = [v_x^-, v_y^-]^T$.

Априорная оценка скорости движения ТС определяется путем интегрирования показаний акселерометра $\mathbf{a}_k^g$:

$$\mathbf{v}_k^- = \mathbf{v}_{k-1}^+ + \mathbf{a}_k^g \Delta t = \begin{bmatrix} v_x^+ + a_x^g \Delta t \\ v_y^+ + a_y^g \Delta t \end{bmatrix}; \quad (43)$$

где $\mathbf{a}_k^g = [a_x^g, a_y^g]^T$; $\mathbf{v}_{k-1}^+ = [v_x^+, v_y^+]^T$; для случая на плоскости учитываются только координаты x и y .

Априорная оценка вектора состояния $\mathbf{x}_k^- = [\mathbf{q}_k^-, \mathbf{p}_k^-, \mathbf{v}_k^-]$ для момента времени k (5) построена на основе апостериорной оценки $\mathbf{x}_{k-1}^+ = [\mathbf{q}_{k-1}^+, \mathbf{p}_{k-1}^+, \mathbf{v}_{k-1}^+]$ в предыдущий момент времени $k-1$.

Априорная ковариационная матрица ошибок оценки вектора состояния в момент времени k (6) вычисляется по формуле:

$$\mathbf{P}_k^- = \mathbf{F}_k \mathbf{P}_{k-1}^+ \mathbf{F}_k^T + \mathbf{G}_k \mathbf{U} \mathbf{G}_k^T; \quad (44)$$

где $\mathbf{P}_{k-1}^+$ – апостериорная ковариационная матрица ошибок оценки вектора состояния в предыдущий момент времени $k-1$; $\mathbf{F}_k$ – матрица частных производных уравнений модели априорной оценки (30, 42, 43) по элементам вектора состояния (29); $\mathbf{U}$ – диагональная матрица ошибок процесса, учитывающая шум измерений акселерометра и гироскопа; T – операция транспонирования; $\mathbf{G}_k$ – матрица частных производных уравнений модели априорной оценки (30, 42, 43) по элементам входного воздействия $[w_x, w_y, w_z, a_x, a_y, a_z]$. В общем случае матрицы $\mathbf{F}$ и $\mathbf{G}$ зависят от значений $\mathbf{a}_k^g$, $\mathbf{w}_k$ и $\mathbf{q}_k^+$ в момент времени k , но далее для упрощения индексы не показаны.

Матрица частных производных уравнений модели априорной оценки (30, 42, 43) по элементам вектора состояния определяется выражением:

$$\mathbf{F} = \begin{bmatrix} \mathbf{F}_{4 \times 4}^{\frac{\partial \mathbf{q}^-}{\partial \mathbf{q}}} & \mathbf{F}_{4 \times 2}^{\frac{\partial \mathbf{q}^-}{\partial \mathbf{p}}} & \mathbf{F}_{4 \times 2}^{\frac{\partial \mathbf{q}^-}{\partial \mathbf{v}}} \\ \mathbf{F}_{2 \times 4}^{\frac{\partial \mathbf{p}^-}{\partial \mathbf{q}}} & \mathbf{F}_{2 \times 2}^{\frac{\partial \mathbf{p}^-}{\partial \mathbf{p}}} & \mathbf{F}_{2 \times 2}^{\frac{\partial \mathbf{p}^-}{\partial \mathbf{v}}} \\ \mathbf{F}_{2 \times 4}^{\frac{\partial \mathbf{v}^-}{\partial \mathbf{q}}} & \mathbf{F}_{2 \times 2}^{\frac{\partial \mathbf{v}^-}{\partial \mathbf{p}}} & \mathbf{F}_{2 \times 2}^{\frac{\partial \mathbf{v}^-}{\partial \mathbf{v}}} \end{bmatrix} = \begin{bmatrix} 1 & \frac{\partial q_1^-}{\partial q_2} & \frac{\partial q_1^-}{\partial q_3} & \frac{\partial q_1^-}{\partial q_4} & 0 & 0 & 0 & 0 \\ \frac{\partial q_2^-}{\partial q_2} & 1 & \frac{\partial q_2^-}{\partial q_3} & \frac{\partial q_2^-}{\partial q_4} & 0 & 0 & 0 & 0 \\ \frac{\partial q_3^-}{\partial q_1} & \frac{\partial q_3^-}{\partial q_2} & 1 & \frac{\partial q_3^-}{\partial q_4} & 0 & 0 & 0 & 0 \\ \frac{\partial q_4^-}{\partial q_1} & \frac{\partial q_4^-}{\partial q_2} & \frac{\partial q_4^-}{\partial q_3} & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & \Delta t & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & \Delta t \\ \frac{\partial v_x^-}{\partial q_1} & \frac{\partial v_x^-}{\partial q_2} & \frac{\partial v_x^-}{\partial q_3} & \frac{\partial v_x^-}{\partial q_4} & 0 & 0 & 1 & 0 \\ \frac{\partial v_y^-}{\partial q_1} & \frac{\partial v_y^-}{\partial q_2} & \frac{\partial v_y^-}{\partial q_3} & \frac{\partial v_y^-}{\partial q_4} & 0 & 0 & 0 & 1 \end{bmatrix}, \quad (45)$$

где $\mathbf{q}_k^- = [q_1^-, q_2^-, q_3^-, q_4^-]^T$ – кватернион априорной оценки ориентации ТС (30); $\mathbf{p}_k^- = [p_x^-, p_y^-]$ – априорная оценка МП ТС (42); $\mathbf{v}_k^- = [v_x^-, v_y^-]^T$ – априорная оценка скорости ТС.

Далее рассмотрим составные подматрицы матрицы (45).

$\mathbf{F}_{4 \times 4}^{\frac{\partial \mathbf{q}^-}{\partial \mathbf{q}}}$ – матрица частных производных уравнений априорной оценки ориентации (30) по компонентам кватерниона $\mathbf{q}$, размерность $\mathbb{R}^{4 \times 4}$ (далее слово размерность будем подразумевать):

$$\mathbf{F}_{4 \times 4}^{\frac{\partial \mathbf{q}^-}{\partial \mathbf{q}}} = \begin{bmatrix} \frac{\partial q_1^-}{\partial q_1} & \frac{\partial q_1^-}{\partial q_2} & \frac{\partial q_1^-}{\partial q_3} & \frac{\partial q_1^-}{\partial q_4} \\ \frac{\partial q_2^-}{\partial q_1} & \frac{\partial q_2^-}{\partial q_2} & \frac{\partial q_2^-}{\partial q_3} & \frac{\partial q_2^-}{\partial q_4} \\ \frac{\partial q_3^-}{\partial q_1} & \frac{\partial q_3^-}{\partial q_2} & \frac{\partial q_3^-}{\partial q_3} & \frac{\partial q_3^-}{\partial q_4} \\ \frac{\partial q_4^-}{\partial q_1} & \frac{\partial q_4^-}{\partial q_2} & \frac{\partial q_4^-}{\partial q_3} & \frac{\partial q_4^-}{\partial q_4} \end{bmatrix} = \begin{bmatrix} 1 & -\frac{\Delta t}{2} w_x & -\frac{\Delta t}{2} w_y & -\frac{\Delta t}{2} w_z \\ \frac{\Delta t}{2} w_x & 1 & \frac{\Delta t}{2} w_z & -\frac{\Delta t}{2} w_y \\ \frac{\Delta t}{2} w_y & -\frac{\Delta t}{2} w_z & 1 & \frac{\Delta t}{2} w_x \\ \frac{\Delta t}{2} w_z & \frac{\Delta t}{2} w_y & -\frac{\Delta t}{2} w_x & 1 \end{bmatrix}. \quad (46)$$

$\mathbf{F}_{4 \times 2}^{\frac{\partial \mathbf{q}^-}{\partial \mathbf{p}}}$ – матрица частных производных уравнений априорной оценки ориентации (30) по компонентам вектора координат $\mathbf{p}$, $\mathbb{R}^{4 \times 2}$:

$$\mathbf{F}_{4 \times 2}^{\frac{\partial \mathbf{q}^-}{\partial \mathbf{p}}} = \mathbf{0}_{4 \times 2}. \quad (47)$$

$\mathbf{F}_{4 \times 2}^{\frac{\partial \mathbf{q}^-}{\partial \mathbf{v}}}$ – матрица частных производных уравнений априорной оценки ориентации (30) по компонентам вектора скорости $\mathbf{v}$, $\mathbb{R}^{4 \times 2}$:

$$\mathbf{F}_{4 \times 2}^{\frac{\partial \mathbf{q}^-}{\partial \mathbf{v}}} = \mathbf{0}_{4 \times 2}. \quad (48)$$

$\mathbf{F}_{2 \times 4}^{\frac{\partial \mathbf{p}^-}{\partial \mathbf{q}}}$ – матрица частных производных уравнений априорной оценки положения (42) по компонентам кватерниона $\mathbf{q}$, $\mathbb{R}^{2 \times 4}$:

$$\mathbf{F}_{2 \times 4}^{\frac{\partial \mathbf{p}^-}{\partial \mathbf{q}}} = \mathbf{0}_{2 \times 4}. \quad (49)$$

$\mathbf{F}_{2 \times 2}^{\frac{\partial \mathbf{p}^-}{\partial \mathbf{p}}}$ – матрица частных производных уравнений априорной оценки положения (42) по компонентам вектора координат $\mathbf{p}$, $\mathbb{R}^{2 \times 2}$:

$$\mathbf{F}_{2 \times 2}^{\frac{\partial \mathbf{p}^-}{\partial \mathbf{p}}} = \begin{bmatrix} \frac{\partial p_x^-}{\partial x} & 0 \\ 0 & \frac{\partial p_y^-}{\partial y} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \quad (50)$$

$\mathbf{F}_{2 \times 2}^{\frac{\partial p^-}{\partial v^-}}$ – матрица частных производных уравнений априорной оценки положения (42) по компонентам вектора скорости $\mathbf{v}$, $\mathbb{R}^{2 \times 2}$:

$$\mathbf{F}_{2 \times 2}^{\frac{\partial p^-}{\partial v^-}} = \begin{bmatrix} \frac{\partial p^-}{\partial v_x^-} & 0 \\ 0 & \frac{\partial p^-}{\partial v_y^-} \end{bmatrix} = \begin{bmatrix} \Delta t & 0 \\ 0 & \Delta t \end{bmatrix}. \quad (51)$$

$\mathbf{F}_{2 \times 4}^{\frac{\partial v^-}{\partial q^-}}$ – матрица частных производных уравнений априорной оценки скорости (43) по компонентам кватерниона $\mathbf{q}$, $\mathbb{R}^{2 \times 4}$:

$$\mathbf{F}_{2 \times 4}^{\frac{\partial v^-}{\partial q^-}} = \begin{bmatrix} \frac{\partial v_x^-}{\partial q_1^-} & \frac{\partial v_x^-}{\partial q_2^-} & \frac{\partial v_x^-}{\partial q_3^-} & \frac{\partial v_x^-}{\partial q_4^-} \\ \frac{\partial v_y^-}{\partial q_1^-} & \frac{\partial v_y^-}{\partial q_2^-} & \frac{\partial v_y^-}{\partial q_3^-} & \frac{\partial v_y^-}{\partial q_4^-} \end{bmatrix}. \quad (52)$$

Элементы матрицы (52) определяются выражениями:

$$\begin{aligned} \frac{\partial v_x^-}{\partial q_1^-} &= \Delta t(-2q_1a_x + 2q_4a_y - 2q_3a_z), \\ \frac{\partial v_y^-}{\partial q_1^-} &= \Delta t(-2q_4a_x - 2q_1a_y + 2q_2a_z), \\ \frac{\partial v_x^-}{\partial q_2^-} &= \Delta t(-2q_2a_x - 2q_3a_y - 2q_4a_z), \\ \frac{\partial v_y^-}{\partial q_2^-} &= \Delta t(2q_2a_y + 2q_3a_x - 2q_1a_z), \\ \frac{\partial v_x^-}{\partial q_3^-} &= \Delta t(-2q_2a_y + 2q_3a_x - 2q_1a_z), \\ \frac{\partial v_y^-}{\partial q_3^-} &= \Delta t(2q_2a_y + 2q_3a_x - 2q_1a_z), \\ \frac{\partial v_x^-}{\partial q_4^-} &= \Delta t(-2q_2a_y + 2q_3a_x - 2q_1a_z), \\ \frac{\partial v_y^-}{\partial q_4^-} &= \Delta t(2q_4a_x - 2q_1a_y - 2q_2a_z), \\ \frac{\partial v_x^-}{\partial q_4^-} &= \Delta t(-2q_1a_x + 2q_4a_y - 2q_3a_z). \end{aligned} \quad (53)$$

$\mathbf{F}_{2 \times 2}^{\frac{\partial p^-}{\partial p^-}}$ – матрица частных производных уравнений априорной оценки скорости (43) по компонентам вектора координат $\mathbf{p}$, $\mathbb{R}^{2 \times 2}$:

$$\mathbf{F}_{2 \times 2}^{\frac{\partial p^-}{\partial p^-}} = \mathbf{0}_{2 \times 2}. \quad (54)$$

$\mathbf{F}_{2 \times 2}^{\frac{\partial v^-}{\partial v^-}}$ – матрица частных производных уравнений априорной оценки скорости (43) по компонентам вектора скорости $\mathbf{v}$, $\mathbb{R}^{2 \times 2}$:

$$\mathbf{F}_{2 \times 2}^{\frac{\partial v^-}{\partial v^-}} = \begin{bmatrix} \frac{\partial v^-}{\partial v_x^-} & 0 \\ 0 & \frac{\partial v^-}{\partial v_y^-} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \quad (55)$$

Матрица частных производных уравнений модели априорной оценки (30, 42, 43) по элементам входного воздействия $[w_x, w_y, w_z, a_x, a_y, a_z]$ определяется по выражению:

$$\mathbf{G} = \begin{bmatrix} \mathbf{G}_{4 \times 3}^{\frac{\partial q^-}{\partial w^-}} & \mathbf{G}_{4 \times 3}^{\frac{\partial q^-}{\partial a^-}} \\ \mathbf{G}_{2 \times 3}^{\frac{\partial p^-}{\partial w^-}} & \mathbf{G}_{2 \times 3}^{\frac{\partial p^-}{\partial a^-}} \\ \mathbf{G}_{2 \times 3}^{\frac{\partial v^-}{\partial w^-}} & \mathbf{G}_{2 \times 3}^{\frac{\partial v^-}{\partial a^-}} \end{bmatrix} = \begin{bmatrix} \frac{\partial q_1^-}{\partial w_x^-} & \frac{\partial q_1^-}{\partial w_y^-} & \frac{\partial q_1^-}{\partial w_z^-} & 0 & 0 & 0 \\ \frac{\partial q_2^-}{\partial w_x^-} & \frac{\partial q_2^-}{\partial w_y^-} & \frac{\partial q_2^-}{\partial w_z^-} & 0 & 0 & 0 \\ \frac{\partial q_3^-}{\partial w_x^-} & \frac{\partial q_3^-}{\partial w_y^-} & \frac{\partial q_3^-}{\partial w_z^-} & 0 & 0 & 0 \\ \frac{\partial q_4^-}{\partial w_x^-} & \frac{\partial q_4^-}{\partial w_y^-} & \frac{\partial q_4^-}{\partial w_z^-} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{\partial v_x^-}{\partial a_x^-} & \frac{\partial v_x^-}{\partial a_y^-} & \frac{\partial v_x^-}{\partial a_z^-} \\ 0 & 0 & 0 & \frac{\partial v_y^-}{\partial a_x^-} & \frac{\partial v_y^-}{\partial a_y^-} & \frac{\partial v_y^-}{\partial a_z^-} \end{bmatrix}. \quad (56)$$

Далее рассмотрим составные подматрицы матрицы $\mathbf{G}$ в (56).

$\mathbf{G}_{4 \times 3}^{\frac{\partial q^-}{\partial w^-}}$ – матрица частных производных уравнений априорной оценки ориентации (30) по компонентам вектора угловой скорости $\mathbf{w}$, $\mathbb{R}^{4 \times 3}$:

$$\mathbf{G}_{4 \times 3}^{\frac{\partial q^-}{\partial w^-}} = \begin{bmatrix} \frac{\partial q_1^-}{\partial w_x^-} & \frac{\partial q_1^-}{\partial w_y^-} & \frac{\partial q_1^-}{\partial w_z^-} \\ \frac{\partial q_2^-}{\partial w_x^-} & \frac{\partial q_2^-}{\partial w_y^-} & \frac{\partial q_2^-}{\partial w_z^-} \\ \frac{\partial q_3^-}{\partial w_x^-} & \frac{\partial q_3^-}{\partial w_y^-} & \frac{\partial q_3^-}{\partial w_z^-} \\ \frac{\partial q_4^-}{\partial w_x^-} & \frac{\partial q_4^-}{\partial w_y^-} & \frac{\partial q_4^-}{\partial w_z^-} \end{bmatrix} = \begin{bmatrix} -\frac{\Delta t}{2}q_2 & -\frac{\Delta t}{2}q_3 & -\frac{\Delta t}{2}q_4 \\ \frac{\Delta t}{2}q_1 & -\frac{\Delta t}{2}q_4 & \frac{\Delta t}{2}q_3 \\ \frac{\Delta t}{2}q_4 & \frac{\Delta t}{2}q_1 & -\frac{\Delta t}{2}q_2 \\ -\frac{\Delta t}{2}q_3 & \frac{\Delta t}{2}q_2 & \frac{\Delta t}{2}q_1 \end{bmatrix}. \quad (57)$$

$\frac{\partial q^-}{\partial a}$ – матрица частных производных уравнений априорной оценки ориентации (30) по компонентам вектора ускорения $\mathbf{a}$, $\mathbb{R}^{4 \times 3}$:

$$\mathbf{G}_{4 \times 3}^{\frac{\partial q^-}{\partial a}} = \mathbf{0}_{4 \times 3}. \quad (58)$$

$\frac{\partial p^-}{\partial w}$ – матрица частных производных уравнений априорной оценки положения (42) по компонентам вектора угловой скорости $\mathbf{w}$, $\mathbb{R}^{2 \times 3}$:

$$\mathbf{G}_{2 \times 3}^{\frac{\partial p^-}{\partial w}} = \mathbf{0}_{2 \times 3}. \quad (59)$$

$\frac{\partial p^-}{\partial a}$ – матрица частных производных уравнений априорной оценки положения (42) по компонентам вектора ускорения $\mathbf{a}$, $\mathbb{R}^{2 \times 3}$:

$$\mathbf{G}_{2 \times 3}^{\frac{\partial p^-}{\partial a}} = \mathbf{0}_{2 \times 3}. \quad (60)$$

$\frac{\partial v^-}{\partial w}$ – матрица частных производных уравнений априорной оценки скорости (43) по компонентам вектора угловой скорости $\mathbf{w}$, $\mathbb{R}^{2 \times 3}$:

$$\mathbf{G}_{2 \times 3}^{\frac{\partial v^-}{\partial w}} = \mathbf{0}_{2 \times 3}. \quad (61)$$

$\frac{\partial v^-}{\partial a}$ – матрица частных производных уравнений априорной оценки скорости (43) по компонентам вектора ускорения $\mathbf{a}$, размерности $\mathbf{a}$, $\mathbb{R}^{2 \times 3}$:

$$\mathbf{G}_{2 \times 3}^{\frac{\partial v^-}{\partial a}} = \begin{bmatrix} \frac{\partial v_x^-}{\partial a_x} & \frac{\partial v_x^-}{\partial a_y} & \frac{\partial v_x^-}{\partial a_z} \\ \frac{\partial v_y^-}{\partial a_x} & \frac{\partial v_y^-}{\partial a_y} & \frac{\partial v_y^-}{\partial a_z} \end{bmatrix}. \quad (62)$$

Элементы матрицы (62) определяются выражениями:

$$\begin{aligned} \frac{\partial v_x^-}{\partial a_x} &= -\Delta t(q_1^2 + q_2^2 - q_3^2 - q_4^2), \\ \frac{\partial v_y^-}{\partial a_x} &= -\Delta t(2q_1q_4 + 2q_2q_3), \\ \frac{\partial v_x^-}{\partial a_y} &= \Delta t(2q_1q_4 - 2q_2q_3), \\ \frac{\partial v_y^-}{\partial a_y} &= -\Delta t(q_1^2 - q_2^2 + q_3^2 - q_4^2), \\ \frac{\partial v_x^-}{\partial a_z} &= -\Delta t(2q_1q_3 - 2q_2q_4), \\ \frac{\partial v_y^-}{\partial a_z} &= \Delta t(2q_1q_2 - 2q_3q_4). \end{aligned} \quad (63)$$

Диагональная матрица ошибок процесса $\mathbf{U}$, учитывающая шум измерений акселерометра и гироскопа, формализует модель инструментальных погрешностей бесплатформенной ИНС и определяется выражением:

$$\mathbf{U} = \begin{bmatrix} \sigma_w^2 & 0 & 0 & 0 & 0 & 0 \\ 0 & \sigma_w^2 & 0 & 0 & 0 & 0 \\ 0 & 0 & \sigma_w^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma_a^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sigma_a^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & \sigma_a^2 \end{bmatrix}, \quad (64)$$

где $\sigma_{wx}^2 = \sigma_{wy}^2 = \sigma_{wz}^2 = \sigma_w^2$ – дисперсия аддитивного белого гауссового шума (АБГШ) измерений гироскопа, $\sigma_{ax}^2 = \sigma_{ay}^2 = \sigma_{az}^2 = \sigma_a^2$ – дисперсия измерений акселерометра.

На этапе обновления апостериорная оценка вектора состояния $\mathbf{x}_k^+$ и ковариационной матрицы ошибок $\mathbf{P}_k^+$ формируется на основе $\mathbf{x}_k^-$, $\mathbf{P}_k^-$ и ДМ/УМ измерений, полученных в момент k . Этап обновления описывается уравнениями:

$$\mathbf{x}_k^+ = \mathbf{x}_k^- + \mathbf{K}\mathbf{y}; \quad (65)$$

$$\mathbf{y} = [\mathbf{h}_{TOA}; \mathbf{h}_{AOA}]; \quad (66)$$

$$\mathbf{P}_k^+ = (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{P}_k^-; \quad (67)$$

$$\mathbf{K} = \mathbf{P}_k^- \mathbf{H}^T \mathbf{S}^{-1}; \quad (68)$$

$$\mathbf{S} = \mathbf{H}\mathbf{P}_k^- \mathbf{H}^T + \mathbf{R}; \quad (69)$$

где $\mathbf{K}$ – матрица коэффициентов усиления РФК; $\mathbf{y}$ – вектор обновления; $\mathbf{S}$ – ковариационная матрица ошибок вектора $\mathbf{y}$; $\mathbf{H}$ – матрица частных производных вектора измерений по составляющим вектора состояния; $\mathbf{R}$ – ковариационная матрица ошибок измерений; $\mathbf{I}$ – единичная матрица.

Вектор обновления включает в себя два вектора $\mathbf{h}$, которые описывают разность прямых измерений дальности и угла прихода с соответствующими величинами, полученными на основе априорной оценки $\mathbf{x}_{k+1}^-$.

Вектор обновления ДМ измерений

$$\mathbf{h}_{TOA} = [d_0 - r_0, d_1 - r_1, \dots, d_{n-1} - r_{n-1}]; \quad (70)$$

где d_i – измеренная дальность от ТС до i -го ПП; $i = 0 \dots n-1$, n – число ПП, измерения дальности от которых доступны в момент времени k ; r_i – оценка дальности от ТС до i -го ПП на основе $\mathbf{x}_k^-$:

$$r_i = \|\mathbf{p}_k^- - \mathbf{s}_i\| = \sqrt{(x - x_i)^2 + (y - y_i)^2}; \quad (71)$$

где $\mathbf{p}_k^- = [x, y]$ – априорная ОК ТС; $\mathbf{s}_i = [x_i, y_i]$ – координаты ПП $_i$.

Вектор обновления УМ измерений

$$\mathbf{h}_{AOA} = [\varphi_0 - \hat{\varphi}_0, \varphi_1 - \hat{\varphi}_1, \dots, \varphi_{n-1} - \hat{\varphi}_{n-1}]; \quad (72)$$

где φ_i – измеренный угол прихода сигнала от ПП $_i$; $i = 0 \dots n-1$, n – число ПП, измерения угла прихода от которых доступны в момент времени k ; $\hat{\varphi}_i$ – оценка угла прихода сигнала от ПП $_i$ на основе $\mathbf{x}_k^-$:

$$\hat{\varphi}_i = \angle(\mathbf{p}_k^-, \mathbf{s}_i) = \arctan\left(\frac{y - y_i}{x - x_i}\right). \quad (73)$$

Матрица наблюдений $\mathbf{H}$ – матрица частных производных векторов $\mathbf{r} = [r_0, r_1, \dots, r_{n-1}]$ и $\hat{\varphi} =$

$= [\hat{\varphi}_0, \hat{\varphi}_1, \dots, \hat{\varphi}_{n-1}]$ по элементам вектора состояний $\mathbf{x}$:

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}_{n \times 4}^{\frac{\partial r}{\partial q}} & \mathbf{H}_{n \times 2}^{\frac{\partial r}{\partial p}} & \mathbf{H}_{n \times 2}^{\frac{\partial r}{\partial v}} \\ \mathbf{H}_{n \times 4}^{\frac{\partial \varphi}{\partial q}} & \mathbf{H}_{n \times 2}^{\frac{\partial \varphi}{\partial p}} & \mathbf{H}_{n \times 2}^{\frac{\partial \varphi}{\partial v}} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 & \frac{\partial r_0}{\partial x} & \frac{\partial r_0}{\partial y} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{\partial r_1}{\partial x} & \frac{\partial r_1}{\partial y} & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & \frac{\partial r_{n-1}}{\partial x} & \frac{\partial r_{n-1}}{\partial y} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{\partial \theta_0}{\partial x} & \frac{\partial \theta_0}{\partial y} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{\partial \theta_1}{\partial x} & \frac{\partial \theta_1}{\partial y} & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & \frac{\partial \theta_{n-1}}{\partial x} & \frac{\partial \theta_{n-1}}{\partial y} & 0 & 0 \end{bmatrix}. \quad (74)$$

Далее рассмотрим составные подматрицы матрицы $\mathbf{H}$ в (74).

$\mathbf{H}_{n \times 4}^{\frac{\partial r}{\partial q}}$ – матрица частных производных уравнения дальности (71) по компонентам кватерниона $\mathbf{q}$, $\mathbb{R}^{n \times 4}$:

$$\mathbf{H}_{n \times 4}^{\frac{\partial r}{\partial q}} = \mathbf{0}_{n \times 4}. \quad (75)$$

$\mathbf{H}_{n \times 2}^{\frac{\partial r}{\partial p}}$ – матрица частных производных уравнения дальности (71) по компонентам вектора координат $\mathbf{p}$, $\mathbb{R}^{n \times 2}$ (76).

$\mathbf{H}_{n \times 2}^{\frac{\partial r}{\partial v}}$ – матрица частных производных уравнения дальности (71) по компонентам вектора скорости $\mathbf{v}$, $\mathbb{R}^{n \times 2}$:

$$\mathbf{H}_{n \times 2}^{\frac{\partial r}{\partial v}} = \mathbf{0}_{n \times 2}. \quad (77)$$

$\mathbf{H}_{n \times 4}^{\frac{\partial \varphi}{\partial q}}$ – матрица частных производных уравнения угла прихода сигнала (73) по компонентам кватерниона $\mathbf{q}$, $\mathbb{R}^{n \times 4}$:

$$\mathbf{H}_{n \times 4}^{\frac{\partial \varphi}{\partial q}} = \mathbf{0}_{n \times 4}. \quad (78)$$

$\mathbf{H}_{n \times 2}^{\frac{\partial \varphi}{\partial p}}$ – матрица частных производных уравнения угла прихода сигнала (73) по компонентам вектора координат $\mathbf{p}$, $\mathbb{R}^{n \times 2}$ (79).

$\mathbf{H}_{n \times 2}^{\frac{\partial \varphi}{\partial v}}$ – матрица частных производных уравнения угла прихода сигнала (73) по компонентам вектора скорости $\mathbf{v}$, $\mathbb{R}^{n \times 2}$:

$$\mathbf{H}_{n \times 2}^{\frac{\partial \varphi}{\partial v}} = \mathbf{0}_{n \times 2}. \quad (80)$$

Ковариационная матрица ошибок измерений $\mathbf{R}$ равна:

$$\mathbf{R} = \begin{bmatrix} \sigma_{r_0}^2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \sigma_{r_1}^2 & 0 & 0 & 0 & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \sigma_{r_{n-1}}^2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sigma_{\varphi_0}^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \sigma_{\varphi_1}^2 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sigma_{\varphi_{n-1}}^2 \end{bmatrix}. \quad (81)$$

где $\sigma_{r_0}^2 = \sigma_{r_1}^2 = \sigma_{r_{n-1}}^2 = \sigma_r^2 = (\sigma_t c)^2$ – дисперсия ошибки измерения дальности; σ_t – СКО измерения времени прихода; c – скорость света; $\sigma_{\varphi_0}^2 = \sigma_{\varphi_1}^2 = \dots = \sigma_{\varphi_{n-1}}^2 = \sigma_{\varphi}^2$ – дисперсия ошибки измерения угла прихода.

Далее в разделе 5 рассмотрим реализацию представленной математической модели в имитационной модели (ИМ). Аналогично [9] примем шум первичных РТИ измерений ДМ/УМ АБГШ, а точность измерений зададим через СКО: для дальности, измеряемой через время прихода, СКО $\sigma_t = 50$ нс, для угла СКО $\sigma_{\theta} = 2^\circ$.

$$\mathbf{H}_{n \times 2}^{\frac{\partial r}{\partial p}} = \begin{bmatrix} \frac{\partial r_0}{\partial x} & \frac{\partial r_0}{\partial y} \\ \frac{\partial r_1}{\partial x} & \frac{\partial r_1}{\partial y} \\ \dots & \dots \\ \frac{\partial r_{n-1}}{\partial x} & \frac{\partial r_{n-1}}{\partial y} \end{bmatrix} = \begin{bmatrix} \frac{x - x_0}{\sqrt{(x - x_0)^2 + (y - y_0)^2}} & \frac{y - y_0}{\sqrt{(x - x_0)^2 + (y - y_0)^2}} \\ \frac{x - x_1}{\sqrt{(x - x_1)^2 + (y - y_1)^2}} & \frac{y - y_1}{\sqrt{(x - x_1)^2 + (y - y_1)^2}} \\ \dots & \dots \\ \frac{x - x_{n-1}}{\sqrt{(x - x_{n-1})^2 + (y - y_{n-1})^2}} & \frac{y - y_{n-1}}{\sqrt{(x - x_{n-1})^2 + (y - y_{n-1})^2}} \end{bmatrix}, \quad (76)$$

$$\mathbf{H}_{n \times 2}^{\frac{\partial \varphi}{\partial p}} = \begin{bmatrix} \frac{\partial \varphi_0}{\partial x} & \frac{\partial \varphi_0}{\partial y} \\ \frac{\partial \varphi_1}{\partial x} & \frac{\partial \varphi_1}{\partial y} \\ \dots & \dots \\ \frac{\partial \varphi_{n-1}}{\partial x} & \frac{\partial \varphi_{n-1}}{\partial y} \end{bmatrix} = \begin{bmatrix} -\frac{y - y_0}{(x - x_0)^2 + (y - y_0)^2} & \frac{x - x_0}{(x - x_0)^2 + (y - y_0)^2} \\ -\frac{y - y_1}{(x - x_1)^2 + (y - y_1)^2} & \frac{x - x_1}{(x - x_1)^2 + (y - y_1)^2} \\ \dots & \dots \\ -\frac{y - y_{n-1}}{(x - x_{n-1})^2 + (y - y_{n-1})^2} & \frac{x - x_{n-1}}{(x - x_{n-1})^2 + (y - y_{n-1})^2} \end{bmatrix}. \quad (79)$$

5. ИМИТАЦИОННАЯ МОДЕЛЬ КОМПЛЕКСИРОВАНИЯ ДМ/УМ/ИНС В РФК

В исследуемой ИМ шум первичных ИНС измерений, как и РТИ измерений, также принимается АБГШ: спектральная плотность мощности шума гироскопа $\sigma_w^2 = 1,25 \cdot 10^{-4}$ рад/с/ $\sqrt{\text{Гц}}$, мощность шума акселерометра $\sigma_a^2 = 2 \cdot 10^{-4}$ м/с²/ $\sqrt{\text{Гц}}$. Значения меняющихся ошибок бесплатформенной ИНС можно считать постоянными [22]; примем смещение измерений гироскопа $3,5 \cdot 10^{-4}$ рад/с, акселерометра $5 \cdot 10^{-3}$ м/с².

Рисунок 4 иллюстрирует работу РФК с комплексированием TOA/AOA/INS в сравнении с РФК на основе только измерений TOA/AOA: рисунок 4а иллюстрирует полную траекторию движения ТС в городе по модели сценария METIS Madrid Simulation Scenario, которая была разработана в [9]; рисунок 4б иллюстрирует оценку траектории ТС на отдельно взятом повороте.

Качественный анализ оценок траекторий (рисунок 4б) позволяет сделать вывод о том, что комплексирование РТИ-измерений с измерениями ИНС позволяет повысить точность позиционирования ТС на перекрестке. В рассматриваемом сценарии ИМ VUE движется со скоростью $v \approx 5$ м/с, интервал сбора ДМ/УМ измерений равен $T = 0,1$ с; интервал сбора ИНС измерений равен $\Delta t = 0,01$ с.

Рисунок 5 детализирует количественно точность ОК ТС для каждой точки траектории его движения; ось x показывает пройденное ТС расстояние; вертикальными штриховыми линиями показано примерное положение начала поворота в траектории движения ТС. Анализ точности ОК ТС для каждой точки траектории его движения (рисунок 5) показывает, что использование ИНС позволяет существенно сократить ошибку на участках траектории с изменением направления движения ТС.

Рисунок 6 показывает оценку траектории ТС на основе только измерений ИНС, т. е. при отсутствии этапа обновления на основе ДМ/УМ измерений.

Анализ точности ОК ТС только на основе измерений ИНС подтверждает накопление ошибки, так как для ИНС характерен дрейф измерений отдельных датчиков, что приводит к невозможности оценки координат ТС методом двойного интегрирования измерений на длительном интервале времени. Так, в алгоритме TOA/AOA/INS по измерениям ИНС строится оценка на относительно коротком интервале времени между этапами обновления на основе дальномерных/угломерных измерений, которые необходимы для периодической корректировки.

В ИМ предполагается, что измерения доступны в каждой точке траектории движения ТС. Ошибка оценки траектории ТС RMSE (аббр. от англ. Root Mean Square Error) вычисляется как среднее зна-

чение евклидова расстояния между истинными координатами и их оценкой по совокупности K точек измерений:

$$RMSE = \frac{1}{K} \sum_{k=1}^K \sqrt{(x_{io} - x_i)^2 + (y_{io} - y_i)^2}, \quad (82)$$

где $[x_k, y_k]$ – k -я оценка координат ТС; $[x, y]$ – истинные значения координат ТС; K – количество оценок координат.

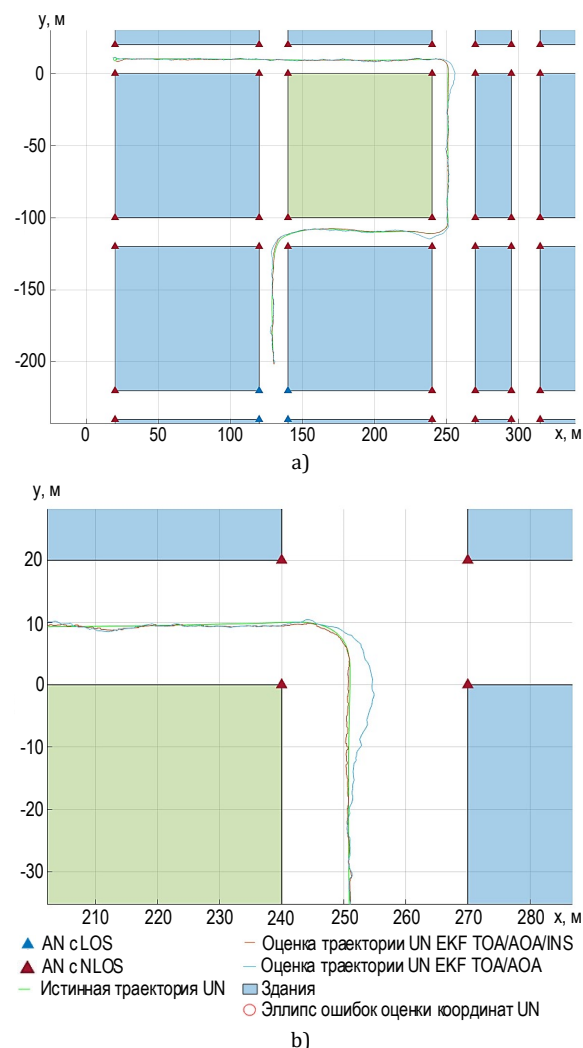


Рис. 4. Оценка траектории ТС: а) полная траектория; б) участок поворота

Fig. 4. Estimation of the VUE Trajectory: a) Full Trajectory; b) Intersection

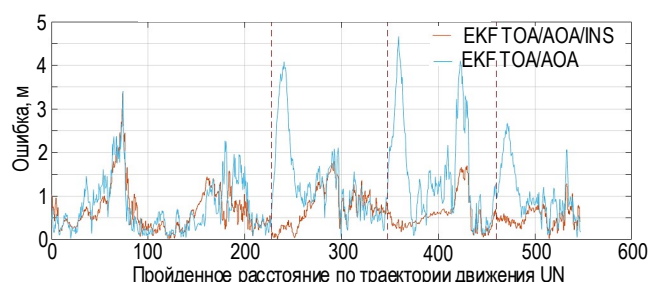


Рис. 5. Точность ОК в каждой точке траектории движения ТС

Fig. 1. Location Estimation Accuracy at Each Point of the VUE Trajectory

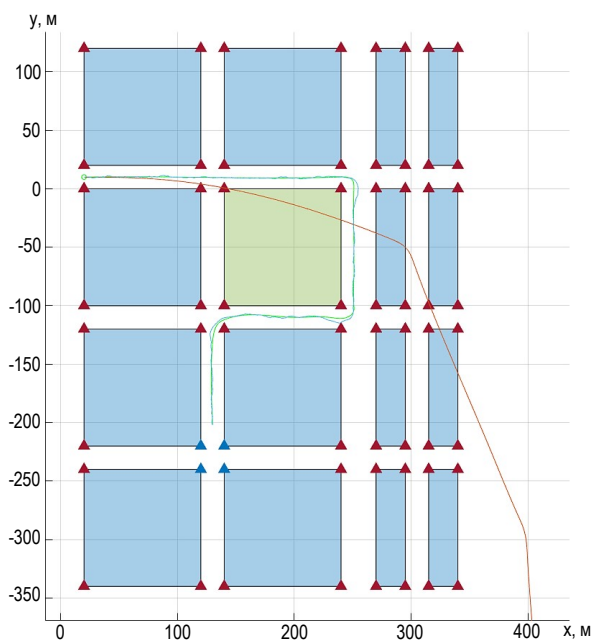


Рис. 6. Оценка траектории ТС только на основе измерений ИНС

Fig. 6. Estimation of the VUE Trajectory Based on INS Measurements Only

В ИМ предполагается, что измерения доступны в каждой точке траектории движения ТС. Ошибка оценки траектории ТС RMSE (Root Mean Square Error) вычисляется как среднее значение евклидова расстояния между истинными координатами и их оценкой по совокупности K точек измерений при движении ТС:

$$RMSE = \frac{1}{K} \sum_{k=1}^K \sqrt{(x_{io} - x_i)^2 + (y_{io} - y_i)^2}, \quad (82)$$

где $[x_k, y_k]$ – k -я оценка координат ТС; $[x, y]$ – истинные значения координат ТС; K – количество оценок координат.

При усреднении оценки траектории движения ТС в городе по модели METIS Madrid ошибка ОК составила порядка 0,7 м для ДМ/УМ/ИНС и 1,15 м для ДМ/УМ без ИНС; выигрыш от комплексирования РТИ-измерений с ИНС особенно проявляется на поворотах при маневрировании ТС.

6. ВЫВОДЫ

Настоящая работа, являясь продолжением исследования [9], развивает методологию комплексирования радиотехнических дальномерных и угломерных измерений с измерениями инерциальной навигационной системы с использованием расширенного фильтра Калмана. Разработанная математическая и имитационная модели, включая преобразование инерциальной системы координат акселерометра и гироскопа в связанную систему координат транспортного средства с использованием кватернионов, подтверждают возможность повышения точности позиционирования транспортного средства при изменении направления движения на перекрестках с точностью до одного метра.

ИСТОЧНИК ФИНАНСИРОВАНИЯ

Исследование выполнено в рамках исполнения Государственного контракта № ПЗ3-1-26/9.

Список используемых источников

1. Фокин Г.А. Технологии сетевого позиционирования: монография. СПб.: СПбГУТ, 2020. 558 с.
2. Фокин Г.А., Кучерявый А.Е. Сетевое позиционирование в экосистеме 5G // Электросвязь. 2020. № 9. С. 51–58. DOI:10.34832/ELSV.2020.10.9.006
3. Киреев А.В., Фокин Г.А. Позиционирование объектов в сетях LTE посредством измерения времени прохождения сигналов // Труды учебных заведений связи. 2016. Том 2. № 1. С. 68–72.
4. Киреев А.В., Фокин Г.А. Пеленгация источников радиоизлучения LTE мобильным пунктом радиоконтроля с круговой антенной решеткой // Труды Научно-исследовательского института радио. 2015. № 2. С. 68–71.
5. Sivers M., Fokin G., Dmitriev P., Kireev A., Volgushev D., Al-Odhari A. Indoor Positioning in WiFi and NanoLOC Networks // Proceedings of the 16th International on Next Generation Wired/Wireless Networking (NEW2AN 2016) and 9th Conference on Conference on Internet of Things and Smart Spaces (ruSMART 2016), St. Petersburg, Russia, 26–28 September 2016. Lecture Notes in Computer Science. Cham: Springer, 2016. Vol. 9870. PP. 465–476. DOI:10.1007/978-3-319-46301-8_39
6. Sivers M., Fokin G., Dmitriev P., Kireev A., Volgushev D., Al-odhari A.H.A. Wi-Fi Based Indoor Positioning System Using Inertial Measurements // Proceedings of the 17th International Conference on International Conference on Next Generation Wired/Wireless Networking (NEW2AN 2017), 10th Conference on Conference on Internet of Things and Smart Spaces (ruSMART 2017), Third Workshop on International Workshop on Nano-scale Computing and Communications (NsCC 2017), St. Petersburg, Russia, 28–30 August 2017. Lecture Notes in Computer Science. Cham: Springer, 2017. Vol. 10531. PP. 734–744. DOI:10.1007/978-3-319-67380-6_69
7. Киреев А.В., Фокин Г.А. Оценка точности локального позиционирования мобильных устройств с помощью радиокарт и инерциальной навигационной системы // Труды учебных заведений связи. 2017. Том 3. № 4. С. 54–62.
8. Духовницкий О.Г., Рагило М.А., Сиверс М.А., Фокин Г.А. Применение фильтра Калмана в задачах позиционирования // Электросвязь. 2016. № 1. С. 78–81.
9. Фокин Г.А., Владыко А.Г. Позиционирование транспортных средств в сверхплотных сетях радиодоступа V2X/5G с использованием расширенного фильтра Калмана // Труды учебных заведений связи. 2020. Т. 6. № 4. С. 45–59. DOI:10.31854/1813-324X-2020-6-4-45-59

10. Kalman R.E. A New Approach to Linear Filtering and Prediction Problems // Journal of Basic Engineering. 1960. Vol. 82. Iss. 1. PP. 35–45. DOI:10.1115/1.3662552
11. Фомин В.Н. Рекуррентное оценивание и адаптивная фильтрация. М.: Наука, 1984. 286 с.
12. Балакришнан А.В. Теория фильтрации Калмана. М.: Мир, 1988. 168 с.
13. Эльясберг П.Е. Определение движения по результатам измерений. М.: Наука, 1976. 416 с.
14. Браммер К., Зиффлинг Г. Фильтр Калмана-Бьюси. Детерминированное наблюдение и стохастическая фильтрация. М.: Наука, 1982. 200 с.
15. Ройтенберг Я.Н. Автоматическое управление. М.: Наука, 1992. 576 с.
16. Кузовков Н.Т., Салычев О. С. Инерциальная навигация и оптимальная фильтрация. М.: Машиностроение, 1982. 216 с.
17. Андреев В.Д. Теория инерциальной навигации. Корректируемые системы. М.: Наука, 1967. 648 с.
18. Мак-Клур К.Л. Теория инерциальной навигации. М.: Наука, 1964. 300 с.
19. Ишлинский А.Ю. Ориентация, гироскопы и инерциальная навигация. М.: Наука, 1976. 670 с.
20. Горенштейн Н.А., Шульман И.А., Сафарян А.С. Инерциальная навигация. М.: Советское радио, 1962. 248 с.
21. Доннел О. Инерциальная навигация. Анализ и проектирование. М.: Наука, 1969. 592 с.
22. Васильев П.В., Мелешко А.В., Пятков В.В. Повышение точности корректируемой инерциальной навигационной системы // Известия высших учебных заведений. Приборостроение. 2014. Т. 57. № 12. С. 15–21.
23. Трефилов П.М. Сравнительный анализ улучшения точностных характеристик инерциальных навигационных систем // XIII Всероссийское совещание по проблемам управления ВСПУ-2019 (Москва, 17–20 июня 2019). М.: Институт проблем управления им. В.А. Трапезникова РАН, 2019. С. 470–474. DOI:10.25728/vspu.2019.0470
24. Munguia R. A GPS-aided Inertial Navigation System in Direct Configuration // Journal of Applied Research and Technology. 2014. Vol. 12. Iss. 4. PP. 803–814.
25. Lai Q., Yuan H., Wei D., Wang N., Li Z., Ji X. A Multi-Sensor Tight Fusion Method Designed for Vehicle Navigation // Sensors. 2020. Vol. 20. Iss. 9. P. 2551. DOI:10.3390/s20092551
26. Zhao F., van Wachem B.G.M. A novel Quaternion integration approach for describing the behaviour of non-spherical particles // Acta Mechanica. 2013. Vol. 224. PP. 3091–3109. DOI:10.1007/s00707-013-0914-2
27. Chen X., Song S., Xing J. A ToA/IMU indoor positioning system by extended Kalman filter, particle filter and MAP algorithms // Proceedings of the 27th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC, Valencia, Spain, 4–8 September 2016). IEEE, 2016. DOI:10.1109/PIMRC.2016.7794980
28. Chen X., Chu S., Li F., Li F., Chu G. Hybrid ToA and IMU indoor localization system by various algorithms // Journal of Central South University. 2019. Vol. 26. PP. 2281–2294. DOI:10.1007/s11771-019-4173-9
29. Lategahn J., Müller M., Röhrig C. Extended Kalman filter for a low cost TDoA/IMU pedestrian localization system // Proceedings of the 11th Workshop on Positioning, Navigation and Communication (WPNC, Dresden, Germany, 12–13 March 2014). IEEE, 2014. DOI:10.1109/WPNC.2014.6843307
30. Sensor Fusion and Tracking Toolbox™ Getting Started Guide. MathWorks. 2018. URL: https://www.mathworks.com/help/releases/R2020b/pdf_doc/fusion/fusion_gs.pdf (дата обращения 30.06.2021)
31. Estimate pose with nonholonomic constraints. MathWorks. 2021. URL: <https://www.mathworks.com/help/fusion/ref/insfilternonholonomic.html> (дата обращения 30.06.2021)
32. Orientation, Position, and Coordinate. MathWorks. 2021. URL: <https://www.mathworks.com/help/fusion/gspatial-representation-coordinate-systems-and-conventions.html> (дата обращения 30.06.2021)
33. Andrieu M.S., Crassidis J.L. Geometric Integration of Quaternions // Journal of Guidance Control and Dynamics. 2013. Vol. 36. Iss. 6. PP. 1762–1767. DOI:10.2514/1.58558
34. Solà J. Quaternion kinematics for the error-state Kalman filter // Preprint arXiv:1711.02508. 2017.
35. Ханукаев Ю.И. Введение в теоретическую механику. М.: МФТИ, 2017. 240 с.
36. Ефремов А.П. Кватернионы: алгебра, геометрия и физические теории // Гиперкомплексные числа в геометрии и физике. 2004. Т. 1. № 1-1. С. 111–127.
37. Гордеев В.Н. Кватернионы и трехмерная геометрия. Киев: Сталь, 2012. 60 с.

* * *

Positioning of Vehicles with the Fusion of Time of Arrival, Angle of Arrival and Inertial Measurements in the Extended Kalman Filter

G. Fokin¹ , A. Vladenko¹ 

¹The Bonch-Bruевич State University of Telecommunications,
St. Petersburg, 193232, Russian Federation

Article info

DOI:10.31854/1813-324X-2021-7-2-51-67

Received 8th May 2021

Accepted 31st May 2021

For citation: Fokin G., Vladiko A. Positioning of Vehicles with the Fusion of Time of Arrival, Angle of Arrival and Inertial Measurements in the Extended Kalman Filter. *Proc. of Telecom. Universities*. 2021;7(2):51–67. (in Russ.) DOI:10.31854/1813-324X-2021-7-2-51-67

Abstract: This work is devoted to the study of models and methods for improving positioning accuracy in ultra-dense V2X/5G radio access networks for vehicles during maneuvers by combining range and angle primary measurements with measurements of inertial navigation systems in the extended Kalman filter. Onboard platformless inertial navigation system is represented by three-axis accelerometer and gyroscope modules. Integration of primary inertial measurements of acceleration and angular velocity with primary radio measurements of range and angle is carried out by converting the inertial coordinate system of the accelerometer and gyroscope into coordinate system of vehicle using quaternions. Secondary processing of inertial and radio measurements is carried out in the extended Kalman filter. The integration results show an increase in the accuracy of estimating the trajectory of a vehicle from several meters to one meter when turning at an intersection.

Keywords: positioning, ultra-dense 5G networks, vehicle to everything, range & angle measurements, inertial navigation system, extended Kalman filter, quaternions.

FUNDING

This research was performed as part of the State of the contract number P33-1-26 / 9.

References

1. Fokin G.A. *Technologies of Network Positioning*. St. Petersburg: The Bonch-Bruевич State University of Telecommunications Publ.; 2020. 558 p. (in Russ.)
2. Fokin G.A., Koucheryavy A.Y. Network Positioning in 5G Ecosystem. *Electrosvyaz*. 2020;9:51–58. (in Russ.) DOI:10.34832/ELSV.2020.10.9.006
3. Kireev A., Fokin G. Positioning of Objects in LTE Networks by Measuring Signal Time of Arrival. *Proc. of Telecom. Universities*. 2016;2(1):68–72. (in Russ.)
4. Kireev A., Fokin G. Radio Direction-Finding of LTE Emissions Using Mobile Spectrum Monitoring Station with Circular Antenna Array. *Trudy NIIR*. 2015;2:68–71. (in Russ.)
5. Sivers M., Fokin G., Dmitriev P., Kireev A., Volgushev D., Al-Odhari A. Indoor Positioning in WiFi and NanoLOC Networks. *Proceedings of the 16th International on Next Generation Wired/Wireless Networking, NEW2AN 2016, and 9th Conference on Conference on Internet of Things and Smart Spaces, ruSMART 2016, St. Petersburg, Russia, 26–28 September 2016. Lecture Notes in Computer Science*. Cham: Springer; 2016. vol.9870. p.465–476. DOI:10.1007/978-3-319-46301-8_39
6. Sivers M., Fokin G., Dmitriev P., Kireev A., Volgushev D., Al-odhari A.H.A. Wi-Fi Based Indoor Positioning System Using Inertial Measurements. *Proceedings of the 17th International Conference on International Conference on Next Generation Wired/Wireless Networking (NEW2AN 2017), 10th Conference on Conference on Internet of Things and Smart Spaces (ruSMART 2017), Third Workshop on International Workshop on Nano-scale Computing and Communications (NsCC 2017), 28–30 August 2017, St. Petersburg, Russia. Lecture Notes in Computer Science*. Cham: Springer; 2017. vol.10531. p.734–744. DOI:10.1007/978-3-319-67380-6_69
7. Kireev A.V., Fokin G.A. Accuracy Evaluation of Local Positioning by Radiomap Building and Inertial Navigation Systems. *Proc. of Telecom. Universities*. 2017;3(4):54–62. (in Russ.)
8. Dukhovnitskiy O.G., Ragilo M.A., Sivers M.A., Fokin G.A. Kalman Filtering for Positioning Applications. *Electrosvyaz*. 2016;1:78–81. (in Russ.)
9. Fokin G., Vladiko A. The Vehicles Positioning in Ultra-Dense 5G/V2X Radio Access Networks Using the Extended Kalman Filter. *Proc. of Telecom. Universities*. 2020;6(4):45–59. (in Russ.) DOI:10.31854/1813-324X-2020-6-4-45-59
10. Kalman R.E. A New Approach to Linear Filtering and Prediction Problems. *Journal of Basic Engineering*. 1960;82(1):35–45. DOI:10.1115/1.3662552
11. Fomin V.N. *Recurrent Estimation and Adaptive Filtering*. Moscow: Nauka Publ.; 1984. (in Russ.)
12. Balakrishnan A.V. *Kalman Filtration Theory*. Moscow: Mir Publ.; 1988. (in Russ.)
13. Elyasberg P.E. *Determination of Motion Based on Measurement Results*. Moscow: Nauka Publ.; 1976. (in Russ.)
14. Brammer K., Ziffiling G. *Kalman-Bucy Filter. Deterministic Observation and Stochastic Filtering*. Moscow: Nauka Publ.; 1982. (in Russ.)
15. Roytenberg Y.N. *Automatic Control*. Moscow: Nauka Publ.; 1992. (in Russ.)
16. Kuzovkov N.T., Salychev O.S. *Inertial Navigation and Optimal Filtration*. Moscow: Mashinostroenie Publ.; 1982. (in Russ.)
17. Andreev V.D. *The Theory of Inertial Navigation. Corrected Systems*. Moscow: Nauka Publ.; 1967. (in Russ.)
18. McClure K.L. *The Theory of Inertial Navigation*. Moscow: Nauka Publ.; 1964. (in Russ.)
19. Ishlinsky A.Y. *Orientation, Gyroscopes and Inertial Navigation*. Moscow: Nauka Publ.; 1976. (in Russ.)

20. Gorenstein N.A., Shulman I.A., Safaryan A.S. *Inertial Navigation*. Moscow: Sovetskoe radio Publ.; 1962. (in Russ.)
21. Donnell O. *Inertial Navigation. Analysis and Design*. Moscow: Nauka Publ.; 1969. (in Russ.)
22. Vasiliev P.V., Meleshko A.V., Pyatkov V.V. Accuracy Improvement of Correctable Inertial Navigation System. *Journal of Instrument Engineering*. 2014;57(12):15–21. (in Russ.)
23. Trefilov P.M. Comparative Analysis of Improving the Accuracy Characteristics of Inertial Navigation Systems. *Proceedings of the XIIIth All-Russian Meeting on Control Problems, 17–20 June 2019, Moscow, Russia*. Moscow: V.A. Trapeznikov Institute of Control Sciences RAS Publ.; 2019. p.470–474. (in Russ.) DOI:10.25728/vspu.2019.0470
24. Munguía R. A GPS-aided Inertial Navigation System in Direct Configuration. *Journal of Applied Research and Technology*. 2014;12(4):803–814.
25. Lai Q., Yuan H., Wei D., Wang N., Li Z., Ji X. A Multi-Sensor Tight Fusion Method Designed for Vehicle Navigation. *Sensors*. 2020;20(9):2551. DOI:10.3390/s20092551
26. Zhao F., van Wachem B.G.M. A novel Quaternion integration approach for describing the behaviour of non-spherical particles. *Acta Mechanica*. 2013;224:3091–3109. DOI:10.1007/s00707-013-0914-2
27. Chen X., Song S., Xing J. A ToA/IMU indoor positioning system by extended Kalman filter, particle filter and MAP algorithms. *Proceedings of the 27th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications, PIMRC, 4–8 September 2016, Valencia, Spain*. IEEE; 2016. DOI:10.1109/PIMRC.2016.7794980
28. Chen Xc., Chu S., Li F., Li F., Chu G. Hybrid ToA and IMU indoor localization system by various algorithms. *Journal of Central South University*. 2019;26:2281–2294. DOI:10.1007/s11771-019-4173-9
29. Latgahn J., Müller M., Röhrig C. Extended Kalman filter for a low cost TDoA/IMU pedestrian localization system. *Proceedings of the 11th Workshop on Positioning, Navigation and Communication, WPNC, 12–13 March 2014, Dresden, Germany*. IEEE; 2014. DOI:10.1109/WPNC.2014.6843307
30. MathWorks. Sensor Fusion and Tracking Toolbox™ Getting Started Guide. 2018. Available from: https://www.mathworks.com/help/releases/R2020b/pdf_doc/fusion/fusion_gs.pdf [Accessed 15th June 2021]
31. MathWorks. Estimate pose with nonholonomic constraints. 2021. URL: <https://www.mathworks.com/help/fusion/ref/insfilternonholonomic.html> [Accessed 15th June 2021]
32. MathWorks. Orientation, Position, and Coordinate. 2021. URL: <https://www.mathworks.com/help/fusion/gs/spatial-representation-coordinate-systems-and-conventions.html> [Accessed 15th June 2021]
33. Andrieu M.S., Crassidis J.L. Geometric Integration of Quaternions. *Journal of Guidance Control and Dynamics*. 2013;36(6):1762–1767. DOI:10.2514/1.58558
34. Solà J. Quaternion kinematics for the error-state Kalman filter. *Preprint arXiv:1711.02508*. 2017.
35. Khanukaev Y.I. *Introduction to Theoretical Mechanics*. Moscow: Moscow Institute of Physics and Technology Publ.; 2017. (in Russ.)
36. Efremov A.P. Quaternions: Algebra, Geometry and Physical Theories. *Hypercomplex Numbers in Geometry and Physics*. 2004;1(1-1):111–127. (in Russ.)
37. Gordeev V.N. *Quaternions and 3D Geometry*. Kiev: Stal' Publ.; 2012. (in Russ.)

Сведения об авторах:

ФОКИН
Григорий Алексеевич

кандидат технических наук, доцент, доцент кафедры радиосвязи и вещания Санкт-Петербургского государственного университета телекоммуникаций им. проф. М.А. Бонч-Бруевича, grihafokin@gmail.com
 <https://orcid.org/0000-0002-5358-1895>

ВЛАДЫКО
Андрей Геннадьевич

кандидат технических наук, директор НИИ «Технологии связи» Санкт-Петербургского государственного университета телекоммуникаций им. проф. М.А. Бонч-Бруевича, vladyko@sut.ru
 <https://orcid.org/0000-0002-8852-5607>

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ И ТЕЛЕКОММУНИКАЦИИ

**2.3.1 – Системный анализ,
управление и обработка
информации**

**2.3.6 – Методы и системы защиты
информации, информационная
безопасность**

Развитие теоретических основ синтеза космических адаптивных видеоинформационных систем

А.И. Бобровский¹ 

¹Государственный научно-исследовательский институт прикладных проблем

Санкт-Петербург, 191167, Российская Федерация

*Адрес для переписки: albob@mail.ru

Информация о статье

Поступила в редакцию 28.05.2021

Принята к публикации 21.06.2021

Ссылка для цитирования: Бобровский А.И. Развитие теоретических основ синтеза космических адаптивных видеоинформационных систем // Труды учебных заведений связи. 2021. Т. 7. № 2. С. 69–78. DOI:10.31854/1813-324X-2021-7-2-69-78

Аннотация: Рассмотрены основные положения и принципы синтеза адаптивных видеоинформационных систем, предназначенных для регистрации и анализа изображений динамических сцен на этапах дальнего обнаружения искусственных космических объектов, баллистического и управляемого сближения с ними в средней и ближней зоне кооперации. Исходя из принципа доминантной информации, показано, что основой развития методов и алгоритмов обработки видеоинформации в бортовых системах наблюдения является совместная адаптивная обработка видеоинформации в фотоприемных матрицах и в цифровых процессорах с целью максимизации качества информации, выдаваемой потребителю. Это делается за счет подавления фоновой и шумовой информации путем адаптивного гистерезисного управления кадровой частотой, разрешением кадра и временем накопления сигнала в фотоприемной матрице, а также использование параметрических нейросетевых обнаружителей-классификаторов и адаптивных кодеров изображений доминантного объекта и фона.

Ключевые слова: синтез, видеоинформационная система, адаптация, доминантная, фоновая и шумовая информация, кадровая частота, разрешение, время накопления, кодирование, параметрическая оптимизация.

Введение

Современные космические видеоинформационные системы (телевизионные камеры и цифровые фотоаппараты), эксплуатирующиеся на борту космических аппаратов (КА) и орбитальных станций, предназначены для регистрации и выдачи изображений наблюдаемых сцен в определенный контур автоматического или автоматизированного управления. Традиционно считается, что требуемое качество сформированных изображений, поступающих в контур управления, определяется целью управления и требованиями, предъявляемыми к показателям качества этого управления. Научной школой НИИ телевидения [1–5] предложено использовать информационный подход к оценке качества функционирования телевизионных (видеоинформационных) систем, опирающийся на принцип доминантной информации [6]. Существо данного принципа может быть кратко определено следующим образом.

Сигнал, принимаемый видеоинформационной системой, в общем случае включает сигналы от целевых объектов (доминантных объектов), сигналы от фоновых объектов и шумовые сигналы. В определенных случаях среди наблюдаемых фоновых объектов можно выделить «фейковые» (ложные) объекты, а остальные отнести к «истинно» фоновым.

Априорно все компоненты входного сигнала равноправны и только разработчик видеоинформационной системы, определивший алгоритмы ее работы, или потребитель формируемых изображений, управляющий оперативно изменяемыми параметрами видеоинформационной системы, решают, какую из компонент следует считать доминантным сигналом, а какие – фоновыми, шумовыми или фейковыми (ложными) сигналами. Заметим, что далее в настоящей работе фейковые объекты и фейковая информация не рассматриваются, поскольку являются предметом других специальных исследований.

Соответственно, сигнал, зарегистрированный фотоприемной матрицей («сырое» изображение/кадр) несет (содержит) *доминантную, фоновую и шумовую информацию*.

Качество информации Q , содержащейся в зарегистрированном сигнале, предлагается оценивать через информационный риск R_I потери части доминантной информации $\otimes I_D$ и сохранения части фоновой I_F и шумовой $I_{\text{ш}}$ информации. Количество доминантной, фоновой и шумовой информации оценивается по Шеннону, а качество информации математически определяется как отношение количества *потенциальной доминантной* I_D информации к указанному информационному риску [4]:

$$Q = I_D / R_I = \frac{I_D}{c_0 \Delta I_D + c_1 (I_F + I_{\text{ш}})},$$

где c_0 и c_1 – весовые коэффициенты, отражающие относительные стоимость потери доминантной информации и стоимость имеющихся (оставшихся) фоновой и шумовой информации в сигнале, выдаваемом фотоприемной матрицей, процессором обработки видеоинформации или видеоинформационной системой в целом.

Важным следствием применения принципа доминантной информации является информационное объяснение причины использования «усечения» спектра сигнала В.А. Котельниковым и К. Шенноном в теории связи – максимизация качества информации [4]. Опираясь на принцип доминантной информации, в работах [7–15] была предложена концепция адаптации космической видеоинформационной системы наблюдения к этапам наблюдения для повышения качества выдаваемой информации, расширения горизонта чувствительности системы и приближения границы своевременной реакции потребителя видеоинформации на события наблюдаемых сцен. Данная концепция сформирована и использована в ходе развития теоретических основ синтеза методов и алгоритмов адаптивной обработки видеоинформации в системах наблюдения динамических сцен в космическом пространстве.

В настоящей работе рассмотрены основные положения и принципы синтеза таких адаптивных видеоинформационных систем (ВИС).

Краткий системный анализ космических ВИС с позиций обеспечения адаптивности к условиям их применения

Важнейшими подсистемами (устройствами) современных космических ВИС являются объектив, фотоприемное устройство (ФПУ), видеопроцессор, устройство сопряжения с внешними абонентами и устройство управления режимами и параметрами ВИС. Каждое из перечисленных устройств в ходе разработки (проектирования, синтеза) ВИС характеризуется набором параметров, которые можно разделить на важнейшие (выбираемые в ходе

структурного и параметрического синтеза) и условно второстепенные (выбираемые в ходе конструкторско-технологического проектирования). Множество важнейших параметров конкретной ВИС, предназначенной для определенных условий применения, в свою очередь можно разделить на два подмножества: регулируемые и нерегулируемые параметры на этапе эксплуатации.

В адаптивных ВИС регулируемые параметры принимают определенные значения (оптимальные или субоптимальные) в результате работы контуров автоматического управления, реализованных в таких ВИС. Решение об управляющем воздействии на регулируемый параметр принимается системой управления ВИС на основе обработки сигналов, поступающих в общем случае из ФПУ, датчиков состояния устройств, датчиков внешних воздействий и внешних командных интерфейсов ВИС.

Объектив является первым устройством в видеотракте ВИС, определяющим характеристики сигнала (изображения), поступающего в ФПУ и находящегося (с определенными допущениями) в фокальной плоскости объектива. Регулируемыми параметрами объектива, в зависимости от его конструкции, могут быть фокусное расстояние, диафрагма (ширина апертуры, диаметр входного зрачка), компенсация механических вибраций и спектральная полоса пропускания. Управление этими параметрами широко известно и реализовано во многих современных объективах. Традиционной целью управления этими параметрами является максимизация качества изображения для потребителя (фотографа, кинооператора, оператора) по критерию, определяемому потребителем или производителем объектива.

Вторым устройством видеотракта ВИС, определяющим характеристики «сырого» изображения (сигнала), выдаваемого потребителю, является фотоприемная матрица (ФПУ, содержащее управляемые оптоэлектронные элементы). Важнейшими регулируемыми параметрами твердотельного приемника оптического сигнала являются кадровая частота F_k , число элементов разложения N (число пикселей в кадре, разрешение фотоприемника, определяющее четкость статического изображения), время накопления T_n сигнала кадра ($T_n \leq 1/F_k$). Минимальный размер пиксела $\otimes_p$ относится к важнейшим параметрам ФПУ, влияющим на качество информации в сформированном кадре, и выбирается на этапе синтеза этого устройства для ВИС определенного класса. Реализация автоматического управления параметрами F_k , N , T_n по результатам обработки формируемых изображений динамических сцен в видеопроцессоре с целью максимизации качества информации, выдаваемой потребителю, *существенно развивает* свойство (функционал) адаптивности ВИС по сравнению с

системами, в которых реализовано автоматическое управление параметрами объектива.

Видеопроцессор космических ВИС реализует алгоритмы обработки кадров, считанных с ФПУ, для решения задач обнаружения изображений целевых (доминантных) объектов, их классификации, сопровождения и оценивания пространственно-временных координат этих объектов для оперативной передачи этих данных, включая изображение наблюдаемой сцены, потребителю (оператору или в вышестоящую систему управления, для которой адаптивная ВИС является интеллектуальным датчиком видеoinформации). Именно в видеопроцессоре реализуются алгоритмы управления параметрами ФПУ, объектива и параметрами алгоритмов обнаружения-классификации доминантных объектов, оценивания их пространственно-временных координат, а также алгоритма кодирования (сжатия) изображения для передачи его потребителю по каналу с ограниченной пропускной способностью.

Несомненно, что методы и алгоритмы адаптации ВИС к условиям наблюдения космических динамических сцен должны учитывать этапность (сценарий) применения их функционала и другую априорно известную информацию, необходимую для их «асимптотического» синтеза.

Рассматриваемый класс ВИС ориентирован на решение задач обнаружения доминантного объекта – искусственного космического объекта (ИКО) или КА на больших дальностях, получения оценок пространственно-временных координат для реализации управляемого сближения, передачи изображений наблюдаемой сцены с максимальным качеством информации, используемой потребителем для контроля за сближением и осуществления безаварийной стыковки или корректного обслуживания. Теоретические основы синтеза адаптивных ВИС рассматриваемого класса включают концепцию адаптации и систему принципов, реализуемых в методах и алгоритмах управления параметрами ФПУ и кодера выдаваемых изображений на основе согласованной обработки информации в видеотракте для максимизации качества выдаваемой потребителю информации, также используемые при синтезе модели методы и алгоритмы обработки информации.

Концепция адаптации космической ВИС к этапам наблюдения

Так как все реальные сигналы от наблюдаемых космических объектов характеризуются априорной неопределенностью и нестационарностью, оптимальная обработка информации при наблюдении динамических сцен неизбежно должна быть адаптивной к изменяющимся статистикам. Необходимо достижение в реальном масштабе времени высокого качества выдаваемой потребителю информации, обеспечивающего высокую достовер-

ность обнаружения и точность измерения пространственно-временных координат объектов в широком диапазоне дальностей и видимых скоростей. Теоретические основы адаптации видеoinформационных систем опираются на теории информации, решений, управления, систем [9] и включают следующую концепцию адаптации системы наблюдения к этапам наблюдения.

Рассматриваются два класса адаптации: адаптацию цифровой обработки и кодирования при конкретных фиксированных параметрах фотоприемника и адаптацию накопления информации в фотоприемной матрице. Если адаптация обработки уже сформированного сигнала предназначена для наилучшего согласования системы с каналом связи, то адаптация фотоприемной матрицы предназначена для наилучшего согласования системы с оптическими свойствами наблюдаемых динамических сцен. Методологическая общность адаптации этих двух подсистем ВИС определяется тем, что выбор наилучшей кадровой частоты (времени кадра) является компромиссом, аналогичным компромиссу между длиной (надежностью) кода и скоростью передачи в теории кодирования.

Для расширения горизонта чувствительности системы (актуальной на дальней дистанции) и уменьшения границы своевременной реакции потребителя на события, наблюдаемые на сцене, (что актуально на малой дистанции) необходима *адаптация кадровой частоты* космической ВИС к этапам наблюдения, характеризующимися различиями в удаленности и видимой скорости доминантного объекта, а также оставшимся временем до запланированной терминальной точки сближения. С позиции теории космического телевидения разные этапы характеризуются различием в статистиках принимаемого сигнала.

В отличие от многих систем автоматического регулирования, контур адаптации кадровой частоты ВИС нецелесообразно описывать с использованием дифференциальных уравнений. Наиболее подходящее математическое описание таких контуров дает теория конечных автоматов, поскольку множество состояний контура конечно (технически эффективно реализуемое множество возможных значений F_k , N , T_n) и переходы между состояниями происходят в дискретном времени. При этом ВИС, хотя и имеет замкнутые контуры управления параметрами накопления информации на основе оценок статистик сигнала, не является замкнутой, имеет не только оптический вход, на который подается многомерный сигнал $S(x, y, t, \lambda)$, но и на выходе цифровой код в виде электромагнитного сигнала в определенного компьютерного или сетевого интерфейса.

Использование гистерезиса в контуре автоматического управления в реальном времени параметрами ФПУ концептуально связано с двумя концеп-

циями, относящимися к фундаментальной теории проверки гипотез, но акцентирующих внимание не просто на достоверности принятого решения, но и на времени, необходимом для его принятия.

Во-первых, это теория последовательных решений [16, 17], в которой А. Вальд ввел принципиальный отказ от анализа заданного интервала времени, используемого в классической теории (например, [18–21]. Первым важным результатом этой теории явилось использование не одного порога принятия решения, а двух – верхнего γ_v и нижнего γ_n . При превышении Z_k верхнего порога принимается одна гипотеза, при значении сформированной статистики Z_k ниже нижнего порога принимается альтернативная гипотеза, при нахождении этой статистики между порогами принимается решение о необходимости продолжения наблюдений. В течение этого времени наблюдений осуществляется накопление информации о наблюдаемом процессе, реализуемое с помощью фильтра с импульсной характеристикой h_k с опорой на интегрирование наблюдений.

Второй родственной концепцией, использующей гистерезис, является задача Колмогорова – Ширяева об обнаружении разладки [17]. В ее классической постановке минимизируется время принятия решения об изменении значения выбранной статистики (например, математического ожидания случайного процесса) при заданной вероятности ложной тревоги. Здесь так же, как в теории последовательных решений, осуществляется накопление информации о наблюдаемом процессе, реализуемое с помощью операций фильтрации с опорой на интегрирование наблюдений. Вместе с концептуальной общностью с задачей о разладке в адаптивной системе перестройки кадровой частоты, благодаря дополнительной опоре на принцип максимизации времени нахождения системы в одном из возможных устойчивых состояний ($\max T$, [22]), исключается неизбежная неопределенность, связанная с назначением вероятности ложной тревоги.

Концепция адаптации ВИС имеет как общие черты с адаптивными системами управления, так и свою специфику. К свойствам адаптивных ВИС, общим с другими адаптивными системами, относится ряд широко известных принципов адаптации [23, 24]:

- адаптивное управление рассматривается как управление состоянием ВИС в условиях случайных (коррелированных) входных сигналов с неизвестной функцией распределения;

- основу «механизма» адаптации составляет теория итераций с последовательным использованием измерительных данных, что является логической основой перехода к системам с дискретными состояниями и с гистерезисом;

- адаптивная система относится к классу существенно нелинейных замкнутых систем управления, реализация которых невозможна при использова-

нии линейных методов теории управления, а должна опираться на пороговое обнаружение достаточного изменения анализируемых статистик, в том числе характеризующей относительную динамику сцены.

Наличие гистерезиса означает, что системы адаптации параметров при их перестройке реализуют принцип трихотомии, выполняя в каждый момент дискретного времени, определяемого кадровой частотой, одну из трех возможных операций:

- перейти вверх (увеличить на единицу номер состояния системы, например, увеличить вдвое кадровую частоту);

- остаться в том же состоянии;

- перейти вниз (увеличить на единицу номер состояния системы, например, понизить вдвое кадровую частоту).

Концепция адаптации ВИС включает новую трактовку принципа равенства кадровых статистик дисперсий межэлементной и межкадровой разности, введенной в работах [5, 24]. Этот принцип используется не только на ближней дистанции, но и на дальней. При этом, в силу возможной малости сигнала доминантного объекта – наблюдаемого ИКО, алгоритм адаптации к дистанции до него работает устойчивее при переходе к другому, связанному с ним критерию – поддержанию в заданных границах другой кадровой статистики: смаза фоновых звезд. В результате формируются три вида статистик и три контура управления параметрами фотоприемника.

Главная идея адаптации кадровой частоты и времени накопления: для расширения диапазона рабочих дальностей нужно кривые яркости и динамики «сгладить». Для этого на дальней дистанции, где значения статистики яркости и динамики сцены малы, следует понижать кадровую частоту, а на ближней дистанции, где статистики динамики велики, следует повышать кадровую частоту, что дополняется на средней и ближней дистанции управлением временем накопления внутри кадра.

С повышением кадровой частоты при фиксированном разрешении кадра возрастает вычислительная нагрузка на видеопроцессор, что, в условиях значительных ресурсных ограничений и не превышения требуемого времени выдачи информации потребителю, принимающему решения по управлению сближением, обуславливает целесообразность обмена роста F_k на уменьшение величины N при сохранении достаточного качества выдаваемой информации. Таким образом, производится адаптация разрешения формируемых изображений для сохранения (при определенных условиях – для приближения) границы своевременной реакции потребителя видеoinформации на события наблюдаемой сцены, т. е. для обеспечения результативного управления динамическим объектом (процессом) за счет сохранения скорости выдачи информации с интеллектуального видеосенсора.

Изложенная концепция является отражением новой парадигмы в теории прикладных видеоинформационных систем, пришедшей на смену старой парадигмы пассивного учета уменьшения разрешающей способности при движении наблюдаемого объекта. При этом, если на начальном этапе перехода от старой концепции (Я.А. Рыфтин) к новой концепции, учет динамики сцены осуществлялся лишь при синтезе устройств обработки и кодирования изображений (Н.Н. Красильников, Л.П. Ярославский, И.И. Цуккерман, Д.С. Лебедев, Я. Ричардсон и др.), то новый этап учета параметров наблюдаемой сцены опирается на совместную адаптивную обработку изображений динамических сцен в фотоприемной матрице и цифровом вычислителе в реальном масштабе времени.

Принципы, положенные в основу синтеза космических ВИС, адаптивных к этапам и условиям их применения

Выше были упомянуты только некоторые важнейшие принципы, на которые опирается концепция адаптации ВИС к этапам наблюдения. Вся совокупность принципов, используемых для синтеза адаптивных космических ВИС, следующая.

Принцип оптимальности, включающий статистический итерационный синтез методов адаптации на основе априорной информации, математических моделей, критериев и ограничений для оптимизируемых систем. Под статистическим синтезом понимается математический аппарат теории решений и теории информации; итерационный подход, чередующий методы структурного и параметрического синтеза, является одной из методологических основ адаптации информационных систем. Этот принцип после нахождения параметров системы, обеспечивающих искомый условный оптимум, позволяет обоснованно считать синтез завершенным.

Принцип доминантной информации, который обосновывает цель синтеза по критерию максимума качества информации, определяемого отношением потенциальной доминантной информации к информационному риску, включающему потерянную доминантную информацию и остаточные фоновую и шумовую информации. Он дает единую основу параметрического синтеза фоточувствительной и вычислительной подсистем обработки информации.

Принцип итерационного наблюдения и управления, включающий изменение параметров системы на разных этапах наблюдения по определенному правилу перехода к новой итерации. Он позволяет минимизировать совокупность аппаратных и/или временных затрат для достижения экстремальных значений выбранных критериев.

Принцип динамического выравнивания дисперсий ошибок оценивания параметров изображений наблюдаемых сцен (совокупности объектов), являющийся одним из важных следствий принципа оптимальности, и опирающийся на методы вариационного исчисления и динамического программирования, в том числе:

- выравнивание дисперсий приращений сигнала, формируемого фотоприемной матрицей по полю и по времени [5, 7–9, 25];
- выравнивание дисперсий ошибок кодирования источника.

Принцип системного согласования информационных параметров частей системы после ее декомпозиции, обеспечивающий методологическое единство проектирования системы с учетом принципа доминантной информации, в частности единства синтеза трех важнейших узлов – объектива, фотоприемника и устройства компьютерной обработки – а не их попарного согласования.

Принцип использования инерционности и гистерезиса в контуре отрицательной обратной связи, обеспечивающий в совокупности с третьим и седьмым принципами устойчивость системы к малым (шумовым) изменениям управляющих переменных – отсутствие в ней «дребезга» управления (колебательного процесса управления параметрами F_k, N, T_n с частотой близкой к кадровой) [7–9, 26].

Принцип максимизации времени нахождения системы в заданной зоне фазового пространства, характеризуемой оптимальными пороговыми (верхними и нижними) относительными значениями управляющих величин, который в совокупности с третьим и шестым принципом обеспечивает устойчивость управления параметрами фотоприемной матрицы.

Основные используемые методы адаптации ВИС

Основными методами адаптации бортовых видеоинформационных систем, в которых теоретические принципы дополняются учетом сложности моделей сигналов, сложности управления носителем бортовой видеоинформационной системы и сложности ее эксплуатации, являются [9]:

- адаптация чувствительности (автоматическое регулирование времени накопления) при фиксированных кадровой частоте и четкости к уровню освещенности [1];
- адаптация порога принятия решения [3] и четкости фотоприемных матриц астрономических систем к интенсивности сигналов звезд и этапам обнаружения–оценивания;
- адаптация областей считывания с фотоприемника [2, 3], и их цифрового кодирования [15] при разделимости сцены на объект и фон, на порядок увеличивающие коэффициент сжатия передаваемого изображения без потери доминантной информации;

– адаптация кадровой частоты и разрядности сигнала к динамике сцен с разными каналами считывания фотоприемника для высоких и низких кадровых частот [5], на порядок расширяющая горизонт чувствительности в системах орбитального обслуживания при селекции объектов по скоростному смазу;

– адаптация кадровой частоты с обменом на четкость матричного фотоприемника к динамике сцены [5, 7–11, 25], на порядок расширяющая диапазон дистанций с минимальной ошибкой оценивания пространственно-временных координат объекта;

– адаптация спектральных (по длине волны принимаемого фотоприемной матрицей излучения) характеристик оптических систем и фотоприемников к характеристикам излучения и к свойствам атмосферы [5].

Реализованные методы и алгоритмы управления параметрами ВИС для обнаружения и оценки наблюдаемых параметров сигналов кооперируемого КА в широком диапазоне дальностей и динамических свойств наблюдаемой сцены развивают принцип итерационного контроля сближения КА, сформулированный С.П. Королевым в 1962 г.

Предложенная систематизация принципов теории синтеза адаптивных бортовых систем компьютерного зрения, работающих в реальном времени, учитывает влияние твердотельной технологии получения изображений на изменение методов системного анализа и синтеза видеосистем на основе применения принципа доминантной информации.

Дальнейшее развитие методов адаптации видеoinформационных систем должно привести к влиянию этого направления в русло теории распознавания образов и искусственного интеллекта.

Пример параметрической оптимизации сверточной нейронподобной сети

Перспективным направлением построения космических адаптивных ВИС является использование нейросетевых технологий для анализа изображений наблюдаемых сцен. Высокая вычислительная и структурная сложность нейронподобных сетей, эффективно решающих задачи обнаружения-классификации изображений различных объектов, приводит к необходимости поиска решений, обеспечивающих возможность их реализации при жестких ресурсных ограничениях по энергопотреблению, габаритам и массе. В работе [27] выделен класс нейросетевых архитектур типа YOLO [28], который на сегодняшний день является наиболее подходящим кандидатом для реализации на борту. Важнейшими варьируемыми параметрами архитектуры YOLO для заданного множества различаемых классов изображений, которые определяют вычислительную и структурную сложность сети,

являются: число сверточных слоев (L), информационный размер изображения ($N_x \times N_y \times p$, байт), подаваемого на вход сети. Значимыми параметрами архитектуры YOLO, определяющими ресурсоемкость сети, являются параметры нейронподобных элементов сверточных слоев (вид функции активации, разрядность и способ кодирования входных и выходных переменных, весовых коэффициентов, значения которых формируются в процессе обучения), слоев субдискретизации и двух выходных полносвязных слоев, формирующих окончательный результат. На сегодняшний день отсутствуют сведения о проведенных полномасштабных исследованиях, посвященных параметрической оптимизации нейросетевых архитектур типа YOLO. Данный факт, по-видимому, объясняется большой размерностью задачи, неразработанностью методов аналитического решения (даже «асимптотического») подобных задач и высокой трудоемкостью цикла «обучение сети – оценка качества обучения на множестве тестовых изображений», лежащего в основе действенного, но научно неинтересного, метода перебора (пусть даже «направленного») возможных значений оптимизируемых параметров архитектуры. В работе [29] представлены первичные результаты экспериментальной оценки влияния числа слоев сети типа TinyYOLO ver.2-416 на показатели качества решения задачи обнаружения-классификации-оценивания пространственного положения изображений объектов, имеющих простейшую форму (трек-смаз, «точка»-«круглое пятно»). Сеть данного типа в указанном исследовании использовалась для анализа кадров существенно большего разрешения, чем 416×416 пикселей в соответствии с методом последовательного анализа пересекающихся прямоугольных фрагментов кадра. При этом размер пересечения l_c фрагментов изображений был согласован с максимальной длиной l_s треков (смазов) – изображений «точечных» светящихся объектов, проекция вектора скорости V_f которых на фокальную плоскость ВИС, вместе с временем накопления сигнала T_n и диаметром D неподвижного изображения «точечного» объекта определяла величину смаза (в пикселах):

$$l_s = \left\lceil V_f \frac{T_n}{p} + \frac{D_o}{p} \right\rceil,$$

$$l_c = \max_{V_f, T_n} \left(\left\lceil \frac{V_f T_n + D_o}{\Delta p} \right\rceil \right),$$

где Δp – линейный размер пиксела; $\lceil a \rceil$ – округление нецелого числа a до ближайшего большего целого.

Экспериментальные исследования с использованием нейронной сети TinyYOLO v2-416 и обучающей выборки из совокупности наблюдений семи геостационарных КА при широком наборе кадровых частот ВИС (время накопления кадра 0,25–8 с) и яркости фона от глубокой ночи до сумерек [5] показали,

что существует минимально достаточное число слоев сети [12], обеспечивающих максимальное (предельно возможное) качество информации (результата решения задачи обнаружения-классификации), определяемое показателем информационного риска, в котором учитываются не только информационные сигналы, обуславливающие наличие ошибок обнаружения-классификации, но рост задержки в выдаче решения обнаружения-классификации с увеличением числа слоев.

На рисунке 1 приведены графики качества информации (с учетом задержки в выдаче решения), построенные по результатам экспериментальных исследований сетей типа TinyYOLO v2-416 с разным числом слоев и оптимального (несетевого) обнаружителя-классификатора изображений точечных объектов по признаку скоростного смаза, синтезированного для статистик полученного (заданного) набора кадров с наблюдаемыми КА, звездами и шумами.

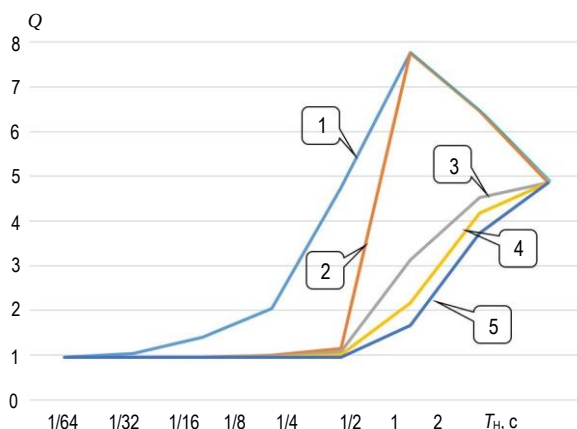


Рис. 1. Зависимость качества информации от времени накопления сигнала (кадра) для разного числа слоев нейроразличной сети типа Tiny YOLO v2-416 (графики 2–5) и оптимального обнаружителя классификатора (график 1), имеющего предельное качество; график 2 – сеть имеет 9 сверточных слоев, график 3 – 8 слоев, график 4 – 7 слоев, график 5 – 6 слоев

Fig. 1. The Dependence of the Quality of Information on the Time of Signal Accumulation (Frame) for a Different Number of Layers a Neural Network of the Tiny YOLO v2-416 Type (Graphs 2–5) and an Optimal Classifier Detector (Graph 1), which has the Maximum Quality; Graph 2 - the Network Has 9 Convolutional Layers, Graph 3 – 8 Layers, Graph 4 – 7 Layers, Graph 5 – 6 Layers

Получаемое качество информации имеет выраженный максимум, так как при слишком малом времени кадра мало и отношение сигнал/шум доминантного объекта, и мало расстояние между кластерами объекта и фоновых сигналов; при слишком большом времени кадра качество информации падает из-за слишком больших задержек оценивания координат объекта. Для данного случая оптимальным оказалась число слоев, равное 9. Увеличение числа слоев свыше при $T_n \geq 1$ с не повысит максимального значения качества информации. Уменьшение числа слоев с 9 до 6 и менее существенно снижает качество информации, выдаваемой сетью.

Следует отметить, что рассмотренный случай характеризуется простотой формы обнаруживаемых и классифицируемых изображений объектов, малым числом классов изображений объектов, детектируемых сетью. Именно эти обстоятельства обусловили возможность использования сети с архитектурой YOLO с небольшим числом слоев. Для случаев с существенно большим числом детектируемых классов изображений и большей «вариативностью» изображений одного класса несомненно потребуются использовать сети типа YOLO с существенно большим числом слоев. И поиск оптимального числа слоев для подобных случаев является существенно более трудоемкой задачей.

Использование нейросетевых решений, оптимизированных для заданных типовых условий применения, является одним из магистральных направлений развития космических адаптивных ВИС.

Наличие показанного на рисунке 1 максимума иллюстрирует конструктивные возможности предложенного подхода к синтезу космических ВИС на основе использованных моделей сигналов и фонов в виде квазидетерминированных сигналов с параметрической неопределенностью и модификации информационного риска введением компонента, отражающей задержку получения оцениваемых координат объекта.

Адаптивное кодирование изображений доминантного объекта и фона

Одним из следствий использования сверточных нейроразличных сетей типа YOLO является определение в реальном масштабе времени параметров прямоугольников, в которых находятся обнаруженные и отклассифицированные изображения объектов. Координаты прямоугольников формирует выходной слой сети. Такие прямоугольники являются прямоугольными стробами или возможными зонами интереса (внимания) оператора. Построение алгоритма выделения среди нескольких возможных зон интереса одной или нескольких, содержащих изображения доминантных (целевых) объектов, является достаточно простой задачей, решение которой позволяет «разделить» кадр, передаваемый в контур управления оператора, на компактную область с изображением доминантного объекта и на оставшуюся область, содержащую изображения фона.

Описанная идея легла в основу метода селективного (семантического) сжатия видеокadra: область с доминантным объектом обрабатывается алгоритмом сжатия с минимальными или нулевыми потерями, а оставшаяся область обрабатывается алгоритмом с максимально возможным коэффициентом сжатия и с допустимыми потерями качества восстановленного изображения. Проведенные экспериментальные исследования показали [15], что

предложенный метод селективного (семантического) сжатия позволяет на порядок увеличить коэффициент сжатия передаваемого изображения без потери доминантной информации.

Заключение

Из приведенных положений ясно, что развитие теоретических основ синтеза космических видеосистем проведено в направлении осуществления совместной адаптивной обработки изображений динамических сцен в фотоприемной матрице и бортовом вычислителе на основе применения принципа доминантной информации для повышения и последующего сохранения качества информации, используемой в контурах управления КА. При этом (как следствие) достигаются важные прагматические цели: повышение дальности обнаружения и

точности измерения пространственно-временных координат наблюдаемых космических объектов, а также сохранение скорости выдачи информации достаточного качества, используемой потребителем для результативного управления динамическим объектом (процессом).

Стержнем (основной идеей) предлагаемых теоретических положений является переход от концепции обработки информации при наблюдении динамических сцен космическими видеоинформационными системами, традиционно использующими фиксированные параметры разложения изображения, к новой концепции адаптации параметров формирования и обработки изображений к динамическим свойствам наблюдаемой сцены.

Список используемых источников

1. Хромов Л.И., Лебедев Н.В., Цыцулин А.К., Куликов А.Н. Твердотельное телевидение. М.: Радио и связь, 1986. 184 с.
2. Хромов Л.И., Цыцулин А.К., Куликов А.Н. Видеоинформатика. М.: Радио и связь, 1991. 192 с.
3. Березин В.В., Умбиталиев А.А., Фахмин Ш.С., Цыцулин А.К., Шипилов Н.Н. Твердотельная революция в телевидении: телевизионные системы на основе приборов с зарядовой связью, систем на кристалле и видеосистем на кристалле. М.: Радио и связь, 2006. 312 с.
4. Цыцулин А.К., Адамов Д.Ю., Манцветов А.А., Зубакин И.А. Твердотельные телекамеры: накопление качества информации. СПб.: СПбГЭТУ «ЛЭТИ», 2014. 272 с.
5. Умбиталиев А.А., Цыцулин А.К., Левко Г.В., Пятков В.В., Кузичкин А.В., Дворников С.В. и др. Теория и практика космического телевидения. СПб: АО «НИИ телевидения», 2017. 368 с.
6. Хромов Л.И., Цыцулин А.К. Основания космической видеоинформатики // Вопросы радиоэлектроники. Серия: Техника телевидения. 2011. № 1. С. 6–31.
7. Бобровский А.И. Адаптация прикладной телевизионной системы к динамике сюжета // Цифровая обработка сигналов. 2018. № 3. С. 3–9.
8. Бобровский А.И. Обработка видеоинформации в адаптивной системе контроля сближения космических аппаратов // Информация и космос. 2018. № 4. С. 149–153.
9. Бобровский А.И. Теоретические основы синтеза космических адаптивных видеоинформационных систем // Вопросы радиоэлектроники. Серия: Техника телевидения. 2019. № 4. С. 42–46.
10. Цыцулин А.К., Фахми Ш.С., Адамов Д.Ю., Бобровский А.И., Зубакин И.А., Черногоубов А.В. Принцип доминантной информации – концептуальная основа проектирования видеоинформационных систем // Датчики и системы. 2018. № 11. С. 3–9.
11. Цыцулин А.К., Бобровский А.И., Морозов А.В. Синтез космической системы наблюдения при широком диапазоне изменения дальности до объекта // Вопросы радиоэлектроники. Серия: Техника телевидения. 2020. № 2. С. 19–37.
12. Цыцулин А.К., Павлов В.А., Бобровский А.И., Морозов А.В., Рогачев В.А. Помехоустойчивость классификации объектов по признаку скоростного смаза // Вопросы радиоэлектроники. Серия: Техника телевидения. 2020. № 3. С. 53–67.
13. Цыцулин А.К., Павлов В.А., Бобровский А.И., Морозов А.В. Информационные оценки в задачах обнаружения–оценивания–передачи сигнала в космическом телевидении // Вопросы радиоэлектроники. Серия: Техника телевидения. 2019. № 3. С. 61–74.
14. Цыцулин А.К., Девяткин А.В., Бобровский А.И., Морозов А.В., Горшанов Д.Л., Павлов В.А. Адаптация кадровой частоты к этапам наблюдения в системе контроля сближения космических аппаратов // Вопросы радиоэлектроники. Серия: Техника телевидения. 2019. № 1. С. 31–38.
15. Цыцулин А.К., Павлов В.А., Бобровский А.И., Морозов А.В., Зубакин И.А. Адаптивное кодирование изображений, разделимых на доминантный объект и фон // Вопросы радиоэлектроники. Серия: Техника телевидения. 2019. № 3. С. 75–85.
16. Вальд А. Последовательный анализ. Пер. с англ. М.: Физматгиз, 1960. 328 с.
17. Ширяев А.Н. Статистический последовательный анализ. М.: Наука, 1976. 272 с.
18. Вудворд Ф.М. Теория вероятностей и теория информации с приложением к радиолокации. Пер. с англ. М.: Сов. радио, 1955. 128 с.
19. Фрэнкс Л. Теория сигналов. Пер. с англ. М.: Мир, 1974. 344 с.
20. Ван Трис Г. Теория обнаружения, оценивания и модуляции. В 4-х томах. Пер. с англ. М.: Радио и связь, 1972–1977.
21. Хелстром К. Статистическая теория обнаружения сигналов. Пер. с англ. М.: Изд-во иностр. лит., 1963. 432 с.
22. Шамис А.Л. Вектор эволюции. Жизнь, эволюция, мышление с точки зрения программиста. М.: Книжный дом «ЛИБРОКОМ», 2013. 200 с.

23. Беллман Р. Процессы регулирования с адаптацией. Пер. с англ. М.: Наука, 1964. 359 с.
24. Васильев К.К., Крашенинников В.Р. Статистический анализ последовательностей изображений. М.: Радиотехника, 2017. 248 с.
25. Умбиталиев А.А., Пятков В.В., Цыцулин А.К., Левко Г.В., Манцветов А.А., Рычажников А.Е. и др. Статистический анализ и синтез прикладных телевизионных систем // Вопросы радиоэлектроники. Серия: Техника телевидения. 2015. № 6. С. 4–14.
26. Красносельский М.А., Покровский А.В. Системы с гистерезисом. М.: Наука, 1983. 271 с.
27. Бобровский А.И., Морозов А.В., Чепелев А.Г., Павлов В.А., Галанов В.В. Выбор архитектуры нейронной сети для обнаружения и классификации изображений космических объектов // Вопросы радиоэлектроники. Серия: Техника телевидения. 2020. № 1. С. 64–73.
28. Redmon J., Farhadi A. YOLO9000: Better, Faster, Stronger // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR, Honolulu, USA, 21–26 July 2017). IEEE, 2017. PP. 7263–7271. DOI:10.1109/CVPR.2017.690
29. Цыцулин А.К., Бобровский А.И., Морозов А.В., Павлов В.А., Галеева М.А. Применение свёрточных нейронных сетей для автоматической селекции малоразмерных искусственных космических объектов на оптических изображениях звёздного неба // Оптический журнал. 2019. Т. 86. № 10. С. 30–38. DOI:10.17586/1023-5086-2019-86-10-30-38

* * *

Theoretical Foundations Development of Space Adaptive Video Information Systems Synthesis

A. Bobrovskiy¹ 

¹State Research Institute of Applied Problems
St. Petersburg, 191167, Russian Federation

Article info

DOI:10.31854/1813-324X-2021-7-2-69-78

Received 28th May 2021

Accepted 21st June 2021

For citation: Bobrovskiy A. Theoretical Foundations Development of Space Adaptive Video Information Systems Synthesis. *Proc. of Telecom. Universities*. 2021;7(2):69–78. (in Russ.) DOI:10.31854/1813-324X-2021-7-2-69-78

Abstract: *The main provisions and principles of the synthesis of adaptive video information systems designed for recording and analyzing images of dynamic scenes at the stages of long-range detection of artificial space objects, ballistic and controlled approach to them in the middle and near zone of cooperation are considered. Based on the principle of dominant information, it is shown that the basis for the development of methods and algorithms for processing video information in on-board surveillance systems is the joint adaptive processing of video information in photodetector matrices and in digital processors in order to maximize the quality of information issued to the consumer by suppressing background and noise information by adaptive hysteresis control of the frame frequency, frame resolution and signal accumulation time in the photodetector matrix, as well as the use of parametric neural network detectors-classifiers.*

Keywords: *synthesis, video information system, adaptation, dominant, background and noise information, frame rate, resolution, accumulation time, encoding, parametric optimization.*

References

1. Chromov L.I., Lebedev N.V., Tsytsulin A.K., Kulikov A.N. *Solid State Television*. Moscow: Radio i Svyaz Publ.; 1986. 184 p. (in Russ.)
2. Chromov L.I., Tsytsulin A.K., Kulikov A.N. *Video Informatics*. Moscow: Radio i Svyaz Publ.; 1991. 192 p. (in Russ.)
3. Berezin V.V., Umbitaliev A.A., Fakhmin Sh.S., Tsytsulin A.K., Shipilov N.N. *The Solid State Revolution in Television: CCD, System-on-Chip and Video-on-Chip Television Systems*. Moscow: Radio i Svyaz Publ.; 2006. 312 p. (in Russ.)
4. Tsytsulin A.K., Adamov D.Yu., Mantsvetov A.A., Zubakin I.A. *Solid State Cameras: Accumulating Information Quality*. St. Petersburg: Saint Petersburg Electrotechnical University "LETI" Publ.; 2014. 272 p. (in Russ.)
5. Umbitaliev A.A., Tsytsulin A.K., Levko G.V., Pyatkov V.V., Kuzichkin A.V., Dvornikov S.V., et al. *Theory and Practice of Space Television*. St. Petersburg: JSC Television Research Institute Publ.; 2017. 368 p. (in Russ.)

6. Hromov L.I., Tsytsulin A.K. Space Videoinformation Basis. *Voprosy radioelektroniki. Seriya: Tekhnika televideniya*. 2011;1:6–31. (in Russ.)
7. Bobrovskij A.I. Adaptation of the Applied Television System to the Dynamics of the Plot. *Digital Signal Processing*. 2018;3:3–9. (in Russ.)
8. Bobrovskij A.I. Processing of Video Information in the Adaptive Control System for the Approach of Spacecraft. *Information and Space*. 2018;4:149–153. (in Russ.)
9. Bobrovskij A.I. Heoretical Bases of Synthesis of Space Adaptive Video Information Systems. *Voprosy radioelektroniki. Seriya Tekhnika televideniya*. 2019;4:42–46. (in Russ.)
10. Tsytsulin A.K., Fahmi S.H.S., Adamov D.Yu., Bobrovskij A.I., Zubakin I.A., Chernogubov A.V. Principle of Dominant Information – a Conceptual Basis for Designing Videoinformation Systems. *Datchiki & Systemi (Sensors & Systems)*. 2018;11:3–9. (in Russ.)
11. Tsytsulin A.K., Bobrovskij A.I., Morozov A.V. Synthesis of a Space Videoinformation System with a Wide Range of Changes in the Distance to the Object. *Voprosy radioelektroniki. Seriya: Tekhnika televideniya*. 2020;2:19–37. (in Russ.)
12. Tsytsulin A.K., Pavlov V.A., Bobrovskij A.I., Morozov A.V., Rogachev V.A. Objects Classification by the Speed-Smearing Feature Noise Resistance. *Voprosy radioelektroniki. Seriya: Tekhnika televideniya*. 2020;3:53–67. (in Russ.)
13. Tsytsulin A.K., Pavlov V.A., Bobrovskij A.I., Morozov A.V. Information Evaluations of Signal Detection-Estimation-Transfer Task in Space Television Systems. *Voprosy radioelektroniki. Seriya: Tekhnika televideniya*. 2019;3:61–74. (in Russ.)
14. Tsytsulin A.K., Devyatkin A.V., Bobrovskij A.I., Morozov A.V., Gorshanov D.L., Pavlov V.A. Adaption of Frame Rates for Observation Phases in the Control System of a Spacecraft. *Voprosy radioelektroniki. Seriya: Tekhnika televideniya*. 2019;1:31–38. (in Russ.)
15. Tsytsulin A.K., Pavlov V.A., Bobrovskij A.I., Morozov A.V., Zubakin I.A. Adaptive Coding of Images Divided into Dominant Object and Background. *Voprosy radioelektroniki. Seriya: Tekhnika televideniya*. 2019;3:75–85. (in Russ.)
16. Wald A. *Sequential Analysis*. Translate from English. Moscow: Fizmatgiz Publ.; 1960. 328 p. (in Russ.)
17. Shiryaev A.N. *Statistical Sequential Analysis*. Moscow: Nauka Publ.; 1976. 272 p. (in Russ.)
18. Vudvord F.M. *Probability Theory and Information Theory with an Application to Radar*. Translate from English. Moscow: Sov. Radio Publ.; 1955. 128 p. (in Russ.)
19. Franks L.E. *Signal Theory*. Translate from English. Moscow: Mir Publ.; 1974. 344 p. (in Russ.)
20. Van Trees L. *Detection, Estimation and Modulation Theory*. In 4 vol. Translate from English. Moscow: Radio i Svyaz Publ.; 1972–1977. (in Russ.)
21. Helstrom C.W. *Statistical Theory of Signal Detection*. Oxford: Pergamon Press; 1960.
22. Shamis A.L. *Evolution Vector. Life, Evolution, Thinking from the Point of View of a Programmer*. Moscow: LIBROKOM Publ.; 2013. 200 p. (in Russ.)
23. Bellman R. *Adaptive Control Processes: a Guided Tour*. Moscow: Nauka Publ.; 1964. 359 p. (in Russ.)
24. Vasil'ev K.K., Krashenninnikov V.R. *Statistical Analysis of Image Sequences*. Moscow: Radiotekhnika Publ.; 2017. 248 p. (in Russ.)
25. Umbitaliev A.A., Pyatkov V.V., Tsytsulin A.K., Levko G.V., Mantsvetov A.A., Rychazhnikov A.E. Statistical Analysis and Synthesis of Applied Television Systems. *Voprosy radioelektroniki. Seriya: Tekhnika televideniya*. 2015;6:4–14. (in Russ.)
26. Krasnosel'skij M. A., Pokrovskij A.V. *Systems with Hysteresis*. Moscow: Nauka Publ.; 1983. 271 p. (in Russ.)
27. Bobrovskij A.I., Morozov A.V., Chepelev A.G., Pavlov V.A., Galanov V.V. Selection of Architecture of Neural Network for Space Objects Images' Detection and Classification. *Voprosy radioelektroniki. Seriya: Tekhnika televideniya*. 2020;1:3–13. (in Russ.)
28. Redmon J., Farhadi A. YOLO9000: Better, Faster, Stronger. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, CVPR, 21–26 July 2017, Honolulu, USA*. IEEE; 2017. p.7263–7271. DOI:10.1109/CVPR.2017.690
29. Tsytsulin A.K., Bobrovskij A.I., Morozov A.V., Pavlov V.A., Galeeva M.A. Using Convolutional Neural Networks to Automatically Select Small Artificial Space Objects on Optical Images of a Starry Sky. *Journal of Optical Technology*. 2019;86(10):627. DOI:10.17586/1023-5086-2019-86-10-30-38

Сведения об авторе:

БОБРОВСКИЙ
Алексей Иванович

кандидат технических наук, старший научный сотрудник, начальник научно-исследовательского центра Государственного научно-исследовательского института прикладных проблем, albob@mail.ru
 <https://orcid.org/0000-0002-4344-9305>

Разработка метода использования цифровых водяных знаков для защиты от атаки клонирования бумажных сертификатов

В.И. Коржик¹ , В.С. Старостин¹, Д.А. Флакسمан² 

¹Санкт-Петербургский государственный университет телекоммуникаций им. проф. М.А. Бонч-Бруевича,

Санкт-Петербург, 193232, Российская Федерация

²ООО «Научно-производственное предприятие Новые Технологии Телекоммуникаций»,

Санкт-Петербург, 195256, Российская Федерация

*Адрес для переписки: flxdima4951@gmail.com

Информация о статье

Поступила в редакцию 06.03.2021

Принята к публикации 29.04.2021

Ссылка для цитирования: Коржик В.И., Старостин В.С., Флакسمан Д.А. Разработка метода использования цифровых водяных знаков для защиты от атаки клонирования бумажных сертификатов // Труды учебных заведений связи. 2021. Т. 7. № 2. С. 79–84. DOI:10.31854/1813-324X-2021-7-2-79-84

Аннотация: Рассматриваются бумажные (или пластиковые) сертификаты различных изделий для защиты оригинальности продуктов и доказательства прав их собственников. Для этого обычно используются линейные или двумерные штрих-коды. Однако этой меры может оказаться недостаточно, особенно для защиты от таких злоумышленных действий, как клонирование и изготовление копий сертификатов и последующего их использования с контрафактной продукцией. В настоящей работе для защиты от такой атаки, которую мы называем «клонированием» сертификатов, предлагается использовать вложение в них цифровых водяных знаков. Приводится алгоритм обнаружения факта такого клонирования и теоретически рассчитываются вероятности ошибки обнаружения и ложной тревоги для данной атаки, в зависимости от параметров системы.

Ключевые слова: сертификаты продукции, клонирование, цифровые водяные знаки, вероятности пропуска и ложной тревоги.

Введение

Защита различных изделий от подделок или фальсификаций их характеристик и действительных производителей является важной задачей, понимаемой в широком смысле информационной безопасности. Для решения этой задачи значительное распространение получил метод бумажных (или пластиковых) сертификатов – попытка использовать штрих-код (например, Data Matrix). Однако не для всех атак злоумышленников защита может оказаться успешной. Так, при выполнении атаки «клонирования сертификата» злоумышленник считывает (фотографирует) сертификат, а затем производит новый поддельный (возможно и измененный) сертификат, который затем прикрепляется к изделию пониженного качества, при этом его содержание говорит о подлинности продукта и, следовательно, требует его повышенной цены. Известен метод повышения надежности сертификатов при помощи вложения в них цифровых водяных знаков (ЦВЗ), подлинность которых подтвер-

ждается использованием специального конфиденциального цифрового ключа, который имеется только у собственника этого продукта [1]. Однако такой подход не может защитить от атаки «клонированием сертификатов». В работе [2] предлагается система ЦВЗ для графических QR-кодов, основанная на искажении формы отдельных блоков, входящих в состав кода, однако данный метод требует изменения структуры самого кода.

В настоящей работе предлагается метод защиты сертификатов от подобной атаки, который использует факт увеличения мощности шума в цифровой копии сертификата, вследствие появления дополнительных операций считывания и печати, выполняемых злоумышленником.

Алгоритм обнаружения факта клонирования

Вложение ЦВЗ осуществляется в частотной области, т. е. после применения к покрываемому объекту дискретно-косинусного преобразования (ДКП).

Вложение одного бита ЦВЗ b_i в N_0 частотных (ДКП) коэффициентов матрицы производится по правилу:

$$C_w^{bi}(n) = C(n) + \delta(-1)^{b_i} \times \pi(n), \quad (1)$$

$$n = 1, 2, \dots, N_0,$$

где $C(n)$ – отсчет оригинального сообщения; δ – глубина погружения; $b_i \in (0, 1)$ – бит ЦВЗ с номером i ; $\pi(n)$ – псевдослучайная последовательность для одного бита широкополосного сигнала (ШПС); N_0 – длина ШПС для одного бита.

Приведенная выше операция применяется для всех бит ЦВЗ, при этом величины $C_w^{bi}(n)$, $n = 1, 2, \dots, N_0$, $i = 1, 2, \dots, S$, необходимо сохранять в памяти для последующего использования. Если всего вкладывается S бит, т. е. $i = 1, 2, \dots, S$, то общая «длина» ЦВЗ для покрывающего объекта будет $N = S \times N_0$.

Рассмотрим математическую модель клонирования сертификата. Обозначим отсчеты ЦВЗ в тесте, проводимом легитимным пользователем, через $C_t^i(n)$.

Тогда, если клонирования не было, то:

$$C_t^{i,i}(n) = C_w^{bi}(n) + N_{p1}(n) + N_{s1}(n), \quad (2)$$

$$n = 1, 2, \dots, N, i = 1, 2, \dots, S,$$

где $N_{p1}(n)$ – шумы при печати цифрового сертификата легитимным пользователем; $N_{s1}(n)$ – шумы, возникающие при сканировании сертификата легитимным пользователем в процессе тестирования.

Если же было выполнено клонирование, то тогда получаем:

$$C_t^{i,i}(n) = C_w^{bi}(n) + N_{p1}(n) + N_{s2}(n) + N_{p2}(n) + N_{s1}'(n), \quad (3)$$

$$n = 1, 2, \dots, N, i = 1, 2, \dots, S,$$

где $N_{s2}(n)$ – шумы при сканировании сертификата злоумышленником; $N_{p2}(n)$ – шумы при печати клона злоумышленником; $N_{s1}'(n)$ – шумы при сканировании сертификата легитимным пользователем при тестировании.

Сравнивая выражения (2) и (3), видим, что при клонировании добавляются две новые компоненты аддитивного шума N_{p2} , N_{s2} , которые возникают из-за того, что злоумышленник должен дополнительно считать и отпечатать сертификат. В нашей модели все шумы полагаются гауссовскими случайными величинами с нулевыми средними и дисперсиями (Var) одинаковыми для принтера и сканера, но разными для легитимного пользователя и злоумышленника:

$$Var(N_{p1}) = Var(N_{s1}) = Var(N_{s1}') = \sigma^2,$$

$$Var(N_{s2}) = Var(N_{p2}) = \frac{\sigma^2}{r}, r \geq 1,$$

где r – некоторый коэффициент, который показывает, что злоумышленник может иметь мощность шума в r раз меньше, чем легитимный пользователь (последнее условие дает некоторую фору злоумышленнику по сравнению с легитимным пользователем).

Ранее предполагалось, что легитимный пользователь сохраняет в своем ПК дата-код сертификата с вложением ЦВЗ, т. е. $C_w^{bi}(n)$, $n = 1, 2, \dots, N$. Затем он, получив от сканера результаты тестирования $C_t(n)$, $n = 1, 2, \dots, N$ находит величины $\lambda(n) = C_t(n) - C_w^{bi}(n)$, $n = 1, 2, \dots, N$.

Если клонирования не было, то проверяющий вычисляет:

$$\lambda'(n) = C_t'(n) - C_w^{bi}(n) = C_w^{bi}(n) + N_{p1}(n) + N_{s1}(n) - C_w^{bi}(n) = N_{p1}(n) + N_{s1}(n), \quad (4)$$

$$n = 1, 2, \dots, N.$$

Если же клонирование было, то проверяющий получит:

$$\lambda''(n) = C_t^{i,i}(n) - C_w^{bi}(n) = C_w^{bi}(n) + N_{p1}(n) + N_{s2}(n) + N_{p2}(n) + N_{s1}'(n) - C_w^{bi}(n) = N_{p1}(n) + N_{s2}(n) + N_{p2}(n) + N_{s1}'(n). \quad (5)$$

Тогда суммарная мощность шумов будет:

$$E\{\lambda_{nc}^2(n)\} = 2\sigma^2 \text{ – для случая без клонирования,}$$

$$E\{\lambda_c^2(n)\} = 2\sigma^2 r' \text{ – для случая с клонированием,}$$

где $r' = \frac{r+1}{r}$; r – это величина, которая показывает во сколько раз дисперсия шумов у атакующего меньше дисперсии шумов у легального пользователя.

Далее измеряется нормированная мощность шумов:

$$\Omega = \frac{1}{N} \sum_{n=1}^N \lambda^2(n).$$

Решение о наличии клонирования принимается по правилу:

$$\begin{cases} \Omega \geq \Omega_0 \Rightarrow \text{клонирование есть} \\ \Omega < \Omega_0 \Rightarrow \text{клонирования нет} \end{cases} \quad (6)$$

где Ω_0 – некоторый заранее заданный порог.

Расчет эффективности обнаружения атаки клонирования

При принятии решения могут появиться два вида ошибок: P_m – вероятность пропуска клонирования, когда оно в действительности было, но это не обнаружено; P_{fa} – вероятность ложной тревоги, когда клонирования не было, но по правилу (6) принимается решение о его наличии. Определим полную вероятность ошибки как $P_e = \frac{1}{2}(P_m + P_{fa})$ и

будем называть оптимальным порогом такую величину $\Omega = \Omega_0$, которая обеспечивает минимум P_e .

Примем гауссовскую аппроксимацию Ω , т. е. когда $\Omega \in N(E\{\Omega\}, Var\{\Omega\})$, что будет верно согласно центральной предельной теореме [3] при больших N . Рассчитаем величины $E\{\Omega\}, Var\{\Omega\}$ без клонирования и с ним.

Без клонирования:

$$\begin{aligned} E\{\Omega\} &= \frac{1}{N} \sum_{n=1}^N E\{\lambda_n^2\} = 2\sigma^2, \\ Var\{\Omega\} &= \frac{1}{N} \sum_{n=1}^N Var\{\lambda^2(n)\} = \frac{8}{N} \sigma^4, \end{aligned} \quad (7)$$

поскольку λ_n – это гауссовские случайные величины с нулевым средним и дисперсией $2\sigma^2$ и тогда без клонирования $Var\{\lambda^2(n)\} = 8\sigma^4$ [3].

С клонированием:

$$\begin{aligned} E\{\Omega\} &= 2\sigma^2 r', \\ Var\{\Omega\} &= \frac{8}{N} \sigma^4 (r')^2. \end{aligned} \quad (8)$$

Тогда вероятности ошибок P_m, P_{fa} и P_e будут иметь выражения:

$$\begin{aligned} P_m &= \int_{-\infty}^{\Omega_0} w_{\Omega}^c(x) dx; \\ P_{fa} &= \int_{\Omega_0}^{\infty} w_{\Omega}^{nc}(x) dx; P_e = \frac{1}{2} (P_m + P_{fa}). \end{aligned} \quad (9)$$

Эти вероятности можно выразить (10) при помощи функции Лапласа [4]:

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x \exp\left(-\frac{t^2}{2}\right) dt.$$

Для нахождения оптимального порога Ω_0 необходимо вычислить производную $P_e'(N)$ в (10), а затем приравнять ее к нулю:

$$\begin{aligned} \frac{r\sqrt{N}}{2\sqrt{2}\sigma^2(r+1)} \varphi\left(\sqrt{N} \frac{\Omega_0 - 2\sigma^2 \frac{r+1}{r}}{\sigma^2 2\sqrt{2}} \frac{r}{r+1}\right) - \\ - \frac{\sqrt{N}}{2\sqrt{2}\sigma^2} \varphi\left(\sqrt{N} \frac{\Omega_0 - 2\sigma^2}{\sigma^2 2\sqrt{2}}\right) = 0, \end{aligned} \quad (11)$$

где $\varphi(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}$.

Уравнение (11) эквивалентно следующему уравнению:

$$\begin{aligned} \frac{r}{r+1} \exp\left(-\frac{N}{16\sigma^4} \left(\frac{r}{r+1}\right)^2 \left(\Omega_0 - 2\sigma^2 \frac{r+1}{r}\right)^2\right) = \\ = \exp\left(-\frac{N}{16\sigma^4} (\Omega_0 - 2\sigma^2)^2\right). \end{aligned} \quad (12)$$

Дальнейшее упрощение (12) приводит к следующему уравнению:

$$\begin{aligned} \Omega_0 \left(1 - \frac{r}{r+1}\right) \left(\Omega_0 \left(1 + \frac{r}{r+1}\right) - 4\sigma^2\right) = \\ = \frac{16\sigma^4}{N} \ln\left(1 + \frac{1}{r}\right). \end{aligned} \quad (13)$$

Решая уравнение (13) относительно Ω_0 , получаем следующее квадратное уравнение:

$$\omega^2 \frac{2r+1}{(r+1)^2} - \omega \frac{4}{r+1} - \frac{16}{N} \ln\left(1 + \frac{1}{r}\right) = 0, \quad (14)$$

где введено обозначение $\omega = \frac{\Omega_0}{\sigma^2}$.

Вещественный положительный корень этого уравнения может быть найден по известной формуле:

$$\omega = \frac{r+1}{2r+1} \left(2 + \sqrt{4 + (2r+1) \frac{16}{N} \ln\left(1 + \frac{1}{r}\right)} \right).$$

Тогда окончательно находим величину оптимального порога как:

$$\Omega_0 = \sigma^2 \frac{r+1}{2r+1} \left(2 + \sqrt{4 + \frac{16}{N} (2r+1) \ln\left(1 + \frac{1}{r}\right)} \right) \quad (15)$$

При $N \gg 1$ (что, как будет показано далее, типично для практики), получим для Ω_0 по (15) следующее приближение:

$$\Omega_0 \approx 4\sigma^2 \frac{r+1}{2r+1}. \quad (16)$$

В таблице 1 показаны результаты расчета по (16) для различных r .

$$\begin{aligned} P_m &= \Pr(\Omega < \Omega_0 | H_1) = 0,5 + \Phi\left(\sqrt{N} \frac{\Omega_0 - 2\sigma^2 \frac{r+1}{r}}{\sigma^2 2\sqrt{2}} \frac{r}{r+1}\right), \\ P_{fa} &= \Pr(\Omega \geq \Omega_0 | H_0) = 0,5 - \Phi\left(\sqrt{N} \frac{\Omega_0 - 2\sigma^2}{\sigma^2 2\sqrt{2}}\right), \\ P_e &= \frac{P_m + P_{fa}}{2} = \frac{1}{2} \left(1 + \Phi\left(\sqrt{N} \frac{\Omega_0 - 2\sigma^2 \frac{r+1}{r}}{\sigma^2 2\sqrt{2}} \frac{r}{r+1}\right) - \Phi\left(\sqrt{N} \frac{\Omega_0 - 2\sigma^2}{\sigma^2 2\sqrt{2}}\right) \right). \end{aligned} \quad (10)$$

ТАБЛИЦА 1. Расчет оптимального порога Ω_0 по (16)TABLE 1. Calculation of the Optimal Threshold Ω_0 by (16)

r	1	2	4	8	16
Ω_0	$2,66\sigma^2$	$2,4\sigma^2$	$2,22\sigma^2$	$2,06\sigma^2$	$1,94\sigma^2$

Для расчета величины P_e при оптимальном пороге Ω_0 , найденном по (16), предварительно рассчитаем аргументы функции $\Phi(\cdot)$ в (11):

$$\sqrt{N} \frac{\Omega_0 - 2\sigma^2 \frac{r+1}{r}}{\sigma^2 2\sqrt{2}} \frac{r}{r+1} =$$

$$= r\sqrt{N} \frac{\frac{r+1}{2r+1} \left(2 + \sqrt{4 + \frac{16}{N} (2r+1) \ln \left(1 + \frac{1}{r} \right)} \right) - 2 \frac{r+1}{r}}{2\sqrt{2}(r+1)}, \quad (17)$$

$$\sqrt{N} \frac{\Omega_0 - 2\sigma^2}{\sigma^2 2\sqrt{2}} =$$

$$= \sqrt{N} \frac{\frac{r+1}{2r+1} \left(2 + \sqrt{4 + \frac{16}{N} (2r+1) \ln \left(1 + \frac{1}{r} \right)} \right) - 2}{2\sqrt{2}}. \quad (18)$$

При $N \gg 1$ выражения (17), (18) приводятся, соответственно к следующему виду:

$$-\frac{\sqrt{N}}{\sqrt{2}(2r+1)}, \quad \frac{\sqrt{N}}{\sqrt{2}(2r+1)} \quad (19)$$

Подставляя (19) в (10), окончательно находим при $N \gg 1$:

$$P_e = \frac{P_m + P_{fa}}{2} \approx$$

$$\approx \frac{1}{2} \left(1 + \Phi \left(-\frac{\sqrt{N}}{\sqrt{2}(2r+1)} \right) - \Phi \left(\frac{\sqrt{N}}{\sqrt{2}(2r+1)} \right) \right) = \quad (20)$$

$$= \frac{1}{2} - \Phi \left(\frac{\sqrt{N}}{\sqrt{2}(2r+1)} \right).$$

Из выражения (20) видно, что при $r \rightarrow \infty$, $P_e \rightarrow 0,5$, т. е. алгоритм обнаружения факта клонирования эквивалентен «случайному угадыванию», что соответствует «физическому смыслу», поскольку случай $r \rightarrow \infty$ соответствует пренебрежимо малому шуму злоумышленника, добавление которого, конечно, оказывается невозможно обнаружить.

С другой стороны, если $N \rightarrow \infty$, то $P_e \rightarrow 0$ при любых конечных значениях r , что также совпадает с нашей интуицией, обеспечивая надежное обнаружение факта клонирования.

Для того, чтобы рассчитать вероятность ошибки P_e при конечных величинах N и r , можно воспользоваться следующим приближением функции $\Phi(\cdot)$ [4]:

$$\Phi(x) \approx \frac{1}{2} - \frac{1}{\sqrt{2\pi}} \frac{e^{-\frac{x^2}{2}}}{x}. \quad (21)$$

Подставляя (21) в (20), получаем:

$$P_e \approx \Phi \left(\frac{1}{2\pi} \frac{e^{-\frac{x^2}{2}}}{x} \right), \quad (22)$$

где $x = \frac{\sqrt{N}}{\sqrt{2}(2r+1)}$.

На рисунке 1 показана зависимость P_e от N для различных величин $r = 1, 2, 4, 8, 16$.

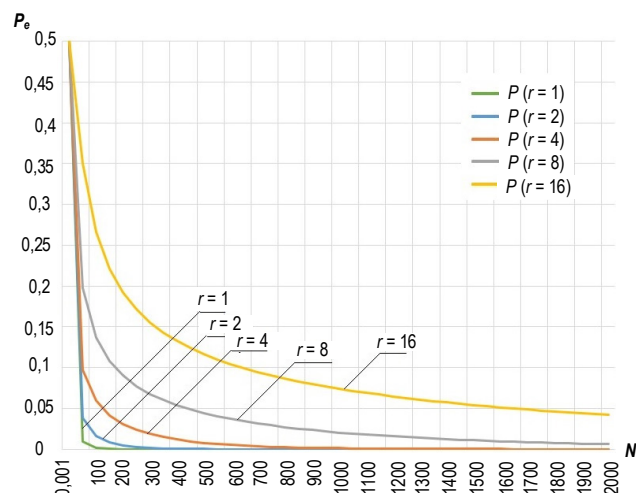
Рис. 1. Зависимости полной вероятности ошибки обнаружения атаки клонирования от длины ЦВЗ N

Fig. 1. Dependences of the Total Error Probability of Detecting a Cloning Attack on the Length N of the Digital Watermark

В таблице 2 показаны результаты расчета P_e для различных значений r и N .

ТАБЛИЦА 2. Расчет P_e при $N = 600, 800, 1000, 1200, 1500$ и $r = 1, 2, 4, 8, 16$ TABLE 2. Calculation of P_e when $N = 600, 800, 1000, 1200, 1500$ and $r = 1, 2, 4, 8, 16$

$r \backslash N$	600	800	1000	1200	1500
1	0	0	0	0	0
2	0,000046	0,0000053	0,00000065	0,000000079	0,0000000036
4	0,0052	0,0024	0,0012	0,00058	0,00021
8	0,037	0,027	0,021	0,016	0,011
16	0,11	0,087	0,074	0,065	0,054

Заключение

В работе был предложен и теоретически исследован метод обнаружения атаки клонирования. Видно, что он позволяет решить эту задачу при некоторых длинах вкладываемых ЦВЗ и ограничениях на шумы сканера и принтера атакующего. Однако окончательное подтверждение возможности такого вывода требует дополнительных экспериментальных исследований, которые, правда, уже частично проводились – для анализа возможности считывания ЦВЗ с бумажных и пластиковых носителей в работе [5], а также частично рассматривались в работе [6].

Дальнейшие исследования, особенно относящиеся к обоснованию выбора параметра r , авторы предполагают выполнить в ближайшем будущем. Наилучшим вариантом стало бы проведение множества экспериментов, включающих в себя печать и сканирование сертификата, подделку сертификата, печать и сканирование поддельного сертификата. Однако при проведении таких экспериментов, кроме основных параметров (размер сертификата, длина ШПС и глубина вложения) стоит учесть также множество второстепенных параметров и настроек алгоритмов вложения/извлечения ЦВЗ, таких как расположение оптимальной области вложения в отчеты ДКП, размеры DataMatrix и др. Это в свою очередь, потребует проведения чрезмерно большого количества операций, проводимых вручную (печать, сканирование), а также серьезные материальные вложения.

В связи с этим предполагается разбить проводимые эксперименты на два этапа: предварительный и основной.

Предварительный этап будет включать в себя выявление образцов шума присущих используемому оборудованию (принтеру и сканеру) и проведение программных (цифровых) экспериментов, в которых реальная печать заменена на виртуальную. Проведение экспериментов виртуально позволит легко автоматизировать процесс, произвести множество экспериментов и подобрать оптимальные настройки используемых алгоритмов.

На основном этапе будут проводиться реальные эксперименты. Вначале необходимо будет подтвердить предварительные результаты путем проведения выборочной печати сертификатов из множества предварительных экспериментов. После чего будет проводиться накопление статистических данных о работе предложенного метода в реальных условиях.

БЛАГОДАРНОСТИ

Авторы благодарят Букушина Ивана Дмитриевича за помощь при проведении расчетов и построении графиков.

Список используемых источников

1. Cox I.J., Miller M.L., Bloom J.A. Digital Watermarking. Elsevier, 2002. 542 p. DOI:10.1016/B978-1-55860-714-9.X5000-7
2. Tkachenko I. Generation and analysis of graphical codes using textured patterns for printed document authentication. D.Sc Thesis. Montpellier: Université de Montpellier, 2015.
3. Ван дер Варден Б.Л. Математическая статистика. Пер. с нем. М.: Издательство иностранной литературы, 1960. 435 с.
4. Финк Л.М. Теория передачи дискретных сообщений. М.: Советское радио, 1970.
5. Ho A.T.S., Shu F. A print-and-scan resilient digital watermark for card authentication // Proceedings of the Fourth International Conference on Information, Communications and Signal Processing and The Fourth Pacific Rim Conference on Multimedia (Singapore, 15–18 December 2003). Vol. 2. IEEE, 2003. PP. 1149–1152. DOI:10.1109/ICICIS.2003.1292640
6. Коржик В.И., Флакман Д.А. Система цифровых водяных знаков с возможностью их извлечения из бумажных копий цифровых документов // Труды учебных заведений связи. 2019. Т. 5. № 3. С. 75–85. DOI:10.31854/1813-324X-2019-5-3-75

* * *

Elaboration of Digital Watermarking Method for a Protection of Cloning Attack on Paper Certificates

V. Korzhik¹ , V. Starostin¹, D. Flaksman² 

¹The Bonch-Bruевич Saint-Petersburg State University of Telecommunications,
St. Petersburg, 193232, Russian Federation

²Research and Production Enterprise "Novye Tekhnologii Telekommunikatsii", Ltd,
St. Petersburg, 195256, Russian Federation

Article info

DOI:10.31854/1813-324X-2021-7-2-79-84

Received 6th March 2021

Accepted 29th April 2021

For citation: Korzhik V., Starostin V., Flaksman D. Elaboration of Digital Watermarking Method for a Protection of Cloning Attack on Paper Certificates. *Proc. of Telecom. Universities*. 2021;7(2):79–84. (in Russ.) DOI:10.31854/1813-324X-2021-7-2-79-84

Abstract: Paper (or plastic) certificates are considered as a means against different forgery of product quality and brand falsification. It is commonly to use barcodes or data matrices to solve this problem. However such approach does not work usually against such sophisticated attacks as cloning of certificate copies. In the current paper we propose to use a digital watermarking and estimation of inner noises of scanners and printers in order to detect cloning attack effectively. Algorithm of cloning attack detecting is presented. The probabilities of attack missing and false alarm are proved.

Keywords: certificate of product, cloning attack, digital watermarking, the probabilities of attack missing and false alarm.

ACKNOWLEDGMENT

The authors are grateful to Ivan Dmitrievich Bukshin for help with calculations and graphing.

References

1. Cox I.J., Miller M.L., Bloom J.A. *Digital Watermarking*. Elsevier; 2002. 542 p. DOI:10.1016/B978-1-55860-714-9.X5000-7
2. Tkachenko I. *Generation and analysis of graphical codes using textured patterns for printed document authentication*. D.Sc Thesis. Montpellier: Université de Montpellier; 2015.
3. Van der Waerden B.L. *Math Statistics*. Translated from German. Moscow: Inostrannaia literatura Publ.; 1960. 435 p. (in Russ.)
4. Fink L.M. *Theory of Discrete Signal Transmission*. Moscow: Sovetskoe radio Publ.; 1970. (in Russ.)
5. Ho A.T.S., Shu F. A print-and-scan resilient digital watermark for card authentication. *Proceedings of the Fourth International Conference on Information, Communications and Signal Processing and The Fourth Pacific Rim Conference on Multimedia, 15–18 December 2003, Singapore*. IEEE; 2003. vol.2. p.1149–1152. DOI:10.1109/ICICS.2003.1292640
6. Korzhik V., Flaksman D. Digital Watermark System with an Ability of its Extraction from Hard Copies of Data. *Proc. of Telecom. Universities*. 2019;5(3):75–85. (in Russ.) DOI:10.31854/1813-324X-2019-5-3-75-85

Сведения об авторах:

КОРЖИК
Валерий Иванович

доктор технических наук, профессор, профессор кафедры защищенных систем связи Санкт-Петербургского государственного университета телекоммуникаций им. проф. М. А. Бонч-Бруевича, val-korzhik@yandex.ru
 <https://orcid.org/0000-0002-8347-6527>

СТАРОСТИН
Владимир Сергеевич

кандидат физико-математических наук, доцент кафедры высшей математики Санкт-Петербургского государственного университета телекоммуникаций им. проф. М. А. Бонч-Бруевича, star_vs_47@mail.ru

ФЛАКСМАН
Дмитрий Алексеевич

программист ООО «Научно-производственное предприятие Новые Технологии Телекоммуникаций», flxdima4951@gmail.com
 <https://orcid.org/0000-0002-0326-4592>

Схема цифровой подписи со скрытой группой, обладающей двухмерной цикличностью

Д.Н. Молдовян¹, Р.Ш. Фахрутдинов¹, А.Ю. Мирин¹, А.А. Костина¹

¹Санкт-Петербургский Федеральный исследовательский центр Российской академии наук,
Санкт-Петербург, 199178, Российская Федерация

*Адрес для переписки: fahr@cobra.ru

Информация о статье

Поступила в редакцию 29.04.2021

Принята к публикации 23.05.2021

Ссылка для цитирования: Молдовян Д.Н., Фахрутдинов Р.Ш., Мирин А.Ю., Костина А.А. Схема цифровой подписи со скрытой группой, обладающей двухмерной цикличностью // Труды учебных заведений связи. 2021. Т. 7. № 2. С. 85–93. DOI:10.31854/1813-324X-2021-7-2-85-93

Аннотация: Предложен способ построения схем цифровой подписи, основанных на скрытой задаче дискретного логарифмирования, удовлетворяющих общему критерию постквантовой стойкости, который обеспечивает сравнительно малые размеры открытого ключа и подписи. На основе способа разработана практическая схема цифровой подписи, в которой операция экспоненцирования в скрытой группе с двухмерной цикличностью является базовым криптографическим примитивом. Алгебраический носитель крипто-схемы представляет собой четырехмерную конечную некоммутативную алгебру с ассоциативной операцией умножения. Благодаря заданию алгебры по прореженной таблице умножения базисных векторов, обеспечивается повышение производительности процедур генерации и проверки подлинности подписи. Открытым ключом является тройка четырехмерных векторов, вычисляемых как образы элементов скрытой группы, отображаемые с помощью маскирующих операций двух типов: 1) взаимно коммутативных с операцией экспоненцирования и 2) не обладающих этим свойством.

Ключевые слова: защита информации, компьютерная безопасность, криптография, цифровая подпись, постквантовая криптосхема, задача дискретного логарифмирования, конечная ассоциативная алгебра, некоммутативная алгебра, скрытая группа.

Введение

Одной из актуальных текущих проблем теоретической и практической криптографии является разработка двухключевых алгоритмов и протоколов [1, 2], стойких к атакам с использованием квантовых компьютеров. Криптосхемы такого типа называются постквантовыми. Всемирный конкурс на разработку постквантовых схем цифровой подписи и открытого согласования ключа, объявленный Национальным институтом стандартов и технологий США (НИСТ), ориентирован на получение к 2024 г. черновых версий постквантовых криптографических стандартов. Однако общим минусом предложенных кандидатов на постквантовые схемы электронной подписи является их недостаточная практичность, в частности большие размеры открытого ключа и подписи.

Сравнительно новым и многообещающим подходом к разработке практичных постквантовых схем цифровой подписи является использование скрытой задачи дискретного логарифмирования (СЗДЛ), задаваемой в конечных некоммутативных

ассоциативных алгебрах (КНАА), в качестве базового примитива [3–6]. Обоснование использования СЗДЛ, как постквантового криптографического примитива, объясняется тем, что задание по открытым параметрам криптосхемы периодической функции с периодом, зависящим от значения дискретного логарифма, приводит к тому, что эта функция принимает значения из многочисленных различных циклических групп, содержащихся в КНАА. Благодаря этому обстоятельству обеспечивается стойкость к квантовым атакам (атаки с использованием квантового компьютера), использующим известные полиномиальные по времени квантовые алгоритмы нахождения длины периодов периодических функций, значения которых лежат в одной явно заданной конечной циклической группе [7–9].

В первых постквантовых схемах цифровой подписи [4, 5], базирующихся на СЗДЛ и допускающих задание периодических функций, включающих период, зависящий от значения дискретного логарифма в скрытой группе, использован следующий

критерий построения: *вычислительно невыполнимо построение периодической функции на основе открытых параметров схемы цифровой подписи, содержащих период с длиной, зависящей от значения дискретного логарифма, и с достаточно высокой вероятностью, принимающих значения в некоторой фиксированной циклической группе.*

В расширенной постановке вопроса обеспечения достаточно высокой стойкости к квантовым атакам (постквантовой стойкости) [10] следует учесть потенциальную возможность разработки новых квантовых алгоритмов нахождения длины периода периодических функций, принимающих значения в различных коммутативных группах, содержащихся в фиксированной КНАА, используемой в качестве алгебраического носителя схемы цифровой подписи. В работе [10] предложен общий критерий постквантовой стойкости схем цифровой подписи, базирующихся на СЗДЛ: *по открытым параметрам схемы цифровой подписи вычислительно невыполнима задача построения периодической функции, содержащей период, зависящий от значения логарифма в скрытой циклической группе.*

При выполнении общего критерия постквантовой стойкости, разрабатываемая схема цифровой подписи будет защищенной от атак, использующих как известные, так и новые квантовые алгоритмы нахождения длины периода периодических функций. В работе [10] предложен механизм реализации общего критерия, включающий использование коммутативной группы, обладающей двухмерной циклическостью (частный тип групп с двумя порождающими элементами, характеризующийся тем, что порождающие элементы имеют одинаковый порядок, см., например, [11]) в качестве скрытой группы.

Постановка задачи

Известные схемы цифровой подписи [10, 12], удовлетворяющие общему критерию постквантовой стойкости, построены с удвоением открытого ключа и удвоением проверочного уравнения, а также использованием вектора, в качестве дополнительного элемента подписи. Это приводит к удвоению размера открытого ключа и увеличению длины подписи в 3 раза при одновременном снижении производительности криптосхемы по сравнению со схемами цифровой подписи, основанными на первом критерии.

В настоящей статье предлагается способ реализации схем цифровой подписи, удовлетворяющих общему критерию постквантовой стойкости и свободных от недостатков их аналогов, предложенных ранее в работах [10, 12].

Используемый алгебраический носитель

В качестве алгебраического носителя алгоритмов цифровой подписи, основанных на скрытой задаче дискретного логарифмирования, обычно применяются КНАА четной размерности $m \geq 4$, в которых содержится достаточно большое число различных подмножеств элементов, представляющих собой конечные коммутативные группы. Конечные алгебры задаются путем определения в конечном векторном пространстве с операциями сложения векторов и скалярного умножения дополнительной операции векторного умножения, обладающей свойствами дистрибутивности слева и справа относительно операции сложения. Пусть некоторое m -мерное векторное пространство задано над простым конечным полем $GF(p)$, и имеется некоторый его базис в виде набора векторов $e_1, \dots, e_i, \dots, e_m$. Некоторый вектор A можно записывать в виде:

1) упорядоченного множества его координат $a_i \in GF(p)$: $A = (a_0, a_1, \dots, a_{m-1})$;

2) суммы $A = \sum_{i=0}^{m-1} a_i e_i$, где $a_i e_i$ – компоненты вектора A .

Операция умножения вектора $A = \sum_{i=0}^{m-1} a_i e_i$ на вектор $B = \sum_{j=0}^{m-1} b_j e_j$ определяется по следующей формуле, в которой всевозможные произведения вида $e_i \cdot e_j$ заменяются по некоторому правилу на однокомпонентные векторы:

$$A \cdot B = \sum_{i=0}^{m-1} \sum_{j=0}^{m-1} a_i b_j (e_i \cdot e_j), \quad (1)$$

Указанное правило удобно представить в виде таблицы умножения базисных векторов (ТУБВ), предполагая, что вместо каждого из произведений $e_i \cdot e_j$ подставляется на некоторый вектор λe_k , указанный в ячейке на пересечении i -й строки и j -го столбца. Если $\lambda = 1$, то в ячейках ТУБВ указывается базисный вектор e_k . Если $\lambda \neq 1$, то значение λ называется структурным коэффициентом.

Для задания КНАА разрабатывается ТУБВ, которая определяет операцию векторного умножения, обладающую свойствами некоммутативности и ассоциативности. При реализации последнего свойства для произвольной тройки векторов A, B и C выполняется следующее равенство:

$$(A \cdot B) \cdot C = A \cdot (B \cdot C). \quad (2)$$

С учетом формул (1) и (2) легко показать, что ТУБВ, которая задает выполнимость равенства для всевозможных троек базисных векторов, определяет ассоциативное векторное умножение:

$$(e_i \cdot e_j) \cdot e_k = e_i \cdot (e_j \cdot e_k).$$

С целью повышения производительности алгоритмов генерации и проверки подлинности цифровой подписи представляет интерес использование четырехмерных КНАА, в которых операция умножения задана по прореженным ТУБВ, т. е. по ТУБВ,

в ячейках которых присутствует структурный коэффициент $\lambda = 0$.

Таблица 1 дает пример прореженной ТУБВ и используется для задания алгебраического носителя в разрабатываемой схеме цифровой подписи, основанной на скрытой задаче дискретного логарифмирования и удовлетворяющей общему критерию постквантовой стойкости.

В случае, если равенство $A \cdot B = B \cdot A$ выполняется для произвольной пары векторов A и B , операция умножения и задаваемая алгебра называются коммутативными, в противном случае – некоммутативными. В последнем случае имеем КНАА.

ТАБЛИЦА 1. Прореженная ТУБВ, задающая четырехмерную КНАА с глобальной единицей $(\mu^{-1}, \lambda^{-1}, 0, 0)$, где $\mu \neq 0; \lambda \neq 0$ [12]

TABLE 1. A Sparse Basis Vector Multiplication Table that Defines a Four-Dimensional Non-Commutative Associative Algebra with Global Unit $(\mu^{-1}, \lambda^{-1}, 0, 0)$, where $\mu \neq 0; \lambda \neq 0$ [12]

$\cdot$	e_0	e_1	e_2	e_3
e_0	μe_0	0	0	μe_3
e_1	0	λe_1	λe_2	0
e_2	μe_2	0	0	μe_1
e_3	0	λe_3	λe_0	0

Конечные группы с многомерной циклическостью

Конечной группой с μ -мерной циклическостью называется коммутативная группа с μ порождающими элементами, имеющими одинаковый порядок [11, 13]. В работах [11, 13] приведены примеры групп с μ -мерной циклическостью ($\mu \geq 2$) и обсуждается их применение для построения схем цифровой подписи. Рассмотрим, например, двухмерную конечную ассоциативную алгебру с операцией умножения, заданной по таблице 2. Ее мультипликативная группа имеет строение, зависящее от структурного коэффициента λ . Если λ является квадратичным невычетом в $GF(p)$, то указанная группа является циклической и имеет порядок $\Omega = p^2 - 1$ [13]. Если λ является квадратичным вычетом в $GF(p)$, то она является группой с двухмерной циклическостью и имеет порядок $\Omega = (p - 1)^2$ [13].

ТАБЛИЦА 2. Задание двухмерной ассоциативной алгебры над полем $GF(p)$ ($\lambda \neq 0$)

TABLE 2. Setting Two-Dimensional Associative Algebra Over Field $GF(p)$ ($\lambda \neq 0$)

$\cdot$	e_0	e_1
e_0	e_0	e_1
e_1	e_1	λe_0

В настоящей работе разрабатывается схема цифровой подписи, основанная на СЗДЛ и использующая коммутативную группу с двухмерной циклическостью в качестве скрытой группы. Важным мо-

ментом является задание скрытой группы, содержащей элементы достаточно большого простого порядка. Это может быть реализовано при выборе характеристики p поля $GF(p)$, равной простому числу вида $p = 2q + 1$, где q – 256-битное простое число. При таком выборе и значении λ , являющимся квадратичным вычетом, мультипликативная группа двухмерной алгебры, заданной по таблице 2, имеет двухмерное циклическое строение и содержит подгруппу $\Gamma_{\langle G, Q \rangle}$ порядка q^2 , генерируемую некоторым базисом $\langle G, Q \rangle$, где каждый из двухмерных векторов G и Q имеет порядок, равный простому числу q . Легко установить, что группа $\Gamma_{\langle G, Q \rangle}$ включает единичный вектор $(1, 0)$ и $q^2 - 1$ различных векторов, имеющих одно и тоже значение порядка, равное простому числу q . Группа $\Gamma_{\langle G, Q \rangle}$ включает $q + 1$ циклических подгрупп порядка q . Любой фиксированный вектор, отличный от $(1, 0)$, содержится в единственной циклической подгруппе.

Свойства алгебраического носителя

В работе [14] показано, что двухсторонней глобальной единицей четырехмерной КНАА, заданной по таблице 1, является вектор $E = (\mu^{-1}, \lambda^{-1}, 0, 0)$, а порядок мультипликативной группы этой алгебры равен значению:

$$\Omega = p(p - 1)(p^2 - 1). \quad (3)$$

При этом для произвольного заданного вектора $Q = (q_0, q_1, q_2, q_3)$ доказаны следующие положения:

- 1) условие обратимости вектора Q выражается формулой $q_0 q_1 \neq q_2 q_3$ [14];
- 2) условие необратимости вектора Q выражается формулой $q_0 q_1 = q_2 q_3$ [14];
- 3) множество всех векторов, перестановочных с Q , описывается следующей формулой [14]:

$$X = (x_0, x_1, x_2, x_3) = \left(d, \frac{\mu q_2 d + (\lambda q_1 - \mu q_0) h}{\lambda q_2}, h, \frac{q_3}{q_2} h \right), \quad (4)$$

где $d, h = 0, 1, \dots, p - 1$.

Множество (4) включает p^2 различных векторов, в том числе вектор Q , нулевой элемент $(0, 0, 0, 0)$ и единицу E .

Утверждение 1 [14]. Любые два элемента V и W , входящие в множество (4), являются взаимно перестановочными, т.е. имеет место равенство $W \cdot V = V \cdot W$.

Утверждение 2 [15]. Подмножество обратимых векторов, входящих в множество (4) образуют мультипликативную коммутативную группу с единицей E .

Утверждение 3 [15]. Координаты ненулевых необратимых векторов, содержащихся в множестве (4) удовлетворяют следующему равенству:

$$h = \left(\frac{\lambda q_1 - \mu q_0}{2\lambda q_3} \pm \sqrt{\frac{(\lambda q_1 - \mu q_0)^2}{4\lambda^2 q_3^2} + \frac{\mu q_2}{\lambda q_3}} \right) d. \quad (5)$$

Используя равенство (5), можно подсчитать число необратимых и обратимых векторов, содержащихся в множестве (4).

Рассмотрим следующие два случая.

Случай 1. Вектор Q является обратимым элементом. Если координаты Q таковы, что подкоренное выражение в (5) является квадратичным невычетом в поле $GF(p)$, то все ненулевые векторы, входящие в (4), являются обратимыми. Легко заметить, что в этом случае операции умножения и сложения векторов являются замкнутыми в (4) и удовлетворяют свойствам операций конечного поля характеристики p , т. е. множество векторов (4) образует поле $GF(p^2)$, мультипликативная группа которого является циклической и имеет порядок $p^2 - 1$. Для построения схемы цифровой подписи важным является существование обратимых векторов, порядок которых равен $p^2 - 1$ и произвольному заданному делителю $p^2 - 1$.

Если подкоренное выражение в (5) является квадратичным вычетом в поле $GF(p)$, то формула (4) для каждого из значений $d = 0, 1, \dots, p - 1$ при двух различных значениях h , определяемых формулой (5), задает пару ненулевых необратимых векторов, т. е. с учетом нулевого вектора в множестве (4) имеем $2p - 1$ необратимых векторов и $p^2 - (2p - 1) = (p - 1)^2$ обратимых. Значение порядка группы обратимых векторов $\Omega = (p - 1)^2$ подсказывает, что они образуют конечную группу с двухмерной циклическостью. Пусть α — примитивный элемент поля $GF(p)$, тогда вектор-скаляр αE входит в указанную группу и имеет порядок $p - 1$. Всевозможные степени вектора αE генерируют $p - 1$ различных векторов. Пусть обратимый вектор V (например, $V = Q$), входящий в множество (4) и непредставимый в виде степени вектора αE имеет порядок $p - 1$. Тогда легко показать, что базис $\langle \alpha E, V \rangle$ генерирует коммутативную группу $\Gamma_{\langle \alpha E, V \rangle}$ порядка $(p - 1)^2$, имеющую двухмерное циклическое строение. Единицей группы $\Gamma_{\langle \alpha E, V \rangle}$ является вектор E . Так как $\Gamma_{\langle \alpha E, V \rangle}$ включает только обратимые векторы, перестановочные с Q , то $\Gamma_{\langle \alpha E, V \rangle}$ является группой обратимых векторов множества (4). Для построения предлагаемой схемы цифровой подписи важным является наличие в используемом алгебраическом носителе достаточно большого числа коммутативных групп с двухмерной циклическостью как подмножеств алгебраических элементов. (В общем случае различные векторы Q задают различные группы с двухмерной циклическостью.)

Случай 2. Вектор Q является необратимым элементом. Допустим, что координаты вектора Q

удовлетворяют условию $q_0 q_1 = q_2 q_3$, т. е. он является необратимым. Из условия необратимости получаем $q_2 = q_3^{-1} q_0 q_1$ и представление формулы (5) в следующем виде:

$$h = \begin{cases} \frac{q_1}{q_3} d \\ -\frac{\mu q_0}{\lambda q_3} d' \end{cases}, \quad (6)$$

где $d = 1, 2, \dots, p - 1$. Из (6) следует, что множество (4) включает $2(p - 1)$ необратимых векторов, отличных от нулевого вектора, и $p^2 - 2(p - 1) - 1 = (p - 1)^2$ обратимых векторов. Множество последних включает все вектор-скаляры iE , где $i = 1, 2, \dots, p - 1$, и образует конечную коммутативную группу порядка $(p - 1)^2$, обладающую двухмерной циклическостью.

Таким образом, группу с двухмерной циклическостью (ее базис) можно задать, генерируя необратимый вектор Q и вычисляя по формулам (4) и (6) некоторый вектор X порядка $p - 1$ (для различных значений d проверяем значение порядка вектора X), т. е. имеем еще один способ задания группы с двухмерной циклическостью. В работе [15] доказано, что в случае необратимого вектора Q множество (4) действительно содержит обратимый вектор, отличный от вектор-скаляра и имеющий значение порядка $p - 1$.

Способ выполнения критерия постквантовой стойкости

Для построения схемы цифровой подписи, удовлетворяющей общему критерию постквантовой стойкости, представляет интерес использовать идею, предложенную ранее в работе [10] и заключающуюся в применении маскирующего умножения элементов скрытой циклической группы простого порядка q на перестановочные с ними элементы из другой циклической группы, также имеющей порядок q . В принципе данная идея может быть реализована при задании скрытой группы, имеющей двухмерное или многомерное циклическое строение. При этом могут быть применены различные способы формирования открытых параметров схемы цифровой подписи с использованием указанного маскирующего умножения. Однако практический интерес представляют способы, которые обеспечивают существенное уменьшение размера подписи по сравнению со схемой цифровой подписи [10]. С этой целью предлагается следующий способ формирования открытого ключа при использовании КНАА, заданной по таблице 1, в качестве алгебраического носителя:

1) сформировать базис $\langle G, Q \rangle$ коммутативной группы $\Gamma_{\langle G, Q \rangle}$, включающий векторы порядка q ; обратимые векторы G и Q являются секретными и задают скрытую группу $\Gamma_{\langle G, Q \rangle}$, обладающую двухмерной циклическостью;

2) сгенерировать случайные обратимые четырехмерные векторы $\mathbf{A}$ и $\mathbf{B}$ порядка $p^2 - 1$;

3) сгенерировать случайное натуральное число $x < q$ и вычислить вектор $\mathbf{U} = \mathbf{A} \cdot \mathbf{G}^x \cdot \mathbf{B}^{-1}$ (умножение слева на $\mathbf{A}$ и умножение справа на $\mathbf{B}^{-1}$ маскируют группу $\Gamma_{\langle \mathbf{G}, \mathbf{Q} \rangle}$, в которой выполняется операция экспоненцирования в степень x ; при этом отображение вектора $\mathbf{G}^x$ в вектор $\mathbf{U}$ не является коммутативным с операцией экспоненцирования);

4) вычислить вектор $\mathbf{Y} = \mathbf{B} \cdot \mathbf{G} \cdot \mathbf{Q} \cdot \mathbf{B}^{-1}$ (эта формула задает автоморфное отображение вектора $\mathbf{G} \cdot \mathbf{Q}$, которое является взаимно коммутативным с операцией экспоненцирования);

5) сгенерировать случайное натуральное число $u < q$ ($u \neq x$) и вычислить вектор $\mathbf{Z} = \mathbf{B} \cdot \mathbf{Q}^u \cdot \mathbf{A}^{-1}$.

Левое умножение на $\mathbf{B}$ и правое на $\mathbf{A}^{-1}$ задает маскирующее отображение вектора $\mathbf{Q}^u$, которое не является коммутативным с операцией экспоненцирования. Однако это отображение и отображение, используемое в ходе применения указанного выше способа формирования открытого ключа, позволяют построить композиционные отображения, обладающие свойством взаимной коммутативности с операцией экспоненцирования. Последнее используется в процедуре проверки подлинности цифровой подписи и для обеспечения корректности разработанной схемы цифровой подписи. Открытым ключом является тройка векторов $\mathbf{U}, \mathbf{Y}$ и $\mathbf{G}$. Размер открытого ключа $(\mathbf{U}, \mathbf{Y}, \mathbf{Z})$ равен 768 байт.

Первый этап формирования открытого ключа состоит в генерации базиса $\langle \mathbf{G}, \mathbf{Q} \rangle$ скрытой коммутативной группы с двухмерной циклическостью. Для задания группы $\Gamma_{\langle \mathbf{G}, \mathbf{Q} \rangle}$ предлагается следующий алгоритм генерации двух независимых порождающих векторов $\mathbf{G}$ и $\mathbf{Q}$.

Шаг 1.1. Сгенерировать случайный необратимый вектор $\mathbf{R}$.

Шаг 1.2. Выбирая случайное значение d ($1 < d < p - 1$) и используя формулы (4) и (6), вычислить обратимый вектор $\mathbf{X}$.

Шаг 1.3. Проверить, является ли число q порядком вектора $\mathbf{X}$. Если порядок вектора $\mathbf{X}$ не равен q или $\mathbf{X}$ является вектор-скаляром, то перейти к шагу 1.2. (с учетом того, что $p = 2q + 1$, легко показать, что в среднем потребуется выполнить ≈ 4 раза шаги 1.2 и 1.3 для перехода к шагу 1.4).

Шаг 1.4. Взять в качестве первого элемента базиса скрытой группы вектор $\mathbf{G} = \mathbf{X}$.

Шаг 1.5. Сгенерировать натуральное число β , где ($1 < \beta < p - 1$), и порядок которого по модулю p равен значению q ; затем вычислить второй элемент $\mathbf{Q}$ базиса скрытой группы: $\mathbf{Q} = \beta \mathbf{E} \cdot \mathbf{X}$.

Второй этап формирования открытого ключа состоит в генерации случайных обратимых векторов $\mathbf{A}$ и $\mathbf{B}$ порядка $p^2 - 1$ и выполняется следующим образом.

Шаг 2.1. Сгенерировать случайный обратимый вектор $\mathbf{R} = (r_0, r_1, r_2, r_3)$ с координатами $r_2 \neq 0$ и $r_3 \neq 0$.

2.2. Вычислить значения (см. формулу (5)):

$$\delta = \left(\frac{(\lambda r_1 - \mu r_0)^2}{4\lambda^2 r_3^2} + \frac{\mu r_2}{\lambda r_3} \right) \bmod p,$$

$$\varepsilon = \delta^q \bmod p.$$

Шаг 2.3. Если $\varepsilon = p - 1$, то значение подкоренного выражения формулы (5), записанной для вектора $\mathbf{R}$, является квадратичным невычетом (в соответствии с критерием Эйлера) и множество векторов, перестановочных с $\mathbf{R}$, содержат векторы порядка $p^2 - 1$.

Шаг 2.4. Если $\varepsilon = 1$, то перейти к шагу 2.1.

Шаг 2.5. Сгенерировать пару случайных значений d ($1 < d < p - 1$) и h ($1 < h < p - 1$).

Шаг 2.6. В соответствии с формулой (3) вычислить вектор:

$$\mathbf{X} = \left(d, \frac{\mu r_2 d + (\lambda r_1 - \mu r_0)h}{\lambda r_2}, h, \frac{r_3}{r_2} h \right).$$

Шаг 2.7. Если одновременно выполняются неравенства $\mathbf{X}^2 \neq \mathbf{E}$ и $\mathbf{X}^q \neq \mathbf{E}$, то в качестве вектора порядка $p^2 - 1$ взять вектор $\mathbf{V} = \mathbf{X}$.

Все случайно генерируемые векторы и числа, использованные в ходе процесса вычисления открытого ключа, относятся к секретным параметрам. Открытыми параметрами схемы цифровой подписи являются элементы открытого ключа и параметры задания четырехмерной алгебры, используемой в качестве алгебраического носителя крипто-схемы.

Соответствие общему критерию стойкости к квантовым атакам и корректность разработанной схемы цифровой подписи обеспечивается комбинированием согласованных маскирующих операций двух различных типов: 1) свободных от свойства взаимной коммутативности с операцией экспоненцирования и 2) обладающих этим свойством. Без знания секретных векторов $\mathbf{A}, \mathbf{B}, \mathbf{G}$ и $\mathbf{Q}$ предположительно задание периодической функции, в которой содержатся периоды с длиной, зависящей от значений x и u , является вычислительно невыполнимой задачей (нахождение x и u представляет собой задачу дискретного логарифмирования в скрытой группе).

Так, рассмотрим периодическую функцию:

$$F_1(i, j) = (\mathbf{U} \cdot \mathbf{Y} \cdot \mathbf{Z})^i (\mathbf{U} \cdot \mathbf{Z})^j =$$

$$= \mathbf{A} \cdot \mathbf{G}^{xi+u+xj} \cdot \mathbf{Q}^{i+ui+uj} \cdot \mathbf{A}^{-1}.$$

Пусть длина периода этой функции равна (δ_i, δ_j) . Тогда в силу независимости векторов $\mathbf{G}$ и $\mathbf{Q}$ при условии $\Delta = (1+x)u - (1+u)x \neq 0$ (Δ — главный детерминант следующей системы линейных уравнений с неизвестными δ_i и δ_j) имеем:

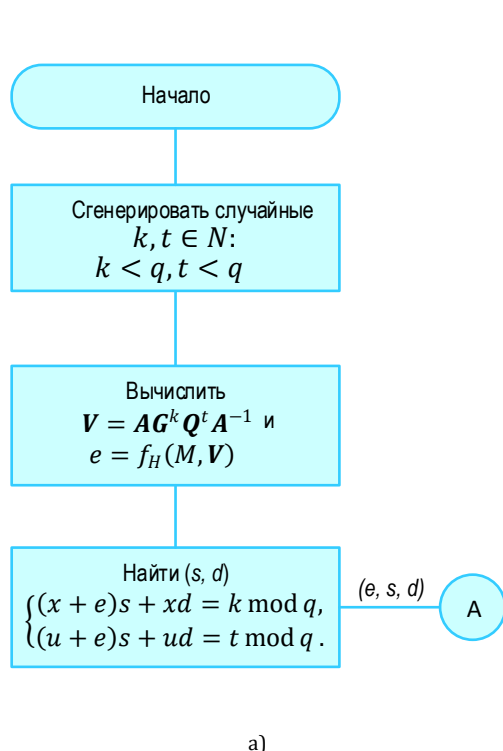
$$\begin{cases} x\delta_i + \delta_i + x\delta_j \equiv 0 \pmod{q} \\ \delta_i + u\delta_i + u\delta_j \equiv 0 \pmod{q} \end{cases} \Rightarrow \begin{cases} \delta_i \equiv 0 \pmod{q} \\ \delta_j \equiv 0 \pmod{q} \end{cases},$$

т. е. функция $F_1(i, j)$ содержит только периоды, длина которых равны значениям (kq, tq) при целочисленных значениях k и t (вероятность выполнения условия $\Delta = 0$ пренебрежимо мала). Легко показать, что последнее относится и к следующим двум периодическим функциям:

$$\begin{aligned} F_2(i, j) &= (\mathbf{Z} \cdot \mathbf{U})^i \mathbf{Y}^j = \mathbf{B} \cdot \mathbf{Q}^{ui+j} \cdot \mathbf{G}^{xi+j} \cdot \mathbf{B}^{-1}, \\ F_3(i, j) &= (\mathbf{U} \cdot \mathbf{Z})^i \cdot \mathbf{U} \cdot \mathbf{Y}^j = \mathbf{A} \cdot \mathbf{G}^{xi+x+j} \cdot \mathbf{Q}^{ui+j} \cdot \mathbf{B}^{-1}. \end{aligned}$$

Процедуры генерации и проверки подлинности цифровой подписи

Вычисление подписи к электронному документу M предлагается выполнять по следующему алгоритму (рисунок 1а):



а)

Рис. 1. Алгоритм вычисления (а) и проверка (б) цифровой подписи

Fig. 1. Calculation Algorithm (a) and Verification Algorithm (b) of Digital Signature

Подписью является тройка 256-битных чисел (e, s, d) и имеет длину 96 байт. Вычислительная сложность процедуры вычисления подписи приближенно может быть оценена как две операции возведения четырехмерных векторов в 256-битную степень.

Проверка подлинности цифровой подписи (рисунок 1б) осуществляется по открытому ключу $(\mathbf{U}, \mathbf{Y}, \mathbf{Z})$, используя следующий алгоритм:

1) вычислить значение $\tilde{\mathbf{V}}$:

$$\tilde{\mathbf{V}} = (\mathbf{U} \cdot \mathbf{Y}^e \cdot \mathbf{Z})^s \cdot (\mathbf{U} \cdot \mathbf{Z})^d;$$

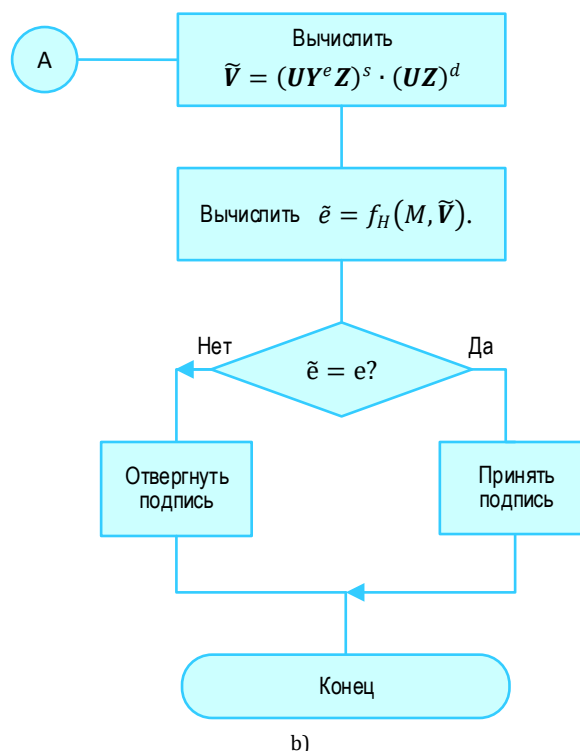
2) сгенерировать случайные натуральные числа $k < q$ и $t < q$ и вычислить вектор $\mathbf{V}$:

$$\mathbf{V} = \mathbf{A} \cdot \mathbf{G}^k \cdot \mathbf{Q}^t \cdot \mathbf{A}^{-1};$$

3) вычислить первый элемент подписи в виде двоичного числа $e = f_H(M, \mathbf{V})$, где f_H – некоторая специфицированная 256-битная хэш-функция, удовлетворяющая требованию строгой коллизии стойкости;

3) вычислить второй s и третий d элементы цифровой подписи как решение следующей системы уравнений первой степени:

$$\begin{cases} (x + e)s + xd = k \pmod{q}; \\ (u + e)s + ud = t \pmod{q}. \end{cases}$$



б)

2) вычислить значение хэш-функции от документа с присоединенным к нему вектором $\tilde{\mathbf{V}}$:

$$\tilde{e} = f_H(M, \tilde{\mathbf{V}});$$

3) при выполнении равенства $\tilde{e} = e$ подпись признается подлинной (в противном случае подпись признается ложной).

Докажем корректность предложенной схемы цифровой подписи. Допустим, что подпись (e, s, d) была вычислена в соответствии с алгоритмом генерации подписи. Тогда на первом шаге проверочной процедуры будет вычислен вектор:

$$\begin{aligned}
\tilde{V} &= (U \cdot Y^e \cdot Z)^s \cdot (U \cdot Z)^d = \\
&= (A \cdot G^x \cdot B^{-1} \cdot B \cdot G \cdot Q \cdot B^{-1} \cdot B \cdot Q^u \cdot A^{-1})^s \cdot \\
&\quad \cdot (A \cdot G^x \cdot B^{-1} \cdot B \cdot Q^u \cdot A^{-1})^d = \\
&= A \cdot G^{xs} \cdot G^{es} \cdot Q^{es} \cdot Q^{us} \cdot G^{xd} \cdot Q^{ud} \cdot A^{-1} = \\
&= A \cdot G^{xs+es+xd} \cdot Q^{es+us+ud} \cdot A^{-1} = \\
&= A \cdot G^{(x+e)s+xd} \cdot Q^{(e+u)s+ud} \cdot A^{-1} = \\
&= A \cdot G^k \cdot Q^t \cdot A^{-1} = V.
\end{aligned}$$

Поскольку $\tilde{V} = V$, то на втором шаге проверочной процедуры будет вычислено значение $\tilde{e}$, удовлетворяющие равенству $\tilde{e} = e$, т. е. подпись признается подлинной, что доказывает корректность разработанной схемы цифровой подписи.

Сопоставление с известными постквантовыми схемами цифровой подписи

С точки зрения компромисса между производительностью и размером открытого ключа и подписи предпочтительными представляются следующие два кандидата на постквантовый стандарт цифровой подписи, которые предложены в ходе конкурса НИСТ: Falcon [16] и Dilithium [17]. Сравнение разработанной постквантовой схемы цифровой подписи с указанными кандидатами и схемой цифровой подписи из работы [10] представлено в таблице 3 (сравнение выполнено для версий для версий алгоритмов Falcon и Dilithium, соответствующих уровню 128-битной стойкости). Схемы цифровой подписи, разработанные на основе СЗДЛ, обладают значительно меньшими размерами подписи и открытого ключа. При этом предложенная

схема цифровой подписи является более производительной.

ТАБЛИЦА 3. Сравнение разработанной авторами схемы цифровой подписи с известными аналогами

TABLE 3. A Rough Comparison of the Developed by the Authors Signature Scheme with Known Analogues

Схема цифровой подписи	Длина, байт		Скорость, отн. ед.	
	цифровой подписи	открытого ключа	генерации цифровой подписи	верификации цифровой подписи
[16]	657	897	5	3
[17]	2044	1184	2	1
[10]	192	768	4	3
Авторская схема	96	384	15	10

Заключение

Предложен новый способ построения схем цифровой подписи, удовлетворяющих общему критерию постквантовой стойкости, который позволяет существенно сократить размер подписи по сравнению со способом, описанным в работе [10]. На основе способа разработана конкретная постквантовая схема цифровой подписи, которая представляется более практичной по сравнению с известными кандидатами на постквантовый стандарт цифровой подписи, предложенными в ходе конкурса НИСТ [2] по разработке постквантовых двухключевых криптосхем.

ИСТОЧНИК ФИНАНСИРОВАНИЯ

Работа выполнена при финансовой поддержке РФФИ в рамках научного проекта № 21-57-54001-Вьет_а).

Список используемых источников

1. Ding J., Steinwand R. Post-Quantum Cryptography // Revised Selected Papers of the 10th International Conference (PQCrypto 2019, Chongqing, China, 8–10 May 2019). Lecture Notes in Computer Science. Security and Cryptology. Springer, 2019. Vol. 11505. 418 p.
2. Announcing Request for Nominations for Public-Key Post-Quantum Cryptographic Algorithms // Federal Register. 2016. Vol. 81. No. 244. Available from: <https://www.gpo.gov/fdsys/pkg/FR-2016-12-20/pdf/2016-30615.pdf> [Accessed 24th May 2021]
3. Молдовян А.А., Молдовян Д.Н. Постквантовые протоколы цифровой подписи на основе скрытой задачи дискретного логарифмирования // Вопросы защиты информации. 2019. № 2(125). С. 23–32.
4. Moldovyan N.A., Moldovyan A.A. Finite Non-commutative Associative Algebras as carriers of Hidden Discrete Logarithm Problem // Вестник ЮрГУ. Серия «Математическое моделирование и программирование». 2019. Т. 12. № 1. С. 66–81. DOI:10.14529/mmp190106
5. Молдовян Н.А., Абросимов И.К. Постквантовая схема ЭЦП на основе скрытой задачи дискретного логарифмирования в четырехмерной конечной алгебре // Вопросы защиты информации. 2019. № 2. С. 18–22.
6. Moldovyan A.A., Moldovyan N.A. Post-quantum signature algorithms based on the hidden discrete logarithm problem // Computer Science Journal of Moldova. 2018. Vol. 26. No. 3(78). PP. 301–313.
7. Shor P.W. Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on Quantum Computer // SIAM Review. 1999. Vol. 41. Iss. 2. PP. 303–332. DOI:10.1137/S0036144598347011
8. Ekert A., Jozsa R. Quantum computation and Shor's factoring algorithm // Reviews of Modern Physics. 1996. Vol. 68. PP. 733. DOI:10.1103/RevModPhys.68.733
9. Jozsa R. Quantum algorithms and the Fourier transform // Proceedings of the Royal Society A. Mathematical, Physical and Engineering Sciences. 1998. Vol. 454. Iss. 1969. PP. 323–337. DOI:10.1098/rspa.1998.0163
10. Moldovyan N.A., Moldovyan A.A. Candidate for practical post-quantum signature scheme // Вестник Санкт-Петербургского университета. Прикладная математика. Информатика. Процессы управления. 2020. Т. 16. №. 4. С. 455–461. DOI:10.21638/11701/spbu10.2020.410

11. Moldovyan N.A. Signature Schemes on Algebras, Satisfying Enhanced Criterion of Post-quantum Security // Buletinul Academiei de Stiinte a Republicii Moldova. Matematica. 2020. No. 2(93). PP. 62–67.
12. Moldovyan N.A. Fast signatures based on non-cyclic finite groups // Quasigroups and Related Systems. 2010. Vol. 1. No. 18. PP. 83–94.
13. Moldovyan N.A., Moldovyanu P.A. New primitives for digital signature algorithms // Quasigroups and Related Systems. 2009. Vol. 2. No. 17. PP. 271–282.
14. Фахрутдинов Р.Ш., Мирин А.Ю., Молдовян Д.Н., Костина А.А. Схемы открытого согласования ключей на основе скрытой задачи дискретного логарифмирования // Информационные технологии. 2020. Т. 26. № 10. С. 577–585. DOI:10.17587/it.26
15. Гурьянов Д.Ю., Молдовян Д.Н., Молдовян А.А. Постквантовые схемы цифровой подписи: задание скрытой группы с двухмерной циклическостью // Информатизация и связь. 2020. № 4. С. 75–82. DOI:10.34219/2078-8320-2020-11-4-75-82
16. Fast-Fourier Lattice-Based Compact Signatures over NTRU // Falcon. URL: <https://falcon-sign.info> (дата обращения 24.05.2021)
17. Dilithium Home // CRYSTALS. Cryptographic Suite for Algebraic Lattices. URL: <https://pq-crystals.org/dilithium/index.shtml> (дата обращения 24.05.2021)

* * *

Digital Signature Scheme with Hidden Group Possessing Two-Dimensional Cyclicity

D. Moldovyan¹, R. Fahrutdinov¹, A. Mirin¹, A. Kostina¹

¹St. Petersburg Federal Research Center of the Russian Academy of Sciences,
St. Petersburg, 199178, Russian Federation

Article info

DOI:10.31854/1813-324X-2021-7-2-85-93

Received 29th April 2021

Accepted 23rd May 2021

For citation: Moldovyan D., Fahrutdinov R., Mirin A., Kostina A. Digital Signature Scheme with Hidden Group Possessing Two-Dimensional Cyclicity. *Proc. of Telecom. Universities*. 2021;7(2):85–93. (in Russ.) DOI:10.31854/1813-324X-2021-7-2-85-93

Abstract: A method is proposed for constructing digital signature schemes based on the hidden discrete logarithm problem, which meet a general criterion of post-quantum resistance. The method provides a relatively small size of the public key and signature. Based on the method, a practical digital signature scheme has been developed, in which the exponentiation operation in a hidden group with two-dimensional cyclicity is the basic cryptographic primitive. The algebraic support of a cryptoscheme is a four-dimensional finite non-commutative algebra with associative multiplication operation. By specifying algebra using a basis vector multiplication table with half of empty cells, the performance of signature generation and authentication procedures is improved. A public key is a triple of four-dimensional vectors calculated as images of elements of a hidden group which are mapped using two types of masking operations: 1) mutually commutative with the exponentiation operation and 2) not having this property.

Keywords: information protection, computer security, cryptography, digital signature, post-quantum cryptoscheme, discrete logarithm problem, finite associative algebra, non-commutative algebra, hidden group.

FUNDING

This research was funded by RFBR according to the research project 21-57-54001 Viet_a.

References

1. Ding J., Steinwand R. Post-Quantum Cryptography. *Revised Selected Papers of the 10th International Conference, PQCrypto 2019, Chongqing, China, 8–10 May 2019. Lecture Notes in Computer Science. Security and Cryptology*. Springer; 2019. vol.11505. 418 p.
2. Announcing Request for Nominations for Public-Key Post-Quantum Cryptographic Algorithms. *Federal Register*. 2016;81(244). Available from: <https://www.gpo.gov/fdsys/pkg/FR-2016-12-20/pdf/2016-30615.pdf> [Accessed 24th May 2021]

3. Moldovyan A.A., Moldovyan D.N. Post-Quantum Digital Signature Protocols Based on the Hidden Discrete Logarithm Problem. *Information Security Questions* (Voprosy zašity informacii). 2019;2(125):23–32.
4. Moldovyan N.A., Moldovyan A.A. Finite Non-commutative Associative Algebras as carriers of Hidden Discrete Logarithm Problem. *Bulletin of the South Ural State University. Ser. Mathematical Modelling, Programming & Computer Software*. 2019;12(1):66–81. DOI:10.14529/mmp190106
5. Молдовьян Н.А., Молдовьян Д.Н. Постквантовая схема ЭЦП на основе скрытой задачи дискретного логарифмирования в четырехмерной конечной алгебре // Вопросы защиты информации. 2019. № 2. С. 18–22.
6. Moldovyan A.A., Moldovyan N.A. Post-quantum signature algorithms based on the hidden discrete logarithm problem. *Computer Science Journal of Moldova*. 2018;26(3(78)):301–313.
7. Shor P.W. Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on Quantum Computer. *SIAM Review*. 1999;41(2):303–332. DOI:10.1137/S0036144598347011
8. Ekert A., Jozsa R. Quantum computation and Shor's factoring algorithm. *Reviews of Modern Physics*. 1996;68:733. DOI:10.1103/RevModPhys.68.733
9. Jozsa R. Quantum algorithms and the Fourier transform. *Proceedings of the Royal Society A. Mathematical, Physical and Engineering Sciences*. 1998;454(1969):323–337. DOI:10.1098/rspa.1998.0163
10. Moldovyan N.A., Moldovyan A.A. Candidate for practical post-quantum signature scheme. *Vestnik of St Petersburg University. Applied Mathematics. Computer Science. Control Processes*. 2020;16(4):455–461. DOI:10.21638/11701/spbu10.2020.410
11. Moldovyan N.A. Signature Schemes on Algebras, Satisfying Enhanced Criterion of Post-quantum Security. *Buletinul Academiei de Stiinte a Republicii Moldova. Matematica*. 2020;2(93):62–67.
12. Moldovyan N.A. Fast signatures based on non-cyclic finite groups. *Quasigroups and Related Systems*. 2010;1(18):83–94.
13. Moldovyan N.A., Moldovyan P.A. New primitives for digital signature algorithms. *Quasigroups and Related Systems*. 2009;2(17):271–282.
14. Fahrutdinov R.S., Mirin A. Yu., Moldovyan D.N., Kostina A.A. Public Key – Agreement Schemes Based on the Hidden Discrete Logarithm Problem. *Information Technologies* (Informacionnye Tehnologii). 2020;26(10):577–585. DOI:10.17587/it.26
15. Guryanov D.Yu., Moldovyan D.N., Moldovyan A.A. Post-Quantum Digital Signature Schemes: Setting a Hidden Group with Two-Dimensional Cyclicity. *Informatizatsiia i sviaz*. 2020;4:75–82. DOI:10.34219/2078-8320-2020-11-4-75-82
16. Fast-Fourier Lattice-Based Compact Signatures over NTRU. *Falcon*. URL: <https://falcon-sign.info> [Accessed 24th May 2021]
17. Dilithium Home. *CRYSTALS. Cryptographic Suite for Algebraic Lattices*. URL: <https://pq-crystals.org/dilithium/index.shtml> [Accessed 24th May 2021]

Сведения об авторах:

МОЛДОВЯН
Дмитрий Николаевич

кандидат технических наук, научный сотрудник лаборатории кибербезопасности и постквантовых криптосистем Санкт-Петербургского федерального исследовательского центра Российской академии наук, mdn.spectr@mail.ru
 <https://orcid.org/0000-0001-5039-7198>

ФАХРУТДИНОВ
Роман Шафкатович

кандидат технических наук, заведующий лабораторией кибербезопасности и постквантовых криптосистем Санкт-Петербургского Федерального исследовательского центра Российской академии наук, fahr@cobra.ru

МИРИН
Анатолий Юрьевич

кандидат технических наук, старший научный сотрудник лаборатории кибербезопасности и постквантовых криптосистем Санкт-Петербургского Федерального исследовательского центра Российской академии наук, mirin@cobra.ru

КОСТИНА
Анна Александровна

научный сотрудник лаборатории кибербезопасности и постквантовых криптосистем Санкт-Петербургского Федерального исследовательского центра Российской академии наук, anna-kostina1805@mail.ru

ТРУДЫ МОЛОДЫХ УЧЕНЫХ

*Основой всей научной работы служит убеждение,
что мир представляет собой упорядоченную
и познаваемую сущность...*

Альберт Эйнштейн

1.2.2

2.2.6

2.2.13

2.2.14

2.2.15

2.2.16

2.3.1

2.3.6

Способ определения потенциала нарушителя безопасности информации и реализуемых им уязвимостей программного обеспечения

Р.В. Жук¹ 

¹Филиал «Макрорегион Юг» ООО ИК «СИБИНТЕК»,
Краснодар, 350063, Российская Федерация

*Адрес для переписки: goonerkrd@gmail.com

Информация о статье

Поступила в редакцию 27.01.2021

Принята к публикации 22.06.2021

Ссылка для цитирования: Жук Р.В. Способ определения потенциала нарушителя безопасности информации и реализуемых им уязвимостей программного обеспечения // Труды учебных заведений связи. 2021. Т. 7. № 2. С. 95–101. DOI:10.31854/1813-324X-2021-7-2-95-101

Аннотация: В настоящее время внедрено множество методических документов, регламентирующих подходы к разработке моделей угроз безопасности информации для информационных систем, обрабатывающих информацию различного характера. Существуют разные методики разработки угроз и построения модели нарушителя, предлагаемые регуляторами в области информационной безопасности, в зависимости от направления их деятельности. Для поддержки принятий решений в процессе моделирования угроз разработан банк данных угроз безопасности информации. Однако в существующих подходах имеется ряд противоречий, при этом методики определения угроз и построения модели нарушителя, в большинстве случаев, предполагают привлечение экспертов для оценки факторов и условий возникновения угроз. В существующих методиках отсутствует взаимосвязь между нарушителем безопасности информации, и уязвимостями программного обеспечения в информационных системах, что не позволяет построить адекватную модель угроз без привлечения квалифицированных экспертов. Целью данной работы является определение потенциала нарушителя безопасности информации в зависимости от его возможностей и оценка влияния данного потенциала на реализацию уязвимостей программного обеспечения в информационных системах.

Ключевые слова: потенциал, нарушитель безопасности информации, возможность, уязвимость, программное обеспечение, несанкционированный доступ.

Введение

Основопологающим этапом при проектировании подсистемы защиты информации является разработка модели угроз безопасности информации, которая должна включать в себя модель нарушителя безопасности информации (далее – нарушитель). В общем случае, механизм определения угроз, представленный в [1–4], состоит из следующих этапов:

- определение и подготовка перечня источников угроз, включая нарушителей, оценку их возможностей и потенциала;
- подготовка перечня уязвимостей программного обеспечения информационной системы и ее активов для последующего анализа;
- определение способов (сценариев) реализации угроз;
- оценка последствий от реализации угроз и

степень влияния выбранных угроз на свойства безопасности информации, обрабатываемой в информационной системе.

Несмотря на различное количество нарушителей, представленное в методических документах регуляторов в области информационной безопасности и защиты информации, их категорирование осуществляется с использованием мотивации, целей и типов, зависящих от наличия физического доступа внутрь контролируемой зоны, в которой размещены информационные системы. При этом, перечисленные в [5–7] нарушители обладают различными возможностями и потенциалом.

Наряду с этим, выбор способов (сценариев) реализации угроз нарушителем необходимо осуществлять с использованием банка угроз безопасности информации, что способствует подготовке перечня угроз с использованием типа и потенциа-

ла нарушителя, а также выбранных экспертным путем последствий реализации угрозы. При этом один из ключевых этапов процесса определения угроз безопасности информации, связанный с подготовкой перечня уязвимостей программного обеспечения, исключается. Дополнительно, нерешенной проблемой является отсутствие параметра потенциала для нарушителей, представленных в [5]. Таким образом, целью исследования является разработка способа установления потенциала нарушителей на основании их категорирования и подготовка алгоритма оценки влияния выбранных нарушителей на уязвимости программного обеспечения в информационной системе.

Анализ способов разработки моделей нарушителя и его влияния на угрозы безопасности информации

Впервые категорирование нарушителей по типам и возможностям было приведено в [5], однако в рамках предметной области приведенные характеристики нарушителя никак не связаны с процессом определения угроз безопасности информации. В [6] к возможностям нарушителей была добавлена мотивация (возможные цели), а также каждому нарушителю был присвоен потенциал, который позволяет для определения перечня угроз применять соответствующий банк угроз безопасности информации. Основным нововведением предлагаемого в [6] подхода является смена параметра вероятности возникновения угрозы на потенциал нарушителя при определении возможности реализации угрозы. При этом сама оценка возможности реализации угрозы представлена в виде таблицы с качественными значениями. Дополнительно, в [6] приводится способ определения потенциала нарушителя при использовании угроз безопасности информации, данные по которым не приведены в банке угроз безопасности информации. Данный способ предполагает экспертную оценку следующих возможностей на этапах идентификации и эксплуатации уязвимостей программного обеспечения: затрачиваемое время; техническая компетентность; знание проекта и возможность доступа к информационной системе; оснащенность нарушителя.

В [7] перечень видов нарушителей с присвоением потенциала представлен в зависимости от их возможностей по реализации угроз. Одной из характеристик угрозы является наличие хотя бы одного сценария ее реализации нарушителем с заданным потенциалом. Сценарий реализации угроз определяется, как совокупность тактик и техник, которые могут применяться нарушителем. Однако взаимосвязь между потенциалом и возможностями нарушителя с представленными в [7] тактиками является косвенной и должна быть установлена экспертным путем.

Дополнительно, при проектировании защиты информационных систем персональных данных (ИСПДн) необходимо руководствоваться методикой [8], которая заключается в дополнении функционала нарушителя возможностями, направленными на реализацию угроз, связанных с функционированием средств криптографической защиты информации.

В каждом из вышеперечисленных документов модель нарушителя является абстрактным понятием и может применяться справочно экспертами в ходе разработки модели угроз безопасности информации.

Наряду с регуляторами процесс построения модели нарушителя широко обсуждается в научных кругах. В 2017 г. С.О. Савченко и Н.В. Капчук в своей статье [9] изложили подход к построению модели нарушителя с использованием соотношения стратегий защитника, направленных на нейтрализацию стратегий нарушителя, реализующих множества угроз. В предлагаемом подходе для построения модели нарушителя и выбора стратегий оппонентов необходимо провести анализ рисков, ранжирование угроз безопасности информации, классифицировать систему и средства защиты; при этом механизмы и методы, которыми необходимо для этого руководствоваться, не приведены.

Ранее, в 2012 г. В.Г. Жуков, М.Н. Жукова и А.П. Стефаров в своей статье [10] предложили дополнить модели нарушителя уровнями несанкционированного доступа (НСД) к защищаемой информации в информационной системе. Классификация уровней была проведена с использованием сетевой модели ТСР/ІР. Однако для определения актуальности каждой угрозы и каждому нарушителю экспертным путем необходимо присвоить уровень НСД.

В работе [11] Е.А. Максимовой предложено использовать пятиуровневую модель доступа нарушителя к информации и компонентам объектов критической информационной инфраструктуры. Основными параметрами нарушителя при этом являются: возможность физического доступа; возможность логического доступа; компетенция; оснащенность; мотивация. При этом возможности физического и логического доступа, а также мотивацию предложено относить к базовому потенциалу, а оснащенность и компетенцию оценить экспертным путем с использованием качественной и количественной шкал.

Рассмотренные методы и подходы к разработке модели нарушителя, несмотря на свою значимость, не в полной мере описывают механизм оценки возможности использования им уязвимостей программного обеспечения информационной системы. В представленных работах и методиках отсутствует простой и удобный механизм выбора.

Также в них отсутствует определенный перечень нарушителей, а оценка потенциала нарушителей представлена крайне субъективно. Рассмотрим в этом контексте выбор нарушителей на примере ИСПДн.

Определение потенциала нарушителей безопасности информации

На первом этапе, согласно [5], необходимо присвоить потенциал нарушителям. Для этого воспользуемся методом анализа иерархий [12] Т. Саати. Оценка влияния критериев матрицы отношений будет осуществляться экспертным путем с использованием шкалы отношений, представленной в таблице 1.

ТАБЛИЦА 1. Шкала отношений

TABLE 1. Relationship Scale

Степень важности	Значимость
1	Одинаковая
3	Слабая
5	Сильная
7	Очень сильная
9	Абсолютная
2, 4, 6, 8	Промежуточные значения между соседними значениями шкалы

Согласно методологии Т. Саати, необходимо построить иерархические модели для трех категорий нарушителей безопасности информации – внешний, внутренний и внутренний со специальными возможностями.

Пример иерархической модели для внешнего нарушителя представлен на рисунке 1, где: НСД-I – НСД через автоматизированные рабочие места в сетях общего пользования; НСД-II – НСД с использованием специального программного обеспечения и вирусов; НСД-III – НСД через компоненты за пределами контролируемой зоны; НСД-IV – НСД через смежные информационные системы.

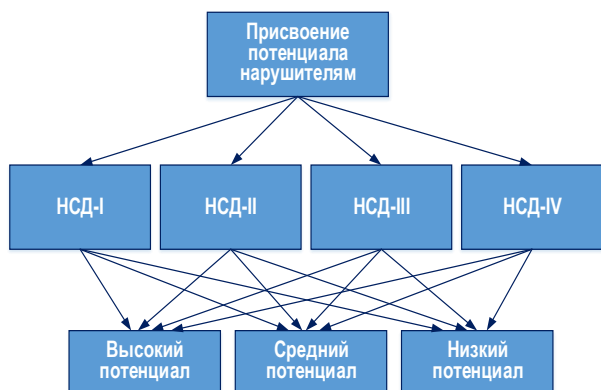


Рис. 1. Иерархическая модель определения потенциала внешнего нарушителя безопасности информации

Fig. 1. Hierarchical Model for Determining the Potential of an External Information Security Intruder

Частично матрицы отношений возможностей нарушителей к потенциалу заполняются экспертным путем с применением значений из таблицы 1. Остальные значения критериев заполняются с использованием следующей формулы:

$$x_{nm} = \frac{1}{x_{nm}}, \quad (1)$$

где x_{nm} – элемент в строке n столбца m .

Аналогичные матрицы разрабатываются для существующих альтернатив. На следующем этапе осуществляется нормирование составленных матриц по формуле:

$$X_{ij} = \frac{x_{nj}}{S_j}, \quad (2)$$

где X_{ij} – элемент i -й строки j -го столбца нормированной матрицы отношений; x_{nj} – элемент n -й строки j -го столбца матрицы отношений; S_j – сумма элементов x в j -м столбце.

Для каждой нормированной матрицы рассчитывается вес в долях по формуле:

$$D = \frac{\sum_j^1 x_i}{j}, \quad (3)$$

где x_i – элемент j -го столбца.

Примеры нормированных матриц для критериев и альтернатив потенциала нарушителей безопасности информации представлены в таблицах 2 и 3.

ТАБЛИЦА 2. Парное сравнение критериев возможностей нарушителей

TABLE 2. Pairwise Comparison of Intruder's Capability Criteria

	НСД-I	НСД-II	НСД-III	НСД-IV	Вес в долях
НСД-I	1	0,2	3	4	0,30
НСД-II	5	1	7	0,33	0,49
НСД-III	0,33	0,14	1	0,2	0,06
НСД-IV	0,25	0,25	2	1	0,13

ТАБЛИЦА 3. Парное сравнение альтернатив по критерию

TABLE 3. Pairwise Comparison of Alternatives by Criterion

НСД-I	Низкий	Средний	Высокий	Вес в долях
Низкий	0,13	0,14	0,10	0,12
Средний	0,50	0,57	0,60	0,56
Высокий	0,38	0,29	0,30	0,32

Аналогичные таблицы строятся для каждого оставшегося критерия (возможности нарушителя).

Потенциал нарушителя предлагается присваивать в зависимости от полученного результата при вычислении весов альтернатив по формуле:

$$A = (z_{ij}) * (y_j), \quad (4)$$

где z_{ij} – вес в долях альтернатив для каждого критерия (на примере таблицы 3); y_j – вес в долях для критериев.

Рассчитаем потенциал нарушителя на примере данных таблицы 2:

$$A = \begin{pmatrix} 0,12 & 0,1 & 0,1 & 0,1 \\ 0,56 & 0,57 & 0,57 & 0,57 \\ 0,32 & 0,33 & 0,33 & 0,33 \end{pmatrix} * \begin{pmatrix} 0,31 \\ 0,49 \\ 0,06 \\ 0,13 \end{pmatrix} = \begin{pmatrix} 0,11 \\ 0,56 \\ 0,33 \end{pmatrix}. \quad (5)$$

Видно, что перечисленным в [5] внешним нарушителям с весовой долей 0,56, вычисленной по формуле (5), может быть присвоен средний потенциал (см. таблицу 3).

По результатам использования метода при минимальном участии эксперта потенциал может быть присвоен любому нарушителю в зависимости от его возможностей.

Однако потенциала нарушителя недостаточно для определения перечня угроз безопасности информации, т. к. другим не менее значимым фактором угрозы НСД является наличие уязвимости программного обеспечения. При использовании возможностей или потенциала нарушителей для подготовки перечня таких уязвимостей, которые позволят реализовать угрозы НСД в рассматриваемой ИСПДн, возникает необходимость привлечения экспертов. Сам процесс выбора и сопоставления уязвимостей и нарушителей в информационных системах не регламентирован.

Для организации взаимосвязи уязвимостей программного обеспечения и потенциала нарушителя информационной безопасности рекомендуется спроецировать базовые метрики вектора уязвимости программного обеспечения на возможности потенциала нарушителя. При этом для описания потенциала нарушителя предлагается использовать унифицированные возможности, аналогично приведенным в [11], которые могут быть спроецированы на метрики уязвимостей с использованием метода из [12]:

- техническая компетентность;
- знание проекта и информационной системы;
- возможность доступа к информационной системе;
- оснащенность нарушителя.

Иерархическая модель проецирования метрик вектора уязвимости программного обеспечения на возможности нарушителя представлена на рисунке 2.

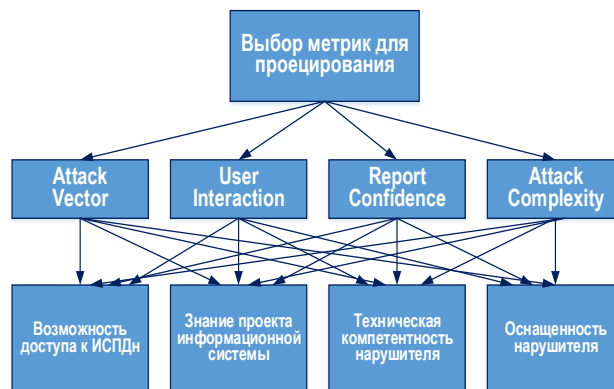


Рис. 2. Иерархическая модель проецирования метрики вектора уязвимости программного обеспечения на возможности потенциала нарушителя

Fig. 2. Hierarchical Model of Projection of the Vulnerability Metric of the Security Potential of the Attacker

Количественные значения параметрам данных метрик присваиваются с применением градации [13] и далее могут выступать, как условие возможности использования существующего в информационной системе уязвимости программного обеспечения выбранным ранее нарушителем с установленным потенциалом (таблица 4).

ТАБЛИЦА 4. Проецирование параметров вектора уязвимостей на возможности потенциала нарушителя

TABLE 4. Projecting the Parameters of the Vulnerability Vector onto the Capabilities of the Intruder's Potential

Метрика уязвимости программного обеспечения	Возможности нарушителя	Наибольший весовой коэффициент
Attack Vector (AV)	Возможность доступа к ИСПДн	0,52
User Interaction (UI)	Знание проекта информационной системы	0,41
Report Confidence (RC)	Техническая компетентность	0,32
Attack Complexity (AC)	Оснащенность	0,53

Таким образом, спроецированным значениям нарушителя экспертным путем могут быть присвоены количественные значения из верхнего диапазона, представленного в [13]. Для оценки возможности использования уязвимости программного обеспечения нарушителями с заданным потенциалом может быть разработана продукционная модель прямого вывода [14], состоящая из следующих правил:

– *реализуемая уязвимость*: если «AV уязвимости актива» <= «AV нарушителя» и «UI уязвимости актива» <= «UI нарушителя» и «RC уязвимости актива» <= «RC нарушителя» и «AC уязвимости актива» <= «AC нарушителя», то «Уязвимость №» = «Актуальная уязвимость»;

– *нереализуемая уязвимость*: если «AV уязвимости актива» > «AV нарушителя» или «UI уязвимости актива» > «UI нарушителя» или «RC уязвимости актива» > «RC нарушителя» или «AC уязвимости актива» > «AC нарушителя».

сти актива» > «RC нарушителя» или «АС уязвимости актива» > «АС нарушителя», то «Уязвимость №» = «Не Актуальная уязвимость».

На основании вышеперечисленных правил может быть построен алгоритм определения возможности реализации уязвимости программного обеспечения нарушителем с заданным потенциалом (рисунок 3).

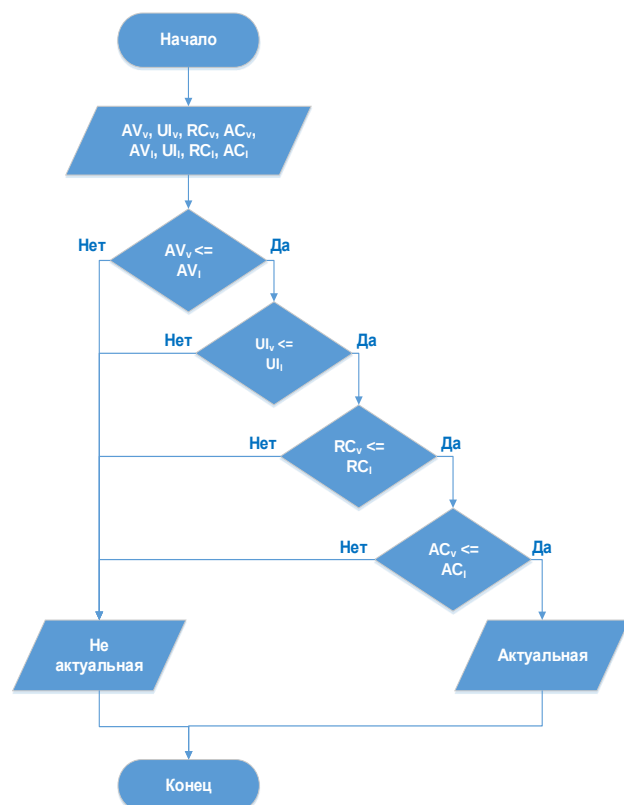


Рис. 3. Блок-схема алгоритма

Fig. 3. Flowchart of the Algorithm

Работа алгоритма состоит в следующем. На вход, с одной стороны, подаются количественные значения базовых метрик вектора уязвимостей программного обеспечения, перечень которых подготовлен на этапе определения активов информационной системы. С другой стороны, заранее определенные экспертным путем количественные значения возможностей нарушителя. Далее, по представленным выше правилам осуществляется попарное сравнение количественных значений базовых метрик вектора уязвимостей программного обеспечения информационной системы и возможностей нарушителя (см. таблицу 4). При положительном исходе, в результате выполнения всех приведенных в алгоритме условий, уязвимость программного обеспечения признается возможной к использованию выбранным нарушителем (т.е. актуальной); при невыполнении хотя бы одного из установленных условий – неактуальной.

Заключение

Предлагаемый способ определения потенциала нарушителя и реализуемых им уязвимостей программного обеспечения может применяться независимо от вида информационной системы и информации, обрабатываемой в ней. Основным преимуществом предлагаемого способа является возможность динамического и своевременного внесения изменений в модель угроз безопасности информации и переоценке возможностей нарушителя, что обеспечивается использованием комбинации удачно выбранного классического метода экспертной оценки и оригинального авторского алгоритма.

Использование метода анализа иерархий для установления потенциала нарушителей, а также при проецировании потенциала на параметры вектора уязвимости программного обеспечения, является простым и понятным инструментом, которым дополнительно могут руководствоваться эксперты при разработке моделей угроз безопасности для информационных систем, обрабатывающих информацию различного характера.

Данный метод позволяет исключить сложные вычислительные операции и сократить временные затраты экспертов, при этом позволяет подготовить исчерпывающий перечень уязвимостей программного обеспечения, которые могут быть реализованы выбранными нарушителями с заданным потенциалом.

Использование алгоритма определения возможности реализации уязвимости программного обеспечения информационной системы нарушителем с заданным потенциалом позволяет организовать взаимосвязь между его возможностями и характеристиками уязвимости. Основное достоинство данного алгоритма заключается в организации попарного сравнения количественных значений уязвимостей программного обеспечения, заранее установленных банком угроз безопасности информации, и установленных экспертным путем количественных значений возможностей нарушителя, что упрощает процесс подготовки перечня актуальных уязвимостей программного обеспечения в информационной системе.

Предлагаемый способ может быть дополнен путем построения взаимосвязи между уязвимостями программного обеспечения и угрозами безопасности информации, представленными в банке угроз безопасности информации.

Список используемых источников

1. Федеральный закон от 27.07.2006 № 152-ФЗ «О персональных данных».
2. Приказ ФСТЭК России от 11.02.2013 № 17 «Об утверждении требований о защите информации, не составляющей государственную тайну, содержащейся в государственных информационных системах».
3. Приказ ФСТЭК России от 14.03.2014 № 31 «Об утверждении требований к обеспечению защиты информации в автоматизированных системах управления производственными и технологическими процессами на критически важных объектах, потенциально опасных объектах, а также объектах, представляющих повышенную опасность для жизни и здоровья людей и для окружающей природной среды».
4. Приказ ФСТЭК России от 18.02.2013 № 21 «Об утверждении состава и содержания организационных и технических мер по обеспечению безопасности персональных данных при их обработке в информационных системах персональных данных».
5. Базовая модель угроз персональных данных при их обработке в информационных системах персональных данных. Москва, 2008. URL: <https://fstec.ru/tekhnicheskaya-zashchita-informatsii/dokumenty/114-spetsialnye-normativnye-dokumenty/379-bazovaya-model-ugroz-bezopasnosti-personalnykh-dannykh-pri-ikh-obrabotke-v-informatsionnykh-sistemakh-personalnykh-dannykh-vypiska-fstek-rossii-2008-god> (дата обращения 06.01.2021)
6. Методический документ ФСТЭК России «Методика определения угроз безопасности информации в информационных системах» URL: <https://fstec.ru/normotvorcheskaya/poisk-po-dokumentam/114-tekhnicheskaya-zashchita-informatsii/dokumenty/spetsialnye-normativnye-dokumenty/380-metodika-opredeleniya-aktualnykh-ugroz-bezopasnosti-personalnykh-dannykh-pri-ikh-obrabotke-v-informatsionnykh-sistemakh-personalnykh-dannykh-fstek-rossii-2008-god> (дата обращения 06.01.2021)
7. Методический документ ФСТЭК России «Методика моделирования угроз безопасности информации». URL: <https://fstec.ru/tekhnicheskaya-zashchita-informatsii/dokumenty/149-proekty/2070-metodicheskij-dokument> (дата обращения 06.01.2021)
8. Методические рекомендации по разработке нормативных правовых актов, определяющих угрозы безопасности персональных данных, актуальные при обработке персональных данных в информационных системах персональных данных, эксплуатируемых при осуществлении соответствующих видов деятельности. Утверждены руководством 8 Центра ФСБ России № 149/7/2/6-432 от 31.03.2015.
9. Савченко С.О., Капчук Н.В. Алгоритм построения модели нарушителя в системе информационной безопасности с применением теории игр // Динамика систем, механизмов и машин. 2017. Т. 5. № 4. С. 84–89. DOI:10.25206/2310-9793-2017-5-4-84-89
10. Жуков В.Г., Жукова М.Н., Стефаров А.П. Модель нарушителя прав доступа в автоматизированной системе // Программные продукты и системы. 2012. № 2. С. 75–78.
11. Максимова Е.А. Когнитивное моделирование деструктивных злоумышленных воздействий на объектах критической информационной инфраструктуры // Труды учебных заведений связи. 2020. Т. 6. № 4. С. 91–103. DOI:10.31854/1813-324X-2020-6-4-91-103.
12. Саати Т. Принятие решений. Метод анализа иерархий. Пер. с англ. М.: Радио и связь, 1993. 278 с.
13. Калькулятор метрик уязвимостей. URL: <https://www.first.org/cvss/calculator/cvsscalc30.js>, (дата обращения: 08.01.2020)
14. Хомоненко А.Д., Бубнов В.П., Басыров А.Г. Модели и методы исследования информационных систем. Монография. СПб.: Изд-во Лань, 2019. 204 с.

* * *

Method for Determining the Potential of an Information Security Intruder and Realizable Software Vulnerabilities

R. Zhuk¹ 

¹Branch "Macroregion Yug" LLC IC "SIBINTEK",
Krasnodar, 350063, Russian Federation

Article info

DOI:10.31854/1813-324X-2021-7-2-95-101

Received 27th January 2021

Accepted 22nd June 2021

For citation: Zhuk R. A Method for Determining the Potential of an Information Security Intruder and Realizable Software Vulnerabilities. *Proc. of Telecom. Universities*. 2021;7(2):95–101. (in Russ.) DOI:10.31854/1813-324X-2021-7-2-95-101

Abstract: Currently, many methodological documents have been developed that regulate approaches to the development of models of threats to information security. for information systems that process information of a different nature. There are different methods of threat development and intruder model building proposed by information security regulators, depending on the direction of their activity. To support decision-making in the process of threat modeling, a databank of information security threats has been developed. However, there are a number of contradictions in existing approaches, while the methods for identifying threats and building a model of an intruder, in most cases, involve the involvement of experts to assess the factors and conditions for the emergence of threats. In the existing methods, there is no relationship between the violator of information security. and software vulnerabilities in information systems, which does not allow building an adequate threat model without the involvement of qualified experts. The purpose of this work is to determine the potential of an information security violator. depending on its capabilities and assessing the impact of this potential on the implementation of software vulnerabilities in information systems.

Keywords: potential, information security intruder, opportunity, vulnerability, software, unauthorized access.

References

1. Federal Law No. 152-FZ of 27.07.2006 "On Personal Data". (in Russ.)
2. Order of the Federal Service for Technical and Export Control of Russia No. 17 of 11.02.2013 "On Approval of Requirements for the Protection of Information that Does Not Constitute a State Secret Contained in State Information Systems" (in Russ.)
3. Order of the Federal Service for Technical and Export Control of Russia No. 31 of 14.03.2014 "On Approval of Requirements for Ensuring the Protection of information in automated control systems for Production and Technological Processes at Critical Facilities, Potentially Dangerous Facilities, as well as Objects that pose an Increased danger to Human Life and health and to the environment" (in Russ.)
4. Order of the Federal Service for Technical and Export Control of Russia № 21 of 18.02.2013 "On Approval of Composition and Content of Organizational and Technical Measures for Ensuring Personal Data Security During their Processing in Personal Data Information Systems" (in Russ.)
5. *The Basic Model of Threats to Personal Data During their Processing in Personal Data Information Systems*. Moscow, 2008. Available from: <https://fstec.ru/tekhnicheskaya-zashchita-informatsii/dokumenty/114-spetsialnye-normativnye-dokumenty/379-bazovaya-model-ugroz-bezopasnosti-personalnykh-dannykh-pri-ikh-obrabotke-v-informatsionnykh-sistemakh-personalnykh-dannykh-vypiska-fstek-rossii-2008-god> [Accessed 6th January 2021] (in Russ.)
6. Methodological document of the Federal Service for Technical and Export Control of Russia "Methodology for Determining Information Security Threats in Information Systems". Available from: <https://fstec.ru/normotvorcheskaya-poisk-podokumentam/114-tekhnicheskaya-zashchita-informatsii/dokumenty/spetsialnye-normativnye-dokumenty/380-metodika-opredeleniya-aktualnykh-ugroz-bezopasnosti-personalnykh-dannykh-pri-ikh-obrabotke-v-informatsionnykh-sistemakh-personalnykh-dannykh-fstek-rossii-2008-god> [Accessed 6th January 2021] (in Russ.)
7. Methodological document of the Federal Service for Technical and Export Control of Russia "Methodology for Modeling Information Security Threats" Available from: <https://fstec.ru/tekhnicheskaya-zashchita-informatsii/dokumenty/149-proekty/2070-metodicheskij-dokument> [Accessed 6th January 2021] (in Russ.)
8. Methodological Recommendations for the Development of Regulatory Legal Acts Defining Threats to the Security of Personal Data that are Relevant in the Processing of Personal Data in Personal Data Information Systems Operated in the Implementation of Relevant Activities. No. 149/7/2/6-432 of 31.03.2015 (in Russ.)
9. Savchenko S.O., Kapchuk N.V. Algorithm for Constructing the Intruder Model in the Information Security System Using Game Theory. *Dynamics of Systems, Mechanisms and Machines*. 2017;5(4):84–89 (in Russ.). DOI:10.25206/2310-9793-2017-5-4-84-89
10. Zhukov V.G., Zhukova M.N., Stefarov A.P. Model of an Access Rights Intruder in an Automated System. *Programmnye produkty i sistemy*. 2012;2:75–78 (in Russ.)
11. Maximova E. Cognitive Modeling of Destructive Malicious Impacts on Critical Information Infrastructure Objects. *Proc. of Telecom. Universities*. 2020;6(4):91–103 (in Russ.) DOI:10.31854/1813-324X-2020-6-4-91-103
12. Saati T. *Decision Making. Hierarchy Analysis Method*. Translated from English. Moscow: Radio i sviaz Publ.; 1993. 278 p. (in Russ.)
13. Vulnerability Metrics Calculator. URL: <https://www.first.org/cvss/calculator/cvsscalc30.js> (Accessed 8 January 2020)
14. Khomonenko A.D., Bubnov V.P., Basyrov A.G. *Models and Research Methods for Information Systems*. St. Petersburg: Lan Publishing House; 2019. 204 p. (in Russ.)

Сведения об авторе:

ЖУК
Роман Владимирович

начальник отдела информационной безопасности филиала «Макрорегион Юг» ООО ИК «СИБИНТЕК», goonerkr@gmail.com
 <https://orcid.org/0000-0002-6604-9443>

Анализ гомоморфных криптосистем Бенало и Пэ́йе для построения системы электронного голосования

В.Д. Салман¹^{*}

¹Санкт-Петербургский государственный университет телекоммуникаций им. проф. М.А. Бонч-Бруевича,

Санкт-Петербург, 193232, Российская Федерация

^{*}Адрес для переписки: wasan.salman@mail.ru

Информация о статье

Поступила в редакцию 16.04.2021

Принята к публикации 28.05.2021

Ссылка для цитирования: Салман В.Д. Анализ гомоморфных криптосистем Бенало и Пэ́йе для построения системы электронного голосования // Труды учебных заведений связи. 2021. Т. 7. № 2. С. 102–109. DOI:10.31854/1813-324X-2021-7-2-102-109

Аннотация: Проведен анализ двух криптографических алгоритмов (Пэ́йе и Бенало) с целью их применения для построения систем электронного голосования. Дано описание каждого алгоритма и их гомоморфных свойств. Сформулированы требования к системам электронного голосования и приведено их построение при использовании этих криптоалгоритмов. Сравнительный анализ систем голосования на основе схем Пэ́йе и Бенало показал, что схема Пэ́йе является лучшим и более простым методом построения безопасных электронных систем голосования, в то время как схема Бенало является более сложной.

Ключевые слова: гомоморфное шифрование, криптосистема Пэ́йе, криптосистема Бенало, электронная система голосования.

Введение

Под гомоморфным шифрованием понимается криптографический примитив, представляющий собой функцию шифрования, удовлетворяющую дополнительному требованию гомоморфности относительно каких-либо алгебраических операций над открытыми сообщениями [1]. Гомоморфное шифрование используется во многих современных коммуникационных архитектурах. Они наиболее перспективны для построения безопасных систем электронного голосования [2]. Электронное голосование (ЭГ) – термин, определяющий различные виды голосования, охватывающий как электронные средства голосования, так и электронные средства подсчета голосов [2]. Технология электронного голосования позволяет ускорить процесс подсчета голосов, а также позволяет голосовать людям с ограниченными возможностями [3]. Важность применения гомоморфного шифрования заключается в том, что оно позволяет безопасно передавать, хранить и, главное, обрабатывать данные в зашифрованном виде без ущерба для конфиденциальности информации.

Первые гомоморфные системы были предложены Бенало в 1996 г. [4] и Пэ́йе в 1999 г. [5]. Эти схемы практичны, надежны и безопасны. Эти криптосистемы основаны на обычной модульной

арифметике. В работе [6] представлен обзор гомоморфных методов шифрования с учетом времени шифрования, памяти и безопасности. Обобщены многие типы гомоморфных схем, в том числе Пэ́йе, Бенало и др., по их параметрам и свойствам. Вывод этой работы состоит в том, что эти схемы не всегда подходят для любого применения и необходимо учитывать их характеристики в каждом конкретном случае. В работе [7] анализируются частично и полностью гомоморфные криптосистемы: RSA, Пэ́йе и Эль-Гамала. Подчеркивается преимущество гомоморфного шифрования, состоящее в возможности сохранения конфиденциальности при работе с криптограммами, а также отмечаются их недостатки: сложность, увеличенные размеры криптограмм и то, что некоторые из этих систем уязвимы для вредоносных программ. В работе [8] сравниваются гомоморфные схемы шифрования, основанные на схемах Голдвассера – Микали, Бенало, Наккаша – Штерна, Эль-Гамала и Пэ́йе. Сравнение проведено по времени генерации ключа, времени шифрования, времени дешифрования и общему размеру криптограммы. Сравнение осуществлялось на основе шифрования 100 случайных чисел. Результаты показали, что схема Пэ́йе дает наиболее эффективные результаты среди всех алгоритмов.

Протокол Гелиос (Helios) в [9] предложен для построения системы голосования членов студенческого совета и для небольших выборов. Основывается на подходе Бенало к голосованию. Helios – это первая доступная реализация системы веб-голосования с открытым аудитом. В ней используется неинтерактивное доказательство нулевого знания для подтверждения голосования. Helios – это веб-приложение, написанное на языке программирования Python, работающее внутри сервера приложений CherryPy 3.0 с веб-сервером Lighttpd. Все данные хранятся в базе данных PostgreSQL. Недостатки этого протокола: требуется Firefox 2 (или более поздняя версия), а также сложность пользовательского интерфейса.

В работе [10] была реализована и протестирована двойная система голосования, основанная на традиционном и электронном голосовании, на двух выборах в студенческий совет с участием более 2000 избирателей и на выборах лидера политической партии в 2012 г. В этой работе использовалась криптосистема Бенало. Эта система удовлетворяет следующим требованиям безопасности: конфиденциальность, целостность, неразличимость, любой избиратель может проверить, что его голос был правильно записан и включен в подсчет, а также любой может проверить, что все записанные голоса правильно подсчитаны. Недостатком данной работы является частично используемая бумажная система.

В работах [11–15] предложена система электронного голосования, основанная на гомоморфном шифровании с использованием криптосистемы Пэе и других методов. Свойство гомоморфного шифрования позволяет добавлять голоса в зашифрованном виде. Эта система удовлетворяет таким требованиям безопасности, как приемлемость, конфиденциальность, уникальность, целостность и точность.

В данной работе проведен анализ криптосистем Бенало и Пэе, так как они являются наиболее часто используемыми гомоморфными криптосистемами и имеют высокие гарантии безопасности. Основным результатом данной работы является анализ и сравнение двух криптосистем применительно к построению системы электронного голосования.

Анализ принципов построения криптосистем Бенало и Пэе

Криптосистема Бенало

В 1994 г. Д. Бенало [4], предложил гомоморфную схему шифрования, которая является вероятностным асимметричным алгоритмом для криптографии с открытым ключом.

Рассмотрим пошагово и поэтапно основные преобразования, характеризующие эту криптосистему: генерирование ключей, шифрование, расшифрование.

Генерирование ключей

Шаг 1. Выбираются два больших простых числа p, q и размер блока сообщения r (r -максимальное значение сообщения, представленного в виде числа). Число r выбирается таким образом, чтобы выполнялись условия:

$$(p-1) \bmod r = 0, \\ \gcd\left(r, \left(\frac{p-1}{r}\right)\right) = 1, \gcd(r, q-1) = 1.$$

Шаг 2. Вычисляется $n = pq$.

Шаг 3. Выбирается $y \in Z_n^*$, при $y^{\varphi/r} \not\equiv 1 \bmod n$, $y^{\varphi/r} \not\equiv 1$, где $\varphi(n) = (p-1)(q-1)$ – функция Эйлера от n .

Шаг 4. Вычисляется $x = y^{\varphi/r} \bmod n$.

Полагаем, что числа (y, n) являются открытым ключом, а (φ, x) – закрытым.

Шифрование

Если сообщение $m \in Z_r$, то шифрование производится путем выбора произвольного числа $u \in Z_n^*$ и вычисления криптограммы $c = y^m u^r \bmod n$.

Расшифрование

Расшифрование полученной криптограммы $c \in Z_n^*$ проходит в два этапа:

- вычисление $a = c^{\varphi/r} \bmod n$;
- подбор такого числа m , чтобы $m = \log_x a$,

Действительно, для любых $m \in Z_r, u \in Z_n^*$ можно записать выражение:

$$a = (c)^{\varphi/r} \equiv (y^m u^r)^{\varphi/r} \equiv (y^m)^{\varphi/r} (u^r)^{\varphi/r} \equiv \\ \equiv (y^{\varphi/r})^m (u)^{\varphi} \equiv (x)^m (u)^0 \equiv x^m \bmod n.$$

Криптосистема Бенало гомоморфна относительно операции сложения открытых тестов [4]:

$$\varepsilon(x_1) \times \varepsilon(x_2) = (g^{x_1} u_1^r)(g^{x_2} u_2^r) = g^{x_1+x_2} (u_1 u_2)^r \\ = \varepsilon(x_1 + x_2) \bmod r,$$

где $\varepsilon(x)$ – функцией шифрования от сообщения x .

Рассмотрим пример построения схемы Бенало на основе вышеприведенных соотношений. (В приведенных ниже примерах случайные числа будем получать от датчика чисел в программе Mathcad. В реальных системах предполагается использование физического датчика).

Генерирование ключей

Шаг 1. Выберем размер блока сообщения $r < 100$ и два больших простых числа $p = 397, q = 191$.

Если $r = 99$, тогда выполняются условия: $(397-1) \bmod 99 = 0, \gcd(99, \left(\frac{397-1}{99}\right)) = 1, \gcd(99, 191-1) = 1$.

Шаг 2. Вычислим $n = 397 \times 191 = 75827$, отсюда $\varphi(n) = 75240$.

Шаг 3. Выберем случайное целое число $y = 13213$, соответствующее требованию $y \in Z_n$ и проверим условие $y^{\varphi/r} \neq 1 \bmod n$.

Шаг 4. Определим x по выражению:

$$x = y^{\varphi/r} \bmod n = 13213^{75240/99} \bmod 75827 = 24640.$$

Таким образом, найден открытый (13213, 75827) и закрытый (75240, 24640) ключ.

Шифрование

Пусть $m \in Z_r$, $m = 78$. С целью шифрования сообщения выберем произвольное $u \in Z_n^*$: $u = 66183$. Тогда $c = 13213^{78} 66183^{99} \bmod 75827 = 47158$.

Расшифрование

Расшифруем криптограмму $c \in Z_n^*$:

$$- a = 67158^{75240/99} \bmod 75827 = 47178,$$

- построим таблицу значений $x^m = a \bmod n$ для $m = 0, 1, 2, \dots, 77, 78, \dots, 98$ и проверим выполнение условия ($x^m = a \bmod n$?):

$$m = 0, 24640^0 \bmod 75827 = 1 \neq a$$

$$m = 1, 24640^1 \bmod 75827 = 24640 \neq a$$

$$m = 2, 24640^2 \bmod 75827 = 58638 \neq a$$

$$m = 78, 24640^{78} \bmod 75827 = 47178 = a$$

Проверка показала, что восстановлен открытый текст $m = 78$.

Криптосистема Пэе

Криптосистема Пэе [5], является вероятностным асимметричным алгоритмом для криптографии с открытым ключом. Она основана на задаче вычисления n -го класса вычетов, которая считается трудновыполнимой. Основные преобразования, характеризующие эту криптосистему, состоят в следующем.

Генерирование ключей

Шаг 1. Выбираются два больших простых числа p и q , удовлетворяющие условию:

$$\gcd(pq, (p-1)(q-1)) = 1.$$

Шаг 2. Вычисляются числа n и λ по выражениям:

$$n = pq, \lambda = \text{lcm}(p-1, q-1),$$

где $\lambda(n)$ – функция Кармайкла от n ; $\lambda = \frac{(p-1)(q-1)}{\gcd(p-1, q-1)}$; lcm – наименьшее общее кратное.

Шаг 3. Выбирается $y \in Z_n^*$.

Шаг 4. Вычисляется $x = \left(L(y^\lambda \bmod n^2) \right)^{-1} \bmod n$, где $L(u) = \left\lfloor \frac{u-1}{n} \right\rfloor$; а выражение в скобках определяется, как – наибольшее целое s число, удовлетворяющее условию $u - 1 \geq s \cdot n$.

После вычислений, полагаем: (y, n) – открытый ключ, (λ, x) – закрытый ключ.

Шифрование

Сообщение $m \in Z_n$ шифруется посредством выполнения двух действий:

- выбора произвольного числа $u \in Z_n^*$;
- вычисления криптограммы $c = y^m u^r \bmod n^2$.

Расшифрование

Полученная криптограмма $c \in Z_{n^2}^*$ расшифровывается путем вычисления по формуле:

$$m = L(c^\lambda \bmod n^2) \times x \bmod n.$$

Криптосистеме Пэе присущи гомоморфные свойства [5], а именно:

– произведение двух шифротекстов будет расшифровано как сумма соответствующих им открытых текстов:

$$D(E(m_1, r_1) \cdot E(m_2, r_2)) \bmod n^2 = m_1 + m_2 \bmod n;$$

– шифротекст, возведенный в степень, равную другому шифротексту, будет расшифрован как произведение двух открытых текстов:

$$D(E(m_1, r_1)^{(m_2)} \bmod n^2) = m_1 m_2 \bmod n.$$

Рассмотрим пример построения схемы Пэе.

Генерирование ключей:

Шаг 1. Выберем два простых числа $p = 7$, $q = 5$ и проверим соответствие заданному условию:

$$\gcd(pq, (p-1)(q-1)) = 1,$$

Шаг 2. Определим, что $n = pq = 35$, $n^2 = 1225$ и $\lambda = \text{lcm}(6, 4) = 12$.

Шаг 3. Выберем случайное целое число $y = 3$, соответствующее требованию $y \in Z_n^*$.

Шаг 4. Определим, что $x = \left(L(y^\lambda \bmod n^2) \right)^{-1} \times \bmod n = 29$.

Таким образом, найден открытый (3, 35) и закрытый (12, 29) ключи.

Шифрование

Пусть $m \in Z_n$, $m = 8$. Выберем произвольное $u \in Z_n^*$, $u = 9$. Тогда $c = 3^8 \times 9^{35} \bmod 1225 = 939$.

Расшифрование

Расшифрование криптограммы $c \in Z_{n^2}^*$ выполняем по формуле:

$$m = L(939^{12} \bmod 1225) \times 29 \bmod 35 = 8,$$

т. е. открытый текст $m = 8$ восстановлен правильно.

Сравнение криптосистем Пэе и Бенало

Сравнение криптосистем Пэе и Бенало по некоторым свойствам приведено в таблице 1. Из таблицы видно, что обе криптосистемы схожи: имеют примерно одинаковый размер ключей и обладают гомоморфными свойствами. Однако в практическом плане система Пэе является более простой, поскольку при дешифровании криптограм-

мы не нужно решать задачу дискретного логарифмирования для чисел относительно небольшого размера.

ТАБЛИЦА 1. Сравнение криптосистем Пэе и Бенало

TABLE 1. Comparison of Paillier and Benaloh Cryptosystem

Свойства	Схема	
	Пэе	Бенало
Стойкость	Основана на задаче факторизации больших чисел	Основана на трудно-решаемой задаче о вычетах высокой степени
Удобство использования	Особенностью алгоритма является простота понимания и использования	
Время выполнения операций	Для дешифрования работает значительно быстрее	Операция дешифрования требует решения задачи дискретного логарифмирования для чисел относительно небольшого размера, что увеличивает время выполнения операции
Недостатки	Увеличение размера шифротекста по отношению к входному тексту	Является более сложной и трудоемкой для вычисления
Сходство	Используют шифрование с открытым ключом Обладают свойством гомоморфизма	

Электронная система голосования

Это удаленная система, использующая Интернет, мобильные компьютеры, смартфоны для того, чтобы дать возможность избирателям голосовать на выборах дистанционно.

Система электронного голосования состоит из следующих элементов:

- 1) *избиратели* – физические лица, имеющие законное право голоса;
- 2) *кандидаты* – лица, которые стремятся избраться на должность и имеют законное право на нее (кандидат избирается, если он был выбран по наибольшему числу отданных за него голосов);

3) *сервер* выполняет функцию перемножения криптограмм, полученных от избирателей и хранит зашифрованные голоса (рекомендуется использовать сервер с большой памятью и высокоскоростным процессором, поскольку сервер должен обрабатывать большое количество потоков для обслуживания большого количества избирателей за один раз);

4) *избирательная комиссия* (ИК) определяет, что избиратель имеет право на получение права голоса, гарантирует то, что избиратель голосует только один раз на данных выборах, осуществляет подсчет голосов избирателей и объявляет результаты выборов;

5) *наблюдатели* следят за соблюдением закона при проведении голосования и подсчете голосов (в случае выявления нарушений наблюдатель

должен либо постараться их пресечь, либо зафиксировать).

Для построения безопасной системы электронного голосования необходимо выполнить аутентификацию, а также соблюдать тайну голосования, целостность, точность, анонимность, уникальность, подтверждение голоса [13, 16–18].

Схема голосования при использовании системы с аддитивным гомоморфизмом, описанная в [19], в общем виде представляет из себя последовательности действий избирателей и избирательной комиссии, а также операций сервера. Избиратели проходят авторизацию в избирательной комиссии, получают бюллетени и ключ (ключи), зашифровывают свои бюллетени публичным ключом избирательной комиссии, после чего отправляют свои зашифрованные бюллетени на сервер. Сервер обрабатывает все зашифрованные бюллетени и отправляет получившийся результат в ИК. Избирательная комиссия с помощью секретного ключа расшифровывает криптограмму, содержащую сумму бюллетеней и объявляет победителя выборов. На рисунке 1 показана простая модель системы электронного голосования.



Рис. 1. Модель системы электронного голосования

Fig. 1. Electronic Voting System Model

Рассмотрим, как может быть построена избирательная система на криптосхемах Пэе и Бенало.

Примеры построения систем электронного голосования на криптосхемах Бенало и Пэе

Гомоморфная система голосования на основе схемы Пэе

Введем обозначения для описания системы электронного голосования:

N_v – количество избирателей;

N_c – количество кандидатов;

b – основание системы счисления, которое должно выбираться из условия: $b > N_v$, (т. е. должно быть больше, чем число избирателей).

Предположим, что избираются 2 члена для союза студентов из пяти кандидатов. Выбор одного кандидата или оставление бюллетеня пустым также возможен. Пусть $N_v = 9$, $N_c = 5$, $b = 10$, $b > N_v$. Результаты голосования избирателей представлены в таблице 2.

ТАБЛИЦА 2. Результаты голосования избирателей

TABLE 2. Results of Voter's Choice

Избиратели	Кандидаты					Шифрование сообщения
	C_1 10^0	C_2 10^1	C_3 10^2	C_4 10^3	C_5 10^4	
V_1		•				$m = 10^1 = 10$
V_2			•		•	$m = 10^2 + 10^4 = 10100$
V_3						$m = 0$
V_4				•		$m = 10^3 = 1000$
V_5	•			•		$m = 10^0 + 10^3 = 1001$
V_6		•		•		$m = 10^1 + 10^3 = 1010$
V_7			•	•		$m = 10^2 + 10^3 = 1100$
V_8		•		•		$m = 10^1 + 10^3 = 1010$
V_9	•					$m = 10^0 = 1$
Итого:	2	3	2	5	1	

Каждому кандидату C_i присваивается идентификатор – число b^i , где i – номер кандидата в списке. В данном случае выбираем десятичную систему счисления $b = 10$.

Согласно правилу выборов (голосовать можно не более чем за двух кандидатов), максимальное сообщение о голосовании, которое может быть зашифровано избирателем $m_{\max} = 10^3 + 10^4 = 11000$. А максимально возможная сумма голосов всех избирателей $T_{\max}$ определяется как:

$$T_{\max} = N_v \times m_{\max} = 9 \times 11000 = 99000.$$

Поэтому выберем модуль $n > T_{\max}$; $n > 99000$.

Используя шаги, описанные в криптосистеме Пэе для генерации ключей, получим ключи:

– для шифрования бюллетеней – открытый (6497955158, 126869);

– для расшифровки криптограммы избирательной комиссией – закрытый (31536, 53022).

Для шифрования сообщения $m \in Z_n$ каждый избиратель выбирает (программой) произвольное $u \in Z_n^*$ и создает криптограмму c .

Так, криптограмму первого избирателя можно записать в виде выражения:

$$c_i = 6497955158^{m_i} \times u_i^{126869} \bmod 16095743161.$$

Каждый избиратель отправляет криптограмму на сервер. Криптограммы всех избирателей сведены в таблицу 3.

Сервер вычисляет произведение полученных от избирателей криптограмм c_i и отправляет результат в избирательную комиссию:

$$\begin{aligned}
 c &= \prod_{i=1}^{N_v} c_i \bmod n^2 = \\
 &= (13039287935 \times 848742150 \times 7185465039 \times \\
 &\times 8093326 \times 722036441 \times 350667930 \times \\
 &\times 4980449314 \times 7412822644 \times \\
 &\times 3033281324) \bmod 16095743161 = 2747997353.
 \end{aligned}$$

ТАБЛИЦА 3. Криптограммы избирателей в системе Пэе

TABLE 3. Cryptograms of All Voters

Избиратель	Сообщение избирателя	Случайное число u_i	Криптограмма c_i
V_1	$m = 10^1 = 10$	35145	13039287935
V_2	$m = 10^2 + 10^4 = 10100$	74384	848742150
V_3	$m = 0$	96584	7185465039
V_4	$m = 10^3 = 1000$	10966	80933260
V_5	$m = 10^0 + 10^3 = 1001$	17953	722036441
V_6	$m = 10^1 + 10^3 = 1010$	7292	350667930
V_7	$m = 10^2 + 10^3 = 1100$	24819	4980449314
V_8	$m = 10^1 + 10^3 = 1010$	4955	7412822644
V_9	$m = 10^0 = 1$	118037	3033281324
Общая сумма голосов	15232		

Избирательная комиссия, используя закрытый ключ, проводит расшифрование криптограммы произведения:

$$\begin{aligned}
 T_{m \text{ общ.}} &= L(c^{\lambda} \bmod n^2) \times x \bmod n = \\
 &= \left(\frac{(2747997353^{31536} \bmod 16095743161) - 1}{126869} \right) \times \\
 &\times 53022 \bmod 126869 = 15232.
 \end{aligned}$$

Далее расшифрованное сообщение записывается как сумма разрядов в десятичной системе счисления:

$$15232 = 1 \times 10^4 + 5 \times 10^3 + 2 \times 10^2 + 3 \times 10^1 + 2 \times 10^0.$$

Наибольшие значения имеют коэффициенты при степенях 10^3 и 10^4 . Это означает, что победителями выборов являются кандидаты C_2 и C_4 .

Гомоморфная система голосования на основе схемы Бенало

Условия голосования такие же, как при рассмотрении схемы Пэе ($N_v = 9$, $N_c = 5$). Выберем основание системы счисления $b = 10$, $b > N_v$.

Используя шаги, описанные в криптосистеме Бенало избирательная комиссия генерирует ключи:

- открытый - (62369, 20205597437);
- закрытый - (20205310716, 6922019540).

Открытый ключ передается всем избирателям, закрытый остается в ИК.

Выбрав произвольное $u \in Z_n^*$, избиратели шифруют сообщения $m \in Z_r$ и создают криптограммы.

В качестве примера криптограмма первого избирателя представлена в виде:

$$c_i = 15^{m_i} \times u_i^{62369} \bmod 20205597437.$$

Криптограммы всех избирателей приведены в таблице 4.

ТАБЛИЦА 4. Криптограммы избирателей в системе Бенало

TABLE 4. Cryptograms of All Voters

Избиратель	Сообщение избирателя	Случайное число u_i	Криптограмма c_i
V_1	$m = 10^1 = 10$	35145	183946066
V_2	$m = 10^2 + 10^4 = 10100$	74384	10050175893
V_3	$m = 0$	96584	16340434784
V_4	$m = 10^3 = 1000$	10966	8189135103
V_5	$m = 10^0 + 10^3 = 1001$	17953	4202784310
V_6	$m = 10^1 + 10^3 = 1010$	7292	14220855747
V_7	$m = 10^2 + 10^3 = 1100$	24819	7937933779
V_8	$m = 10^1 + 10^3 = 1010$	4955	15466873618
V_9	$m = 10^0 = 1$	118037	1049803439
Общая сумма голосов	15232		

Обработка криптограмм на сервере проводится аналогично обработке криптограмм на основе схемы Пэе по выражению:

$$T = \prod_{i=1}^{N_p} c_i \bmod n^2 =$$

$$= (183946066 \times 10050175893 \times 1634043784 \times 8189135103 \times 4202784310 \times 14220855747 \times 7937933779 \times 15466873618 \times 1049803439) \bmod 20205597437 = 4249761249.$$

С целью расшифровки криптограммы $c \in Z_n^*$ необходимо вычислить a_i :

$$a_i = 4249761249^{323964} \bmod 20205597437 = 2281530556,$$

после чего – для $m = 0, 1, 2, 3, 4, \dots, 15232, 15777, 16123, 20000, 24565, 30567$ построить таблицу значений $x^m \bmod n$, и проверить выполнение сравнения ($x^m \bmod n = a$?).

Из таблицы следует, что восстановленный открытый текст $m = 15232$. Из анализа этого числа $15232 = 1 \times 10^4 + 5 \times 10^3 + 2 \times 10^2 + 3 \times 10^1 + 2 \times 10^0$. Наибольшие значения имеют коэффициенты при степенях 10^3 и 10^1 , из этого следует, что кандидаты C_2 и C_4 являются победителями.

Список используемых источников

1. Коржик В.И., Яковлев В.А. Основы криптографии: учебное пособие. СПб.: ИЦ Интермедия, 2016. 296 с.
2. Jabbar I., Alsaad N.S. Design and Implementation of Secure Remote e-Voting System Using Homomorphic Encryption // International Journal of Network Security. 2017. Vol. 19. № 5. PP. 694–703. DOI:10.6633/IJNS.201709.19(5).06
3. Singh A., Ramakanth K., Cholli P., Nagaraj G. Empowering E-governance with E-voting // Indonesian Journal of Electrical Engineering and Computer Science. 2018. Vol.12. PP. 1081–1086. DOI:10.11591/ijeecs.v12.i3.pp1081-1086
4. Benaloh J.C. Verifiable Secret-Ballot Elections. PhD Thesis. Yale University, 1996. 134 p.
5. Gennaro R., Halevi S., Rabin T. Secure Hash-and-Sign Signatures without the Random Oracle // Proceedings of the International Conference on the Theory and Application of Cryptographic Techniques (EUROCRYPT 1999, Prague Czech Republic, 2–6 May 1999). Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, 1999. PP. 123–139. DOI:10.1007/3-540-48910-X_9
6. Fontaine C., Galand F. A Survey of Homomorphic Encryption for Nonspecialists // EURASIP Journal on Information Security. 2007. Vol. 2007. DOI:10.1155/2007/13801
7. Morris L. Analysis of Partially and Fully Homomorphic Encryption. Rochester: Rochester Institute of Technology, 2013.

ТАБЛИЦА 5. Значение функции x^m для $m = 1, \dots, 15232$ TABLE 5. The Values of the Function x^m для $m = 1, \dots, 15232$

m	$x^m \bmod n$
0	$6922019540^0 \bmod 20205597437 = 1 \neq a$
1	$6922019540^1 \bmod 20205597437 = 6922019540 \neq a$
2	$6922019540^2 \bmod 20205597437 = 12864689861 \neq a$
3	$6922019540^3 \bmod 20205597437 = 17168902137 \neq a$
.....	
15232	$6922019540^{15232} \bmod 20205597437 = 2281530556 = a$

Из приведенных выше примеров, объясняющих применение этих схем при построении защищенной электронной системы голосования, можно сделать вывод, что эти схемы отвечают следующим требованиям:

- конфиденциальность обеспечивается за счет шифрования бюллетеня избирателем;
- анонимность обеспечивается тем, что при расшифровании известна сумма голосов, но не известен голос отдельного избирателя.

Заключение

В работе проведен анализ криптосистем Пэе и Бенало и возможности их применения в системах электронного голосования. Между криптосистемами Пэе и Бенало существует много общего: шифрование с открытым ключом, свойство гомоморфизма, конфиденциальность и анонимность голосования. Схема Бенало требует выполнения большего объема математических вычислений при дешифровании криптограммы, что в свою очередь занимает много времени, поэтому на практике распространение получила криптосистема Пэе.

Дальнейшие исследования следует провести в направлении обеспечения требований безопасности, которые не в полном объеме выполняются в криптосистеме Пэе: аутентификация избирателей, подтверждение поданного голоса, предотвращение раннего дешифрования криптограмм избирательной комиссией и решение вопросов масштабирования этой системы.

8. Bhumika P., Dharmendra B. Homomorphic Encryption: Privacy Preserving Amicable E-voting System // International Journal of Computer Sciences and Engineering. 2019. Vol. 7. Iss. 12. PP. 46–50. DOI:10.26438/ijcse/v7i12.4650
9. Adida B. Helios: Web-based Open-Audit Voting // Proceedings of the 17th USENIX Security Symposium (USA, San Joce, 28 jule–1 august 2008). 2008. PP. 335–348.
10. Ben-Nun J., Fahri N., Llewellyn M., Riva B., Rosen A., Tashma A., Wikström D. A new implementation of a dual (paper and cryptographic) voting system // Proceedings of the 5th International Conference on Electronic Voting (EVOTE2012, Bregenz, Austria, 11–14 July 2012). 2012. PP. 315–329.
11. Hussien H., Aboelnaga H. Design of a secured e-voting system // International Conference on Computer Applications Technology (ICCAT, Tunisia, Sousse, 20–22 January 2013). IEEE, 2013. DOI:10.1109/ICCAT.2013.6521985
12. Sharma T. E-Voting using Homomorphic Encryption Scheme // International Journal of Computer Applications. 2016. Vol. 141. No. 13.
13. Huszti A. A homomorphic encryption-based secure electronic voting scheme // Publicationes Mathematicae. 2011. Vol. 79. PP. 479–496.
14. Varun M., Rahul S., Lawrence S., Kevin Zhu. Apollo. A secure, anonymized voting system using the Paillier cryptosystem // Project Report. 2016. PP. 1–12.
15. Ryan P.Y.A. Prêt à Voter with Paillier encryption // Mathematical and Computer Modelling. 2008. Vol. 48. Iss. 9–10. PP. 1646–1662. DOI:10.1016/j.mcm.2008.05.015
16. Htet N.O., Aye M. A. A Survey of Different Electronic Voting Systems // International Journal of Scientific Engineering and Technology Research. 2014. Vol. 3. PP. 3460–3464.
17. Qadah G.Z., Taha R. Electronic voting systems: Requirements, design, and implementation // Computer Standards and Interfaces. 2007. Vol. 29. Iss. 3. PP. 376–386. DOI:10.1016/j.csi.2006.06.001
18. Hao F., Ryan P.Y.A. Real-World Electronic Voting: Design, Analysis and Deployment. Boca Raton: Taylor & Francis Group, 2016. 461 p.
19. Toapanta S.M.T., Chalén L.J.Ch., Rojas J.G.O., Gallegos L.E.M. A Homomorphic Encryption Approach in a Voting System in a Distributed Architecture // IEEE International Conference on Power, Intelligent Computing and Systems (ICPICS, China, Shenyang, 28–30 July 2020). IEEE, 2020. PP. 206–210. DOI:10.1109/ICPICS50287.2020.9202073

* * *

Analysis of Homomorphic Cryptosystems of Benaloh and Paillier for the Construction of an Electronic Voting System

W. Salman¹ 

¹The Bonch-Bruевич Saint-Petersburg State University of Telecommunications,
St. Petersburg, 193232, Russian Federation

Article info

DOI:10.31854/1813-324X-2021-7-2-102-109

Received 16th April 2021

Accepted 28th May 2021

For citation: Salman W. Analysis of Homomorphic Cryptosystems of Benaloh and Paillier for the Construction of an Electronic Voting System. *Proc. of Telecom. Universities*. 2021;7(2):102–109. (in Russ.) DOI:10.31854/1813-324X-2021-7-2-102-109

Abstract: *The analysis of the performance of two cryptographic algorithms (Paillier and Benaloh), in order to apply them in the construction of electronic voting systems is carried out. A description of each system and their homomorphic properties is given. Electronic voting systems based on these schemes are described. The requirements for the voting system are formulated and a comparative analysis of the voting systems based on the schemes of Paillier and Benaloh is carried out. The analysis showed that the Paillier scheme is the best and simplest method for building secure electronic voting systems, while the Benaloh scheme is more complex and computationally more time-consuming.*

Keywords: *homomorphic encryption, Paillier cryptosystem, Benaloh cryptosystem, electronic voting system.*

References

1. Korzhik V.I., Yakovlev V.A. *Fundamentals of Cryptography*. St. Petersburg: IC Intermedia Publ.; 2016. 296 p. (in Russ.)
2. Jabbar I., Alsaad N.S. Design and Implementation of Secure Remote e-Voting System Using Homomorphic Encryption. *International Journal of Network Security*. 2017;19(5):694–703. DOI: 10.6633/IJNS.201709.19(5).06
3. Singh A., Ramakanth K., Cholli P., Nagaraj G. Empowering E-governance with E-voting. *Indonesian Journal of Electrical Engineering and Computer Science*. 2018;12:1081–1086. DOI:10.11591/ijeecs.v12.i3.pp1081-1086
4. Benaloh J.C. *Verifiable Secret-Ballot Elections*. PhD Thesis. Yale University; 1996. 134 p.
5. Gennaro R., Halevi S., Rabin T. Secure Hash-and-Sign Signatures without the Random Oracle. *Proceedings of the International Conference on the Theory and Application of Cryptographic Techniques, EUROCRYPT 1999, 2–6 May 1999, Prague, Czech Republic. Lecture Notes in Computer Science*. Berlin, Heidelberg: Springer; 1999. p.123–139. DOI:10.1007/3-540-48910-X_9
6. Fontaine C., Galand F. A Survey of Homomorphic Encryption for Nonspecialists. *EURASIP Journal on Information Security*. 2007;2007. DOI:10.1155/2007/13801
7. Morris L. *Analysis of Partially and Fully Homomorphic Encryption*. Rochester: Rochester Institute of Technology; 2013.
8. Bhumika P., Dharmendra B. Homomorphic Encryption: Privacy Preserving Amicable E-voting System. *International Journal of Computer Sciences and Engineering*. 2019;7(12):46–50. DOI:10.26438/ijcse/v7i12.4650
9. Adida B. Helios: Web-based Open-Audit Voting. *Proceedings of the 17th USENIX Security Symposium, 28 July–1 August 2008, USA, San Jose*. 2008. p.335–348.
10. Ben-Nun J., Fahri N., Llewellyn M., Riva B., Rosen A., Tashma A., Wikström D. A new implementation of a dual (paper and cryptographic) voting system. *Proceedings of the 5th International Conference on Electronic Voting, EVOTE2012, 11–14 July 2012, Bregenz, Austria*. 2012. p.315–329.
11. Hussien H., Aboelnaga H. Design of a secured e-voting system. *International Conference on Computer Applications Technology, ICCAT, 20–22 January 2013, Tunisia, Sousse*. IEEE; 2013. DOI:10.1109/ICCAT.2013.6521985
12. Sharma T. E-Voting using Homomorphic Encryption Scheme. *International Journal of Computer Applications*. 2016;141(13).
13. Huszti A. A homomorphic encryption-based secure electronic voting scheme. *Publicationes Mathematicae*. 2011;79:479–496.
14. Varun M., Rahul S., Lawrence S., Kevin Zhu. Apollo. A secure, anonymized voting system using the Paillier cryptosystem. *Project Report*. 2016:1–12.
15. Ryan P.Y.A. Prêt à Voter with Paillier encryption. *Mathematical and Computer Modelling*. 2008;48(9-10):1646–1662. DOI:10.1016/j.mcm.2008.05.015
16. Htet N.O., Aye M. A. A Survey of Different Electronic Voting Systems. *International Journal of Scientific Engineering and Technology Research*. 2014;3:3460–3464.
17. Qadah G.Z., Taha R. Electronic voting systems: Requirements, design, and implementation. *Computer Standards and Interfaces*. 2007;29(3):376–386. DOI:10.1016/j.csi.2006.06.001
18. Hao F., Ryan P.Y.A. *Real-World Electronic Voting: Design, Analysis and Deployment*. Boca Raton: Taylor & Francis Group; 2016. 461 p.
19. Toapanta S.M.T., Chalén L.J.Ch., Rojas J.G.O., Gallegos L.E.M. A Homomorphic Encryption Approach in a Voting System in a Distributed Architecture. *Proceedings of the International Conference on Power, Intelligent Computing and Systems, ICPICS, 28–30 July 2020, China, Shenyang*. IEEE; 2020. p.206–210. DOI:10.1109/ICPICS50287.2020.9202073

Сведения об авторе:

САЛМАН
Васан Давуд

аспирант кафедры защищенных систем связи Санкт-Петербургского государственного университета телекоммуникаций им. проф. М.А. Бонч-Бруевича,
wasan.salman@mail.ru
 <https://orcid.org/0000-0003-4454-7844>

7ND YOUNG RESEARCHERS INTERNATIONAL CONFERENCE ON THE INTERNET OF THINGS AND ITS ENABLERS:
"ARTIFICIAL INTELLIGENCE IN COMMUNICATION NETWORKS"

INTHITEN

7-Я МЕЖДУНАРОДНАЯ НАУЧНО-ТЕХНИЧЕСКАЯ КОНФЕРЕНЦИЯ
СТУДЕНТОВ, АСПИРАНТОВ И МОЛОДЫХ УЧЕНЫХ

ИНТЕРНЕТ ВЕЩЕЙ И ЕГО ПРИЛОЖЕНИЯ.

ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ В СЕТЯХ СВЯЗИ

НАУЧНЫЕ НАПРАВЛЕНИЯ КОНФЕРЕНЦИИ

- Машинное обучение и искусственный интеллект в сетях связи
- Большие данные и методы их обработки
- Концепция Интернета вещей
- Беспроводные сенсорные сети
- Летящие сети, рои беспилотных летающих аппаратов
- Тактильный интернет
- Интернет Навыков
- Дополненная реальность
- Цифровая экономика
- Промышленный Интернет вещей
- Программно-конфигурируемые сети
- Технологии 5G/IMT-2020
- Сверхплотные гетерогенные сети
- Современные сетевые протоколы и архитектуры сигнализации
- Облачные решения для IoT
- Беспилотный автотранспорт
- Сети связи 6G/NET-2030
- Сетевая безопасность устройств и приложений интернета вещей
- Умные устойчивые города: технические аспекты развития
- Блокчейн и технологии распределенного реестра
- COVID-19: технологии телекоммуникаций как средство борьбы с пандемией
- Программная инженерия для сетевых технологий
- Дистанционное управление роботизированными платформами

25–27 ОКТАБРЯ 2021

Конференция INTHITEN проводится как одна из секций 13-го Международного конгресса по ультрасовременным телекоммуникациям и системам управления.

Труды конференции будут проиндексированы в библиографических и реферативных базах данных Scopus, Web of Science и РИНЦ, а также в электронной библиотеке IEEE Xplore.

СПбГУТ)))

Российский университет
дружбы народов
RUON University

BRND
UNIVERSITY
OF TECHNOLOGY

Tampere University
Tampere University of Applied Sciences

IEEE

RONDE & SCHWARZ

Ростелеком

ПОДРОБНОСТИ НА САЙТЕ
INTHITEN.ORG

О конференции

Первая конференция INTHITEN (THE INTERNET OF THINGS AND ITS ENABLERS) проходила в Санкт-Петербурге в июне 2013 года. В это время в России только начинали говорить про Интернет вещей, а тематика научных исследований только начинала формулироваться. Заведующий кафедрой сетей связи и передачи данных профессор Андрей Евгеньевич Кучерявый стоял у истоков создания концепции Интернета Вещей и участвовал в стандартизации первых международных стандартов, в которых описывались термины и определения по данному направлению в Международном Союзе Электросвязи. Именно он и его научная школа одними из первых стали проводить исследования по Интернету вещей, выступать на конференциях, публиковать статьи, читать лекции ещё в 2010 году, когда этот термин в России всерьёз никто не воспринимал. С учетом больших достижений в 2013 года научно-педагогическая школа проф. Кучерявого А. Е. «Интернет вещей и самоорганизующиеся сети» была включена в Реестр ведущих научных и научно-педагогических школ Санкт-Петербурга в соответствии с решением Президиума Научно-технического совета при Правительстве Санкт-Петербурга Комитета по науке и высшей школе. Благодаря этому, а также результатам научных достижений студентов, аспирантов и преподавателей ВУЗов в начале июня 2013 года была проведена первая конференция INTHITEN.

Традиционно, Международная конференция INTHITEN проводится в России в Санкт-Петербурге — месте встречи Востока и Запада, с его потрясающей архитектурой и прекрасными художественными галереями и музеями, а также многими другими достопримечательностями. В этом прекрасном городе можно увидеть и почувствовать историю: Зимний Дворец, Исаакиевский собор, Русский музей и другие достопримечательности на слуху у людей во всем мире, но увидеть их воочию — это настоящее удовольствие!

Программный и организационный комитет конференции состоит из ведущих ученых России и других стран, а участниками выступают молодые ученые, которые представляют результаты своих научных исследований по Интернету вещей и связанным технологиям. В 2021 году фокус конференции ориентирован на теоретические и практические аспекты применения искусственного интеллекта в сетях связи



Выходные данные

Дизайн обложки – ООО «Комильфо»

План издания научной литературы 2021 г., п. 5

Дата выхода в свет	Усл.-печ. л.	Формат	Тираж	Заказ	Свободная цена
30.06.2021	14	60×84 _{1/8}	1000 экз.	№ 1240	

Ответственный редактор **Татарникова И.М.**
Выпускающий редактор **Яшугин Д.Н.**

Адрес типографии:
193232, Санкт-Петербург, пр. Большевиков, 22/1

Учредитель и издатель:

Федеральное государственное бюджетное образовательное учреждение
высшего образования "Санкт-Петербургский государственный университет
телекоммуникаций им. проф. М.А. Бонч-Бруевича"

E-mail: tuzs@spbgut.ru Web: tuzs.sut.ru VK: vk.com/spbtuzs



Подписной индекс в Объединенном каталоге "ПРЕССА РОССИИ" - 59983