ISSN 1684–8853

ИНФОРМАЦИОННО-УПРАВЛЯЮЩИЕ СИСТЕМЫ

НАУЧНЫЙ ЖУРНАЛ

5(96)/2018

doi:10.31799/1684-8853-2018-5 ISSN 1684-8853

SCIENTIFIC JOURNAL

5(96)/2018

INFORMATSIONNO-UPRAVLIAIUSHCHIE SISTEMY (INFORMATION AND CONTROL SYSTEMS)

REFEREED EDITION

Founder
«Information and Control Systems», Ltd.
Publisher
Saint-Petersburg State University of Aerospace Instrumentation
Editor-in-Chief
M. Sergeev
Dr. Sc., Professor, Saint-Petersburg, Russia
Deputy Editor-in-Chief
E. Krouk
Dr. Sc., Professor, Moscow, Russia
Executive secretary
O. Muravtsova
Editorial Board
S. Andreev
PhD, Tampere, Finland
V. Anisimov
Dr. Sc., Professor, Saint-Petersburg, Russia
B. Bezruchko
Dr. Sc., Professor, Saratov, Russia
N. Blaunstein
Dr. Sc., Professor, Beer-Sheva, Israel
C. Christodoulou
PhD, Professor, Albuquerque, New Mexico, USA
A. Dudin
Dr. Sc., Professor, Minsk, Belarus
I. Dumer
PhD., Professor, Riverside, USA
M. Favorskaya
Dr. Sc., Professor, Krasnoyarsk, Russia
L. Fortuna
PhD, Professor, Catania, Italy
A. Fradkov
Dr. Sc., Professor, Saint-Petersburg, Russia
A. Hramov
Dr. Sc., Professor, Saratov, Russia
L. Jain
PhD, Professor, Canberra, Australia
V. Khimenko
Dr. Sc., Professor, Saint-Petersburg, Russia
G. Maltsev
Dr. Sc., Professor, Saint-Petersburg, Russia
G. Matvienko
Dr. Sc., Professor, Tomsk, Russia
V. Melekhin
Dr. Sc., Professor, Saint-Petersburg, Russia
A. Myllari
PhD, Professor, Grenada, West Indies
Y. Podoplyokin
Dr. Sc., Professor, Saint-Petersburg, Russia
J. Seberry
PhD, Professor, Wollongong, Australia
A. Shalyto
Dr. Sc., Professor, Saint-Petersburg, Russia
A. Shelupanov
Dr. Sc., Professor, Tomsk, Russia
A. Shepeta
Dr. Sc., Professor, Saint-Petersburg, Russia
Yu. Shokin
RAS Academician, Dr. Sc., Novosibirsk, Russia
A. Smirnov
Dr. Sc., Professor, Saint-Petersburg, Russia
T. Sutikno
PhD, Associate Professor, Yogyakarta, Indonesia
Z. Yuldashev
Dr. Sc., Professor, Saint-Petersburg , Russia
R. Yusupov
RAS Corr. Member, Dr. Sc., Professor, Saint-Petersburg, Russia
A. Zeifman
Dr. Sc., Professor, Vologda, Russia
Editor: A. Larionova
Proofreader: T. Zvertanovskaia
Design: M. Chernenko, Y. Umnitsina
Layout and composition: Y. Umnitsina

Contact information
The Editorial and Publishing Center, SUAI
67, B. Morskaia, 190000, St. Petersburg, Russia
Website: http://i-us.ru/en, e-mail: ius.spb@gmail.com
Tel.: +7 - 812 494 70 02

Submitted for publication 05.09.18. Passed for printing 22.10.18. Format 60×84 1/8.
Phototype SchoolBookC. Digital printing.

Layout original is made at the Editorial and Publishing Center, SUAI.
67, B. Morskaia, 190000, St. Petersburg, Russia
Printed from slides at the Editorial and Publishing Center, SUAI.
67, B. Morskaia, 190000, St. Petersburg, Russia

The journal is distributed by subscription. Subscription can be made in the Editorial and publishing center, SUAI as well as in any post office based on «Rospechat» catalogue:
№ 15385 — semiannual subscript.

The Journal was registered in the Ministry of Press, Broadcasting and Mass Media of the Russian Federation.
Registration Certificate JD № 77-12412 from April, 19, 2002.
Re-registration in the Federal Service for Supervision in the Sphere of Telecom, Information Technologies and Mass Communications (ROSKOMNADZOR) due to change of the founder: «Information and Control Systems», Ltd., JD № FS77-49181 from March, 30, 2012.
© Corporate authors, 2018

doi:10.31799/1684-8853-2018-5

ISSN 1684-8853

НАУЧНЫЙ ЖУРНАЛ

5(96)/2018

РЕЦЕНЗИРУЕМОЕ ИЗДАНИЕ

ИНФОРМАЦИОННО-УПРАВЛЯЮЩИЕ СИСТЕМЫ

Учредитель
ООО «Информационно-управляющие системы»
Издатель
Санкт-Петербургский государственный университет аэрокосмического приборостроения

Главный редактор
М. Б. Сергеев,
д-р техн. наук, проф., Санкт-Петербург, РФ
Зам. главного редактора
Е. А. Крук,
д-р техн. наук, проф., Москва, РФ
Ответственный секретарь
О. В. Муравцова
Редакционная коллегия:
С. Д. Андреев,
канд. техн. наук, Тампере, Финляндия
В. Г. Анисимов,
д-р техн. наук, проф., Санкт-Петербург, РФ
Б. П. Безручко,
д-р физ.-мат. наук, проф., Саратов, РФ
Н. Блаунштейн,
д-р физ.-мат. наук, проф., Беэр-Шева, Израиль
Л. С. Джайн,
д-р наук, проф., Канберра, Австралия
А. Н. Дудин,
д-р физ.-мат. наук, проф., Минск, Беларусь
И. И. Думер,
д-р наук, проф., Риверсайд, США
А. И. Зейфман,
д-р физ.-мат. наук, проф., Вологда, РФ
К. Кристодолу,
д-р наук, проф., Альбукерке, Нью-Мексико, США
Г. Н. Мальцев,
д-р техн. наук, проф., Санкт-Петербург, РФ
Г. Г. Матвиенко,
д-р физ.-мат. наук, проф., Томск, РФ
В. Ф. Мелехин,
д-р техн. наук, проф., Санкт-Петербург, РФ
А. А. Мюлляри
д-р наук, профессор, Гренада, Вест-Индия
Ю. Ф. Подоплёкин,
д-р техн. наук, проф., Санкт-Петербург, РФ
Т. Сутикноу
д-р наук, доцент, Джокарта, Индонезия
Д. Себерри,
д-р наук, проф., Волонгонг, Австралия
А. В. Смирнов,
д-р техн. наук, проф., Санкт-Петербург, РФ
М. Н. Фаворская,
д-р техн. наук, проф., Красноярск, РФ
Л. Фортуна,
д-р наук, проф., Катания, Италия
А. Л. Фрадков,
д-р техн. наук, проф., Санкт-Петербург, РФ
В. И. Хименко,
д-р техн. наук, проф., Санкт-Петербург, РФ
А. Е. Храмов,
д-р техн. наук, проф., Саратов, РФ
А. А. Шалыто,
д-р техн. наук, проф., Санкт-Петербург, РФ
А. А. Шелупанов,
д-р техн. наук, проф., Томск, РФ
А. П. Шепета,
д-р техн. наук, проф., Санкт-Петербург, РФ
Ю. И. Шокин,
акад. РАН, д-р физ.-мат. наук, проф., Новосибирск, РФ
З. М. Юлдашев,
д-р техн. наук, проф., Санкт-Петербург, РФ
Р. М. Юсупов,
чл.-корр. РАН, д-р техн. наук, проф., Санкт-Петербург, РФ

Редактор: А. Г. Ларионова
Корректор: Т. В. Звертановская
Дизайн: М. Л. Черненко, Ю. В. Умницына
Компьютерная верстка: Ю. В. Умницына

Адрес редакции: 190000, Санкт-Петербург,
Б. Морская ул., д. 67, ГУАП, РИЦ
Тел.: (812) 494-70-02, эл. адрес: ius.spb@gmail.com,
сайт: http://i-us.ru

Журнал зарегистрирован в Министерстве РФ по делам печати, телерадиовещания и средств массовых коммуникаций.
Свидетельство о регистрации ПИ № 77-12412 от 19 апреля 2002 г.
Перерегистрирован в Роскомнадзоре.
Свидетельство о регистрации ПИ № ФС77-49181 от 30 марта 2012 г.

© Коллектив авторов, 2018

Журнал входит в БД SCOPUS, в RSCI на платформе Web of Science и в Перечень рецензируемых научных изданий, в которых должны быть опубликованы основные научные результаты диссертаций на соискание ученой степени кандидата наук, на соискание ученой степени доктора наук.

Сдано в набор 05.09.18. Подписано в печать 22.10.18. Формат 60×84 1/8.
Гарнитура SchoolBookC. Печать цифровая.
Усл. печ. л. 14,5. Уч.-изд. л. 20,1. Тираж 1000 экз (1-й завод 120 экз.). Заказ № 449.

Оригинал-макет изготовлен в редакционно-издательском центре ГУАП.
190000, Санкт-Петербург, Б. Морская ул., 67.

Отпечатано с готовых диапозитивов в редакционно-издательском центре ГУАП.
190000, Санкт-Петербург, Б. Морская ул., 67.

Журнал распространяется по подписке. Подписку можно оформить через редакцию, а также в любом отделении связи по каталогу «Роспечать»: № 15385 — полугодовой индекс.

УДК 004
doi:10.31799/1684-8853-2018-5-2-12

Научные статьи
Articles

Подход к распознаванию стиля вождения водителя транспортного средства на основе использования сенсоров смартфона

И. Б. Лашков[а, б], *научный сотрудник, orcid.org/0000-0001-6418-4660, igor-lashkov@ya.ru*
[а]*Университет ИТМО, Кронверкский пр., 49, Санкт-Петербург, 197101, РФ*
[б]*Санкт-Петербургский институт информатики и автоматизации РАН, 14-я линия В. О., 39, Санкт-Петербург, 199178, РФ*

Введение: *анализ поведения водителя за рулем транспортного средства и оценка его стиля вождения позволяют обратить внимание водителя на навыки управления, снизить процент небезопасного вождения, повысить эффективность эксплуатации транспортного средства и, таким образом, снизить количество дорожно-транспортных происшествий. Однако современные системы содействия водителю ограничены в возможностях персонализации системы для водителя, снижая общую эффективность работы подобных систем и сужая их область применимости.* ***Цель исследования:*** *разработка подхода к анализу и оценке стиля вождения водителя при управлении транспортным средством с использованием фронтальной камеры и сенсоров смартфона, ориентированного на применение в системах повышения безопасности водителя.* ***Результаты:*** *описана схема потоков данных с сенсоров смартфона для системы предупреждения аварийных ситуаций; представлена информационная модель профиля водителя; разработан алгоритм анализа стиля управления транспортным средством индивидуально для водителя, позволяющий повысить качество распознавания опасных состояний в поведении водителя во время вождения и учитывающий предыдущий опыт использования системы, а также предоставляющий возможность оценить стиль вождения каждого водителя в отдельности, тем самым обратив его внимание на навыки управления транспортным средством; разработан прототип системы оценки стиля вождения водителя при управлении транспортным средством на основе обработки данных с сенсоров смартфона за счет выделения связей между ними и характеристиками управления транспортным средством, а также анализа истории взаимодействия водителя с системой предупреждения аварийных ситуаций.* ***Практическая значимость:*** *полученные результаты позволят повысить точность распознавания опасных состояний и, как следствие, эффективность системы предупреждения аварийных ситуаций, а также найдут применение при формировании и отображении отчетов для представителей страховых компаний, администраторов автопарков и руководителей логистических компаний, осуществляющих наблюдение и контроль за статистикой совершения поездок водителями из штата автопарка.*

Ключевые слова *— современные системы содействия водителю, водитель, транспортное средство, поведение водителя, стиль вождения.*

Цитирование: Лашков И. Б. Подход к распознаванию стиля вождения водителя транспортного средства на основе использования сенсоров смартфона. *Информационно-управляющие системы*, 2018, № 5, с. 2–12. doi:10.31799/1684-8853-2018-5-2-12
Citation: Lashkov I. B. Smartphone-based approach to determining driving style with on-board sensors. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 5, pp. 2–12 (In Russian). doi:10.31799/1684-8853-2018-5-2-12

Введение

Согласно статистике Госавтоинспекции МВД России более 80 % дорожно-транспортных происшествий происходят по вине самих водителей. Факторы риска, соответствующие психофизиологической составляющей поведения водителя, оказывают значительное влияние на возникновение дорожно-транспортных происшествий. К таким факторам повышенного риска наступления дорожно-транспортного происшествия можно отнести усталость и ослабленное внимание водителя, проявляющиеся в ситуациях, в которых водители зачастую сами не осознают наступление признаков того или иного опасного состояния.

Одним из вариантов повышения безопасности водителя является использование систем активной безопасности, направленных на предотвращение аварийных ситуаций на основе мониторинга поведения водителя и своевременного его оповещения о текущей ситуации за счет генерации ему контекстно-ориентированных рекомендаций. Первые исследования ученых по разработке систем мониторинга окружающей обстановки и предупреждения водителя автомобиля об опасности отмечены 1992 годом. Мониторинг опасного поведения за рулем способен помочь водителю обратить внимание на стиль своего вождения и связанные с этим риски, и тем самым снизить процент неосторожного вождения и улучшить навыки безопасного поведения на дороге.

По форме представления и принципу функционирования все системы слежения за состоянием водителя и дорожной обстановкой можно условно разделить на четыре категории: системы предупреждения аварийных ситуаций, существующие в виде аппаратно-программных комплексов и устанавливаемые автопроизводителями [1], мобильные системы генерации рекомендаций, существующие в виде мобильных приложений

для смартфонов [2], видеокамеры, устанавливаемые водителем в кабине транспортного средства (ТС), представленные автомобильными видеорегистраторами и отдельными устройствами видеонаблюдения, направленными на водителя, или на дорогу [3], устройства носимой электроники, надеваемые и носимые водителем перед совершением каждой поездки на автомобиле [4].

Основной целью данной статьи является разработка подхода к анализу поведения водителя, основанного на использовании предыдущего опыта взаимодействия водителя с ТС, включающим контекст внутри кабины ТС, данные с сенсоров смартфона и опасные состояния в поведении водителя [5–7].

Обзор литературы

Алгоритмы, анализирующие поведение водителя на основе его предыдущего опыта взаимодействия с ТС, представлены многочисленными исследованиями и разработками, рассмотрим некоторые, наиболее значимые из них.

На сегодняшний день одним из широко встречающихся подходов для интеллектуального анализа данных является применение различных методов Data Mining [8–12] и машинного обучения [13–17], ориентированных на обучение в условиях решения множества схожих задач. Значительная часть предлагаемых учеными решений в области изучения воздействия поведения водителя на дорожную обстановку ориентирована на использование данных с датчиков и сенсоров ТС (автомобиля) или смартфона водителя. Общим для всех таких систем является определение последовательности совершенных водителем в некоторый момент времени действий, образующих паттерн поведения [18]: алгоритм определения [19] действий водителя на основе данных с датчиков автомобиля посредством использования специализированного автомобильного симулятора для моделирования процесса управления ТС; алгоритм распознавания паттернов поведения [20] водителя за счет дискретизации его действий путем использования не фиксированного ограничения количества поступающих на вход системе событий, а временного интервала, а также другие исследования, решающие задачу анализа и классификации поведения водителя [21–29].

Стоит отметить, что рассмотренные исследования нацелены на решение только задачи классификации поведения водителя и фиксированной оценки его стиля вождения без учета его предыдущей истории взаимодействия с системой предупреждения аварийных ситуаций, включающей распознавание опасных состояний в его поведении (усталость, ослабленное внимание).

Схема потоков данных в системе предупреждения аварийных ситуаций

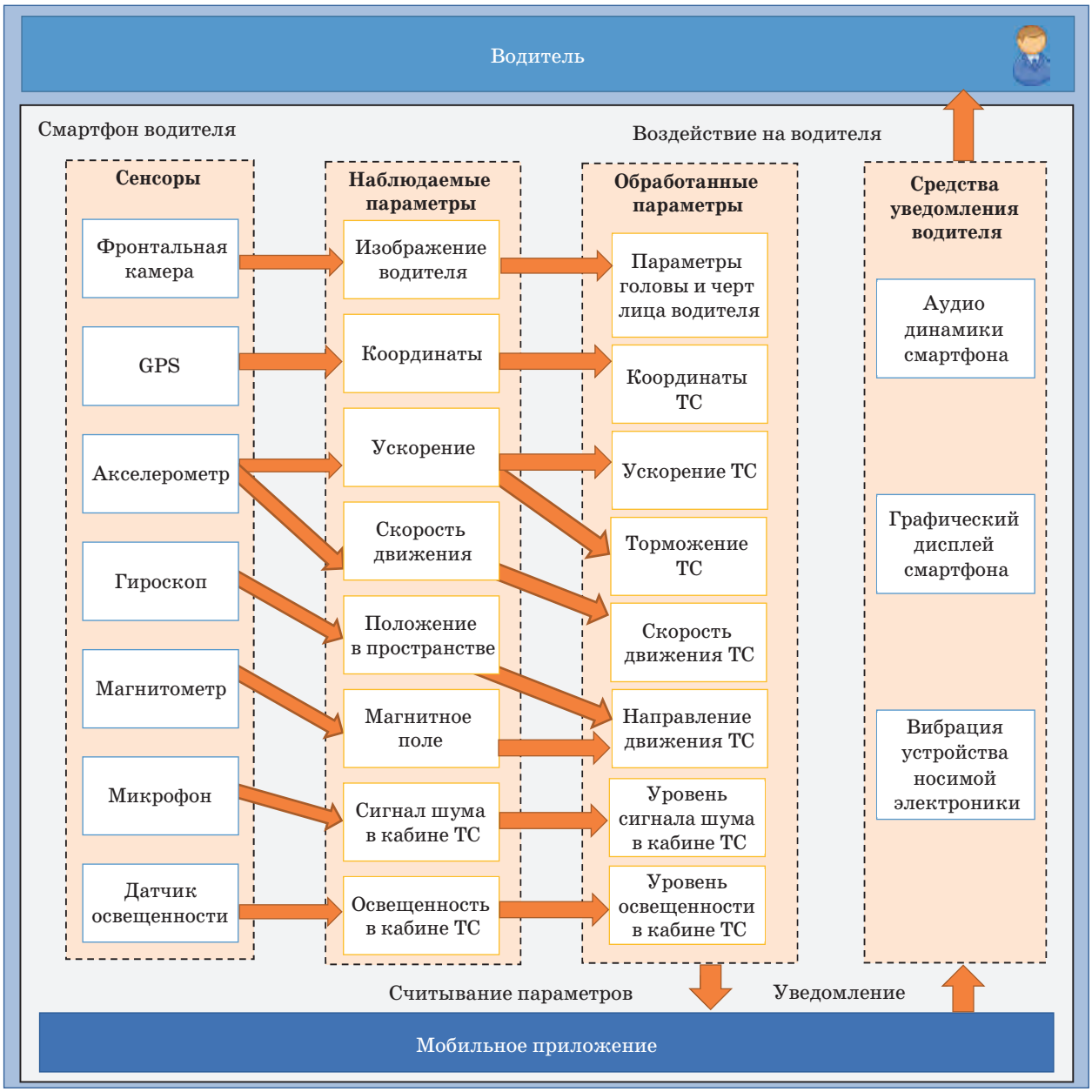
С целью общего представления о взаимодействии водителя с системой предупреждения аварийных ситуаций предложена схема потоков данных, считываемых с сенсоров смартфона и применяемых в дальнейшем при анализе и оценке навыков управления ТС (рис. 1). Исходными данными для системы предупреждения аварийных ситуаций являются показания встроенных в смартфон сенсоров, таких как фронтальная камера, GPS, акселерометр, гироскоп, магнитометр, микрофон и датчик освещенности. Получая изображение водителя с фронтальной камеры, мобильное приложение применяет последовательность программных операций для выделения лицевых характеристик водителя с целью определения его опасного поведения.

Изменения в скорости, ускорении или торможении автомобиля, вычисляемые GPS и акселерометром на основе спутниковых данных и ускорения силы тяжести, позволяют описать поведение водителя при управлении ТС. Положение в пространстве смартфона и направление движения ТС, определяемые гироскопом и магнитометром, позволяют точнее охарактеризовать то или иное поведение водителя. Микрофон, измеряющий уровень сигнала шума в кабине ТС, применяется при генерации рекомендаций водителю для определения ситуаций, когда водитель едет один или с пассажирами. Датчик освещенности смартфона используется при выявлении темного времени суток, когда обработка изображений водителя с фронтальной камеры становится нецелесообразной ввиду невозможности распознавания лицевых характеристик. На основе информации с сенсоров смартфона происходит уведомление водителя о наступлении опасного состояния при помощи рекомендаций, поступающих водителю через графический дисплей, аудио динамики смартфона или вибрации устройства носимой электроники.

Информационная модель профиля водителя

Разработанная информационная модель профиля водителя распределенной системы предупреждения аварийных ситуаций (РСПАС) при управлении ТС показана на рис. 2. Анализ поведения водителя в кабине ТС позволил выделить следующую информацию: общая информация о водителе, его контекст, компетенции водителя и история его взаимодействия с системой.

Раздел *Общая информация о водителе* включает в себя уникальный идентификатор водите-



■ ***Рис. 1.*** Схема объединения сенсоров на основе использования смартфона
■ ***Fig. 1.*** Scheme of combining sensors fusion based on the use of a smartphone

ля, имя, фамилию, пол, дату рождения, семейное положение, стаж вождения, номер мобильного телефона. Данная информация является базовой и первоначально характеризует водителя.

Раздел *Контекст* описывает информацию, характеризующую окружающую обстановку, в которой находятся водитель и ТС, и состоит из двух подразделов *Контекст водителя* и *Контекст ТС*.

Раздел *Контекст водителя* содержит информацию, которая изменяется в зависимости от текущей ситуации в кабине ТС и дорожной обстановки и включает в себя следующие атрибуты:

1) параметры калибровки системы — позволяют лучше подстроиться под конкретного водителя. Они описывают сведения о водителе (например, наклон головы водителя влево/вправо, вперед/назад по отношению к туловищу), настройки и возможности его смартфона (например, присутствие того или иного датчика, необходимого для функционирования РСПАС, уровень громкости предупреждений об опасной ситуации) и транспортного средства (легковой или грузовой автомобиль);

2) используемое оборудование — описывает технические характеристики смартфона, используемого водителем ТС;

*Атрибуты «Контекст» и «Компетенции водителя» соответствуют категориям «Контекст» и «Компетенции водителя» соответственно

■ ***Рис. 2.*** Информационная модель профиля водителя
■ ***Fig. 2.*** Information model of the driver profile

3) используемое программное обеспечение — характеризует программный комплекс, устанавливаемый и настраиваемый на смартфоне водителя и предназначенный для прогнозирования вероятности наступления аварийных ситуаций;

4) психофизиологические особенности водителя — описывает состояние водителя в текущий момент времени, характеризуя его скорость реакции, точность и последовательность действий;

5) симптомы небезопасного поведения — содержит информацию о выявленном небезопасном поведении водителя за рулем автомобиля с целью дальнейшей выработки рекомендаций для предотвращения наступления аварийной ситуации, например, количество часов непрерывного управления ТС;

6) состояние обстановки внутри салона кабины ТС — описывает различные параметры и условия обстановки в кабине ТС, например, уровень освещенности, уровень сигнала шума.

Раздел *Контекст ТС* включает в себя следующие атрибуты:

1) местоположение — содержит информацию о текущем географическом местоположении ТС и близлежащих мест отдыха, используемых при формировании рекомендаций. Сведения о местах отдыха описывают придорожные остановки (кафе, отели, автомобильные заправочные станции), которыми водитель может воспользоваться при наступлении усталости или ослабленного внимания и сделать небольшой отдых в 20–30 минут, выпив тонизирующий напиток, или же воспользоваться полноценным 7–8-часовым сном, восстановив силы перед следующей поездкой;

2) характеристики движения — содержит информацию о направлении и скорости движения водителя в конкретный момент времени;

3) дорожные условия — описывает время в пути, оставшееся время до конечного пункта назначения, текущее время суток и отражает текущую степень загруженности автомобильных дорог.

Раздел *История взаимодействия* включает историю работы водителя с системой предупреждения аварийных ситуаций:

1) статистика использования системы (протоколирование действий водителя), собираемая в основном неявным образом;

2) контекст водителя;

3) компетенции водителя.

Раздел *Компетенции водителя* описывает способность и стремление водителя к действиям на основе его личностных параметров, знаний, навыков и включает в себя:

1) соблюдение правил дорожного движения;

2) паттерны поведения — описывают успешность, последовательность и время выполнения тех или иных действий конкретного водителя в каждый момент времени на основе использования считываемых сенсорных данных и информации, определяющей профиль водителя. Примером паттерна поведения водителя может служить ситуация, описывающая процесс торможения ТС перед красным сигналом светофора на некотором промежутке времени. В данном случае паттерн поведения водителя включает почти неизменное движение ТС в направлении прямо, регистрируемое гироскопом и магнитометром, снижение скорости движения ТС, регистрируемого сенсором GPS, отрицательные значения, считываемые при помощи акселерометра смартфона и свидетельствующие о торможении ТС и, наконец, прекращение движения ТС и установление его скорости, равной нулю;

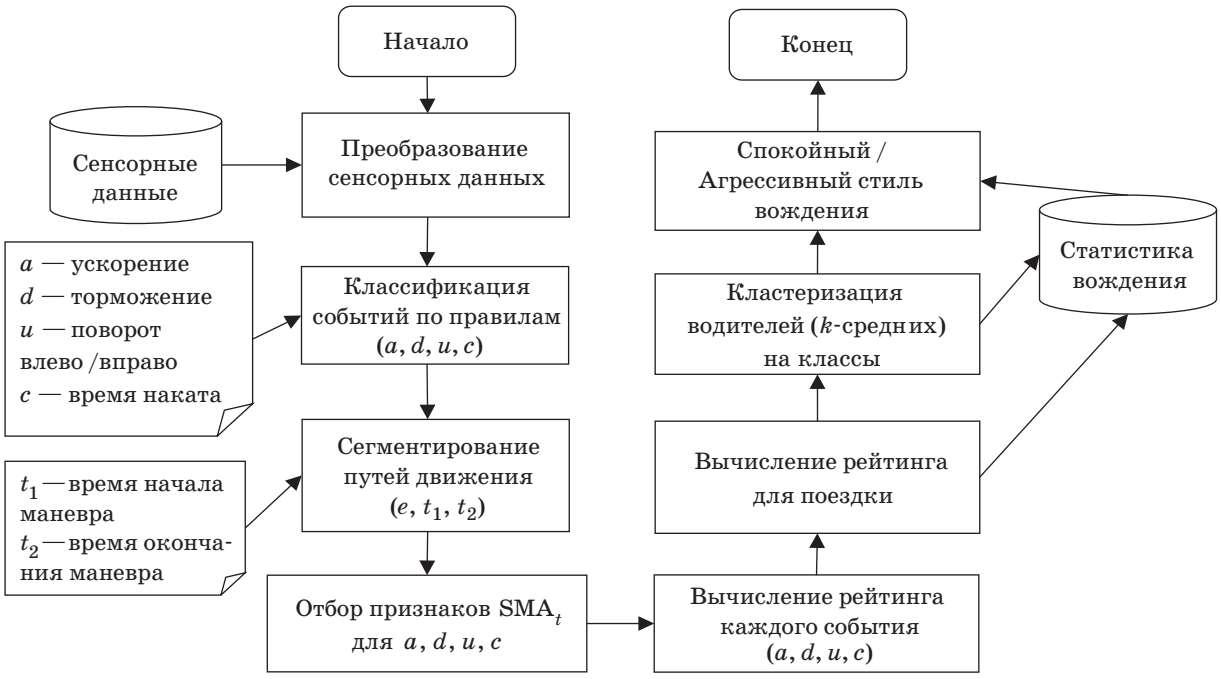
3) группы водителей — включают в себя выявляемую информацию о принадлежности к той или иной группе, к которой система отнесла данного водителя.

Раздел *Компетенции водителей* заполняется системой автоматически на основе истории работы с системой, паттернов поведения и стиля вождения водителя. Формализация компетенций позволяет осуществить автоматическую кластеризацию водителей посредством формирования поведенческих профилей с последующим выявлением групп водителей со схожими характеристиками в управлении ТС (паттерны поведения, стиль вождения). Данная информация используется для дальнейшей синхронизации с облачным сервисом и офлайн-анализа действий водителя и обновления раздела в автоматическом режиме. С целью выделения групп водителей со схожими характеристиками на основе информации об их поведении за рулем ТС используется метод классификации их профилей.

Алгоритм оценки стиля вождения водителя ТС

Информация о паттернах поведения водителя в кабине ТС оказывает большое влияние на распознавание опасных состояний и генерацию рекомендаций, ориентированных на помощь водителю для предотвращения наступления аварийной ситуации. Для решения проблемы неполноты знаний о водителе предлагается решить задачу реализации обучения системы для каждого водителя путем применения методов машинного обучения, работающих чаще всего с численными значениями. Данные методы должны быть предназначены для решения задач распознавания образов на основе использования методов кластеризации и идентификации объектов, характеризующихся конечным набором некоторых признаков. В результате чего строится модель, предсказывающая недостающие атрибуты (признаки) поведения водителя в конкретный момент времени.

В общем случае алгоритм оценки стиля вождения водителя ТС (рис. 3) включает в себя



■ ***Рис. 3.*** Общая схема алгоритма определения стиля вождения водителя

■ ***Fig. 3.*** General scheme of the algorithm for detection of the driving style

решение следующих задач: выделение характерных признаков в поведении водителя в кабине ТС, считывание и преобразование сенсорных данных, классификация событий по правилам, сегментирование путей движения, отбор признаков SMA (Simple Moving Average — простое скользящее среднее) для данных, кластеризация профилей водителей на группы, соотнесение водителя с группой.

Одной из первостепенных задач предлагаемого алгоритма является построение связей между наблюдаемыми характеристиками поведения водителя и обрабатываемыми измерениями контекста водителя и ТС. Входными данными для такого алгоритма являются сенсорные показатели с датчиков смартфона, включающие акселерометр, гироскоп, GPS и магнитометр. Такая связь формируется за счет отображения параметров поведения водителя, к которым относятся:

— лицевые характеристики: угол поворота головы влево и вправо, наклон вперед и назад, продолжительность и частота моргания век, PERCLOS;

— характеристики управления ТС: скорость движения, ускорение и торможение ТС на конкретном дорожном участке;

— список грубых нарушений ПДД: вождение в алкогольном опьянении, превышение скоростного режима, использование мобильного телефона во время движения.

Полученные параметры поведения водителей формируют совокупность атрибутов (дискриминантных признаков), характеризующих стиль вождения того или иного водителя в отличие от стиля вождения остальных участников системы. Данные с сенсоров смартфона классифицируются следующим образом:

— бинарные, ответом на который является «да» или «нет» (1 или 0): пол водителя, усталость, ослабленное внимание;

— порядковые, ответом на который является конкретный класс: степень усталости, степень ослабленного внимания, уровень громкости звуковых сигналов смартфона, уровень освещенности в кабине ТС;

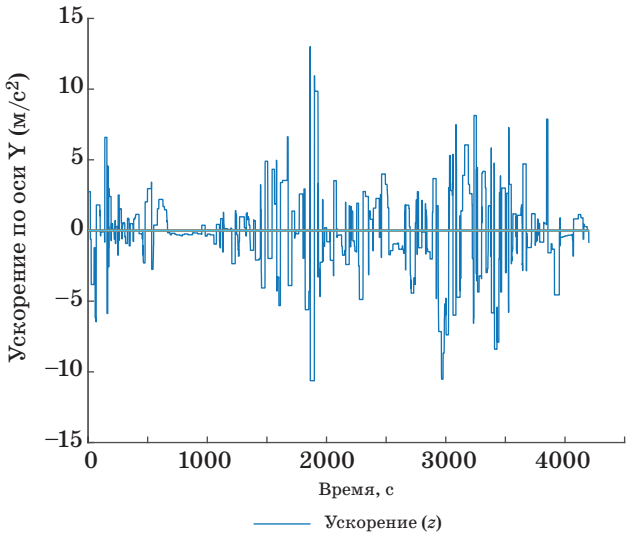
— количественные, ответом на который является число, характеризующее конкретную меру: возраст водителя, стаж вождения, PERCLOS, продолжительность моргания век, открытость рта, угол наклона и поворота головы.

Примером входных данных являются данные с акселерометра смартфона по оси Z, характеризующие продольное ускорение текущего ТС (рис. 4). Положительные значения продольного ускорения позволяют охарактеризовать ускорение ТС, а отрицательные значения — торможение ТС.

Далее, атрибуты и события, выявленные в результате работы алгоритма на первом этапе, проходят процесс фильтрации и предобработки «выбивающихся» значений в рамках того или иного признака. Для решения этой задачи в качестве первого шага используется алгоритм sliding window [30] (рис. 5), позволяющий выбирать информацию за определенный промежуток времени из непрерывного потока данных таким образом, что отфильтровываются наиболее устаревшие наборы таких данных.

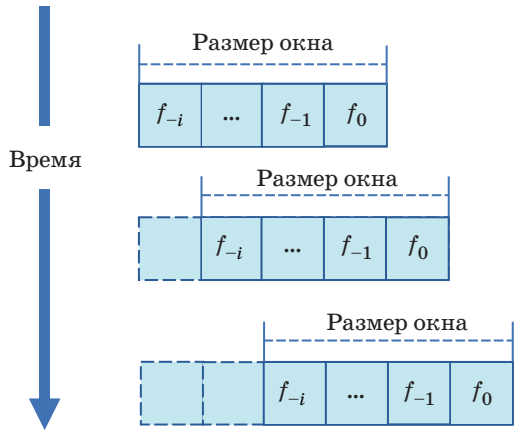
На втором шаге применяется фильтр нижних частот (LPF, low-pass filter) [31], пропускающий спектр сигнала ниже некоторой частоты и подавляющий частоты сигнала выше этой частоты. Алгоритм фильтра LPF описывается следующим образом:

$$\text{output}[i] \rightarrow \text{output}[i] + \alpha^*(\text{input}[i] - \text{output}[i]), \quad (1)$$



■ ***Рис. 4.*** Продольное ускорение водителя на основе полученных данных с акселерометра

■ ***Fig. 4.*** Longitudinal acceleration of the driver based on the data from the accelerometer



■ ***Рис. 5.*** Схема алгоритма sliding window

■ ***Fig. 5.*** Scheme of the sliding window algorithm

где α — коэффициент сглаживания фильтра; input — массив входных значений; output — массив выходных значений.

Атрибуты формируются в виде множества признаков для каждого i водителя, где x — атрибут; Y — категория, показывающая вид распознанного опасного состояния у водителя:

$$X^{(i)} = \{x^{(i)}_1, x^{(i)}_2, ..., x^{(i)}_n\} \text{ и класс } Y^{(i)}.$$

На втором этапе работы алгоритма производятся классификация и отбор событий (ускорение (*a*), торможение (*d*), поворот руля (*u*), время наката (*c*)) по их типу на основе заранее определенных правил:

— $a \rightarrow \text{if } (a_{LN} > 0{,}1) \cap (v > 0)$,
— $d \rightarrow \text{if } (a_{LN} < 0) \cap (v > 0)$,
— $u \rightarrow \text{if } (a_{Lt} > 0) \cap (a_{Lt} > a_{LN}) \cap (v > 0)$,
— $c \rightarrow \text{if } (v = 0)$,

где a_{LN} — продольное ускорение ТС; a_{Lt} — поперечное ускорение ТС; v — скорость ТС.

Как только каждое событие было идентифицировано, на следующем этапе происходит свертка событий одинакового типа и формирование набора $\{e, t_0, t_1, l\}$ для каждого нового объекта, где e — тип события; t_0 — время начала события (мс, прошедшие с полуночи 01.01.1970 по UTC); t_1 — время окончания события (мс, прошедшие с полуночи 01.01.1970 по UTC); l — географические координаты события (долгота, широта и высота). Данная операция обрабатывает число событий, полученное в результате разделения времени поездки на интервалы в 10 секунд.

На следующем шаге алгоритма происходит вычисление SMA-признаков для событий $\{a, d, u, c\}$ с целью нахождения среднего арифметического значений событий на интервале по следующей формуле:

$$SMA_t = \frac{1}{n}\sum_{i=0}^{n-1} e_{t-i}, \quad (2)$$

где SMA_t — значение простого скользящего среднего в точке t; n — количество значений исходной функции для расчета скользящего среднего; e_{t-i} — значение исходной функции в точке $t - i$.

Порог чувствительности для каждого из событий (a_z — продольное ускорение, a_x — поперечное ускорение), указанный в таблице, позволяет охарактеризовать то или иное событие по степени резкости совершенного водителем маневра.

Далее происходит вычисление рейтинга для каждого события по формуле:

$$S_{it} = \frac{1}{DT_{it}}\sum_{i=1}^{n}\sum_{s=0}^{m} \beta_{js} N_{ijst}, \quad (3)$$

где S_{it} — рейтинг события за период времени t; DT — время вождения за весь период; N_{ijst} — количество маневров типа j и опасности (резкости) s, которые водитель совершил; β_{js} — веса (коэффициенты) различных маневров. Вычисление рейтинга для всей совершенной поездки производится по следующей формуле:

$$\bar{S} = \frac{1}{n}\sum_{i=0}^{n} x_i, \quad (4)$$

где $\bar{S}$ — среднее значение рейтингов всех событий; n — количество значений исходной функции для расчета скользящего среднего; x_i — значение рейтинга функции в точке i.

На последнем этапе работы алгоритма производится оценка стиля вождения водителя (в интервале значений [0; 1], где 0 — спокойный стиль вождения, а 1 — агрессивный) на основе препроцессинга агрегированных атрибутов (признаков). Процессинг представляет собой отображение данных (пол, возраст, стаж вождения водителя, тип кузова ТС) в формат, пригодный для обучения модели. Можно выделить две основные операции, производимые на этапе препроцессинга:

1) Создание векторного пространства признаков, где будут расположены примеры обучающей выборки. По сути, это процесс приведения всех данных в числовую форму, позволяющий отойти от категорийных, булевых и остальных не числовых типов.

2) Процесс нормализации данных, в результате которого необходимо изменить шкалу измерений таким образом, чтобы среднее значение каждого признака по всем данным было равно нулю,

■ Порог чувствительности событий
■ Threshold for the intensity of the events

Тип события	Порог чувствительности (м/с2)		
	Низкий	Средний	Высокий
Ускорение ТС	$0{,}1 < a_z < 0{,}5$	$0{,}5 < a_z < 1{,}5$	$a_z > 1{,}5$
Торможение ТС	$-0{,}1 < a_z < -0{,}5$	$-0{,}5 < a_z < -1{,}5$	$a_z < -1{,}5$
Поворот ТС влево/вправо	$0{,}1 < \lvert a_x \rvert < 0{,}5$	$0{,}5 < \lvert a_x \rvert < 1{,}5$	$\lvert a_x \rvert > 1{,}5$

а стандартное отклонение — одному. В общем виде нормализация данных выглядит следующим образом:

$$X = (X - \mu)/\sigma. \quad (5)$$

С целью разбиения всех водителей системы на конечное число групп (кластеров), характеризующих степень схожести водителей по стилю вождения, решим задачу кластеризации. Входными данными для метода кластеризации выступает сформированный на основе накопленных сведений профиль водителя системы, а в качестве класса — стиль вождения водителя (в интервале множества значений [0; 1] от спокойного до агрессивного).

В настоящее время кластеризация является одной из основополагающих задач в области анализа данных и Data Mining. Существует большое количество методов кластеризации, которые можно условно разделить на следующие основные группы: применяющие вероятностный подход (*K*-means, EM-алгоритм), методы искусственного интеллекта (генетические алгоритмы, нейронные сети), использующие иерархические алгоритмы и теоретико-графовые модели. С целью достижения высокой скорости обработки данных, наглядности и простоты реализации метода для выявления групп водителей со схожими компетенциями (User-User, коллаборативная фильтрация) выбран метод кластеризации *k*-средних (*K*-means) [32]. Ключевыми входными параметрами для данного алгоритма являются *n*, обозначающее исходное число объектов (наблюдений) для обработки, и *k* — число кластеров (групп) наблюдения, содержащих похожие элементы внутри группы и непохожие на элементы, не принадлежащие к данной группе. При использовании данного метода возникает необходимость в определении числа кластеров для разбиения. Предварительный анализ исходных данных позволяет найти оптимальное число кластеров для разбиения за счет минимизации суммы внутрикластерных расстояний. Исходя из вышеперечисленного, для кластеризации профилей водителей был выбран алгоритм *k*-means.

Разобьем множество объектов x (всех участников системы РСПАС) на конечное число непересекающихся классов $S_1 \ldots S_m$, каждый из которых соответствует определенному стилю вождения водителя

$$X = \{x_1, \ldots, x_n\} — \text{на } M \text{ непересекающихся классов.} \quad (6)$$

Суть алгоритма *k*-means заключается в том, что он стремится минимизировать суммарное квадратичное отклонение точек кластеров от центров этих кластеров:

$$S = \sum_{j=1}^{k} \sum_{x \in S_i}^{m} (x - \mu_i)^2, \quad (7)$$

где k — число кластеров; S_i — полученные кластеры; μ_i — центры масс всех S_i объектов x из кластера S_i. Последним этапом работы алгоритма кластеризации является сопоставление каждого водителя системы РСПАС с определенным номером кластера, или, иначе говоря, за каждым водителем закрепляется метка кластера, характеризующая группу водителей со схожим стилем вождения ТС. Стоит отметить, что алгоритм кластеризации профилей водителей периодически запускается повторно при изменении числа водителей или поступлении новой информации в облачный сервис о статистике их вождения ТС.

В дальнейшем результаты работы кластеризации используются системой при оценке паттернов поведения того или иного водителя во время вождения. Оценка паттернов поведения водителя, определенная заранее для некоторого контекста, помогает скорректировать мониторинг опасного поведения водителя в кабине ТС и тем самым точнее сформировать контекстно-ориентированные рекомендации, необходимые для принятия мер по предотвращению аварийной ситуации. Так, например, алгоритм выявления опасных состояний в поведении для водителя с более агрессивным стилем вождения работает с видоизмененными входными параметрами (время распознавания, доля опасных состояний к общему числу событий и т. п.).

Реализация алгоритма анализа стиля вождения водителей

Алгоритм анализа поведения водителя, основанный на обработке данных сенсоров смартфона, кластеризации профилей водителей и анализе статистики взаимодействия водителей с системой, реализован с использованием языка программирования Python. Выбор языка Python обусловлен развитой стандартной библиотекой функций, большим множеством сторонних модулей и лаконичным синтаксисом.

Анализ статистики взаимодействия водителей с системой заключается в получении данных со смартфона водителя, их анализ, обработку и фильтрацию с целью дальнейшей записи в базу данных *ClickHouse*. В выбранной базе данных заранее создана таблица *pure_stat_events*, содержащая данные об опасных ситуациях, и таблица *pure_critical_event*, содержащая данные об опасных состояниях для каждого водителя системы РСПАС. Задачей базы данных *ClickHouse* является предоставление хранящихся данных для модуля кластеризации профилей водителей

с целью дальнейшего определения стиля вождения каждого из них. Запрос выборки данных (см. Листинг) позволяет получить необходимые данные для кластеризации профилей водителей.

Листинг. Запрос выборки данных для кластеризации

```
SELECT * FROM ` pure_critical_event` p
INNER JOIN ` pure_stat_events` ps ON
ps.id = p.event_id
```

Заключение

Разработан подход к анализу поведения водителя ТС с использованием информационной модели профиля водителя и алгоритма анализа стиля вождения водителя, который основан на использовании данных с сенсоров смартфона, формировании связей между изначальными данными и характеристиками управления ТС и истории взаимодействия водителя с системой предупреждения аварийных ситуаций. Такой подход позволяет не только обратить внимание водителя на навыки управления и характеристику езды, но и снизить процент неосторожного вождения, повысить эффективность эксплуатации ТС и таким образом снизить количество дорожно-транспортных происшествий на дорогах общего пользования.

Исследования выполнены при финансовой поддержке РНФ (проект № 18-71-10065).

Литература

1. **Abraham H., McAnulty H., Mehler B., Reimer B.** Case study of today's automotive dealerships: introduction and delivery of advanced driver assistance systems. *Transportation Research Record: Journal of the Transportation Research Board*, 2017, vol. 2671, pp. 1–17. doi:10.3141/2660-02
2. **Botzer A., Musicant O., Perry A.** Driver behavior with a smartphone collision warning application — a field study. *Safety Science*, 2017, vol. 91, pp. 361–372. doi:10.1016/j.ssci.2016.09.003
3. **Schofield K., O'brien F., Bingle R. L., Lynam N. R.** Cabin monitoring system for a vehicle. Patent US20170237946A1, 2001.
4. **Choi M., Koo G., Seo M., Kim S. W.** Wearable device-based system to monitor a driver's stress, fatigue, and drowsiness. *IEEE Transactions on Instrumentation and Measurement,* 2018, vol. 67, no. 3, pp. 634–645. doi:10.1109/TIM.2017.2779329
5. **Лашков И. Б.** Анализ поведения водителя при управлении транспортным средством с использованием фронтальной камеры смартфона. Информационно-управляющие системы. 2017. № 4(89). С. 7–17. doi:10.15217/issnl684-8853.2017.4.7
6. **Smirnov A., Kashevnik A., Lashkov I.** Human-smartphone interaction for dangerous situation detection & recommendation generation while driving. *Speech and Computer, LNCS,* 2016, vol. 9811, pp. 346–353. doi:10.1007/978-3-319-43958-7_41
7. **Smirnov A., Kashevnik A., Lashkov I., Baraniuc O., Parfenov V.** Smartphone-based identification of dangerous driving situations: algorithms and implementation. *Proc. of the 18th Conference of Open Innovations Association FRUCT*, 2016, pp. 306–313. doi:10.1109/FRUCT-ISPIT.2016.7561543
8. **Witten I. W., Frank E., Hall M. A., Pal C. J.** Data mining: practical machine learning tools and techniques. San Francisco, Morgan Kaufmann, 2016. 654 p.
9. **Chen Y., Wu J.** Distribution patterns of energy consumed in classified public buildings through the data mining process. *Applied Energy,* 2018, vol. 226, pp. 240–251. doi:10.1016/j.apenergy.2018.05.123
10. **Si G., Zheng K., Zhou Z., Pan C., Zhang Y.** Three-dimensional piecewise cloud representation for time series data mining. *Neurocomputing,* 2018, vol. 316, pp. 78–94. doi:10.1016/j.neucom.2018.07.053
11. **El Mohadab M., Bouikhalene B., Safi S.** Automatic CV processing for scientific research using data mining algorithm. *Journal of King Saud University — Computer and Information Sciences.* 2018, vol. 30, pp. 1–7. doi:10.1016/j.jksuci.2018.07.002
12. **Silva N., Soares J., Shah V., Yasmina Santos M., Rodrigues H.** Anomaly detection in roads with a data mining approach. *Procedia Computer Science*, 2017, vol. 121, pp. 415–422. doi:10.1016/j.procs.2017.11.056
13. **Николенко С. И., Тулупьев А. Л.** Самообучающиеся системы. М.: МНЦМО, 2009. 288 с.
14. **James S. C., Zhang Y., O'Donncha F.** A machine learning framework to forecast wave conditions. *Coastal Engineering*, 2018, vol. 137, pp. 1–10. doi:10.1016/j.coastaleng.2018.03.004
15. **Cai J., Luo J., Wang S., Yang S.** Feature selection in machine learning: a new perspective. *Neurocomputing*, 2018, vol. 300, pp. 70–79. doi:10.1016/j.neucom.2017.11.077
16. **Aksjonov A., Nedoma P., Vodovozov V., Petlenkov E., Herrmann M.** A Novel driver performance model based on machine learning. *IFAC-PapersOnLine*, 2018, vol. 51, iss. 9, pp. 267–272. doi:10.1016/j.ifacol.2018.07.044
17. **Zou X., Long Yue W., Le Vu H.** Visualization and analysis of mapping knowledge domain of road safety studies. *Accident Analysis & Prevention*, 2018, vol. 118, pp. 131–145. doi:10.1016/j.aap.2018.06.010
18. **Bishop C. M.** Pattern recognition and machine learning (information science and statistics): Springer-Verlag Berlin, Heidelberg, 2006. 738 p.

19. **Škrjanc I., Andonovski G., Ledezma Espino A., Sipele O., Iglesias J., Sanchis de Miguel A.** Evolving cloud-based system for the recognition of drivers' actions. *Expert Systems with Applications*, 2018, vol. 99, pp. 231–238. doi:10.1016/j.eswa.2017.11.008
20. **Verwer S., de Weerdt M., Witteveen C.** Learning driving behavior by timed syntactic pattern recognition. *Proceedings of the 22 International Joint Conference on Artificial Intelligence*, 2011, pp. 1529–1534. doi:10.5591/978-1-57735-516-8/IJCAI11-257
21. **Johnson D. A., Trivedi M. M.** Driving style recognition using a smartphone as a sensor platform. *Intelligent Transportation Systems (ITSC) 14th International IEEE Conference*, 2011, pp. 1609–1615. doi:10.1109/ITSC.2011.6083078
22. **Van Ly M., Martin S., Trivedi M. M.** Driver classification and driving style recognition using inertial sensors. *IEEE Intelligent Vehicles Symposium (IV)*, 2013, pp. 1040–1045. doi:10.1109/IVS.2013.6629603
23. **Rolim C., Baptista P.** Comparing drivers' self-perception on driving behaviour changes with real world driving performance data: Lisbon case-study. *Travel Behaviour and Society*, 2018, vol. 11, pp. 86–92. doi:10.1016/j.tbs.2018.02.002
24. **Hong J. H., Margines B., Dey A. K.** A smartphone-based sensing platform to model aggressive driving behaviors. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '14). ACM, New York, NY, USA. 2014. pp. 4047–4056. doi:10.1145/2556288.2557321
25. **Yang L., Ma R., Zhang H. M., Guan W., Jiang S.** Driving behavior recognition using EEG data from a simulated car-following experiment. *Accident Analysis & Prevention*, vol. 116, 2018, pp. 30–40. doi:10.1016/j.aap.2017.11.010
26. **Li G., Eben Li S., Cheng B., Green P.** Estimation of driving style in naturalistic highway traffic using maneuver transition probabilities. *Transportation Research Part C: Emerging Technologies,* 2017, vol. 74, pp. 113–125. doi:10.1016/j.trc.2016.11.011
27. **He H., Sun C., Zhang X.** A Method for identification of driving patterns in hybrid electric vehicles based on a LVQ neural network. *Energies*, 2012, pp. 3363–3380. doi:10.3390/en5093363
28. **Zhang Y., Lin W. C., Chin Y.-K. S.** A Pattern-recognition approach for driving skill characterization. *Intelligent Transportation Systems, IEEE Transactions*, 2010, vol. 4(11), pp. 905–916. doi:10.1109/TITS.2010.2055239
29. **Han W., Wang W., Li X., Xi J.** Statistical-based approach for driving style recognition using Bayesian probability with kernel density estimation. *IET Intelligent Transport Systems*, 2018, pp. 1–8. doi:10.1049/iet-its.2017.0379
30. **Halim Mohd Noor M., Salcic Z., I-Kai Wang K.** Adaptive sliding window segmentation for physical activity recognition using a single tri-axial accelerometer. *Pervasive and Mobile Computing*, 2017, vol. 38, pp. 41–59. doi:10.1016/j.pmcj.2016.09.009
31. **Hammed R. T., Hassan S. H., Ajeel S. L.** New compact low-pass filter (LPF) using cascaded square open loop resonator AEU. *International Journal of Electronics and Communications,* 2018, vol. 92, pp. 93–97. doi:10.1016/j.aeue.2018.05.030
32. **Ismkhan H.** I-k-means−+: An iterative clustering algorithm based on an enhanced version of the k-means. *Pattern Recognition*. 2018, vol. 79, pp. 402–413. doi:10.1016/j.patcog.2018.02.015

UDC 004
doi:10.31799/1684-8853-2018-5-2-12

Smartphone-based approach to determining driving style with on-board sensors

I. B. Lashkov[a, b], Researcher, orcid.org/0000-0001-6418-4660, igor-lashkov@ya.ru
[a]ITMO University, 49, Kronverkskii Pr., 197101, Saint-Petersburg, Russian Federation
[b]Saint-Petersburg Institute for Informatics and Automation of the RAS, 39, 14 Line, V. O., 199178, Saint-Petersburg, Russian Federation

Introduction: Driving behavior analysis and driving style evaluation allow you to pay the drivers' attention to their driving skills and characteristics, reduce reckless driving rate, increase the vehicle usage efficiency and thereby reduce the number of traffic accidents. However, the available driver assistance systems lack personalization, which in its turn decreases the efficiency of these systems and limits the scope of their usage. **Purpose:** Developing an approach to analysis and evaluation of a driving style, based on smartphone features such as a front-facing camera and sensors, and focused on its usage in driver safety systems. **Results:** The article discusses a scheme of combining data from smartphone sensors for a driver assistant system. It presents an information model of a driver's profile. An algorithm has been developed for the analysis of an individual driving style in order to to improve the quality of recognizing dangerous states in one's driving behavior, taking into account the previous experience of using the system. This algorithm allows drivers to evaluate their driving style and thereby to pay more attention to their driving skills. A prototype has been developed for a system evaluating a driving style based on the data from smartphone sensors related to the vehicle operation parameters, taking into account the history of driver's interaction with the assistant system. **Practical relevance:** The obtained results allow you to increase the accuracy of recognizing dangerous states and, thereby, improve the efficiency of the driver assistance system. They also can be used for generating and displaying driving reports for insurance agents, fleet managers and logistic managers who observe and monitor the driving statistics.

Keywords — advanced driver assistance systems, driver, vehicle, driving behavior, driving style.

Citation: Lashkov I. B. Smartphone-based approach to determining driving style with on-board sensors. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 5, pp. 2–12 (In Russian). doi:10.31799/1684-8853-2018-5-2–12

References

1. Abraham H., McAnulty H., Mehler B., Reimer B. Case study of today's automotive dealerships: introduction and delivery of advanced driver assistance systems. *Transportation Research Record: Journal of the Transportation Research Board*, 2017, vol. 2671, pp. 1–17. doi:10.3141/2660-02
2. Botzer A., Musicant O., Perry A. Driver behavior with a smartphone collision warning application — a field study. *Safety Science*, 2017, vol. 91, pp. 361–372. doi:10.1016/j.ssci.2016.09.003
3. Schofield K., O'brien F., Bingle R. L., Lynam N. R. *Cabin monitoring system for a vehicle*. Patent US20170237946A1, 2001.
4. Choi M., Koo G., Seo M., Kim S.W. Wearable device-based system to monitor a driver's stress, fatigue, and drowsiness. *IEEE Transactions on Instrumentation and Measurement,* 2018, vol. 67, no. 3, pp. 634–645. doi:10.1109/TIM.2017.2779329
5. Lashkov I. B. Driver's behavior analysis with smartphone front camera. *Informatsionno-upravliaiushchie sistemy*, 2017, 4(89), pp. 7–17 (In Russian). doi:10.15217/issnl684-8853.2017.4.7
6. Smirnov A., Kashevnik A., Lashkov I. Human-smartphone interaction for dangerous situation detection & recommendation generation while driving. *Speech and Computer, LNCS*, 2016, vol. 9811, pp. 346–353. doi:10.1007/978-3-319-43958-7_41
7. Smirnov A., Kashevnik A., Lashkov I., Baraniuc O., Parfenov V. Smartphone-based identification of dangerous driving situations: algorithms and implementation. *Proc. of the 18th Conference of Open Innovations Association FRUCT*, 2016, pp. 306–313. doi:10.1109/FRUCT-ISPIT.2016.7561543
8. Witten I. W., Frank E., Hall M. A., Pal C. J. *Data mining: practical machine learning tools and techniques*. San Francisco, Morgan Kaufmann, 2016. 654 p.
9. Chen Y., Wu J. Distribution patterns of energy consumed in classified public buildings through the data mining process. *Applied Energy,* 2018, vol. 226, pp. 240–251. doi:10.1016/j.apenergy.2018.05.123
10. Si G., Zheng K., Zhou Z., Pan C., Zhang Y. Three-dimensional piecewise cloud representation for time series data mining. *Neurocomputing*, 2018, vol. 316, pp. 78–94. doi:10.1016/j.neucom.2018.07.053
11. El Mohadab M., Bouikhalene B., Safi S. Automatic CV processing for scientific research using data mining algorithm. *Journal of King Saud University — Computer and Information Sciences.* 2018, vol. 30, pp. 1–7. doi:10.1016/j.jksuci.2018.07.002
12. Silva N., Soares J., Shah V., Yasmina Santos M., Rodrigues H. Anomaly detection in roads with a data mining approach. *Procedia Computer Science*, 2017, vol. 121, pp. 415–422. doi:10.1016/j.procs.2017.11.056
13. Nikolenko S. I., Tulupiev A. L. *Samoobuchajushhiesia sistemy* [Self-learning systems]. Moscow, MNWMO Publ., 2009, 288 p. (In Russian).
14. James S. C., Zhang Y., O'Donncha F. A machine learning framework to forecast wave conditions. *Coastal Engineering*, 2018, vol. 137, pp. 1–10. doi:10.1016/j.coastaleng.2018.03.004
15. Cai J., Luo J., Wang S., Yang S. Feature selection in machine learning: a new perspective. *Neurocomputing*, 2018, vol. 300, pp. 70–79. doi:10.1016/j.neucom.2017.11.077
16. Aksjonov A., Nedoma P., Vodovozov V., Petlenkov E., Herrmann M. A Novel driver performance model based on machine learning. *IFAC-PapersOnLine*, 2018, vol. 51, iss. 9, pp. 267–272. doi:10.1016/j.ifacol.2018.07.044
17. Zou X., Long Yue W., Le Vu H. Visualization and analysis of mapping knowledge domain of road safety studies. *Accident Analysis & Prevention*, 2018, vol. 118, pp. 131–145. doi:10.1016/j.aap.2018.06.010
18. Bishop C. M. *Pattern recognition and machine learning* (information science and statistics): Springer-Verlag Berlin, Heidelberg, 2006, 738 p.
19. Škrjanc I., Andonovski G., Ledezma Espino A., Sipele O., Iglesias J., Sanchis de Miguel A. Evolving cloud-based system for the recognition of drivers' actions. *Expert Systems with Applications*, 2018, vol. 99, pp. 231–238. doi:10.1016/j.eswa.2017.11.008
20. Verwer S., de Weerdt M., Witteveen C. Learning driving behavior by timed syntactic pattern recognition. *Proceedings of the 22 International Joint Conference on Artificial Intelligence*, 2011, pp. 1529–1534. doi:10.5591/978-1-57735-516-8/IJCAI11-257
21. Johnson D. A., Trivedi M. M. Driving style recognition using a smartphone as a sensor platform. *Intelligent Transportation Systems (ITSC) 14th International IEEE Conference*, 2011, pp. 1609–1615. doi:10.1109/ITSC.2011.6083078
22. Van Ly M., Martin S., Trivedi M. M. Driver classification and driving style recognition using inertial sensors. *IEEE Intelligent Vehicles Symposium (IV)*, 2013, pp. 1040–1045. doi:10.1109/IVS.2013.6629603
23. Rolim C., Baptista P. Comparing drivers' self-perception on driving behaviour changes with real world driving performance data: Lisbon case-study. *Travel Behaviour and Society*, 2018, vol. 11, pp. 86–92. doi:10.1016/j.tbs.2018.02.002
24. Hong J. H., Margines B., Dey A. K. A smartphone-based sensing platform to model aggressive driving behaviors. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14).* ACM, New York, NY, USA. 2014. pp. 4047–4056. doi:10.1145/2556288.2557321
25. Yang L., Ma R., Zhang H. M., Guan W., Jiang S. Driving behavior recognition using EEG data from a simulated car-following experiment. *Accident Analysis & Prevention*, vol. 116, 2018, pp. 30–40. doi:10.1016/j.aap.2017.11.010
26. Li G., Eben Li S., Cheng B., Green P. Estimation of driving style in naturalistic highway traffic using maneuver transition probabilities. *Transportation Research Part C: Emerging Technologies,* 2017, vol. 74, pp. 113–125. doi:10.1016/j.trc.2016.11.011
27. He H., Sun C., Zhang X. A Method for identification of driving patterns in hybrid electric vehicles based on a LVQ neural network. *Energies*, 2012, pp. 3363–3380. doi:10.3390/en5093363
28. Zhang Y., Lin W. C., Chin Y.-K. S. A Pattern-recognition approach for driving skill characterization. *Intelligent Transportation Systems, IEEE Transactions*, 2010, vol. 4(11), pp. 905 — 916. doi:10.1109/TITS.2010.2055239
29. Han W., Wang W., Li X., Xi J. Statistical-based approach for driving style recognition using Bayesian probability with kernel density estimation. *IET Intelligent Transport Systems*, 2018, pp. 1–8. doi:10.1049/iet-its.2017.0379
30. Halim Mohd Noor M., Salcic Z., I-Kai Wang K. Adaptive sliding window segmentation for physical activity recognition using a single tri-axial accelerometer. *Pervasive and Mobile Computing*, 2017, vol. 38, pp. 41–59. doi:10.1016/j.pmcj.2016.09.009
31. Hammed R. T., Hassan S. H., Ajeel S. L. New compact low-pass filter (LPF) using cascaded square open loop resonator AEU. *International Journal of Electronics and Communications,* 2018, vol. 92, pp. 93–97. doi:10.1016/j.aeue.2018.05.030
32. Ismkhan H. I-k-means−+: An iterative clustering algorithm based on an enhanced version of the k-means. *Pattern Recognition.* 2018, vol. 79, pp. 402–413. doi:10.1016/j.patcog.2018.02.015

УДК 681.5
doi:10.31799/1684-8853-2018-5-13-21

Научные статьи
Articles

Нечеткий супервизор ПИД-регулятора

М. В. Бураков[a], *канд. техн. наук, доцент, orcid.org/0000-0001-8047-9493, bmv@sknt.ru*
А. С. Коновалов[a], *доктор техн. наук, профессор, orcid.org/0000-0001-6127-6789*
[a]*Санкт-Петербургский государственный университет аэрокосмического приборостроения, Б. Морская ул., 67, Санкт-Петербург, 190000, РФ*

Постановка проблемы: *ПИД-регуляторы являются основным инструментом автоматизации производства. Однако их настройка традиционными средствами требует выполнения опытов с объектом, что снижает эффективность системы управления при изменении условий работы. Для решения этой проблемы необходимо использовать контур адаптации, автоматически изменяющий параметры регулятора при ухудшении оценок качества.* ***Цель:*** *разработка структуры и алгоритма функционирования нечеткого супервизора ПИД-регулятора для повышения качества работы в условиях неконтролируемого изменения параметров объекта.* ***Результаты:*** *предлагаются два варианта организации нечеткого супервизорного управления ПИД-регулятором. Первый вариант предполагает непрерывное изменение параметров контроллера на основании информации о текущей ошибке управления. Обучение супервизора происходит в режиме off line с помощью генетического алгоритма. Такой подход обеспечивает универсальность применения, однако полученная структура является «черным ящиком». Второй вариант использует оценки качества переходного процесса и может применяться при периодическом изменении входного сигнала. Параметры контроллера изменяются с помощью нечетких правил, имеющих ясную семантику. Показано, что системы с нечетким супервизором позволяют сократить перерегулирование и статическую ошибку, которые могут возникать при изменении параметров объекта управления. Выполнено моделирование работы системы средствами MatLab Simulink.* ***Практическая значимость:*** *применение ПИД-регуляторов с нечеткими супервизорами может оказаться полезным при проектировании систем управления широким кругом динамических объектов.*

Ключевые слова *— адаптация, ПИД-регулятор, нечеткая логика, нечеткий супервизор, моделирование, генетический алгоритм.*

Цитирование: Бураков М. В., Коновалов А. С. Нечеткий супервизор ПИД-регулятора. *Информационно-управляющие системы*, 2018, № 5, с. 13–21. doi:10.31799/1684-8853-2018-5-13-21
Citation: Burakov M. V., Konovalov A. S. Fuzzy supervisor for PID controller. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 5, pp. 13–21 (In Russian). doi:10.31799/1684-8853-2018-5-13-21

Введение

Использование ПИД-регуляторов для промышленной автоматизации продолжается уже многие годы благодаря простоте проектирования, низкой стоимости и высокой эффективности этого подхода. Традиционный вариант настройки ПИД-регулятора для конкретного объекта предполагает проведение серии опытов, в результате которых выбираются значения трех коэффициентов, входящих в описание закона управления [1].

Современная технологическая база позволяет реализовывать ПИД-регуляторы с помощью микроконтроллеров, что дает дополнительные преимущества в виде возможности перепрограммирования и оперативной коррекции закона управления [2–4].

Поскольку классический ПИД-регулятор является линейным корректирующим звеном, он не может быть эффективен при управлении объектами с существенными нелинейностями или неопределенными параметрами. Поэтому в последние годы интенсивно развиваются нелинейные модификации ПИД-регуляторов на базе нейросетевых и нечетких технологий [5–9].

Использование искусственной нейронной сети (НС) в качестве супервизора ПИД-регулятора допускает разные варианты исполнения. Например, в работе [10] применена НС прямого распространения и алгоритм обратного распространения ошибки, в [11] — НС Хопфилда, в [12] — радиально-базисная НС. Однако использование НС и нейронечетких систем связано с реализацией алгоритмов обучения, в том числе — алгоритмов эволюционной оптимизации [13, 14]. В результате полученная структура является «черным ящиком», и происходит потеря семантики, характерной для нечетких правил.

Нечеткие супервизоры ПИД-регуляторов образуют важный класс систем прямого адаптивного управления, которые не требуют применения идентификационных процедур. Нечеткие супервизоры ПИД-регуляторов позволяют придать системе управления адаптивные свойства при сохранении простой структуры закона управления. Хотя идея нечеткого супервизора возникла достаточно давно [15, 16], ее конкретные воплощения могут отличаться большим разнообразием, которое проявляется в способах получения информации от объекта управления, в обработке этой информации по нечетким правилам и

в способе коррекции коэффициентов ПИД-регулятора.

Нечеткие супервизоры в настоящее время применяют в таких областях, как электротехника [17], робототехника [18], автомобилестроение [19] и т. д.

В настоящей работе анализируются различные варианты организации нечеткого супервизорного управления. На основании сделанного анализа предлагаются два базовых варианта организации супервизорной системы, которые исследуются с помощью компьютерного моделирования.

Варианты организации нечеткого супервизора

Закон управления классического ПИД-регулятора описывается формулой

$$u(t) = k_p e(t) + k_i \int_0^t e(\tau)\mathrm{d}\tau + k_d \frac{\mathrm{d}e(t)}{\mathrm{d}t},$$

где $u(t)$ и $e(t)$ — сигнал управления и ошибка управления; k_p, k_d и k_i — настраиваемые коэффициенты регулятора.

Идея нечеткого супервизора заключается в описании коэффициентов ПИД-регулятора как функций времени, значения которых определяются нечеткими логическими правилами. В результате возникает двухуровневая система управления.

Нечеткий супервизор образует верхний уровень системы управления, на нижнем уровне которой располагается ПИД-регулятор. Концептуально супервизор является нечетким логическим регулятором (НЛР), содержащим все необходимые компоненты: блок фаззификации, базу правил и блок дефаззификации [20]. Таким образом, при проектировании супервизора необходимо выбрать его вход-выходные переменные, определить их лингвистическое описание и составить нечеткие логические правила для описания закона управления.

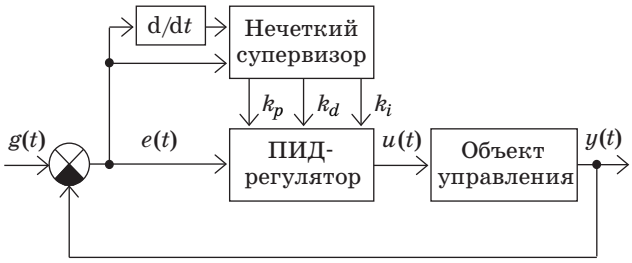
Один из вариантов работы нечеткого супервизора, рассмотренный, например, в статьях [21, 22], непрерывно использует на входе информацию о поведении ошибки управления и ее производной. Таким образом, здесь описывается определение коэффициентов регулятора в виде

$$k_p = F_1(e, \dot{e});\ k_d = F_2(e, \dot{e}),$$

где F_1, F_2 — нечеткие функции.

Структура супервизорной системы представлена на рис. 1.

Для описания каждой из входных переменных супервизора в работе [21] использованы лингвистические переменные (ЛП), имеющие по пять термов с треугольными функциями принадлежности, равномерно распределенными по нормализованным базовым шкалам. База правил состоит из трех частей — отдельно для каждого коэффициента k_p, k_d и k_i, так что всего используется 75 управляющих правил. Управляющие правила основаны на анализе свойств переходного процесса в системе. Подобный подход впервые изложен [22] для описания правил нечеткого регулятора. Например, «если ошибка управления большая и медленно изменяется, то k_p должен быть большим», иначе, «если ошибка управления малая и быстро изменяется, то k_p должен быть малым». Однако если для НЛР управляющие правила имеют ясную семантику, то для супервизора семантика большинства из 75 правил неоднозначна. Еще больше эта ситуация ухудшается при использовании семи термов для описания ошибки и ее производной [23]. Здесь количество правил увеличивается до 343, соответственно, растет неопределенность их формулировки. Кроме того, при управлении коэффициентами регулятора важную роль может играть интеграл ошибки, поэтому необходимо описать зависимость



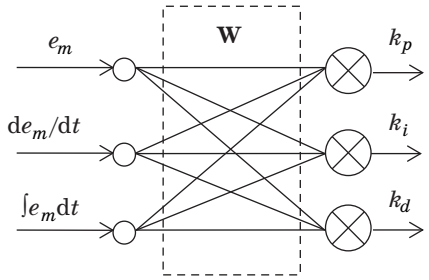
■ ***Рис. 1.*** Вариант организации нечеткого супервизора

■ ***Fig. 1.*** The variant of the organization of the fuzzy supervisor

$$k_i = F_3(e, \dot{e}).$$

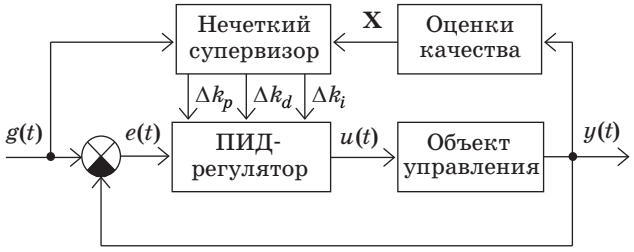
Это еще более усложняет задачу составления правил, делая ее практически невозможной.

Таким образом, обеспечение адекватного выбора закона управления коэффициентами ПИД-регулятора оказывается возможным только при использовании процедур оптимизации. Допустим, что термы лингвистических переменных, описывающих входы и выходы нечеткого супервизора, образуют нечеткое разбиение соответствующих базовых шкал, а для дефаззификации используется метод центра тяжести. В этом случае значения коэффициентов на выходе нечеткой системы линейно зависят от ее входов [24]. Тогда реализовать отображение вход-выход можно с помощью линейной НС, представленной на рис. 2.



■ ***Рис. 2.*** Описание супервизора как линейной нейронной сети

■ ***Fig. 2.*** Description of supervisor as linear neural network



■ ***Рис. 3.*** Нечеткий супервизор с оценкой качества переходного процесса

■ ***Fig. 3.*** Fuzzy supervisor with an assessment of the quality of the transient process

Матрица весов **W** на рис. 2 должна быть найдена в процессе оптимизации, которая может быть выполнена с помощью генетического алгоритма [13, 25] или другого метода глобального поиска.

Второй вариант схемы нечеткого супервизора (рис. 3) реализует определение коэффициентов регулятора в виде

$$k_p = F_1(\mathbf{X});\ k_d = F_2(\mathbf{X});\ k_i = F_3(\mathbf{X}),$$

где **X** — вектор оценок параметров переходного процесса, образующий вход супервизора.

Этот вариант нечеткого супервизора предполагает инкрементальное изменение коэффициентов ПИД-регулятора на основании построения оценок качества переходного процесса, таких как статическая ошибка, перерегулирование и время нарастания. Предполагается, что входное воздействие $g(t)$ носит периодический характер. Целевыми значениями являются малое перерегулирование и малая статическая ошибка при малом времени нарастания.

Как известно, коэффициенты ПИД-регулятора оказывают различное влияние на параметры переходного процесса (табл. 1).

На основании табл. 1 можно сформулировать примерные правила изменения коэффициентов ПИД-регулятора:

■ ***Таблица 1.*** Эффекты при увеличении коэффициентов ПИД-регулятора

■ ***Table 1.*** Effects from increasing PID controller ratios

Коэффициент	Время нарастания $T_н$	Статическая ошибка E	Перерегулирование P, %
k_p	Уменьшает	Уменьшает	Увеличивает
k_d	Не влияет	Не влияет	Уменьшает
k_i	Уменьшает	Устраняет	Увеличивает

— если статическая ошибка большая, то k_i следует увеличивать;

— если статическая ошибка очень большая, то k_p следует увеличивать;

— если перерегулирование большое, то k_d следует увеличивать;

— если перерегулирование очень большое, то k_p следует уменьшать;

— если перерегулирование малое, то k_p следует увеличивать.

Последнее правило позволяет сократить время нарастания.

Будем считать, что объект управления устойчив и переходный процесс гарантированно заканчивается за время T, тогда

$$E = g(T) - y(T);$$

$$P = (y_{\max} - y(T)) \times 100\,\%.$$

Таким образом, по результатам сделанного анализа могут быть рассмотрены два варианта: супервизор непрерывного действия и супервизор периодического действия.

Моделирование работы нечеткого супервизора непрерывного действия

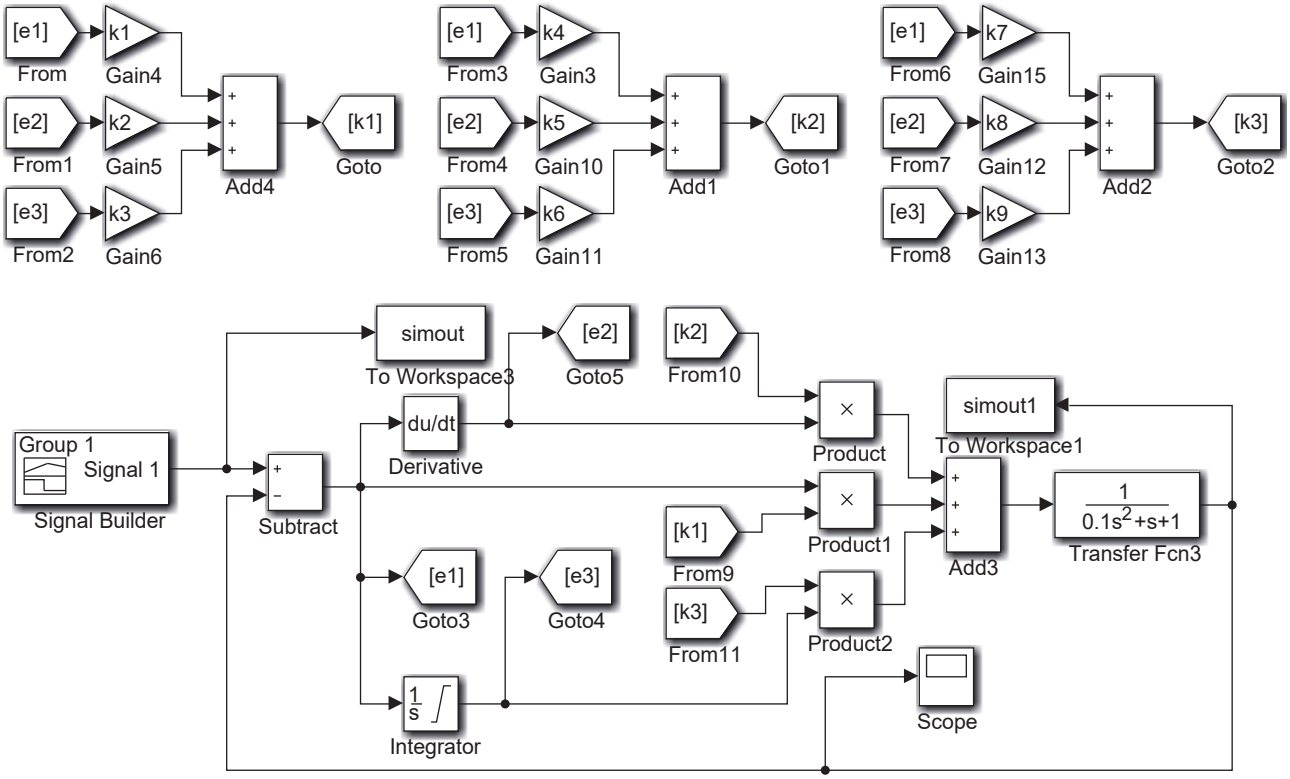
Схема эксперимента в Simulink MatLab представлена на рис. 4 (блоки *simout* используются при генетическом обучении).

При настройке регулятора использовался объект управления, заданный с помощью передаточной функции W_1. При тестировании использовались варианты объектов с передаточными функциями W_2, W_3, где

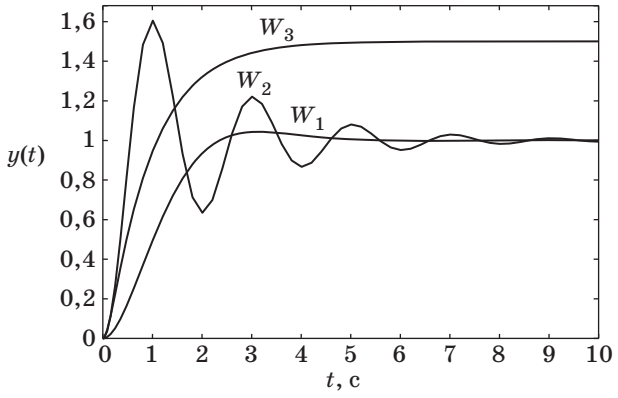
$$W_1(s) = \frac{1}{0{,}1s^2 + s + 1};\ W_2(s) = \frac{1{,}5}{0{,}5s^2 + s + 1};$$

$$W_3(s) = \frac{1}{0{,}1s^2 + 0{,}1s + 1}.$$

Динамические характеристики этих объектов иллюстрирует рис. 5.



■ ***Рис. 4.*** Блок-схема супервизора непрерывного действия
■ ***Fig. 4.*** Block diagram of continuous operation supervisor

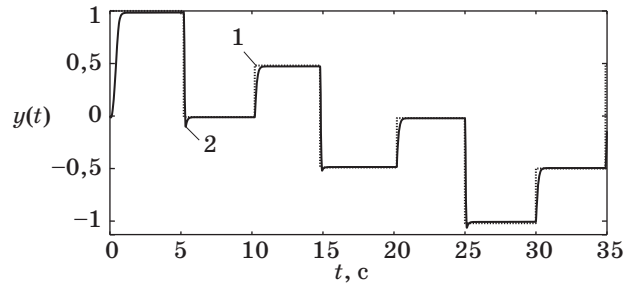


■ ***Рис. 5.*** Переходные характеристики объектов
■ ***Fig. 5.*** Transient processes for plants

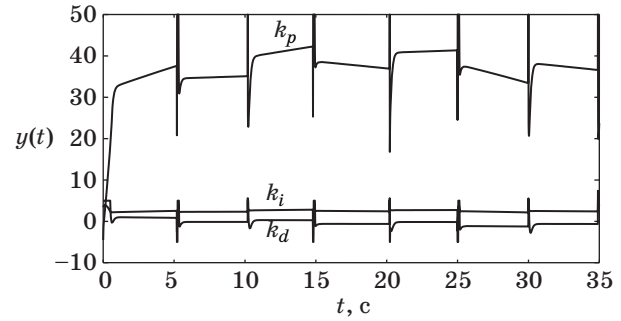
После обучения была получена следующая матрица весовых коэффициентов супервизора:

$$\mathbf{W} = \begin{bmatrix} -4{,}5 & 3{,}42 & 76 \\ 4{,}25 & 0{,}32 & 9{,}5 \\ 73{,}8 & 4{,}88 & 0{,}51 \end{bmatrix}.$$

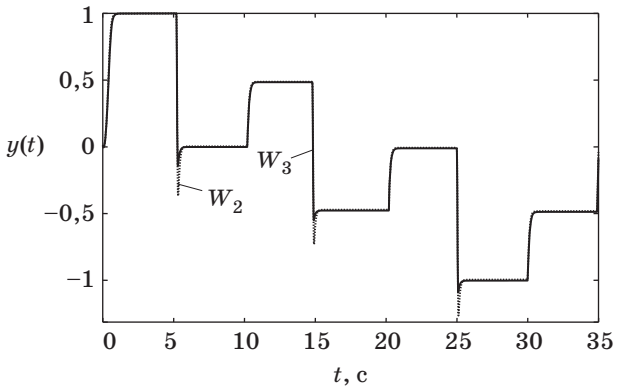
Реакция системы с супервизором для объекта W_1 показана на рис. 6. На рис. 7 представлены графики изменения коэффициентов регулятора



■ ***Рис. 6.*** Переходный процесс в системе с супервизором: *1* — входной сигнал; *2* — выход системы
■ ***Fig. 6.*** Transient processes in system with supervisor: *1* — input signal; *2* — system output



■ ***Рис. 7.*** Изменение коэффициентов ПИД-регулятора
■ ***Fig. 7.*** Changing of PID controller coefficients



■ ***Рис. 8***. Переходные процессы для объектов
■ ***Fig. 8***. Transient processes for plants

во время переходного процесса. На рис. 8 показаны переходные процессы для тестовых объектов W_2 и W_3. Качество работы остается удовлетворительным, хотя динамика объектов управления существенно различна (см. рис. 5).

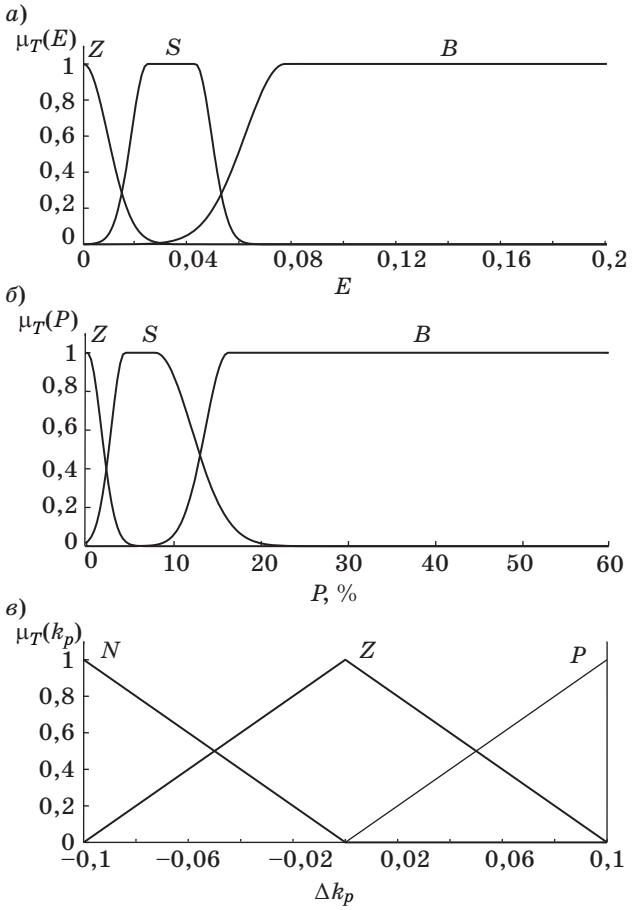
Моделирование работы нечеткого супервизора периодического действия

Лингвистическое описание входных переменных нечеткого супервизора в Simulink MatLab показано на рис. 9, *а* и *б* (где $\mu_T(x)$ — степень принадлежности значения x к терму T соответствующей лингвистической переменной; Z, S и B — сокращения наименований термов *Zero*, *Small*, *Big*).

Выходные переменные описываются с помощью трех термов с треугольной функцией принадлежности и лингвистическими метками N, Z, P (*Negative*, *Zero*, *Positive*). На рис. 9, *в* представлено описание для пропорционального коэффициента. Описание для дифференциального и интегрального коэффициентов аналогично, но размер базовой шкалы — 0,02 и 0,01 соответственно.

Максимальное количество управляющих правил равно мощности декартова произведения терм-множеств посылок правил. Так как для описания E и P использованы по три терма, всего получается девять управляющих правил нечеткого супервизора (табл. 2).

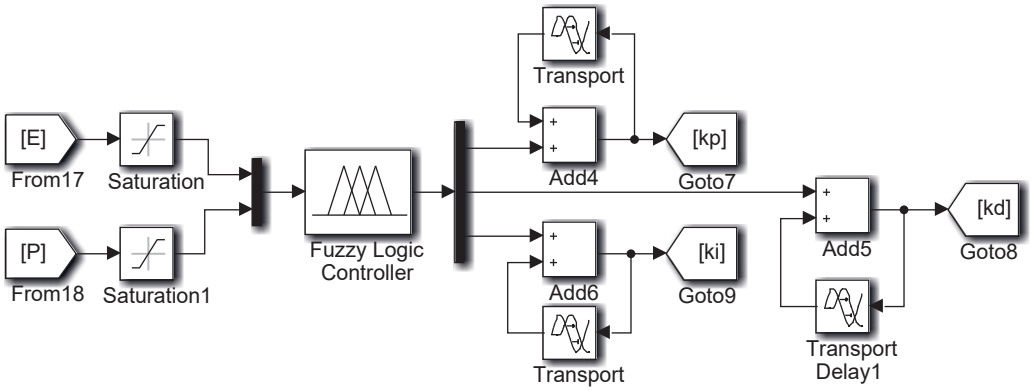
Как показало моделирование, непосредственно использовать оценку времени нарастания переходного процесса нет необходимости, потому что этот параметр зависит от значения максимально допустимого перерегулирования — при малом перерегулировании можно увеличивать пропорциональный коэффициент, уменьшая тем самым время нарастания.



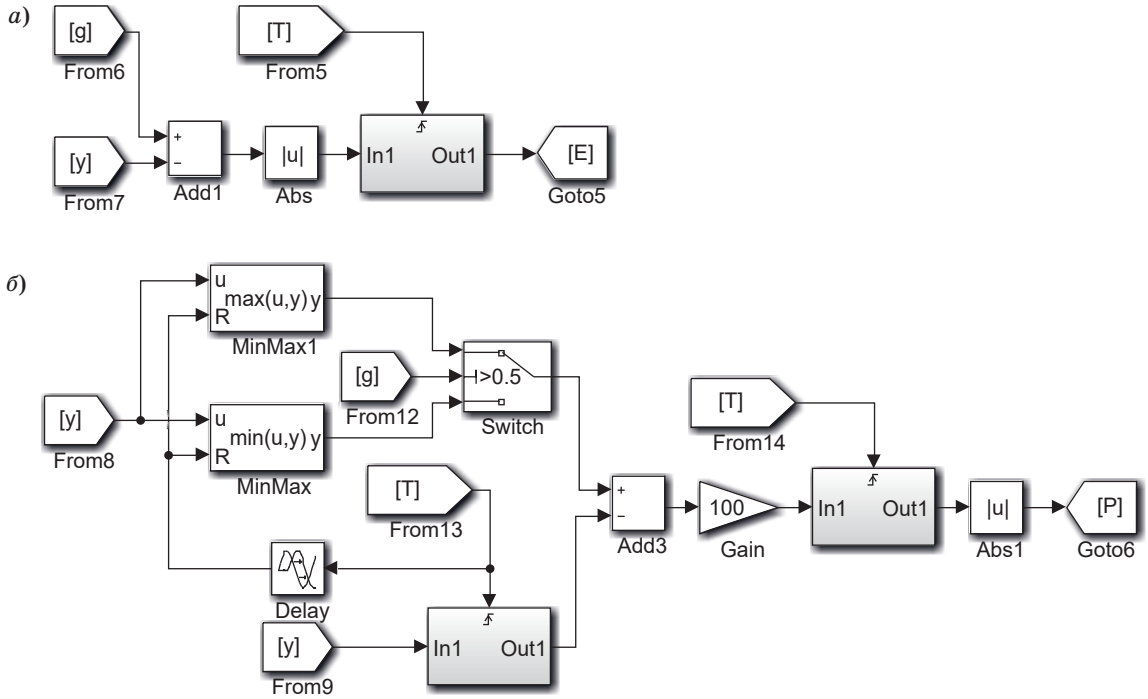
■ ***Рис. 9***. Нечеткое описание статической ошибки (*а*); перерегулирования (*б*); приращения коэффициента ПИД-регулятора (*в*)
■ ***Fig. 9***. Fuzzy description of static error (*a*); of overshoot (*б*); of the increment of the PID controller coefficient (*в*)

■ ***Таблица 2***. Таблица лингвистических правил
■ ***Table 2***. Table of linguistic rules

E	P	k_p	k_d	k_i
Z	Z	Z	Z	Z
Z	S	Z	Z	Z
Z	B	N	P	Z
S	Z	Z	Z	P
S	S	Z	Z	P
S	B	Z	P	P
B	Z	P	Z	P
B	S	P	Z	P
B	B	P	P	P



■ ***Рис. 10.*** Блок-схема нечеткого супервизора
■ ***Fig. 10.*** Block diagram of fuzzy supervisor



■ ***Рис. 11.*** Блок-схема для определения статической ошибки (*а*) и перерегулирования (*б*)
■ ***Fig. 11.*** Block diagram for determining static error (*а*) and overshoot (*б*)

На рис. 10 показана реализация нечеткого супервизора в MatLab. На рис. 11, *а* и *б* приведены блоки извлечения качественных характеристик переходного процесса.

Для оценки качества работы нечеткого супервизора рассмотрим задачу управления объектом, заданным колебательным звеном:

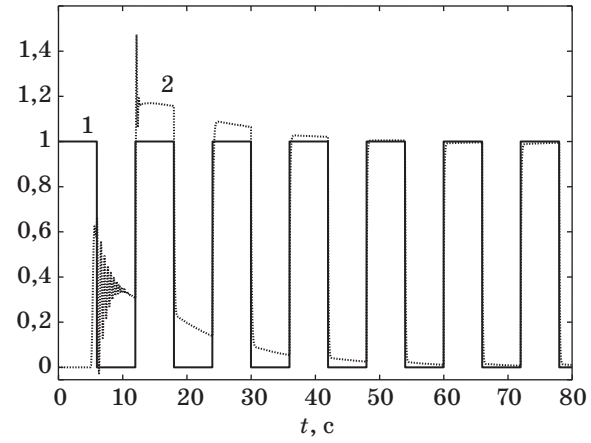
$$W(s) = \frac{0,8}{0,02s^2 + 0,025s + 1}.$$

Коэффициенты ПИД-регулятора должны быть настроены таким образом, чтобы уменьшить статическую ошибку, перерегулирование и колебательность.

Эта задача решается вполне удовлетворительно (рис. 12). Значительные колебания на первых секундах вызваны тем, что начальные значения коэффициентов сильно отличаются от требуемых значений.

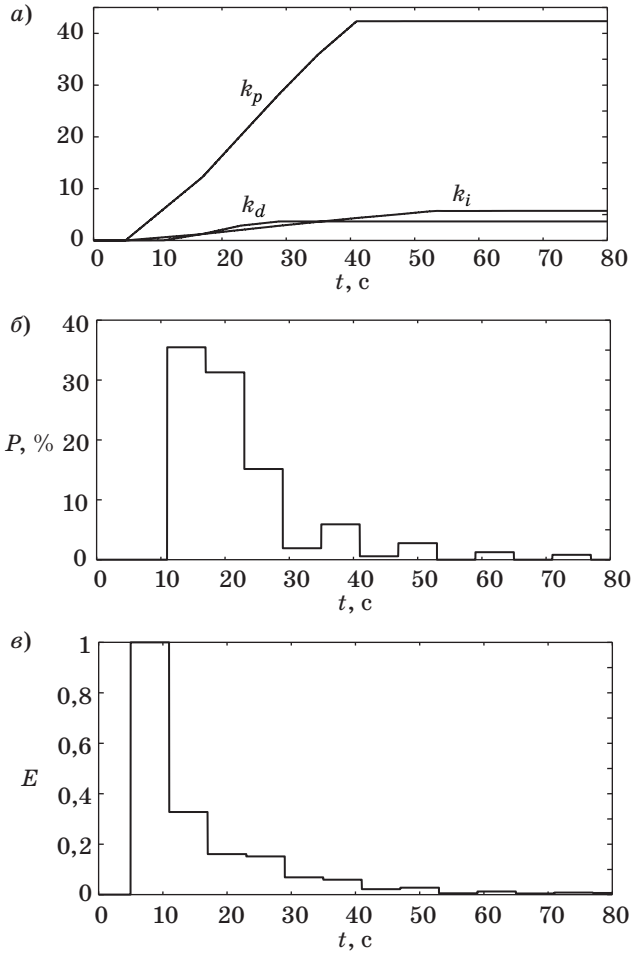
При нулевых начальных значениях коэффициентов регулятора после каждого цикла регулирования происходит их коррекция, которая прекращается по достижении хорошего качества регулирования (рис. 13, *а*).

На рис. 13, *б* и *в* показаны изменения оценок качества во время переходного процесса.



■ ***Рис. 12.*** Переходные процессы в системе с нечетким супервизором: *1* — задающее воздействие; *2* — выход системы

■ ***Fig. 12.*** Transient processes in a system with a fuzzy supervisor: *1* — driving impact; *2* — system output



■ ***Рис. 13.*** Изменение коэффициентов ПИД-регулятора (*а*); оценки перерегулирования (*б*); статической ошибки (*в*)

■ ***Fig. 13.*** Changing of PID controller coefficients (*а*); in the estimation of overshoot (*б*); of static error (*в*)

Периодическое входное воздействие (см. рис. 12) может использоваться на этапе настройки регулятора, так что нечеткий супервизор заменяет здесь традиционные алгоритмы, такие как метод Зиглера — Николса [1].

Заключение

Несмотря на все более активное внедрение нечетких и нейросетевых регуляторов, традиционные ПИД-регуляторы продолжают оставаться наиболее популярным средством автоматизации в промышленности. Использование нечетких супервизоров значительно расширяет возможности ПИД-регуляторов, придавая им адаптивные свойства.

На практике могут быть использованы две схемы супервизорного управления: с непрерывной коррекцией коэффициентов на основании информации об ошибке управления и с периодической коррекцией на основании оценок качества переходного процесса.

Как показали проведенные эксперименты, в схеме с непрерывной коррекцией можно рассматривать гипотезу о линейности нечетких управляющих правил, что позволяет эффективно использовать для их описания линейную НС с генетической настройкой. Такой подход не накладывает ограничений на вид входного сигнала и обеспечивает робастность регулятора.

В случае периодического изменения входного сигнала для работы супервизора могут быть использованы прямые оценки качества переходного процесса. За счет этого оказывается возможным сформулировать правила коррекции параметров регулятора, имеющие ясную семантику. Проведенные вычислительные эксперименты показали хорошее качество работы супервизорной системы.

Применение нечетких супервизоров для ПИД-регуляторов может оказаться полезным при проектировании систем управления широким кругом динамических объектов. Реализация подобной системы на базе микроконтроллера не вызывает технических трудностей и незначительно увеличивает стоимость системы управления.

Работа выполнена при финансовой поддержке РФФИ (проект № 18-01-00076).

Литература

1. **Ziegler J. G., Nichols N. B.** Optimum Settings for Automatic Controllers // Trans. ASME. 1942. Vol. 64. P. 759–768.
2. **Visioli A.** Practical PID Control. — London, UK: Springer, 2006. — 314 p.

3. **Astrom K. J., Hagglund T.** Advanced PID Control. — ISA (The Instrumentation, System, and Automation Society), 2006. — 460 p.
4. **Бураков М. В., Коновалов А. С.** Модификация предиктора Смита для линейного объекта с переменными параметрами // Информационно-управляющие системы. 2017. № 4. С. 25–34. doi:10.15217/issn1684-8853.2017.4.25
5. **Ketata R., Geest D., Titli D. A.** Fuzzy Controller: Design, Evaluation, Parallel and Hierarchical Combination with a PID Controller // Fuzzy Sets and Systems. 1995. Vol. 71. P. 113–229.
6. **Omatu S., Yoshioka M., Fujinaka T.** Neuro-PID Control for Electric Vehicle // Journal of Advanced Computational Intelligence and Intelligent Informatics. 2011. Vol. 15. N 7. P. 846–852.
7. **Burakov M. V., Kurbanov V. G.** Fuzzy PID Controller for Nonlinear Plant // ARPN Journal of Engineering and Applied Sciences. 2016. Vol. 11. N 9. P. 5745–5748.
8. **Burakov M. V., Kurbanov V. G.** Neuro-PID Control for Nonlinear Plants with Variable Parameters // ARPN Journal of Engineering and Applied Sciences. 2017. Vol. 12. N 4. P. 1226–1229.
9. **Karray F., Gueaieb W., Al-Sharhan S.** The Hierarchical Expert Tuning of PID Controllers using Tools of Soft Computing // IEEE Transactions on Systems, Man, and Cybernetics. Part b: Cybernetics. 2002. Vol. 32. N 1. P. 77–90.
10. **Patel R., Kumar V.** Multilayer Neuro PID Controller based on Back Propagation Algorithm // Procedia Computer Science. 2015. Vol. 54. P. 207–214.
11. **Du W., Zhao X., Lv F., Du H.** The Design of PID Controller based on Hopfield Neural Network // TELKOMNIKA Indonesian Journal of Electrical Engineering. 2014. Vol. 12. N 4. P. 3185–3192.
12. **Hasanpour H., Heydari Beni M., Askari M.** Adaptive PID Control based on RBF NN for Quadrotor // International Research Journal of Applied and Basic Sciences. 2017. Vol. 11. N 2. P. 177–186.
13. **Бураков М. В., Коновалов А. С., Яковец О. Б.** Эволюционный синтез нечетких регуляторов // Информационно-управляющие системы. 2015. № 6. С. 28–33. doi:10.15217/issn1684-8853.2015.6.28
14. **Dagher K. A., Al-Araji A. S.** Design of an Adaptive PID Neural Controller for Continuous Stirred Tank Reactor based on Particle Swarm Optimization // Al-Khwarizmi Engineering Journal. 2013. Vol. 9. N 4. P. 46–53.
15. **Lee C. C.** Fuzzy Logic in Control Systems: Fuzzy Logic Controller. Part I // IEEE Transactions on Systems, Man and Cybernetics. 1990. Vol. 20. N 2. P. 404–418.
16. **Zhao Z.-Y., Tomizuka M., Isaka S.** Fuzzy Gain Scheduling of PID Controllers // IEEE Transaction on System, Man and Cybernetics. 1993. Vol. 23. N 5. P. 1392–1398.
17. **Ali M., Khan S., Waleed M.** Application of an Intelligent Self-Tuning Fuzzy PID Controller on DC-DC Buck Converter // International Journal of Advanced Science and Technology. 2012. Vol. 48. P. 139–147.
18. **Qiu C., Huang Y.** The Design of Fuzzy Adaptive PID Controller of Two-Wheeled Self-Balancing Robot // International Journal of Information and Electronics Engineering. 2015. Vol. 5. N 3. P. 193–197.
19. **Ebrahimi N., Gharaveisi A.** Optimal Fuzzy Supervisor Controller for an Active Suspension System // International Journal of Soft Computing and Engineering (IJSCE). 2012. Vol. 2. N 4. P. 36–39.
20. **Passino K. M., Yurkovich S.** Fuzzy Control. — Addison Wesley Longman Inc., 1998. — 500 p.
21. **Song Z., Wang Y., Yuan M.** The Application of Fuzzy-PID Control used in the Control of Car Distance // World Journal of Modelling and Simulation. 2007. Vol. 3. N 2. P. 141–148.
22. **Macvicar-Whelan P. J.** Fuzzy Sets for Man-Machine Interaction // Int. J. Man-Mach. Studies. 1976. Vol. 8. P. 687–697.
23. **Jin J., Huang H., Sun J., Pang Y.** Study on Fuzzy Self-Adaptive PID Control System of Biomass Boiler Drum Water // Journal of Sustainable Bioenergy Systems. 2013. Vol. 3. P. 93–98.
24. **Chen G., Pham T. T.** Introduction to Fuzzy Sets, Fuzzy Logic and Fuzzy Control Systems. — London, UK: CRC Press, 2001. — 316 p.
25. **Бураков М. В., Яковец О. Б.** Нечеткое управление силовым гироскопическим прибором // Изв. вузов. Приборостроение. 2015. Т. 58. № 10. С. 157–166.

UDC 681.5
doi:10.31799/1684-8853-2018-5-13-21

Fuzzy supervisor for PID controller

M. V. Burakov[a], PhD, Tech., Associate Professor, orcid.org/0000-0001-8047-9493, bmv@sknt.ru
A. S. Konovalov[a], Dr. Sc., Tech., Professor, orcid.org/0000-0001-6127-6789
[a]Saint-Petersburg State University of Aerospace Instrumentation, 67, B. Morskaya St., 190000, Saint-Petersburg, Russian Federation

Introduction: PID regulators are an important industrial automation tool. However, the traditional ways of their customization involve experiments with the plant, reducing the control system efficiency when the operating conditions change. To solve this problem, you have to use an adaptation loop which automatically changes the regulator parameters when the quality indicators deteriorate.

Purpose: Developing a structure and an algorithm for a fuzzy supervisor of a PID controller in order to improve the operation quality in the context of uncontrolled changes in the plant parameters. **Results:** We propose two options for organizing a fuzzy supervisor. The first option involves continuous changing of the controller parameters based on the information about the current control error. The supervisor is trained offline with the help of a genetic algorithm. This approach makes the application universal, but the resulting structure is a «black box». The second option is using transient process quality estimates; it can be used when the input signal periodically changes. The controller parameters are changed using fuzzy rules with clear semantics. It is shown that systems with a fuzzy supervisor can reduce the overshoot or static error which both can occur when the control object parameters change. The system operation was simulated using MatLab Simulink. **Practical relevance:** The use of PID controllers with fuzzy supervisors can be useful in the design of control systems for a wide range of dynamic plants.

Keywords — adaptation, PID controller, fuzzy logic, fuzzy supervisor, modeling, genetic algorithm.

Citation: Burakov M. V., Konovalov A. S. Fuzzy supervisor for PID controller. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 5, pp. 13–21 (In Russian). doi:10.31799/1684-8853-2018-5-13-21

References

1. Ziegler J. G., Nichols N. B. Optimum Settings for Automatic Controllers. *Trans. ASME*, 1942, vol. 64, pp. 759–768.
2. Visioli A. *Practical PID Control*. Springer, London, UK, 2006. 314 p.
3. Astrom K. J., Hagglund T. *Advanced PID Control*. ISA (The Instrumentation, System, and Automation Society), 2006. 460 p.
4. Burakov M. V., Konovalov A. S. Modification of Smith Predictor for a Linear Plant with Changeable Parameters. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2017, no. 4, pp. 25–34 (In Russian). doi:10.15217/issn1684-8853.2017.4.25
5. Ketata R., Geest D., Titli D. A. Fuzzy Controller: Design, Evaluation, Parallel and Hierarchical Combination with a PID Controller. *Fuzzy Sets and Systems*, 1995, vol. 71, pp. 113–229.
6. Omatu S., Yoshioka M., Fujinaka T. Neuro-PID Control for Electric Vehicle. *Journal of Advanced Computational Intelligence and Intelligent Informatics*, 2011, vol. 15, no. 7, pp. 846–852.
7. Burakov M. V., Kurbanov V. G. Fuzzy PID Controller for Nonlinear Plant. *ARPN Journal of Engineering and Applied Sciences*, 2016, vol. 11, no. 9, pp. 5745–5748.
8. Burakov M. V., Kurbanov V. G. Neuro-PID Control for Nonlinear Plants with Variable Parameters. *ARPN Journal of Engineering and Applied Sciences*, 2017, vol. 12, no. 4, pp. 1226–1229.
9. Karray F., Gueaieb W., Al-Sharhan S. The Hierarchical Expert Tuning of PID Controllers using Tools of Soft Computing. *IEEE Transactions on Systems, Man, and Cybernetics. Part b: Cybernetics*, 2002, vol. 32, no. 1, pp. 77–90.
10. Patel R., Kumar V. Multilayer Neuro PID Controller based on Back Propagation Algorithm. *Procedia Computer Science*, 2015, vol. 54, pp. 207–214.
11. Du W., Zhao X., Lv F., Du H. The Design of PID Controller Based on Hopfield Neural Network. *TELKOMNIKA Indonesian Journal of Electrical Engineering*, 2014, vol. 12, no. 4, pp. 3185–3192.
12. Hasanpour H., Heydari Beni M., Askari M. Adaptive PID Control Based on RBF NN for Quadrotor. *International Research Journal of Applied and Basic Sciences*, 2017, vol. 11, no. 2, pp. 177–186.
13. Burakov M. V., Konovalov A. S., Yakovets O. B. Evolutionary Design of Fuzzy Controllers. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2015, no. 6, pp. 28–33 (In Russian). doi:10.15217/issn1684-8853.2015.6.28
14. Dagher K. A., Al-Araji A. S. Design of an Adaptive PID Neural Controller for Continuous Stirred Tank Reactor based on Particle Swarm Optimization. *Al-Khwarizmi Engineering Journal*, 2013, vol. 9, no. 4, pp. 46–53.
15. Lee C. C. Fuzzy Logic in Control Systems: Fuzzy Logic Controller. Part I. *IEEE Transactions on Systems, Man and Cybernetics*, 1990, vol. 20, no. 2, pp. 404–418.
16. Zhao Z.-Y., Tomizuka M., Isaka S. Fuzzy Gain Scheduling of PID Controllers. *IEEE Transaction on System, Man and Cybernetics*, 1993, vol. 23, no. 5, pp. 1392–1398.
17. Ali M., Khan S., Waleed M. Application of an Intelligent Self-Tuning Fuzzy PID Controller on DC-DC Buck Converter. *International Journal of Advanced Science and Technology*, 2012, vol. 48, pp. 139–147.
18. Qiu C., Huang Y. The Design of Fuzzy Adaptive PID Controller of Two-Wheeled Self-Balancing Robot. *International Journal of Information and Electronics Engineering*, 2015, vol. 5, no. 3, pp. 193–197.
19. Ebrahimi N., Gharaveisi A. Optimal Fuzzy Supervisor Controller for an Active Suspension System. *International Journal of Soft Computing and Engineering* (*IJSCE*), 2012, vol. 2, no. 4, pp. 36–39.
20. Passino K. M., Yurkovich S. *Fuzzy Control*. Addison Wesley Longman Inc., 1998. 500 p.
21. Song Z., Wang Y., Yuan M. The Application of Fuzzy-PID Control used in the Control of Car Distance. *World Journal of Modelling and Simulation*, 2007, vol. 3, no. 2, pp. 141–148.
22. Macvicar-Whelan P. J. Fuzzy Sets for Man-Machine Interaction *Int. J. Man-Mach. Studies*, 1976, vol. 8, pp. 687–697.
23. Jin J., Huang H., Sun J., Pang Y. Study on Fuzzy Self-Adaptive PID Control System of Biomass Boiler Drum Water. *Journal of Sustainable Bioenergy Systems*, 2013, vol. 3, pp. 93–98.
24. Chen G., Pham T. T. *Introduction to Fuzzy Sets, Fuzzy Logic and Fuzzy Control Systems*. CRC Press, London, UK, 2001. 316 p.
25. Burakov M. V., Yakovets O. B. Fuzzy Control of Power Gyroscopic Unit. *Izvestiia vuzov. Priborostroenie*, 2015, vol. 58, no. 10, pp. 157–166 (In Russian).

УДК 629.78
doi:10.31799/1684-8853-2018-5-22-34

Научные статьи
Articles

Первичная обработка телеметрической информации с использованием динамических моделей изменения параметров и парциальной нелинейной фильтрации

Г. Н. Мальцев[а], *доктор техн. наук, профессор, orcid.org/0000-0002-6755-5700, georgy_maltsev@mail.ru*
В. Л. Якимов[а], *канд. техн. наук, доцент, orcid.org/0000-0001-9721-2453*
С. В. Соловьев[б], *канд. техн. наук, доцент, orcid.org/0000-0002-0391-8728*
Н. В. Лебедева[б], *ассистент, orcid.org/0000-0001-6963-6638*
[а]*Военно-космическая академия им. А. Ф. Можайского, Ждановская наб., 13, Санкт-Петербург, 197198, РФ*
[б]*Московский государственный технический университет им. Н. Э. Баумана, 2-я Бауманская ул., 5, стр. 1, Москва, 105005, РФ*

Постановка проблемы: *на этапе предварительной обработки телеметрической информации о техническом состоянии бортовой аппаратуры космических аппаратов существует необходимость выявления изменений телеметрируемых параметров, важных с точки зрения обнаружения и парирования нештатных и аварийных ситуаций. Решение проблемы осложняется нелинейным характером изменения большинства параметров функционирования бортовой аппаратуры и влиянием возмущений различного рода на изменение анализируемых телеметрируемых параметров.* ***Цель:*** *разработка метода обработки телеметрической информации для повышения достоверности разнородной телеметрической информации космических аппаратов в реальных условиях возмущений и нелинейного характера изменения их параметров на основе фильтрации временных рядов значений телеметрируемых параметров.* ***Результаты:*** *предложен метод обработки телеметрической информации на основе динамических моделей изменения телеметрируемых параметров бортовой аппаратуры космических аппаратов в проекциях фазового пространства и парциальной нелинейной фильтрации анализируемых временных рядов с использованием стохастической аппроксимации функции плотности вероятности распределения вектора состояния. Показаны преимущества первичной обработки телеметрируемых параметров бортовой аппаратуры космических аппаратов в многомерном фазовом пространстве за счет возможности выбора наиболее существенных с точки зрения результатов фильтрации проекций фазового портрета. В разработанный алгоритм парциальной нелинейной фильтрации заложена возможность адаптации к параметрам действующих возмущений. Продемонстрирована работоспособность метода для случая гауссовских и негауссовских помех в условиях, при которых нарушается устойчивость методов линейной фильтрации.* ***Практическая значимость:*** *предлагаемый метод является унифицированным для обработки разнородных телеметрируемых параметров и обеспечивает повышение достоверности оценки значений как функциональных, так и сигнальных телеметрируемых параметров космических аппаратов в реальных условиях функционирования.*

Ключевые слова *— телеизмерения, телеметрируемый параметр, динамический процесс, фазовое пространство, парциальная нелинейная фильтрация.*

Цитирование: Мальцев Г. Н., Якимов В. Л., Соловьев С. В., Лебедева Н. В. Первичная обработка телеметрической информации с использованием динамических моделей изменения параметров и парциальной нелинейной фильтрации. *Информационно-управляющие системы*, 2018, № 5, с. 22–34. doi:10.31799/1684-8853-2018-5-22-34
Citation: Maltsev G. N., Yakimov V. L., Soloviev S. V., Lebedeva N. V. Primary processing of telemetric information using dynamic models of parameter change and partial nonlinear filtration. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 5, pp. 22–34 (In Russian). doi:10.31799/1684-8853-2018-5-22-34

Введение

Первичная обработка телеметрической информации (ТМИ) является важным этапом решения задач контроля технического состояния (ТС) и технического диагностирования бортовой аппаратуры (БА) космических аппаратов (КА). Основной задачей этапа первичной обработки ТМИ является предоставление результатов телеизмерений в требуемой форме и с требуемым качеством для последующей оценки ТС БА КА [1]. Качественное решение данной задачи позволяет повысить достоверность контроля ТС и технического диагностирования БА КА. В то же время на этапе первичной обработки ТМИ могут быть оперативно выявлены изменения телеметрируемых параметров (ТМП), существенные с точки зрения выявления нештатных и аварийных ситуаций на борту КА.

Практика летной эксплуатации КА показывает, что возникновение нештатных ситуаций (НС) на борту КА, как правило, связано с влиянием на БА продолжительных во времени возмущений и факторов космического пространства [2, 3].

Во многих случаях это влияние может быть обнаружено в результате первичной обработки ТМИ и анализа представленных временными рядами значений контролируемых ТМП. При этом несвоевременное обнаружение НС может приводить к срыву выполнения программы полета, целевых задач КА и запланированных космических исследований [2, 3]. Для своевременного выявления возмущений в динамике ТМП в алгоритмы обработки ТМИ и технического диагностирования БА КА должны быть заложены различные модели изменения ТМП, которые, к сожалению, не всегда известны, но могут быть построены на основе накопленной измерительной и экспертной информации, полученной по результатам испытаний БА при подготовке КА к запуску и летной эксплуатации однотипных КА [1, 2]. Использование данной информации позволяет синтезировать модели анализируемых процессов и использовать их для решения задач первичной обработки ТМИ с последующей идентификацией состояния БА и парированием текущих и прогнозируемых НС.

Особенностью решения задач первичной обработки ТМИ является необходимость фильтрации результатов телеизмерений для разделения информативной составляющей динамического процесса изменения анализируемых ТМП и шумов измерений и помех, являющихся его неинформативной составляющей. Выполнение фильтрации предшествует оценке значений анализируемого ТМП и требует использования априорной информации о характере его изменения (поведения) на интервале анализа.

Предлагаемый в статье метод основан на описании динамических процессов изменения ТМП БА КА в фазовом пространстве с использованием унифицированных математических моделей и обработке их временных рядов в каждой проекции фазового пространства на основе парциальной нелинейной фильтрации [4–6].

Динамическая модель изменения телеметрируемых параметров в фазовом пространстве

Временные ряды значений ТМП БА КА в общем случае аппроксимируются нелинейными функциями, а в условиях возмущений характеризуются появлением аномальных отсчетов, нелинейных трендов и скачков значений ТМП [7]. Также во временных рядах ТМП могут наблюдаться постоянные смещения, обусловленные ошибками датчико-преобразующей аппаратуры и телеметрических согласующих устройств. Это позволяет сделать вывод о том, что в общем случае БА КА, характеризуемую совокупностью ТМП, целесообразно рассматривать как нелинейную динамическую систему, функционирующую в условиях воздействия негауссовских помех [7].

Поведение нелинейных динамических систем обычно рассматривают в фазовом пространстве [8, 9]. Под фазовым пространством понимают пространство, в котором каждому состоянию динамической системы соответствует точка этого пространства, определяемая значениями ее параметров. Траекторию этой точки в фазовом пространстве называют фазовой траекторией. Размерность N фазового пространства состояний БА КА может быть определена по реализациям ТМП на основе метода ложных ближайших соседей [10]. Совокупность фазовых траекторий образует фазовый портрет динамической системы. В соответствии с теоремой Такенса поведение динамической системы в фазовом пространстве можно описать по временной реализации параметра системы с использованием метода задержек [8]. При этом каждый временной ряд ТМП динамической системы представляется дискретными отсчетами: x_1, x_2, ..., полученными с интервалом Δt и образующими множество $\{x_k\}$. Обычно выбирается интервал дискретизации Δt, соответствующий моменту первого пересечения нуля автокорреляционной функцией временного ряда. Эту оценку можно считать верхней и при выборе интервала дискретизации Δt осуществлять прореживание временного ряда, исходя из требуемого качества решения целевой задачи. Для цифровых систем значение Δt кратно периоду дискретизации временного ряда ТМП.

Важным достоинством рассмотрения динамических систем в фазовом пространстве является то, что оно позволяет описывать как их локальные свойства, так и глобальное поведение. Так, близлежащие точки фазового портрета в некоторых его проекциях могут находиться на достаточно большом расстоянии друг от друга на временной оси и при этом одинаково характеризовать поведение системы. Это позволяет создавать модели динамических систем на основе кусочно-локальной аппроксимации. А наличие у фазовых портретов динамических систем аттракторов (точек притяжения) позволяет создавать модели их глобального поведения, поскольку при наличии аттракторов все фазовые траектории из его окрестности стремятся к нему с течением времени.

На практике для анализа динамических систем обычно используют метод сечений Пуанкаре и рассматривают двумерные проекции фазовых портретов в N-мерном фазовом пространстве [8]. Такой подход используют для детального исследования динамической системы, выявления закономерностей в ее поведении и получения простых моделей ее динамики в виде одномерных отображений, а также для уменьшения вычис-

лительной сложности алгоритмов анализа, что немаловажно при автоматизированной обработке ТМИ. Исходя из этого можно рассматривать фазовые портреты в каждой из $(N-1)$ двумерных проекций фазового пространства и выбирать наиболее информативные из них с точки зрения задачи фильтрации с последующим объединением результатов обработки.

Для формализации поведения динамической системы в фазовом пространстве целесообразно ввести понятие вектора наблюдения и вектора состояния, характеризующих состояние динамической системы в каждый момент времени. Вектор наблюдения динамической системы в каждой j-й проекции фазового пространства для k-го дискретного отсчета времени определим на основании измеренных значений ТМП следующим образом: $\mathbf{X}_{j,k} = \begin{bmatrix} x_k \\ x_{k+j} \end{bmatrix}$, где x_k, x_{k+j} — сдвинутые отсчеты временного ряда значений ТМП. Под вектором состояния динамической системы в каждой j-й проекции фазового портрета для k-го дискретного отсчета времени будем понимать вектор $\mathbf{Q}_{j,k} = \begin{bmatrix} q_{x,\,j,\,k} \\ q_{y,\,j,\,k} \end{bmatrix}$, где $q_{x,j,k}$, $q_{y,j,k}$ — компоненты вектора состояния, характеризующего области фазового пространства, принадлежащие аттрактору динамической системы.

Одним из способов локализации областей фазового пространства, принадлежащих аттрактору динамической системы, является кластеризация векторов наблюдения $\mathbf{X}_{j,k}$ в каждой j-й проекции. Получаемые кластеры отсчетов значений ТМП характеризуют аттрактор динамической системы с некоторой доверительной вероятностью. Центры таких кластеров можно связать с вектором состояния динамической системы $\mathbf{Q}_{j,k}$ в k-й момент времени в отсутствие возмущений для известных ситуаций штатного функционирования БА КА и рассмотренных нештатных ситуаций. При этом совокупность таких кластеров в каждой j-й проекции можно использовать для описания технического состояния БА КА [11]. При отсутствии статистики измерений по некоторым состояниям БА для получения набора кластеров может быть использована экспертная информация о требуемом поведении ТМП. Рассмотренный подход хорошо согласуется с методами Data mining, используемыми для углубленного анализа динамики сложных технических систем, в том числе КА [12].

Эффективным решением задачи кластеризации точек $\mathbf{X}_{j,k}$ в каждой j-й проекции фазового пространства является аппроксимация функции распределения точек $P(\mathbf{X}_{j,k})$ полигауссовской моделью [4, 11]:

$$P(\mathbf{X}_{j,k}) = \sum_{n=1}^{N_j} \omega_j^{(n)} \eta_j^{(n)}(\mathbf{X}_{j,k} \left| \boldsymbol{\mu}_j^{(n)}, \boldsymbol{\Sigma}_j^{(n)} \right.), \qquad (1)$$

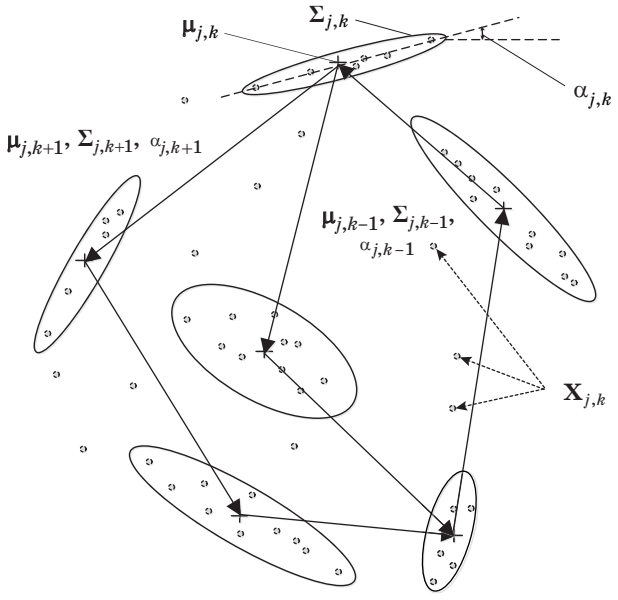
где n — номер гауссовского распределения в j-й проекции фазового пространства, $n = 1...N_j$, $\eta_j^{(n)}(\mathbf{X}_{j,k} \left| \boldsymbol{\mu}_j^{(n)}, \boldsymbol{\Sigma}_j^{(n)} \right.)$ — функция n-го гауссовского распределения в j-й проекции фазового пространства, $\boldsymbol{\mu}_j^{(n)}$ — вектор координат математического ожидания n-го распределения в j-й проекции фазового пространства, $\boldsymbol{\Sigma}_j^{(n)}$ — корреляционная матрица n-го распределения в j-й проекции фазового пространства, $\omega_j^{(n)}$ — вес n-го распределения в j-й проекции фазового пространства, N_j — количество гауссовских распределений в j-й проекции фазового пространства. Для весов $\omega_j^{(n)}$ выполняется условие нормировки $\sum_{n=1}^{N_j} \omega_j^{(n)} = 1$. Локализуемые области (кластеры) имеют форму эллипса со следующими характеристиками: координаты центра определяются вектором $\boldsymbol{\mu}_j^{(n)}$, размеры осей определяются элементами главной диагонали матрицы $\boldsymbol{\Sigma}_j^{(n)}$ и заданной доверительной вероятностью (обычно 0,96), угол наклона эллипса к оси абсцисс на фазовой плоскости — $\alpha_j^{(n)}$. Для создания полигауссовских моделей необходимо иметь представительную выборку значений временных рядов ТМП x_1, x_2, ... в отсутствие сильных возмущений. Количество устойчивых кластеров определяется путем многократной процедуры кластеризации и оптимизации эмпирической целевой функции, например, на основе критерия минимума внутриклассовой дисперсии точек [13].

Недостатком полигауссовских моделей вида (1) является то, что они не учитывают динамику изменения вектора состояния во времени. Учесть динамику изменения вектора состояния во времени можно путем перехода от полигауссовских моделей к динамическим моделям на основе ориентированных графов состояний, дуги которых соответствуют вероятностным переходам между кластерами полигауссовских моделей (1). Такие динамические модели могут повысить достоверность решения задачи идентификации технического состояния БА КА. При этом для решения задачи фильтрации и байесовского оценивания значений временных рядов ТМП достаточно иметь лишь информацию о координатах и параметрах кластеров модели в каждой j-й проекции фазового пространства в текущий и последующий моменты времени. Поэтому в обозначения кластеров и их параметров введем дополнительный индекс k, характеризующий номер дискретного отсчета времени, а верхний индекс n исключим из обозначения, так как текущее состояние

рассматриваемой динамической системы в k-й момент определяется совокупностью параметров $\boldsymbol{\mu}_{j,k}$, $\boldsymbol{\Sigma}_{j,k}$, $\alpha_{j,k}$. На рис. 1 приведен пример отображения такой динамической модели изменения состояния динамической системы в j-й проекции фазового пространства в виде ориентированного графа переходов между кластерами. Параметры кластеров и вероятностные переходы между ними определяют локальные свойства динамической системы в каждой j-й проекции фазового пространства в k-й момент времени, а сам граф характеризует ее глобальное поведение.

Достоинствами рассмотренной динамической модели с точки зрения автоматизированной обработки ТМИ являются: возможность однообразного представления функциональных, сигнальных и обобщенных ТМП, вторичных вычисляемых характеристик, а также время-событийной информации в виде набора унифицированных структур, связанных с ТС БА и подлежащих процедуре распознавания, наличие широкого спектра подходов к решению задачи идентификации технического состояния БА КА. В случае формализованного описания изменения ТМП в фазовом пространстве на основе байесовского подхода множество таких унифицированных динамических моделей можно рассматривать как динамическую байесовскую сеть доверия с циклами [14].

В данной статье рассмотрена возможность использования моделей изменения состояния динамических систем в фазовом пространстве на этапе первичной обработки ТМИ и последующей реализации методов линейной и нелинейной фильтрации ТМП.



■ ***Рис. 1.*** Ориентированный граф изменения состояния динамической системы в проекции фазового пространства

■ ***Fig. 1.*** Oriented graph of change in dynamical system state in projection of phase space

Фильтрация временных рядов значений телеметрируемых параметров с использованием парциальных нелинейных фильтров

Будем рассматривать фильтрацию временных рядов значений ТМП в каждой j-й проекции фазового пространства отдельно с последующей обработкой полученных результатов. Для постановки и решения задачи фильтрации временных рядов ТМП используем описание динамики функционирования БА КА уравнениями состояния и наблюдения в векторной форме [4, 6]:

$$\begin{aligned}\mathbf{Q}_{j,k} &= \mathbf{F}_j\left(\mathbf{Q}_{j,k-1}\right)+\mathbf{S}_{j,k},\\ \mathbf{X}_{j,k} &= \mathbf{H}_j\left(\mathbf{Q}_{j,k}\right)+\mathbf{V}_{j,k},\end{aligned} \qquad (2)$$

где $\mathbf{S}_{j,k}$ и $\mathbf{V}_{j,k}$ — векторы возмущающих воздействий и шумов измерений соответственно, $\mathbf{F}_j$ и $\mathbf{H}_j$ — функционалы, описывающие процесс изменения вектора состояния БА КА и процесс измерения значений ТМП.

С учетом особенностей рассмотренных унифицированных динамических моделей уравнения состояния и наблюдения (2) для каждой j-й проекции фазового пространства можно приближенно представить в следующем виде:

$$\begin{aligned}\mathbf{Q}_{j,k} &= \mathbf{Q}_{j,k-1}+(\boldsymbol{\mu}_{j,k}-\boldsymbol{\mu}_{j,k-1})+\mathbf{S}_k,\\ \mathbf{X}_{j,k} &= \mathbf{Q}_{j,k}+\mathbf{V}_k,\end{aligned} \qquad (3)$$

где $\mathbf{Q}_{j,k}$ — текущий вектор состояния БА в j-й проекции, $\mathbf{S}_{j,k}$ и $\mathbf{V}_{j,k}$ — векторы возмущений и шумов измерений в j-й проекции соответственно, $\mathbf{X}_{j,k}$ — вектор измеренных значений ТМП в j-й проекции. Вектор $\mathbf{S}_{j,k}$ определяется ошибками кластеризации и корреляционной матрицей $\mathbf{D}_{Sj}$, а вектор $\mathbf{V}_{j,k}$ — внешними помеховыми воздействиями и корреляционной матрицей $\mathbf{D}_{Vj}$. Для учета дисперсий компонент вектора $\mathbf{Q}_{j,k}$ при переходе от состояния к состоянию и ориентации кластера точек в проекции фазового портрета выражение (3) преобразуем к виду

$$\begin{aligned}\mathbf{Q}_{j,k} &= \mathbf{r}_{j,k}(\mathbf{Q}_{j,k-1}-\boldsymbol{\mu}_{j,k-1})+\boldsymbol{\mu}_{j,k}+\mathbf{S}_{j,k},\\ \mathbf{X}_{j,k} &= \mathbf{Q}_{j,k}+\mathbf{V}_{j,k},\end{aligned} \qquad (4)$$

где $\mathbf{r}_{j,k} = \begin{bmatrix} \dfrac{\sigma_{x,j,k}}{\sigma_{x,j,k-1}} & 0 \\ 0 & \dfrac{\sigma_{y,j,k}}{\sigma_{y,j,k-1}} \end{bmatrix} \begin{bmatrix} \cos(\Delta\alpha_k) & \sin(\Delta\alpha_k) \\ -\sin(\Delta\alpha_k) & \cos(\Delta\alpha_k) \end{bmatrix}$ —

вектор коэффициентов в каждой j-й проекции, $\sigma^2_{x,j,k}$ и $\sigma^2_{y,j,k}$ — диагональные элементы корреляционных матриц $\mathbf{\Sigma}_{j,k}$, $\Delta\alpha_k = \alpha_k - \alpha_{k-1}$ — угол поворота точек при переходе между состояниями, α_k и α_{k-1} — углы наклона эллипсов $(k-1)$-го и k-го состояний соответственно.

В выражениях (3) и (4)

$$\boldsymbol{\mu}_{j,k} = \begin{bmatrix} \mu_{x,j,k} \\ \mu_{y,j,k} \end{bmatrix}, \ \mathbf{S}_{j,k} = \begin{bmatrix} s_{x,j,k} \\ s_{y,j,k} \end{bmatrix}, \ \mathbf{V}_{j,k} = \begin{bmatrix} v_{x,j,k} \\ v_{y,j,k} \end{bmatrix},$$

$$\mathbf{D}_{Sj} = \begin{bmatrix} \sigma^2_{Sxj} & 0 \\ 0 & \sigma^2_{Syj} \end{bmatrix}, \ \mathbf{D}_{Vj} = \begin{bmatrix} \sigma^2_{Vxj} & 0 \\ 0 & \sigma^2_{Vyj} \end{bmatrix},$$

где $\sigma^2_{s,j} = \{\sigma^2_{Sxj}, \ \sigma^2_{Syj}\}$, $\sigma^2_{v,j} = \{\sigma^2_{Vxj}, \ \sigma^2_{Vyj}\}$ — соответствующие дисперсии шумов кластеризации и измерений для каждой j-й проекции, как правило, выполняется $\sigma_{v,j} >> \sigma_{s,j}$. Элементы векторов $\mathbf{S}_{j,k}$ и $\mathbf{V}_{j,k}$ в общем случае имеют негауссовское распределение. Наличие априорной информации о предельных нижних и верхних значениях ТМП в виде векторов $\mathbf{X}_н$ и $\mathbf{X}_в$ позволяет до этапа обработки ТМИ использовать двустороннее ограничение ТМП. В общем случае величины $\mathbf{X}_н$ и $\mathbf{X}_в$ могут определяться режимами работы БА КА, допусками на значения ТМП, пределами используемой телеметрической шкалы. При этом в уравнениях (4) проявляется определенная нелинейность: $\mathbf{X}_{j,k} = \mathbf{X}_н$, $\forall \mathbf{X}_{j,k} \le \mathbf{X}_н$; $\mathbf{X}_{j,k} = \mathbf{X}_в$, $\forall \ \mathbf{X}_{j,k} \ge \mathbf{X}_в$.

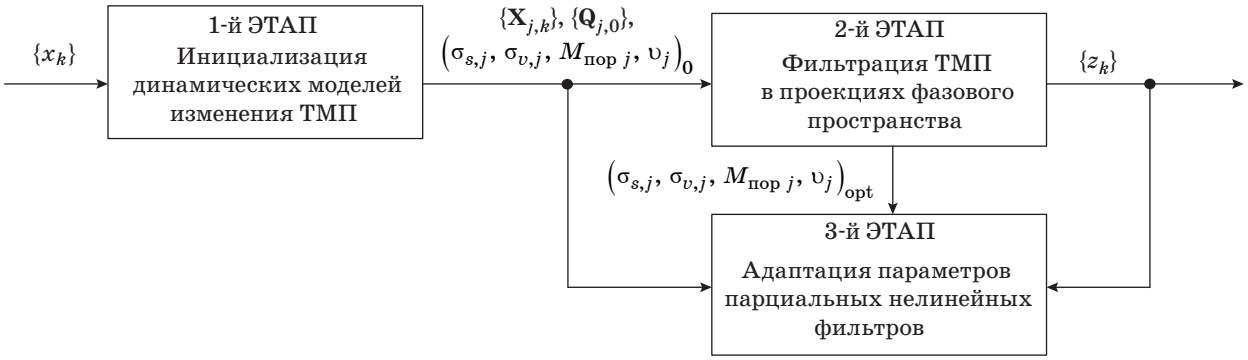
Результатом первичной обработки ТМИ является оценка вектора состояния $\widehat{\mathbf{Q}}_{j,k}$ в каждой проекции фазового пространства и формирование временного ряда обработанных значений ТМП z_1, z_2, ..., образующих множество $\{z_k\}$. Очевидно, что при малых уровнях возмущений обработанные значения временных рядов ТМП z_1, z_2, ..., будут определяться центрами соответствующих кластеров j-х динамических моделей на основе вычисления минимального расстояния между векторами $\mathbf{X}_{j,k}$ и всеми центрами кластеров $\boldsymbol{\mu}_{j,k}$ с использованием определенной метрики, например, эвклидовой или Махаланобиса. Данный метод обработки в дальнейшем будем использовать для сравнения с предлагаемым методом нелинейной парциальной фильтрации. Необходимость решения задачи фильтрации с использованием дополнительных алгоритмических средств возникает при значительных возмущениях, превосходящих размеры кластеров в унифицированных динамических моделях изменения ТМП. Если шумы измерений гауссовские, а модели анализируемых процессов являются линейными, то для решения задачи оптимального оценивания вектора состояния используют фильтр Калмана [15]. Для обработки реальных нелинейных процессов в негауссовских шумах используют модификации фильтра Калмана, а также методы парциальной нелинейной фильтрации [6, 16]. В зарубежной литературе такие фильтры получили название «partical filtr», а в отечественных публикациях — «многочастичный фильтр», «фильтр частиц» или «парциальный фильтр» [4–6, 16, 17]. Последнее название представляется наиболее удачным. Метод парциальной нелинейной фильтрации, основанный на стохастической аппроксимации функции плотности вероятности распределения вектора состояния, был впервые рассмотрен в работе [17]. В статье рассматривается развитие этого метода на основе использования набора из $(N-1)$ таких фильтров для обработки каждого ТМП в каждой j-й проекции фазового пространства в соответствии с унифицированными динамическими моделями и уравнениями вида (4).

Предлагаемый метод обработки ТМИ на основе парциальной нелинейной фильтрации ТМП в фазовом пространстве включает следующие основные этапы: инициализацию динамической модели изменения ТМП, фильтрацию ТМП в каждой проекции фазового портрета, адаптацию параметров фильтров в каждой проекции фазового портрета. Логика выполнения указанных этапов и входные и выходные данные каждого этапа показаны на рис. 2.

На первом этапе происходит преобразование временных рядов ТМП к требуемому для выполнения этапов фильтрации и адаптации виду, задаются начальные значения параметров парциальных нелинейных фильтров. На втором этапе реализуется обработка ТМП с использованием нелинейных парциальных фильтров в проекциях фазового пространства. Адаптация параметров фильтров на третьем этапе происходит итерационно. Для выполнения последующей итерации после выполнения третьего этапа осуществляется возврат к выполнению второго этапа с обновленными параметрами парциальных нелинейных фильтров. Процесс адаптации завершается после заданного числа итераций или на основании анализа текущей погрешности фильтрации. Рассмотрим содержание этапов более подробно.

На этапе инициализации модели происходит запуск процесса формирования из временного ряда ТМП $\{x_k\}$ множества векторов $\{\mathbf{X}_{j,k}\}$ для каждой j-й проекции фазового пространства. В отсутствие возмущений начальный и последующие векторы состояния могут быть определены с помощью унифицированных динамических моделей на основе вычисления минимального расстояния в фазовом пространстве между вектором измерений и центрами всех кластеров в каждой j-й проекции: $\mathbf{Q}_{j,0} = \boldsymbol{\mu}_{j,0}$, $\left\|\mathbf{X}_{j,k} - \boldsymbol{\mu}_{j,0}\right\| = \min\left\|\mathbf{X}_{j,k} - \boldsymbol{\mu}_{j,k}\right\|$.

Векторы $\mathbf{Q}_{j,0}$ образуют множество векторов $\{\mathbf{Q}_{j,0}\}$. На данном этапе также задаются начальные зна-



■ ***Рис. 2.*** Логика выполнения этапов метода обработки ТМИ
■ ***Fig. 2.*** The performing logic of TMI processing method stages

чения настраиваемых параметров парциальных нелинейных фильтров. В условиях возмущений осуществляется запуск процесса фильтрации ТМП в каждой проекции фазового портрета.

На этапе фильтрации ТМП для решения задачи оценивания вектора состояния используется принцип байесовского оценивания, основанный на рекуррентной процедуре предсказания и коррекции вектора состояния. Прогнозное значение вектора состояния определяется в соответствии с уравнением [6]

$$P(\mathbf{Q}_{j,k} / \mathbf{X}_{j,k-1}, \ldots, \mathbf{X}_{j,0}) = \\ = \int P(\mathbf{Q}_{j,k} / \mathbf{Q}_{j,k-1}) P(\mathbf{Q}_{j,k-1} / \mathbf{X}_{j,k-1}, \ldots, \mathbf{X}_{j,0}) \mathrm{d}\mathbf{Q}_{j,k-1}, \quad (5)$$

где $P(\mathbf{Q}_{j,k} / \mathbf{Q}_{j,k-1})$ — переходная плотность вероятности вектора состояния, определяемая уравнением состояния, $P(\mathbf{Q}_{j,k-1} / \mathbf{X}_{j,k-1}, \ldots, \mathbf{X}_{j,0})$ — условная плотность вероятности вектора состояния на предыдущем шаге, $P(\mathbf{Q}_{j,k} / \mathbf{X}_{j,k-1}, \ldots, \mathbf{X}_{j,0})$ — прогнозное значение условной плотности вероятности вектора состояния. После поступления нового вектора измеренных значений ТМП $\mathbf{X}_{j,k}$ экстраполированная плотность вероятности (5) корректируется с использованием правила Байеса [6]:

$$P(\mathbf{Q}_{j,k} / \mathbf{X}_{j,k}, \mathbf{X}_{j,k-1}, \ldots, \mathbf{X}_{j,0}) = \\ = C_j^{-1} P(\mathbf{X}_{j,k} / \mathbf{Q}_{j,k}) P(\mathbf{Q}_{j,k} / \mathbf{X}_{j,k-1}, \ldots, \mathbf{X}_{j,0}), \quad (6)$$

где $P(\mathbf{X}_{j,k} / \mathbf{Q}_{j,k})$ — функция правдоподобия, определяющая вероятность получения вектора значений наблюдаемого процесса $\mathbf{X}_{j,k}$ при заданном значении вектора состояния $\mathbf{Q}_{j,k}$, $C_j = P(\mathbf{X}_{j,k} / \mathbf{X}_{j,k-1}, \ldots, \mathbf{X}_{j,0})$ — нормализующая константа, не зависящая от $\mathbf{Q}_{j,k}$. Знание условной плотности распределения $P(\mathbf{Q}_{j,k} / \mathbf{X}_{j,k-1}, \ldots, \mathbf{X}_{j,0})$ позволяет вычислить оптимальную оценку $\hat{\mathbf{Q}}_{j,k}$. Практически применимые реализации указанного подхода получены лишь для ряда частных случаев, наиболее известным из которых является классический фильтр Калмана [15]. При этом в реальных приложениях используются приближенные и квазиоптимальные методы [6].

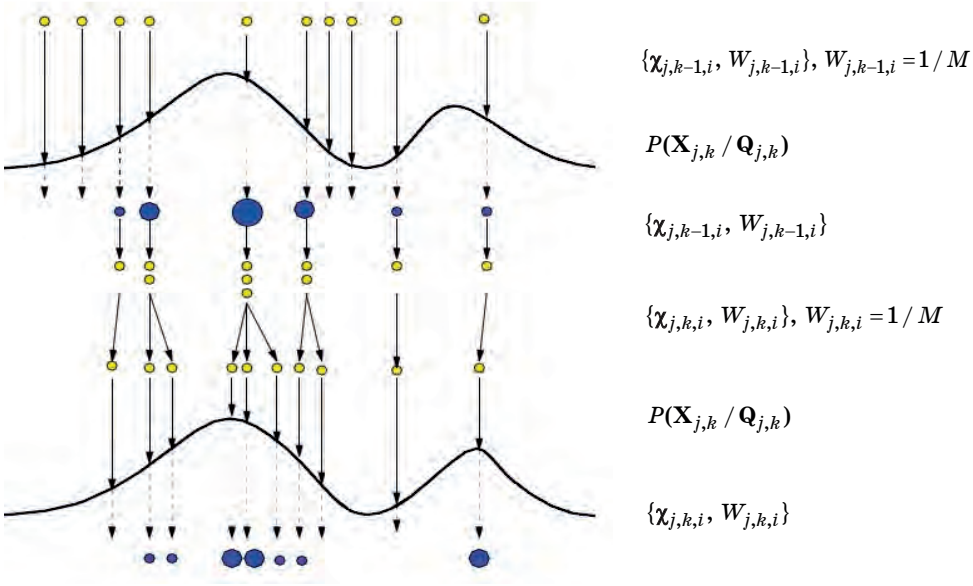
Парциальная нелинейная фильтрация является одним из таких методов и основана на аппроксимации апостериорной плотности вероятности оцениваемого вектора состояния $\mathbf{Q}_{j,k}$ набором M аппроксимирующих векторов («точек», «частиц») $\boldsymbol{\chi}_{j,k,i}$ с весами $W_{j,k,i}$, $i = 1, \ldots, M$. Размерность каждого i-го вектора $\boldsymbol{\chi}_{j,k,i}$ совпадает с размерностью вектора состояния $\mathbf{Q}_{j,k}$. При этом выражение (6) для такого фильтра принимает следующий вид [6]:

$$P(\boldsymbol{\chi}_{j,k,i} / \mathbf{X}_{j,k}, \mathbf{X}_{j,k-1}, \ldots, \mathbf{X}_{j,0}) \approx \\ \approx P(\mathbf{X}_{j,k} / \boldsymbol{\chi}_{j,k,i}) \sum_{i=1}^{M} W_{j,k,i} P(\boldsymbol{\chi}_{j,k,i} / \boldsymbol{\chi}_{j,k-1,i}). \quad (7)$$

Весовые функции $W_{j,k,i}$ нормируются таким образом, чтобы выполнялось условие $\sum_{i=1}^{M} W_{j,k,i} = 1$.

В результате нормировки весовых функций $W_{j,k,i}$ необходимость нормализующих констант C_j в выражении (7) отпадает. В начальный момент времени значения весов $W_{j,k,i}$, $i = 1, \ldots, M$ принимаются равными $1/M$. На этапе адаптации параметров фильтров происходит их изменение. Принцип адаптации «частиц» и их весов при аппроксимации произвольной плотности вероятности распределения вектора состояния поясняет рис. 3 [18]. Он основан на последовательном алгоритме выборки по значимости, в соответствии с которым на каждом шаге осуществляется оценивание функции правдоподобия $P(\mathbf{X}_{j,k} / \boldsymbol{\chi}_{j,k,i})$ и коррекция весовых коэффициентов каждого вектора аппроксимирующей выборки в соответствии со следующим выражением [6]:

$$W_{j,k,i} = W_{j,k-1,i} P(\mathbf{X}_{j,k} / \boldsymbol{\chi}_{j,k,i}). \quad (8)$$



■ ***Рис. 3.*** Принцип аппроксимации функции плотности вероятности распределения вектора состояния на основе последовательной выборки по значимости в *j*-й проекции фазового пространства

■ ***Fig. 3.*** The principle of approximation of state vector probability density function distribution on basis of sequential sampling on importance in *j*-th projection phase space

Однако в ходе работы алгоритма последовательной выборки по значимости наблюдается следующая проблема: после некоторого числа итераций веса всех аппроксимирующих векторов, за исключением одного, становятся пренебрежимо малыми. Данная проблема получила название «вырождение выборки».

Количественно степень вырождения может быть оценена величиной «эффективного размера аппроксимирующей выборки»

$$M_{eff\ j} = 1 / \sum_{i=1}^{M} W_{j,k,i}^2. \tag{9}$$

Для предотвращения вырождения аппроксимирующей выборки используются процедуры регенерации выборки, в процессе которых из аппроксимирующей выборки удаляются векторы с малыми весами, а векторы с большим весом заменяются группами векторов с меньшим весом таким образом, чтобы сохранились объем выборки и выборочная функция распределения [6, 18]. Процедура регенерации выборки осуществляется при равенстве значения $M_{eff\ j}$ некоторому пороговому значению $M_{\text{пор}\ j}$.

Оценка вектора состояния на *k*-м шаге определяется следующим образом:

$$\hat{\mathbf{Q}}_{j,k} = \sum_i W_{j,k,i} \chi_{j,k,i}. \tag{10}$$

Полученные в каждой *j*-й проекции результаты фильтрации $\hat{\mathbf{Q}}_{j,k}$, $\hat{\mathbf{Q}}_{j,k+1}$, …, $\hat{\mathbf{Q}}_{j,k+N-1}$ преобразуются в значения одномерного временного ряда $\mathbf{X}^*_{j,k}$. В ходе преобразования учитываются полученные на каждом *k*-м шаге значения ошибки фильтрации $E_{j,k} = \sqrt{(q_{x,j,k} - \mu_{x,j,k})^2 + (q_{y,j,k} - \mu_{y,j,k})^2}$, которая может рассматриваться в качестве целевой функции. Например, для проекции с номером $j = 2$ получаем следующую матрицу:

$$\begin{bmatrix} q_{x,2,k} & 0 & q_{y,2,k} & 0 & 0 & 0 & & 0 \\ 0 & q_{x,2,k+1} & 0 & q_{y,2,k+1} & 0 & 0 & & 0 \\ 0 & 0 & q_{x,2,k+2} & 0 & q_{y,2,k+2} & 0 & \ldots & 0 \\ 0 & 0 & 0 & q_{x,2,k+3} & 0 & q_{y,2,k+3} & & \\ & & & \ldots & & & & \\ 0 & 0 & 0 & 0 & 0 & 0 & & q_{y,2,k+N-1} \end{bmatrix}. \tag{11}$$

Матрица (11) преобразуется во временной ряд, имеющий следующий вид: $q_{x,2,k}$, $q_{x,2,k+1}$, $q_{y,2,k}$, $q_{x,2,k+3}$, $q_{y,2,k+2}$, $q_{y,2,k+3}$, ..., $q_{y,2,k+N-1}$. Затем полученные в каждой j-й проекции компоненты $\mathbf{X}_j^* = \{x_{j,k}^*\}$ преобразуются во временной ряд $\mathbf{Z} = \{z_k\}$ с учетом значений погрешности фильтрации $E_{j,k}$ в соответствии с выражением $z_k = x_{j,k}^*$, где j — номер проекции, для которой ошибка фильтрации $E_{j,k}$ минимальная, либо с использованием весовых коэффициентов υ_j:

$$z_k = \sum_{j=1}^{N-1} \upsilon_j x_{j,k}^*. \qquad (12)$$

Весовые коэффициенты υ_j нормируются таким образом, чтобы выполнялось условие $\sum_{j=1}^{N-1} \upsilon_j = 1$.

Элементы временного ряда z_1, z_2, ..., определяемые выражением (12), образуют множество $\{z_k\}$ и представляют собой отфильтрованные значения исходного временного ряда x_1, x_2, ..., анализируемого ТМП. Результаты фильтрации $z_k(\sigma_{s,j}, \sigma_{v,j}, M_{\text{пор } j}, \upsilon_j)$ являются функциями параметров $\sigma_{s,j}$, $\sigma_{v,j}$, $M_{\text{пор } j}$, υ_j используемых на текущей итерации парциальных фильтров.

На рис. 4 представлены экспериментальные зависимости в виде диаграмм, позволяющие судить о вкладе различных проекций в результат фильтрации ТМП. На рис. 4, *а* представлены весовые коэффициенты υ_j каждой j-й проекции, обеспечивающие свертку результатов фильтрации (12). На рис. 4, *б* представлена гистограмма распределения количества ситуаций, при которых погрешность фильтрации $E_{j,k}$ в j-й проекции была минимальной.

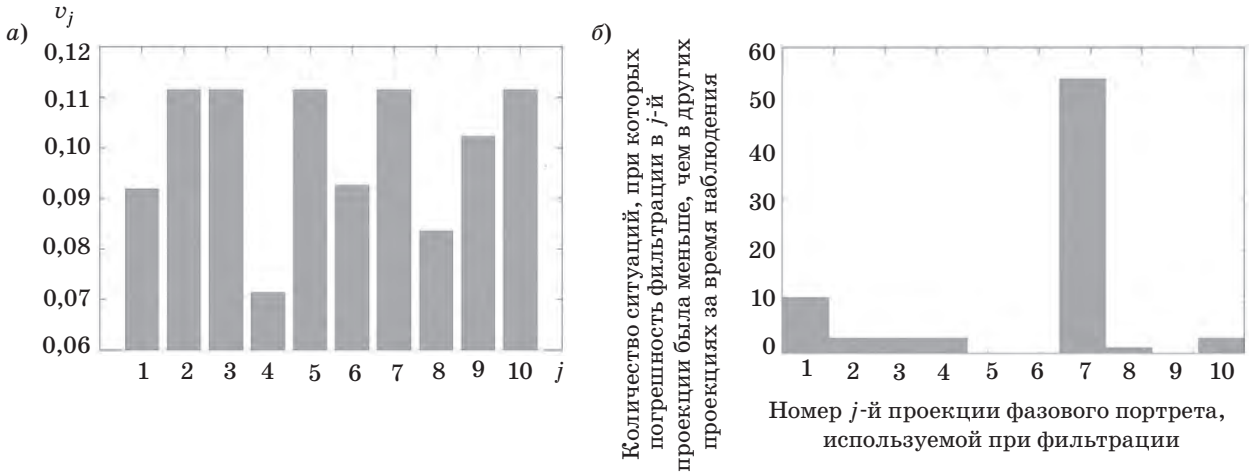
Поскольку в реальных ситуациях анализируемые процессы являются нестационарными, то для решения задачи фильтрации необходима адаптивная настройка параметров фильтров на основе имеющейся измерительной информации в пределах некоторого «окна наблюдения». В рассматриваемом методе парциальной фильтрации адаптация параметров фильтров реализуется на основе алгоритма SPSA [19]. Этап адаптации параметров фильтров реализуется параллельно основному вычислительному процессу фильтрации. В ходе фильтрации и адаптации происходит изменение параметров $\sigma_{s,j}$, $\sigma_{v,j}$, $M_{\text{пор } j}$, υ_j и уменьшается значение целевой функции:

$$E^2\left(\sigma_{s,j}, \sigma_{v,j}, M_{\text{пор } j}, \upsilon_j\right) = \frac{1}{N_{\text{н}} - 1} \times$$

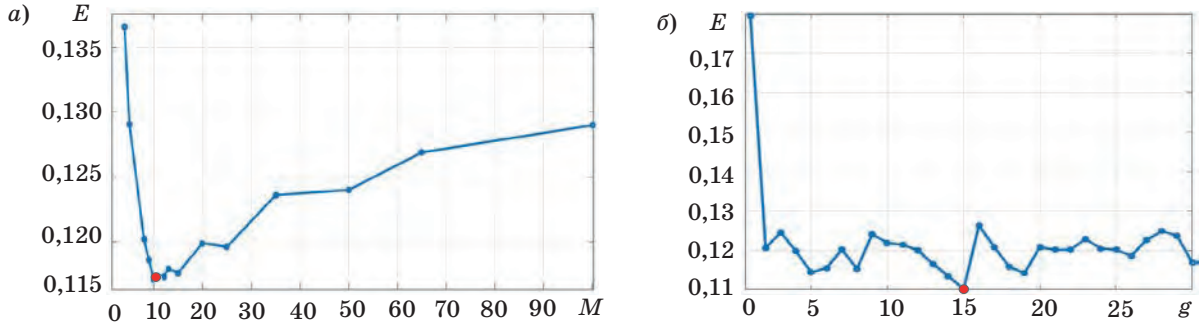
$$\times \sum_{k=1}^{N_{\text{н}}} \left[x_{\text{м}k}(\mu_{j,k}) - z_k\left(\sigma_{s,j}, \sigma_{v,j}, M_{\text{пор } j}, \upsilon_j\right) \right]^2, \qquad (13)$$

где $N_{\text{н}}$ — число отсчетов (размер) «окна наблюдения», $x_{\text{м}1}$, $x_{\text{м}2}$, ... — значения временного ряда, которым соответствуют координаты центров кластеров $\boldsymbol{\mu}_{j,k}$ динамических моделей изменения ТМП в фазовом пространстве. Набор оптимальных параметров $(\sigma_{s,j}, \sigma_{v,j}, M_{\text{пор } j}, \upsilon_j)_{\text{opt}}$ соответствует минимуму целевой функции, определяемой выражением (13) и имеющей смысл дисперсии погрешности фильтрации. Общее число настраиваемых параметров для $(N-1)$ фильтров составляет $4N-4$. Оптимизированные параметры $(\sigma_{s,j}, \sigma_{v,j}, M_{\text{пор } j}, \upsilon_j)_{\text{opt}}$ в дальнейшем используются в основном вычислительном процессе фильтрации.

На рис. 5 представлены экспериментальные зависимости величины E, соответствующей целевой функции E^2, определяемой выражением (13), от числа аппроксимирующих векторов M в каж-



■ ***Рис. 4.*** Диаграммы, иллюстрирующие вклад различных проекций в результат фильтрации ТМП
■ ***Fig. 4.*** Diagrams illustrating the contribution of different projections to the result of filtering TMP



■ ***Рис. 5.*** Погрешность фильтрации в зависимости от числа аппроксимирующих векторов и номера итерации процесса адаптации

■ ***Fig. 5.*** Filtration error depending on number of approximating vectors and adaptation process iteration number

дой j-й проекции и от номера итерации процедуры адаптации фильтров g. Величина E имеет смысл среднеквадратического отклонения погрешности фильтрации. На рис. 5, *а* представлена зависимость величины E, полученной по результатам множества экспериментов, от числа аппроксимирующих векторов M в каждой j-й проекции при заданном количестве итераций процедуры адаптации параметров фильтров $g = 30$. Данный пример иллюстрирует характерную для данной задачи особенность зависимости среднеквадратического отклонения погрешности фильтрации E от числа аппроксимирующих векторов M: начиная с некоторого значения дальнейшее его увеличение не приводит к уменьшению величины E, но требует значительных вычислительных затрат. На рис. 5, *б* представлена зависимость, характеризующая изменение величины E от номера итерации g процедуры адаптации параметров $\sigma_{s,j}$, $\sigma_{v,j}$, $M_{\text{пор } j}$, υ_j с использованием алгоритма SPSA [19]. Данный пример иллюстрирует возможности уменьшения среднеквадратического отклонения погрешности фильтрации E путем адаптивной подстройки параметров парциальных фильтров, а также колебательный характер изменения целевой функции в ходе итерационного процесса подстройки параметров фильтрации и необходимость принятия мер по определению момента остановки процесса адаптации.

Алгоритм парциальной нелинейной фильтрации телеметрируемых параметров

Основой предлагаемого метода первичной обработки ТМИ является алгоритм парциальной нелинейной фильтрации ТМП. Он реализуется на втором этапе рассматриваемого метода обработки ТМИ. Процесс реализации алгоритма включает следующие шаги:

Шаг 1. Вводим значения временного ряда x_1, x_2, ... анализируемого ТМП.

Шаг 2. Задаем размерность фазового пространства N. Вводим начальные значения параметров парциальных нелинейных фильтров: $(\sigma_{s,j}, \sigma_{v,j}, M_{\text{пор } j}, \upsilon_j)_0$, $j = 1, ..., (N - 1)$. Для всех j-х проекций фазового портрета выполняем шаги 3–12.

Шаг 3. Задаем номер дискретного отсчета времени $k = 0$. Загружаем начальные векторы состояния $\mathbf{Q}_{j,0}$, соответствующие координатам центров множества кластеров $\{\boldsymbol{\mu}_{j,k}\}$ в j-х проекциях фазового портрета.

Шаг 4. Задаем число аппроксимирующих векторов M и пороговое значение $M_{\text{пор } j}$. Создаем множество аппроксимирующих векторов $\{\boldsymbol{\chi}_{j,k,i}\}$ путем добавления к начальному вектору состояния $\mathbf{Q}_{j,0}$ значений вектора шумов измерений $\mathbf{V}_{j,k}$ с параметрами $\sigma_{v,j}$ из ковариационной матрицы $\mathbf{D}_{v,j}$: $\boldsymbol{\chi}_{j,k,i} = \mathbf{Q}_{j,k} + \mathbf{V}_{j,k}$. Для всех i-х аппроксимирующих векторов выполняем шаги 5–11, $i = 1, ..., M$.

Шаг 5. Создаем векторы весовых коэффициентов $\mathbf{W}_{j,k} = [W_{j,k,i}]$, $W_{j,k,i} = 1/M$.

Шаг 6. Увеличиваем номер дискретного отсчета времени: $k = k + 1$. Загружаем вектор значений ТМП $\mathbf{X}_{j,k}$. Для всех возможных фазовых переходов в соответствии с динамической моделью изменения ТМП выполняем шаги 7–11.

Шаг 7. Обновляем множество аппроксимирующих векторов $\{\boldsymbol{\chi}_{j,k,i}\}$ с использованием векторов коэффициентов $\mathbf{r}_{j,k}$, векторов возмущений $\mathbf{S}_{j,k}$ с параметрами $\sigma_{s,j}$ из ковариационной матрицы $\mathbf{D}_{s,j}$ и векторов шумов измерений $\mathbf{V}_{j,k}$ с параметрами $\sigma_{v,j}$ из ковариационной матрицы $\mathbf{D}_{v,j}$: $\boldsymbol{\chi}_{j,k,i} = \mathbf{r}_{j,k}(\boldsymbol{\chi}_{j,k-1,i} - \boldsymbol{\mu}_{j,k-1}) + \boldsymbol{\mu}_{j,k} + \mathbf{S}_{j,k} + \mathbf{V}_{j,k}$.

Шаг 8. Определяем функции правдоподобия $P(\mathbf{X}_{j,k}/\boldsymbol{\chi}_{j,k,i})$ и с учетом полученного значения функций правдоподобия пересчитываем вектор весовых коэффициентов $\mathbf{W}_{j,k} = [W_{j,k,i}]$.

Шаг 9. Вычисляем эффективный размер аппроксимирующей выборки $M_{eff\,j}$.

Шаг 10. Если $M_{eff\,j} \le M_{\text{пор } j}$, осуществляем регенерацию множества аппроксимирующих векторов $\{\boldsymbol{\chi}_{j,k,i}\}$, иначе переходим к шагу 11.

Шаг 11. Формируем на основе множества аппроксимирующих векторов $\{\boldsymbol{\chi}_{j,k,i}\}$ оценку вектора состояния $\hat{\mathbf{Q}}_{j,k}$. Вычисляем погрешность $E_{j,k}$ между полученной оценкой вектора состояния $\hat{\mathbf{Q}}_{j,k}$ и координатами центра кластера $\boldsymbol{\mu}_{j,k}$ для всех фазовых переходов динамической модели изменения ТМП. Выбираем наилучшее значение $\hat{\mathbf{Q}}_{j,k}$.

Шаг 12. Рассчитываем значения временного ряда z_1, z_2, ... отфильтрованных значений x_1, x_2, ... анализируемого ТМП.

Шаг 13. Проверяем условие: если временная реализация анализируемого ТМП не закончилась на дискретном отсчете времени k, то переходим на шаг 6, иначе заканчиваем выполнение алгоритма.

На этапе адаптации осуществляется настройка параметров парциальных нелинейных фильтров $\sigma_{s,j}$, $\sigma_{v,j}$, $M_{\text{пор }j}$, υ_j с целью достижения минимального значения целевой функции (13). Полученные на текущей итерации процесса адаптации оптимальные значения параметров $(\sigma_{s,j}, \sigma_{v,j}, M_{\text{пор }j}, \upsilon_j)_{\text{opt}}$ фильтров используются для обработки ТМП на следующей итерации процесса адаптации.

Примеры первичной обработки ТМИ с использованием парциальной нелинейной фильтрации

Рассмотрим результаты обработки модельного ТМП с известными характеристиками, имеющего нелинейный характер изменения во времени. Исследовались два метода обработки в фазовом пространстве анализируемых динамических процессов: 1) с использованием фильтрующих свойств самих динамических моделей изменения ТМП и эвклидовой метрики; 2) с использованием уравнений состояния и наблюдения и предлагаемого метода обработки ТМП на основе парциальных нелинейных фильтров. В качестве возмущений использовался гауссовский шум, а также смесь гауссовского шума, импульсных помех и возмущения типа «постоянное смещение». С учетом ограничения значений ТМП можно считать, что полученная смесь во всех случаях представляет собой негауссовскую помеху.

Моделирование осуществлялось при следующих условиях: отношение среднеквадратического значения шумовой помехи к уровню ТМП не превышало 10 дБ, импульсная помеха задавалась в виде последовательности аномальных отсчетов с вероятностью их возникновения 0,01, уровень помехи типа «постоянное смещение» соответствовал по амплитуде 10 % телеметрической шкалы. Рассчитанная размерность фазового пространства составила $N = 11$, исходя из этого использовалось 10 унифицированных динамических моделей поведения ТМП в двумерных проекциях фазового пространства. Число аппроксимирующих векторов в каждой j-й проекции фазового портрета составило $M = 50$, общее число аппроксимирующих векторов составило 500. По каждой реализации анализируемого ТМП было проведено не менее 10 экспериментов, число итераций процедуры адаптации параметров фильтров в каждом эксперименте было выбрано равным $g = 30$. В ходе решения задачи адаптивной фильтрации и минимизации целевой функции (13) синтезирован $(N - 1)$ парциальный нелинейный фильтр с параметрами $(\sigma_{s,j}, \sigma_{v,j}, M_{\text{пор }j}, \upsilon_j)_{\text{opt}}$. Так как в данной задаче исследовались модельные временные ряды ТМП, то для оценки качества фильтрации была использована среднеквадратическая погрешность фильтрации E_0 между обработанными и истинными значениями ТМП:

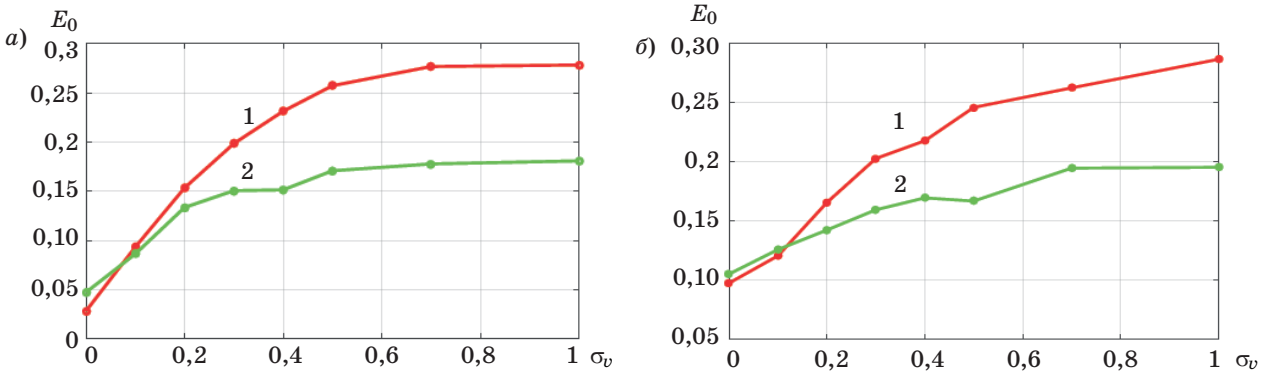
$$E_0\left(\sigma_{s,j}, \sigma_{v,j}, M_{\text{пор }j}, \upsilon_j\right) = \sqrt{\frac{1}{N_{\text{н}} - 1}\sum_{k=1}^{N_{\text{н}}}\left[x_{\text{и}k} - z_k\left(\sigma_{s,j}, \sigma_{v,j}, M_{\text{пор }j}, \upsilon_j\right)\right]^2}, \quad (14)$$

где $x_{\text{и}1}$, $x_{\text{и}2}$, ... — истинные значения временных рядов ТМП. Среднеквадратическая погрешность фильтрации E_0, определяемая выражением (14), отличается от величины E из выражения (13) использованием истинных значений $x_{\text{и}1}$, $x_{\text{и}2}$, ... временных рядов ТМП, которые в данном случае известны, так как ТМП являются модельными.

На рис. 6 представлены зависимости среднеквадратической погрешности E_0 отфильтрованных значений ТМП от среднеквадратического значения помехи σ_υ. Рис. 6, *а* соответствует фильтрации в условиях гауссовской помехи, рис. 6, *б* — фильтрации в условиях негауссовской помехи. Кривые *1* соответствуют фильтрации ТМП с использованием фильтрующих свойств модели и эвклидовой метрики, кривые *2* — фильтрации ТМП с использованием парциальных нелинейных фильтров. Ограничение величины E_0 с ростом σ_υ в обоих случаях объясняется использованием ограничителей амплитуды.

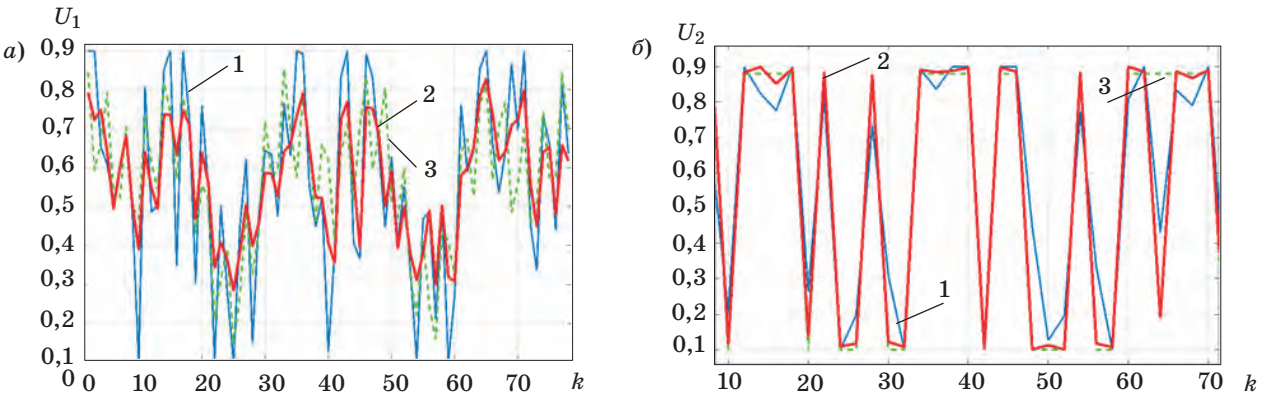
На рис. 7 представлены результаты обработки временных рядов значений типовых ТМП в условиях воздействия гауссовской помехи (при $\sigma_\upsilon = 0{,}2$) в виде отфильтрованных значений ТМП при парциальной нелинейной фильтрации. Рис. 7, *а* соответствует фильтрации функционального ТМП U_1, рис. 7, *б* — фильтрации сигнального ТМП U_2. Кривые *1* соответствуют измеренным значениям ТМП, кривые *2* — отфильтрованным значениям ТМП, кривые *3* — истинным значениям ТМП.

Анализ значений среднеквадратической погрешности фильтрации, достигнутой при обработке временных рядов значений функциональных ТМП, показал, что в рассмотренных условиях



■ ***Рис. 6.*** Зависимости среднеквадратической погрешности фильтрации ТМП в воздействии гауссовской (*а*) и негауссовской (*б*) помехи

■ ***Fig. 6.*** Mean square error of TMP filtering in influence of Gaussian (*a*) and non-Gaussian (*б*) interference



■ ***Рис. 7.*** Результаты парциальной нелинейной фильтрации типовых функционального ТМП (*а*) и сигнального ТМП (*б*)

■ ***Fig. 7.*** Results of partial nonlinear filtering of typical functional TMP (*a*) and signal TMP (*б*)

метод парциальной нелинейной фильтрации позволяет повысить достоверность оценки значений ТМП до 30 %. В качестве показателя достоверности использовалась доверительная вероятность принадлежности оценки ТМП к доверительному интервалу, который определялся среднеквадратической погрешностью фильтрации и соответствовал диапазону ее значений (1...2)E_0. Таким образом, предлагаемый метод позволяет обрабатывать разнородные ТМП в условиях возмущений и шумов измерений, описываемых в виде гауссовских и негауссовских помех. Подобная унификация очень важна для автоматизации первичной обработки ТМИ в интересах оценки состояния БА КА.

Заключение

Предложенный метод первичной обработки ТМИ на основе динамических моделей изменения ТМП и парциальной нелинейной фильтрации позволяет повысить достоверность обработки ТМИ и оперативно выявить изменения ТМП, существенные с точки зрения выявления нештатных и аварийных ситуаций на борту КА. В результате проведенных исследований продемонстрирована работоспособность метода для случая гауссовских и негауссовских помех в условиях, при которых нарушается устойчивость методов линейной фильтрации. Несмотря на то, что данный метод фильтрации требует значительных вычислительных ресурсов, это не является критичным для обработки ТМИ при использовании современных технологий распараллеливания вычислительных процессов [20]. Следует также отметить, что при практической реализации парциальной нелинейной фильтрации на результат обработки ТМП зачастую существенное влияние оказывают всего несколько проекций фазового портрета, что позволяет говорить о развитии предлагаемого метода в направлении создания стратегии выбора наилучших проекций на этапе адаптации параметров фильтрации и тем самым уменьшения его вычислительной сложности.

Литература

1. **Назаров А. В., Козырев Г. И., Шитов И. В.** Современная телеметрия в теории и на практике. СПб.: Наука и техника, 2007. 672 с.
2. **Соловьёв В. А., Лысенко Л. Н., Любинский В. Е.** Управление космическими полетами. Ч. 1. М.: МГТУ, 2009. 476 с.
3. **Соловьев С. В.** Интеллектуальный метод анализа для автоматизированного прогнозирования состояния КА. Инженерный журнал: наука и инновации. 2016. № 2. С. 1–13.
4. **Ширман Я. Д., Багдасарян С. Т., Маляренко А. С.** Радиоэлектронные системы: основы построения и теория. М.: Радиотехника, 2007. 512 с.
5. **Кольчугин Ю. И., Скоробогатов Е. Г.** Применение многочастичных фильтров в задачах оценивания нелинейных негауссовских процессов. Радиотехника. 2012. № 7. С. 91–95.
6. **Микаэльян С. В.** Методы фильтрации на основе многоточечной аппроксимации плотности вероятности оценки в задаче определения параметров движения цели при помощи измерителя с нелинейной характеристикой. Наука и образование. 2011. № 10. С. 1–25.
7. **Мальцев Г. Н., Назаров А. В., Якимов В. Л.** Исследование процесса диагностирования бортовой аппаратуры автоматических космических аппаратов с использованием дискретно-событийной имитационной модели. Труды СПИИРАН. 2018. № 1(56). С. 95–121.
8. **Малинецкий Г. Г., Потапов А. Б.** Современные проблемы нелинейной динамики. М., Едиториал УРСС, 2002. 360 с.
9. **Никульчев Е. В.** Геометрический подход к моделированию нелинейных систем по экспериментальным данным. М.: Изд-во МГУП, 2007. 162 с.
10. **Kugiumtzis D.** State space reconstruction parameters in analysis of chaotic time series — the role of time window length. Physica D: Nonlinear Phenomen. 1996, vol. 95, iss. 1, pp. 13–28.
11. **Lindgren A. C., Johnson M. T., Povinelli R. J.** Speech recognition using reconstructed phase space features. *IEEE International conference on acoustics, speech, and signal processing*. Hong Kong. China. 2003, vol. 1, pp. 61–63.
12. **Iverson D. L.** Inductive system health monitoring. *Proceedings of the 2004 International Conference on Artificial Intelligence* (IC-AI'04). CSREA Press, Las Vegas, NV, 2004.
13. **Граничин О. Н., Шалымов Д. С., Аврос Р., Волкович З.** Рандомизированный алгоритм нахождения количества кластеров. Автоматика и телемеханика. 2011. № 4. С. 86–98.
14. **Бидюк П. И., Терентьев А. Н., Гасанов А. С.** Построение и методы обучения байесовских сетей. Кибернетика и системный анализ. 2005. № 4. С. 133–147.
15. **Браммер К., Зиффлинг Г.** Фильтр Калмана-Бьюси. М.: Наука, 1982. 200 с.
16. **Кудрявцева И. А.** Анализ эффективности расширенного фильтра Калмана, сигма-точечного фильтра Калмана и сигма-точечного фильтра частиц. Научный вестник МГТУ ГА. 2016. № 224. С. 43–52.
17. **Gordon N. J., Salmond D. J., Smith A. F. M.** Novel approach to nonlinear/non-gaussian bayesian state estimation. IEEE Proceedings F, Radar and Signal Processing. IET. 1993, vol. 140, no. 2, pp. 107–113.
18. **Doucet A., De Freitas N., Gordon N.** Sequential monte carlo methods in practice. New York, Springer-Verlag, 2001. 583 p.
19. **Граничин О. Н., Поляк Б. Т.** Рандомизированные алгоритмы оценивания и оптимизации при почти произвольных помехах. М.: Наука, 2003. 291 с.
20. **Гергель В. П.** Теория и практика параллельных вычислений. М.: НОУ «Интуит», 2016. 500 с.

UDC 629.78
doi:10.31799/1684-8853-2018-5-22-34

Primary processing of telemetric information using dynamic models of parameter change and partial nonlinear filtration

G. N. Maltsev[a], Dr. Sc., Tech., Professor, orcid.org/0000-0002-6755-5700, georgy_maltsev@mail.ru
V. L. Yakimov[a], PhD, Tech., Associate Professor, orcid.org/0000-0001-9721-2453
S. V. Soloviev[b], PhD, Tech., Associate Professor, orcid.org/0000-0002-0391-8728
N. V. Lebedeva[b], Assistant, orcid.org/0000-0001-6963-6638
[a]A. F. Mozhaiskii Military Space Academy, 13, Zhdanovskaia Emb., 197198, Saint-Petersburg, Russian Federation
[b]Bauman Moscow State Technical University, 5, Build. 1, 2-nd Bauman St., 105005, Moscow, Russian Federation

Introduction: At the stage of preliminary processing of telemetric information about the technical state of on-board spacecraft equipment, there is a need to detect changes in telemetry parameters which are important from the viewpoint of detecting and fixing emergencies. The solution of this problem is complicated by nonlinear character of change in the majority of on-board equipment functioning parameters and by various perturbations which cause changes in the analyzed telemetric parameters. **Purpose:** Developing a method of telemetry data processing in order to increase the reliability of heterogeneous telemetry information from spacecrafts under real perturbation and nonlinear changes in their parameters, based on filtering the time series of telemetric parameter values.

Results: A method of telemetry information processing is proposed, based on dynamic models of change in telemetric parameters of on-board spacecraft equipment in phase space projections and partial nonlinear filtering of the analyzed time series, using stochastic approximation of state vector distribution probability density function. We discuss the advantages of primary processing of telemetric parameters of spacecraft airborne equipment in multidimensional phase space, provided by the possibility of selecting the most significant phase portrait projections from the viewpoint of filtering results. The developed algorithm of non-linear partial filtration allows you to adapt to the perturbation parameters. The validity of this method has been demonstrated for Gaussian and non-Gaussian perturbations at conditions under which the stability of linear filtration methods is violated. **Practical relevance:** The proposed approach is unified for processing of dissimilar telemetry parameters, providing an increase in estimation reliability of both functional and signal telemetric parameters of spacecrafts under real operating conditions.

Keywords — telemetry, telemetry parameter, dynamic process, phase space, partial nonlinear filtering.

Citation: Maltsev G. N., Yakimov V. L., Soloviev S. V., Lebedeva N. V. Primary processing of telemetric information using dynamic models of parameter change and partial nonlinear filtration. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 5, pp. 22–34 (In Russian). doi:10.31799/1684-8853-2018-5-22-34

References

1. Nazarov A. V., Kozyrev G. I., Shitov I. V. *Sovremennaia telemetriia v teorii i na praktike* [Modern telemetry in theory and in practice]. Saint Petersburg, Nauka i Tekhnika Publ., 2007. 672 p. (In Russian).
2. Soloviev V. A., Lysenko L. N., Liubinskii V. E. *Upravlenie kosmicheskimi poletami* [Mission control]. Part 1. Moscow, MGTU imeni N. E. Baumana Publ., 2009. 476 p. (In Russian).
3. Soloviev S. V. Intellectual Method of Analysis for Automated Forecasting of Spacecraft State. *Inzhenernyj zhurnal: nauka i innovacii*, 2016, no. 2, pp. 1–13 (In Russian).
4. Shirman Ya. D., Baghdasaryan S. T., Malyarenko A. S. *Radioelektronnye sistemy: osnovy postroeniya i teoriya*. [Radio-electronic systems: fundamentals of construction and theory.]. Moscow, Radiotehnika Publ., 2007. 512 p. (In Russian).
5. Kolchugin Yu. I., Skorobogatov E. G. Application of multiparticle filters in problems of nonlinear non-gaussian processes estimation. *Radiotekhnika*, 2012, no. 7, pp. 91–95 (In Russian).
6. Mikaelyan S. V. Methods of filtration based on multipoint approximation of estimation probability density in problem of determining the target motion using parameters a meter with nonlinear characteristic. *Nauka i obrazovanie*, 2011, no.10, pp.1–25 (In Russian).
7. Maltsev G. N., Nazarov A. V., Yakimov V. L. The Simulation of on-board equipment diagnosing process of spacecraft with high level autonomy functioning. *Trudy SPIIRAN*, 2018, vol. 1(56), pp. 95–121 (In Russian).
8. Malinetskii G. G., Potapov A. B. *Sovremennye problemy nelineinoi dinamiki* [Modern problems of nonlinear dynamics]. Moscow, Editorial URSS Publ., 2002. 360 p. (In Russian).
9. Nikulchev E. V. *Geometricheskij podhod k modelirovaniju nelinejnyh sistem po jeksperimental'nym dannym* [Geometric approach to modeling of nonlinear systems from experimental data]. Moscow, MGUP Publ., 2007. 162 p. (In Russian).
10. Kugiumtzis D. State space reconstruction parameters in analysis of chaotic time series — the role of time window length. *Physica D: Nonlinear Phenomen*, 1996, vol. 95, iss. 1, pp. 13–28.
11. Lindgren A. C., Johnson M. T., Povinelli R. J. Speech recognition using reconstructed phase space features. *IEEE International Conference on Acoustics, Speech, and Signal Processing*, Hong Kong, China, 2003, vol. 1, pp. 61–63.
12. Iverson D. L. Inductive system health monitoring. *Proceedings of the 2004 International Conference on Artificial Intelligence* (IC-AI'04). CSREA Press, Las Vegas, NV, 2004.
13. Granichin O. N., Shalymov D. S., Avros R., Volkovich Z. Randomized algorithm for finding the number of clusters. *Avtomatika i telemehanika*, 2011, no. 4, pp. 86–98 (In Russian).
14. Bidyuk P. I., Terentyev A. N., Gasanov A. S. Construction and methods of training Bayesian networks. *Kibernetika i sistemnyi analiz*, 2005, no. 4, pp. 133–147 (In Russian).
15. Brammer K., Siffling G. *Fil'tr Kalmana-B'jusi* [The Kalman-Buschy filter]. Moscow, Nauka Publ., 1982. 200 p. (In Russian).
16. Kudryavtseva I. A. Analysis of the efficiency of the extended Kalman filter, the Kalman sigma-point filter, and the sigma-point particle filter. *Nauchnyj vestnik MGTU GA*, 2016, no. 224, pp. 43–52 (In Russian).
17. Gordon N. J., Salmond D. J., Smith A. F. M. Novel approach to nonlinear/non-gaussian bayesian state estimation. *IEEE Proceedings F, Radar and Signal Processing*, IET, 1993, vol. 140, no. 2, pp. 107–113.
18. Doucet A., De Freitas N., Gordon N. *Sequential monte carlo methods in practice*. New York, Springer-Verlag, 2001, 583 p.
19. Granichin O. N., Polyak B. T. *Randomizirovannye algoritmy ocenivanija i optimizacii pri pochti proizvol'nyh pomehah* [Randomized algorithms for estimation and optimization under almost arbitrary interference]. Moscow, Nauka Publ., 2003. 291 p. (In Russian).
20. Gergel V. P. *Teoriya i praktika parallel`nyx vychislenij* [Theory and practice of parallel computing]. Moscow, NOU «Intuit» Publ., 2016. 500 p. (In Russian).

УДК 004.732
doi:10.31799/1684-8853-2018-5-35-43

Научные статьи
Articles

Статистические методы исследования сетевого трафика

Т. М. Татарникова[a], *доктор техн. наук, доцент, orcid.org/ 0000-0002-6419-0072, tm-tatarn@yandex.ru*
[a]*Санкт-Петербургский государственный университет аэрокосмического приборостроения, Б. Морская ул., 67, Санкт-Петербург, 190000, РФ*

Постановка проблемы: *концепция единой мультисервисной сети, предполагающая интеграцию передачи речи, данных и мультимедиа, побудила интерес к изучению природы сетевого трафика. Исследования трасс трафика, записанного в крупных масштабах времени, показывают наличие в нем самоподобной структуры, что требует пересмотра результатов моделирования инфокоммуникационных сетей в предположении о пуассоновском потоке данных.* ***Цель исследования:*** *с целью получения инструментальных средств для генерации искусственного трафика, адекватно отражающего реальный трафик сети с учетом выявленных свойств самоподобия, изучить последовательность применения методов изучения природы сетевого трафика, выявляющих самоподобную природу трафика в виде статистических оценок и показателя Херста.* ***Результаты:*** *проверены свойства самоподобия рассматриваемого 3G-трафика на разных временных шкалах, полученных агрегацией по 5, 10, 15 и 20 мин на имеющихся суточных данных 3G-трафика. Получена оценка тяжести «хвоста» распределения самоподобного трафика методом построения линии регрессии для дополнительной функции распределения в логарифмическом масштабе. Значение параметра самоподобия, определяемое тяжестью «хвоста» распределения, позволило подтвердить предположение о самоподобии 3G-трафика. Выполнен обзор моделей, имитирующих реальный сетевой трафик с самоподобной структурой. Реализованы инструментальные средства для генерации искусственного трафика в соответствии с рассмотренными моделями. Выполнено сравнение генераторов искусственного сетевого трафика по критерию метода наименьших квадратов аппроксимации точечных значений искусственного трафика аппроксимирующей функцией 3G-трафика. Учтены качественные оценки генераторов искусственного сетевого трафика в виде трудоемкости их программной реализации, что, впрочем, может быть субъективной оценкой. Сравнительные характеристики позволяют выбрать генератор(ы), которые максимально правдоподобно имитируют реальный сетевой трафик.* ***Практическая значимость:*** *предложенная последовательность методов исследования свойств сетевого трафика необходима для понимания его природы и разработки соответствующих моделей, имитирующих реальный сетевой трафик.*

Ключевые слова *— сетевой трафик, самоподобный процесс, долговременная зависимость, медленно затухающая дисперсия, распределение с тяжелым «хвостом», параметр самоподобия, модели имитации самоподобного трафика.*

Цитирование: Татарникова Т. М. Статистические методы исследования сетевого трафика. *Информационно-управляющие системы*, 2018, № 5, с. 35–43. doi:10.31799/1684-8853-2018-5-35-43
Citation: Tatarnikova T. M. Statistical methods for studying network traffic. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 5, pp. 35–43 (In Russian). doi:10.31799/1684-8853-2018-5-35-43

Введение

Моделирование оценивания характеристик обслуживания сетевого трафика остается в настоящее время актуальной научной задачей. Достоверные оценки сетевого трафика необходимы при планировании развития телекоммуникационных сетей, выборе политики дифференцированного обслуживания и характеристик вычислительных ресурсов, гарантирующих требуемое качество обслуживания при соответствующей сетевой нагрузке [1, 2].

Вспыхнувший интерес к изучению природы сетевого трафика объясняется результатами исследований, показавших наличие в трафике долговременных зависимостей или процесса самоподобия. Эти изменения в структуре трафика связаны с реализацией концепции единой мультисервисной сети, предполагающей интеграцию передачи речи, данных и мультимедиа [3, 4].

На сегодняшний день теория самоподобных стохастических процессов не так хорошо развита, как теория пуассоновских процессов. Учитывая известные выводы о самоподобности сетевого трафика, актуальными задачами становятся методы его изучения и разработка инструментальных средств для генерации искусственного трафика, адекватно отражающего реальный неоднородный трафик сети [5, 6].

Свойства и характеристики самоподобных процессов

Самоподобность описывает явление, при котором некоторые статистические характеристики процесса сохраняются при масштабировании времени. При усреднении по шкале времени у самоподобного процесса не происходит быстрого «сглаживания», т. е. сохраняется склонность к всплескам [7, 8].

К свойствам, характеризующим самоподобность процесса, относят такие, как медленно затухающая дисперсия, долговременная зависи-

мость, наличие распределения с тяжелыми «хвостами» [9–11].

Свойство медленно затухающей дисперсии заключается в том, что дисперсия выборочного среднего затухает медленнее, чем величина, обратная размеру выборки, т. е.

$$D\left(X^{(n)}(t)\right)=\sigma^2 n^{2H-2},\ n\to\infty, \qquad (1)$$

где σ^2 — дисперсия процесса $X(t)$; n — объем выборки; H — параметр Херста (параметр самоподобия), $0{,}5 < H < 1$.

Отметим, что для традиционных случайных процессов дисперсия выборочного среднего уменьшается обратно пропорционально объему выборки: $D\left(X^{(n)}(t)\right)=\dfrac{\sigma^2}{n}$.

Наличие долговременной зависимости заключается в том, что самоподобный процесс обладает гиперболически затухающей корреляционной функцией

$$R(k)\cong k^{(2H-2)}L(k),\ \forall k\geq 1,\ k\to\infty, \qquad (2)$$

где $L(k)$ — медленно меняющаяся функция на бесконечности, для которой

$$\lim_{k\to\infty}\frac{L(kx)}{L(k)}=1 \text{ для всех } x>0. \qquad (3)$$

Свойство наличия распределения с тяжелым «хвостом» заключается в том, что случайная величина X имеет распределение с тяжелым «хвостом», если [12]

$$P\left(X>x\right)\sim cx^{-\alpha},\ x\to\infty, \qquad (4)$$

где $0 < \alpha < 2$ — параметр формы распределения; c — некоторая положительная константа.

Методы исследования самоподобного процесса

Существует ряд приемов, позволяющих выполнить проверку свойства самоподобия исследуемого процесса [13–17].

Эффект самоподобия можно наблюдать на графиках (рис. 1–4), иллюстрирующих изменение временной шкалы, при котором структура ряда, полученного усреднением групп элементов, остается такой же, как и структура исходного. Этот факт является предпосылкой для предположения о самоподобной структуре рассматриваемого процесса и основанием для проведения дальнейшего более полного анализа [18].

Рассмотрим суточные данные 3G-трафика от 20.08.2018 г., предоставленные сотовым оператором МТС г. Санкт-Петербурга (рис. 1). Временной ряд состоит из 1440 наблюдений, каждое из которых представляет собой объем эксплуатируемой услуги X_t Гбайт, за время $t=\overline{1{,}1440}$ мин.

Выполним агрегацию временного ряда, уменьшив размер шкалы наблюдений в 5 раз. Значения нового временного ряда получены в соответствии со следующим выражением:

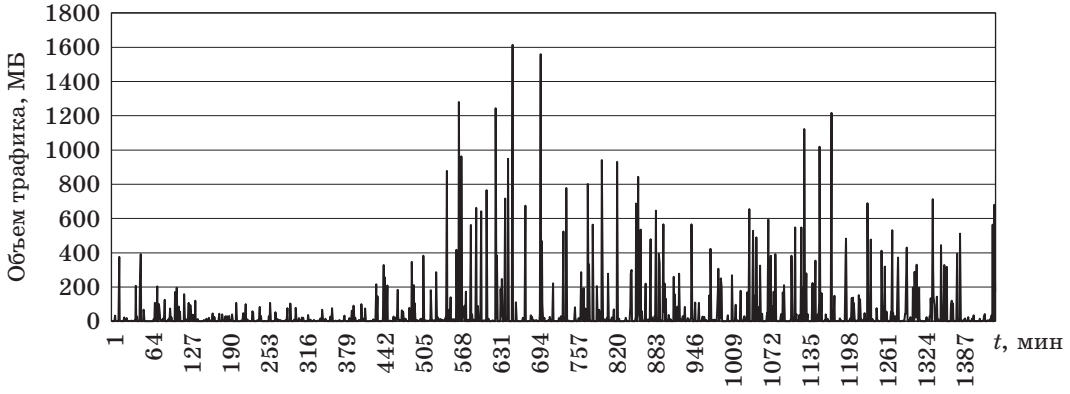
$$X_t=\frac{\sum\limits_{i=(t-1)m+1}^{tm}X_i}{m}, \qquad (5)$$

где m — число усредняемых последовательных членов ряда.

Новый ряд включает 288 событий и приведен на рис. 2.

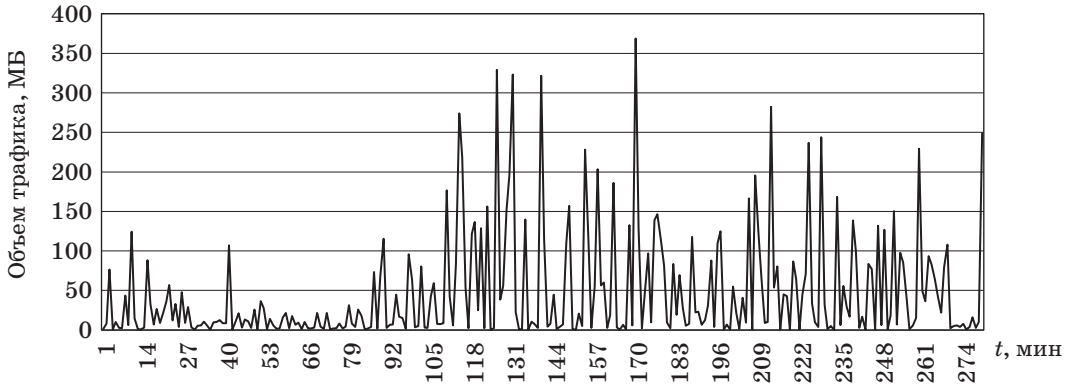
Та же процедура с уменьшением размера временной шкалы исходного ряда наблюдений в 10 раз приводит к результату, представленному на рис. 3.

Уменьшение размера временной шкалы исходного ряда наблюдений в 20 раз приводит к результату, представленному на рис. 4.

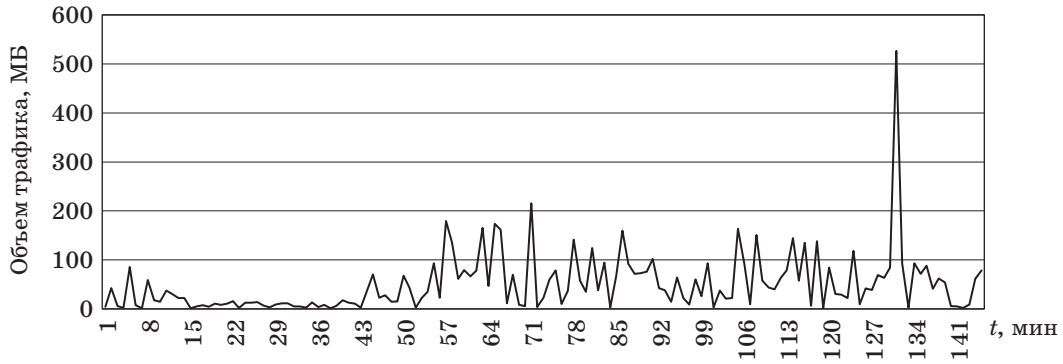


■ ***Рис. 1.*** Исходный 3G-трафик за период 1440 мин

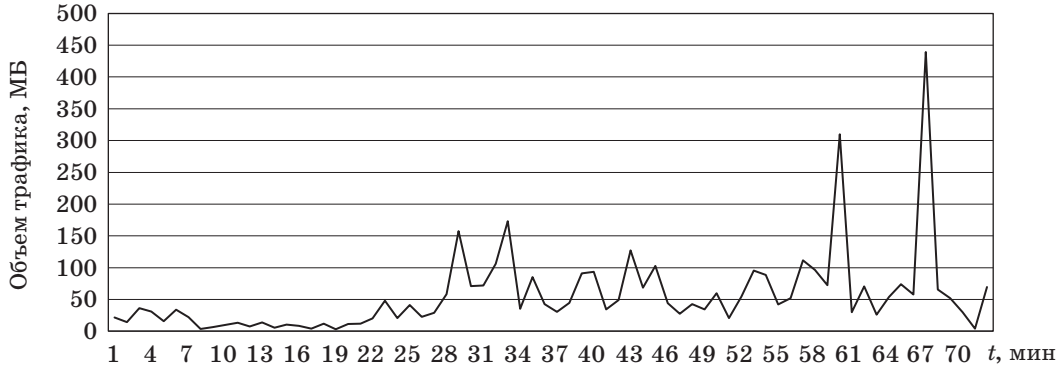
■ ***Fig. 1.*** Initial 3G-traffic period of 1440 min



■ ***Рис. 2.*** Агрегация 3G-трафика по 5 мин за период 1440 мин
■ ***Fig. 2.*** Aggregation 3G-traffic over 5 min for the period of 1440 min



■ ***Рис. 3.*** Агрегация 3G-трафика по 10 мин за период 1440 мин
■ ***Fig. 3.*** Aggregation 3G-traffic over 10 min for the period of 1440 min



■ ***Рис. 4.*** Агрегация 3G-трафика по 20 мин за период 1440 мин
■ ***Fig. 4.*** Aggregation 3G-traffic over 20 min for the period of 1440 min

Визуальное наблюдение агрегированных процессов (см. рис. 1–4) позволяет сделать вывод о сохранении структуры процесса.

Далее необходимо оценить тяжесть «хвоста» распределения — показатель α. С этой целью построим график дополнительного распределения $\bar{F}(x) = 1 - F(x) = P(X > x)$ из непересекающихся групп, объединяющих события исходного процесса, и вычислим частоты попадания в каждый интервал. Построим новый ряд, вычислив среднее значение в каждом интервале, и соответствующую частоту. Полученная функция рас-

пределения позволяет вычислить значения $\bar{F}(x)$ и построить ее график в логарифмической шкале (рис. 5).

Построим линию регрессии для дополнительной функции распределения (рис. 6) и, вычислив тангенс угла наклона к горизонтальной оси, получим значение параметра α

$$\alpha = \frac{\partial \log \bar{F}(x)}{\partial \log x}, \qquad (6)$$

через который выражается параметр самоподобия $H = \frac{3-\alpha}{2}$.

Показатель α может принимать значения на интервале [0, 2]. Чем меньше это значение, тем тяжелее «хвост» распределения [18].

Распределение с тяжелым «хвостом» имеет ряд свойств, которые существенно отличают его от наиболее известных распределений, таких как экспоненциальное, нормальное или пуассоновское [19].

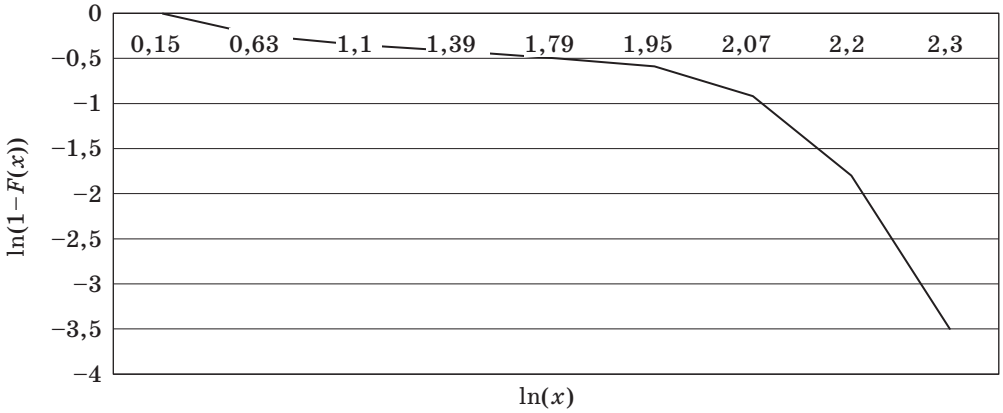
Если $\alpha \le 2$, тогда распределение имеет бесконечную дисперсию.

Если $\alpha \le 1$, то распределение имеет бесконечное среднее.

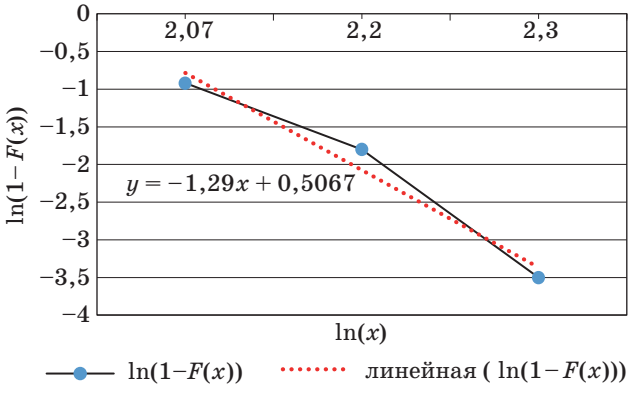
При уменьшении α произвольная большая порция плотности может быть представлена в «хвосте» распределения. Фактически тяжелый «хвост» означает наличие бесконечной дисперсии, иначе говоря, случайная величина может принимать очень большие значения, но с очень маленькой вероятностью.

Выведенное уравнение регрессии показывает, что α принимает значение, равное 1,29, и $\alpha \in [0; 2]$, откуда следует, что распределение данных 3G-трафика имеет свойство тяжелого «хвоста».

Вычислим значение параметра самоподобия $H = (3 - 1{,}29)/2 = 0{,}855$, что также подтверждает свойства самоподобия рассматриваемого процесса, так как $H \in [0{,}5; 1]$. Рассчитанный параметр H будем использовать при увеличении масштаба шкалы. В табл. 1 приведены результаты оценки дисперсии, среднего и коэффициента корреля-



■ ***Рис. 5.*** $\bar{F}(x)$ «хвоста» распределения для непересекающихся групп данных
■ ***Fig. 5.*** $\bar{F}(x)$ tail distribution for disjoint data groups



■ ***Рис. 6.*** Линия регрессии и вывод уравнения регрессии
■ ***Fig. 6.*** The regression line and regression equation

■ ***Таблица 1.*** Результаты оценки свойств 3G-трафика
■ ***Table 1.*** 3G-traffic properties evaluation results

Трафик	$D(X)$	$M(X)$	$R(k)$, $k = 2$
3G исходный	4149,47	37,82	0,0233
3G агрегированный по 5 мин	4341,17	37,25	0,0712
3G агрегированный по 10 мин	3354,93	40,39	0,0678
3G агрегированный по 20 мин	3991,34	38,21	0,0349

ции для исходного и агрегированного трафиков, демонстрирующих свойства распределений с тяжелыми «хвостами».

Модели имитации самоподобного трафика

Свойства, полученные в результате измерений реального сетевого трафика, часто отличаются от свойств трафика, полученного традиционными моделями, в частности моделью пуассоновского потока [20]. Основной причиной такого различия является структура, лежащая в основе зависимости.

Пакетный трафик, представляющий собой интеграцию речи, данных и мультимедиа, охватывает сильно отличающиеся масштабы времени — от микросекунд до секунд и даже минут [1]. В каждом масштабе времени присутствует долговременная зависимость — наличие пульсаций, которые состоят из периодов активности, разделенных менее активными периодами. В традиционных моделях сетевого трафика пульсации сильно сглажены на больших масштабах времени, отчего свойство долговременной зависимости пропадает [21].

В результате оказывается, что традиционный анализ инфокоммуникационных систем, в основе которого лежит предположение о пуассоновском потоке, не может достаточно точно оценить объемы вычислительных ресурсов и производительность системы в условиях пульсирующего трафика.

Необходимы инструментальные средства для генерации искусственного трафика, соответствующего свойствам реального сетевого трафика, которые можно использовать при моделировании процессов передачи, сохранения и обработки сетевого трафика.

Известны всего несколько моделей, предназначенных для имитации самопоподобного трафика. Анализ доступных публикаций по моделированию самоподобного трафика позволяет выделить модели [22], приведенные в табл. 2.

В работе реализованы инструментальные средства для генерации искусственного трафика по моделям, перечисленным в табл. 2. Срав-нительные характеристики позволяют выбрать генераторы, которые максимально правдоподобно имитируют реальный сетевой трафик. При сравнении оценивался критерий S метода наименьших квадратов аппроксимации точечных значений искусственного трафика аппроксимирующей функцией 3G-трафика

$$S = \sum_{i=1}^{N} \left(F(x_i) - y_i \right)^2, \qquad (7)$$

где $F(x_i)$ — значения аппроксимирующей функции в узловых точках x_i искусственного трафика; y_i — заданный массив исходного 3G-трафика в узловых точках x_i.

За узловую точку принята каждая 60-я минута, итого 24 часа — 24 узловые точки.

Кроме количественной оценки, в табл. 2 приведены и качественные оценки в виде трудоемкости реализации программного генератора (число настраиваемых параметров или необходимость обучения). Это субъективная оценка, которую труд-

■ ***Таблица 2***. Модели генераторов самоподобного трафика
■ ***Table 2***. Models of self-similar traffic generators

Модель	Математическое описание модели	Число настраиваемых параметров модели или обучения	S
Фрактальное броуновское движение — процесс движения частицы, совершающей хаотическое перемещение с шагом, задаваемым предысторией движения	$B_H(t) = \frac{1}{\Gamma(H-1/2)} \int_{-\infty}^{t} K(t-t')dB(t')$, где $\Gamma(\cdot)$ — гамма-функция; $dB(t')$ — независимые случайные перемещения частицы в моменты времени t'; $K(t-t')$ — функция памяти системы: $K(t-t') = \begin{cases} (t-t')^{H-1/2}, & 0 \le t' \le t \\ (t-t') - (-t')^{H-1/2}, & t' \le 0 \end{cases}$	Выбор и настройка функции $K(t-t')$	0,324
Фрактальный гауссовский шум — итерационный процесс последовательного деления отрезка единичной длины пополам	$y_c = \frac{(y_1 + y_2)}{2} + h$, где h — случайная величина, распределенная по нормальному гауссову закону с нулевым средним и дисперсией $\sigma_h = r^H$, где $r = (x_1 - x_2)/2$ — расстояние от средней точки x_c рабочего отрезка $[x_1, x_2]$	Настройка параметров генерации h, выбор числа итераций	0,355

■ ***Окончание табл. 2***

■ ***Table 2***

Модель	Математическое описание модели	Число настраиваемых параметров модели или обучения	S
Хаотические отображения — логистическое уравнение	$X_{n+1} = CX_n - C(X_n)^2 = CX_n(1 - X_n)$, где C — параметр скорости распространения случайной величины X, самоподобность проявляется при $3 < C < 3{,}57$	Настройка параметра C и параметров базовой генерации X	0,289
Динамическое моделирование Маркова — автоматы с конечным числом состояний, в которых реализован вероятностный переход из одного состояния в другое [23]	$P_{\text{пер}\,A\to B} = \frac{C_{\text{пер}\,A\to B}}{\sum_i C_i}$, где i — номер счетчика; C_i — значение i-го счетчика; $P_{\text{пер}\,A\to B}$ — вероятность перехода из состояния A в состояние B	Настройка вероятностей перехода	0,243
Модели нечеткой логики — уровни пачечности представлены в виде ступенчатой функции полезности f	$U = f(\lambda, P)$, где U — уровень пачечности трафика; λ — средняя интенсивность трафика на интервале самоподобия t_i, $i = \overline{1, n}$; $P \sim p_i$, $i = \overline{1, n}$ — вероятности разыгрывания значения λ_i, $i = \overline{1, n}$	Настройка ступенчатой функции полезности	0,379
Нейросетевые модели — аппроксимация функций нескольких переменных по обучающей выборке временного ряда [24]	$Y(t) = \hat{F}(Z(t))$, где $\hat{F}(.)$ — оператор нейросетевого отображения; $Z(t) = \{t, X(t), X(t-1), ..., X(t-n)\}$ — временной ряд; $X(t)$ — интенсивность трафика в момент времени t; n — размер пачки	Обучение нейронной сети требует настройки $N = x^2y$ весов, где x — число нейронов в слое; y — число внутренних слоев	0,02
Авторегрессионные модели — модели временных рядов, в которых любые значения временного ряда линейно зависят от предыдущих значений этого же ряда	$X_t = c + \sum_{i=1}^{p} a_i X_{t-1} + \varepsilon_t$, где a — коэффициенты авторегрессии; c — постоянная; p — размер пачки трафика; ε_t — белый шум	Настройка a, c, параметров генератора белого шума ε_t	0,09
ON/OFF-модели — источник пакетов в ON-периоды в течение времени T_0 генерирует пакеты, в OFF-периоды в течение времени T_1 источник пассивен [25]	Периоды T_0, T_1 — случайные переменные с функцией плотности вероятности $w_0(t)$ и $w_1(t)$ соответственно; распределения $w_0(t)$, $w_1(t)$ являются распределениями с тяжелыми «хвостами»	Обучение не требуется	0,05

но оценить, например, как время, затраченное на программирование генератора или сложность алгоритма, поскольку все зависит от размера пачки, времени, затраченного на настройку одного параметра и комплекса параметров, хорошего владения языком программирования и др. Например, несмотря на то, что нейросетевая модель показала лучший результат по критерию S, основное время при ее реализации было потрачено на выбор архитектуры нейронной сети, а затем на настройку модели (3 дня), в то время как модель фрактального гауссовского шума была реализована за 40 мин, но критерий S в 17,5 раз больше, чем у нейросетевой модели. Более того, для имитации трафика с другим показателем Херста вновь потребуется процедура выбора архитектуры нейронной сети и ее обучение.

Анализ приведенных моделей позволяет сконцентрироваться на трех последних, представленных в табл. 2, и использовать их в решении задач моделирования инфокоммуникационных систем и сетей с вытекающими из этого глобальными проблемами: планирование развития инфокоммуникационных сетей, реализация дифференцированного обслуживания, оценка характеристик вычислительных ресурсов, гарантирующих требуемое качество обслуживания соответствующего трафика.

Заключение

В статье представлены результаты исследования 3G-трафика с целью выявления свойства самоподобия. Предположение о самоподобной структуре 3G-трафика основывается на рассмотрении имеющихся данных для различной временной шкалы. Используя метод построения дополнительной функции распределения для логарифмической шкалы, оценены тяжесть «хвоста» распределения и параметр самоподобия. Полученные результаты позволили проверить свойства самоподобия рассматриваемого трафика согласно определению и таким образом подтвердить предположение о самоподобии 3G-трафика.

Подобные исследования необходимы для понимания поведения сетевого трафика и синтеза моделей, имитирующих процесс поступления реального трафика в сети.

Обзор и сравнение существующих моделей имитации самоподобного трафика по результатам аппроксимации 3G-трафика позволили сделать предположение, что настройка модели может быть выполнена по показателю Херста при наличии записанных трасс реального сетевого трафика.

Литература

1. **Tanenbaum A., Wetherall D.** Computer Networks. Prentice Hall, 2010. 960 p.
2. **Кутузов О. И., Татарникова Т. М.** Инфокоммуникационные сети. Моделирование и оценка вероятностно-временных характеристик. СПб.: ГУАП, 2015. 382 с.
3. **Крылов В. В., Самохвалова С. С.** Теория телетрафика и ее приложения. СПб.: БХВ-Петербург, 2005. 288 с.
4. **Кутузов О. И., Татарникова Т. М.** Моделирование систем и сетей телекоммуникаций. СПб.: РГГМУ, 2012. 134 с.
5. **Петров В. В.** То, что вы хотели знать о самоподобных процессах, но стеснялись спросить. М.: Радиотехника, 2003. 112 с.
6. **Addie R. G., Neame T. D., Zukerman M.** Performance evaluation of a queue fed by a poisson pareto burst process. *Computer Networks*, 2002, vol. 40, no. 3, pp. 377–397.
7. **Шелухин О. И., Осин А. В., Смольский С. М.** Самоподобие и фракталы. Телекоммуникационные приложения. М.: Физматлит, 2008. 368 с.
8. **Erramilli A., Narayan O., Willinger W.** Experimental queueing analysis with long-range dependent traffic. *IEEE/ACM Trans. Networking*, 1996, vol. 4, pp. 209–223.
9. **Leland W. E., Taqqu M. S., Willinger W., Wilson D. V.** On the self-similar nature of ethernet traffic. *Proc. ACM SIGCOMM'93*. San-Fransisco, 1993, pp. 183–193.
10. **Шелухин О. И.** Мультифракталы. Инфокоммуникационные приложения. М.: Горячая линия — Телеком, 2011. 576 с.
11. **Willinger W., Taqqu M. S., Sherman R., Wilson D. V.** Self similarity through high-variability: statistical analisys of ethernet LAN Traffic at the source level. *EEE Trans. Networking*, 1997, vol. 5(1), pp. 71–86.
12. **Zwart A. P.** Queueing systems with heavy tails/ eindhoven uneversity of technology, 2001. 227 p.
13. **Костромицкий А. И., Волотка В. С.** Подходы к моделированию самоподобного трафика. Восточно-Европейский журнал передовых технологий. 2010. № 46. С. 46–49.
14. **Yeryomin R., Petersons E.** Transient process relaxation time research under the condition of self-similar traffic input in wireless networks. *Automatic Control and Computer Sciences*, 2009, vol. 43, no. 3, pp. 138–147.
15. **Choi B. D., Kim B., Wee I.** Asymptotic behavior of loss probability in GI/M/1/K queue as K-tends to infinity. *Queueing Systems*, 2000, vol. 36, pp. 437–442.
16. **Guanghui He, Yuan Gao I., Hou Jennifer H., Kihong Park.** A case for exploiting self-similarity of network traffic in TCP congestion control. *The International Journal of Computer and Telecommunications Networking*, 2004, vol. 45, issue 6, pp. 743–766.
17. **Осин А. В.** Сравнительный анализ методик оценки самоподобности телекоммуникационного трафика. Радиоэлектроника и информатика: сб. науч. трудов. М.: МГУС. 2002. С. 37–46.
18. **Crovella M. E., Bestavros A.** Self-similarity in world wide web traffic evidence and possible causes. Proc. of the 1996 ACM SIGMETRICS: Intern. conf. on Measurement and modeling of computer systems, May, 1996, *IEEE/ACM Trans on Networking*, 1997, vol. 5, no. 6, pp. 835–846.
19. **Zadorozhnyi V. N.** *Simulation modeling of fractal queues, in dynamics of systems, mechanisms and machines* (*dynamics*), 2014, pp. 1–4. doi:10.1109 / Dynamics.2014.7005703
20. **Kutuzov O., Tatarnikova T.** Evaluation and comparison of classical and fractal queuing systems. *XV International Symposium "Problems of Redundancy in Information and Control Systems"*, September 26–29, 2016.
21. **Kihong P., Willinger W.** Self-Similar network traffic and performance evaluation. New York, John Wiley & Sons, Inc., 2000. 576 p.
22. **Задорожный В. Н.** О качестве программных генераторов случайных чисел. Омский научный вестник. 2009. № 2 (80). С. 199–205.
23. **Богатырев В. А., Кармановский Н. С., Попцова Н. А., Паршутина С. А., Воронина Д. А., Богатырев С. В.** Имитационная модель поддержки проектирования инфокоммуникационных резервированных

систем. Научно-технический вестник информационных технологий, механики и оптики. 2016. Т. 16. № 5(105). С. 831–838. doi:10.17586/2226-1494-2016-16-5-831-838
24. **Тархов Д. А.** Нейросетевые модели и алгоритмы. М.: Радиотехника, 2014. 349 с.
25. **Привалов А. Ю., Благов А. В.** Об использовании некоторых моделей самоподобного сетевого трафика в имитационном моделировании // Математическое моделирование. 2011. Т. 23. № 7. С. 114–128.

UDC 004.732
doi:10.31799/1684-8853-2018-5-35-43

Statistical methods for studying network traffic

T. M. Tatarnikova[a], Dr. Sc., Tech., Associate Professor, orcid.org/ 0000-0002-6419-0072, tm-tatarn@yandex.ru
[a]Saint-Petersburg State University of Aerospace Instrumentation, 67, B. Morskaia St., 190000, Saint-Petersburg, Russian Federation

Introduction: The single multiservice network concept which involves the integration of voice, data and multimedia communication has prompted interest in studying the nature of network traffic. Studies of traffic traces recorded on a large scale show the presence of a self-similar structure in it, which requires a revision of the results of modeling infocommunication networks under the assumption of a Poisson data flow. **Purpose:** Studying a sequence of application methods for studying the network traffic nature, identifying the self-similar traffic nature in the form of statistical estimates and the Hurst index. This should offer tools for generating artificial traffic which adequately reflects a real network traffic, taking into account the revealed self-similarity properties. **Results:** The self-similarity properties of the considered 3G traffic were checked on different time scales obtained by aggregation of 5, 10, 15 and 20 minutes on the available daily traffic of 3G data. An estimate of the tail severity for self-similar traffic distribution was obtained by constructing a regression line for the additional distribution function on a logarithmic scale. The self-similarity parameter value, determined by the severity of the distribution "tail", made it possible to confirm the assumption of 3G traffic self-similarity. A review of models simulating real network traffic with a self-similar structure was performed. Tools were implemented for generating artificial traffic in accordance with the considered models. Various artificial network traffic generators were compared, according to the least squares method criterion, for approximating the artificial traffic point values by the approximation function of 3G traffic. Qualitative assessments of the traffic generators were taken into account, in the form of the their software implementation complexity, which, however, can be a subjective assessment. Comparative characteristics allow you to choose a generator which most faithfully simulates real network traffic. **Practical relevance:** The proposed sequence of methods to study the network traffic properties is necessary for understanding its nature and for developing appropriate models which simulate real network traffic.

Keywords — network traffic, self-similar process, long-term dependence, slowly decaying dispersion, heavy tail distribution, self-similarity parameter, self-similar traffic simulation models.

Citation: Tatarnikova T. M. Statistical methods for studying network traffic. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 5, pp. 35–43 (In Russian). doi:10.31799/1684-8853-2018-5-35-43

References

1. Tanenbaum A., Wetherall D. *Computer Networks*. Prentice Hall, 2010. 960 p.
2. Kutuzov O. I., Tatarnikova T. M. *Infokommunikacionnye seti. Modelirovanie i ocenka verojatnostno-vremennyh harakteristik* [Infocommunication Network. Simulation and Evaluation of Probability-time Characteristics]. SPb., GUAP Publ., 2013. 148 p. (In Russian).
3. Krylov V. V., Samohvalova S. S. *Teoriya teletrafika i ee prilozheniya* [Teletraffic theory and its applications]. SPb., BHV-Peterburg, 2005. 288 p. (In Russian).
4. Kutuzov O. I., Tatarnikova T. M. *Modelirovanie sistem i setei telekommunikatsi* [Simulation of Systems and Telecommunications Networks]. SPb., RGGMU Publ., 2012. 134 p. (In Russian).
5. Petrov V. V. *Chto vy hoteli znat' o samopodobnyh processah, no stesnjalis' sprosit'* [What You Wanted to Know about Self-similar Processes, but Hesitated to Ask]. Moscow, Radiotehnika Publ., 2003. 112 p. (In Russian).
6. Addie R. G., Neame T. D., Zukerman M. Performance evaluation of a queue fed by a poisson pareto burst process. *Computer Networks*, 2002, vol. 40, no. 3, pp. 377–397.
7. Shelukhin O. I., Osin A. V., Smol'skij S. M. *Samopodobie i fraktaly. Telekommunikacionnye prilozhenija* [Self-similarity and Fractals. Telecommunication Applications]. Moscow, Fizmatlit Publ., 2008. 368 p. (In Russian).
8. Erramilli A., Narayan O., Willinger W. Experimental queueing analysis with long-range dependent traffic. *IEEE/ ACM Trans. Networking*, 1996, vol. 4, pp. 209–223.
9. Leland W. E., Taqqu M. S., Willinger W., Wilson D. V. On the self-similar nature of ethernet traffic. *Proc. ACM SIGCOMM'93*. San-Fransisco, 1993, pp. 183–193.
10. Shelukhin O. I. Mul'tifraktaly. *Infokommunikacionnye prilozheniya* [Multifractals. Infocommunication applications]. Moscow, Goryachaya liniya — Telekom Publ., 2011. 576 p. (In Russian).
11. Willinger W., Taqqu M. S., Sherman R., Wilson D. V. Self similarity through high-variability: statistical analisys of ethernet LAN Traffic at the source level. *EEE Trans. Networking*, 1997, vol. 5(1), pp. 71–86.
12. Zwart A. P. Queueing systems with heavy tails/ eindhoven uneversity of technology, 2001. 227 p.
13. Kostromickij A. I., Volotka V. S. Approaches to modeling self-similar traffic. *Vostochno-Evropeiskii zhurnal peredovykh tekhnologii* [Eastern-European Journal of Enterprise Technologies], 2010, no. 46, pp. 46–49 (In Russian).
14. Yeryomin R., Petersons E. Transient process relaxation time research under the condition of self-similar traffic input in wireless networks. *Automatic Control and Computer Sciences*, 2009, vol. 43, no. 3, pp. 138–147.
15. Choi B. D., Kim B., Wee I. Asymptotic behavior of loss probability in GI/M/1/K queue as K-tends to infinity. *Queueing Systems*, 2000, vol. 36, pp. 437–442.
16. Guanghui He, Yuan Gao I., Hou Jennifer H., Kihong Park. A case for exploiting self-similarity of network traffic in TCP congestion control. *The International Journal of Computer and Telecommunications Networking*, 2004, vol. 45, issue 6, pp. 743–766.
17. Osin A. V. Comparative analysis of methods for assessing the self-similarity of telecommunication traffic. *Vestnik MGUS* [Bulletin of MGUS], 2002. pp. 37–46.
18. Crovella M. E., Bestavros A. Self-similarity in world wide web traffic evidence and possible causes. Proc. of the 1996

ACM SIGMETRICS: Intern. conf. on Measurement and modeling of computer systems, May, 1996, *IEEE/ACM Trans on Networking*, 1997, vol. 5, no. 6, pp. 835–846.
19. Zadorozhnyi V. N. *Simulation modeling of fractal queues, in dynamics of systems, mechanisms and machines* (*dynamics*), 2014, pp. 1–4. doi:10.1109/ Dynamics.2014.7005703
20. Kutuzov O., Tatarnikova T. Evaluation and comparison of classical and fractal queuing systems. *XV International Symposium "Problems of Redundancy in Information and Control Systems"*, September 26–29, 2016.
21. Kihong P., Willinger W. Self-Similar network traffic and performance evaluation. New York, John Wiley & Sons, Inc., 2000. 576 p.
22. Zadorozhnyi V. N. On the quality of program generators of random numbers. *Omskij nauchnyj vestnik* [Omsk Scientific Herald], 2009, no. 2 (80), pp. 199–205 (In Russian).
23. Bogatyrev V. A., Karmanovsky N. S., Poptcova N. A., Parshutina S. A., Voronina D. A., Bogatyrev S. V. Simulation model for design support of infocomm redundant systems. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2016, vol. 16, no. 5, pp. 831–838 (In Russian). doi:10.17586/2226-1494-2016-16-5-831-838/
24. Tarhov D. A. *Nejrosetevye modeli i algoritmy. Spravochnik*. [Neural network models and algorithms. Directory]. Moscow, Radio Engineering Publ. 2014. 349 p. (In Russian).
25. Privalov A. Yu., Blagov A. V. About some self-similar telecommunication traffic models used in traffic simulation. *Matematicheskoe modelirovanie* [Math modeling], 2011, vol. 23, no. 7, pp. 114–128 (In Russian).

Уважаемые авторы!

При подготовке рукописей статей необходимо руководствоваться следующими рекомендациями.

Статьи должны содержать изложение новых научных результатов. Название статьи должно быть кратким, но информативным. В названии недопустимо использование сокращений, кроме самых общепринятых (РАН, РФ, САПР и т. п.).

Объем статьи (текст, таблицы, иллюстрации и библиография) не должен превышать эквивалента в 20 страниц, напечатанных на бумаге формата А4 на одной стороне через 1,5 интервала Word шрифтом Times New Roman размером 13, поля не менее двух сантиметров.

Обязательными элементами оформления статьи являются: индекс УДК, заглавие, инициалы и фамилия автора (авторов), ученая степень, звание (при отсутствии — должность), полное название организации, аннотация и ключевые слова на русском и английском языках, электронные адреса авторов, которые по требованию ВАК должны быть опубликованы на страницах журнала. При написании аннотации не используйте аббревиатур и не делайте ссылок на источники в списке литературы. Предоставляйте подрисуночные подписи и названия таблиц на русском и английском языках.

Статьи авторов, не имеющих ученой степени, рекомендуется публиковать в соавторстве с научным руководителем, наличие подписи научного руководителя на рукописи обязательно; в случае самостоятельной публикации обязательно предоставляйте заверенную по месту работы рекомендацию научного руководителя с указанием его фамилии, имени, отчества, места работы, должности, ученого звания, ученой степени — эта информация будет опубликована в ссылке на первой странице.

Формулы набирайте в Word, не используя формульный редактор (Mathtype или Equation), при необходимости можно использовать формульный редактор; для набора одной формулы не используйте два редактора; при наборе формул в формульном редакторе знаки препинания, ограничивающие формулу, набирайте вместе с формулой; для установки размера шрифта никогда не пользуйтесь вкладкой Other…, используйте заводские установки редактора, не подгоняйте размер символов в формулах под размер шрифта в тексте статьи, не растягивайте и не сжимайте мышью формулы, вставленные в текст; в формулах не отделяйте пробелами знаки: + = −.

Для набора формул в Word никогда не используйте Конструктор (на верхней панели: «Работа с формулами» — «Конструктор»), так как этот ресурс предназначен только для внутреннего использования в Word и не поддерживается программами, предназначенными для изготовления оригинал-макета журнала.

При наборе символов в тексте помните, что символы, обозначаемые латинскими буквами, набираются светлым курсивом, русскими и греческими — светлым прямым, векторы и матрицы — прямым полужирным шрифтом.

Иллюстрации предоставляются отдельными исходными файлами, поддающимися редактированию:

— рисунки, графики, диаграммы, блок-схемы предоставляйте в виде отдельных исходных файлов, поддающихся редактированию, используя векторные программы: Visio (*.vsd, *.vsdx); Coreldraw (*.cdr); Excel (*.xls); Word (*.docx); Adobe Illustrator (*.ai); AutoCad (*.dxf); Matlab (*.ps, *.pdf или экспорт в формат *.ai);

— если редактор, в котором Вы изготавливаете рисунок, не позволяет сохранить в векторном формате, используйте функцию экспорта (только по отношению к исходному рисунку), например, в формат *.ai, *.esp, *.wmf, *.emf, *.svg;

— фото и растровые — в формате *.tif, *.png с максимальным разрешением (не менее 300 pixels/inch).

Наличие подрисуночных подписей обязательно (желательно не повторяющих дословно комментарии к рисункам в тексте статьи).

В редакцию предоставляются:

— сведения об авторе (фамилия, имя, отчество, место работы, должность, ученое звание, учебное заведение и год его окончания, ученая степень и год защиты диссертации, область научных интересов, количество научных публикаций, домашний и служебный адреса и телефоны, e-mail), фото авторов: анфас, в темной одежде на белом фоне, должны быть видны плечи и грудь, высокая степень четкости изображения без теней и отблесков на лице, фото можно представить в электронном виде в формате *.tif, *.png с максимальным разрешением — не менее 300 pixels/inch при минимальном размере фото 40×55 мм;

— экспертное заключение.

Список литературы составляется по порядку ссылок в тексте и оформляется следующим образом:

— для книг и сборников — фамилия и инициалы авторов, полное название книги (сборника), город, издательство, год, общее количество страниц;

— для журнальных статей — фамилия и инициалы авторов, полное название статьи, название журнала, год издания, номер журнала, номера страниц;

— ссылки на иностранную литературу следует давать на языке оригинала без сокращений;

— при использовании web-материалов указывайте адрес сайта и дату обращения.

Список литературы оформляйте двумя отдельными блоками по образцам lit.dot на сайте журнала (http://i-us.ru/paperrules): Литература и References.

Более подробно правила подготовки текста с образцами изложены на нашем сайте в разделе «Оформление статей».

Контакты
Куда: 190000, Санкт-Петербург,
Б. Морская ул., д. 67, ГУАП, РИЦ
Кому: Редакция журнала «Информационно-управляющие системы»
Тел.: (812) 494-70-02
Эл. почта: ius.spb@gmail.com
Сайт: www.i-us.ru

УДК 004.056.5
doi:10.31799/1684-8853-2018-5-44-56

Научные статьи
Articles

Адаптивный алгоритм встраивания информации в сжатые JPEG-изображения на основе операции замены

О. О. Шумская[а], *аспирант, orcid.org/0000-0002-8287-5032, shumskaya.oo@gmail.com*
М. Железны[б], *PhD, заместитель декана, orcid.org/0000-0003-1695-4370*
[а]*Санкт-Петербургский институт информатики и автоматизации РАН, 14-я линия В. О., 39, Санкт-Петербург, 199178, РФ*
[б]*Западночешский университет, Университетская ул, 8, 2732, 301 00, Пльзень, Чешская Республика*

Введение: *один из эффективных подходов к обеспечению конфиденциальности передаваемой и хранимой информации основан на применении методов цифровой стеганографии. Для обеспечения устойчивости перед стегоанализом при встраивании информации в цифровое изображение необходимо, чтобы встраивание не приводило к появлению демаскирующих признаков.* ***Цель исследования:*** *разработка адаптивного алгоритма стеганографического встраивания информации в сжатые JPEG-изображения на основе операции замены с минимизацией вносимых искажений в информативные признаки.* ***Результаты:*** *определена значимость демаскирующих признаков в области стегоанализа и их применения для адаптивности алгоритмов сокрытия секретной информации в цифровых объектах. Приведены основные признаки в пространственной и частотной областях цифровых изображений, применяемые в современных методах стеганографического встраивания. Проведен отбор информативных признаков, исключающий линейно зависимые признаки или признаки, не несущие в себе какую-либо информацию об искажении цифрового объекта при встраивании. Полученный набор позволил повысить точность общей классификации изображений на 19 %. На основе сформированного набора информативных признаков разработана адаптивная модификация алгоритма встраивания информации в сжатые JPEG-изображения на основе замены, обеспечивающая минимизацию искажений изображения-контейнера при встраивании благодаря сформулированной целевой функции. Адаптивность алгоритма заключается в том, что выбор области сокрытия основывается на наборе информативных признаков, который характеризует естественную модель цифрового изображения. Был проведен ряд экспериментов с целью выявления наилучших значений параметров для достижения хорошей емкости встраивания и минимальных искажений признаков. По результатам вычислительных экспериментов разработанный адаптивный алгоритм показал повышенную устойчивость перед стегоанализом при внушительной емкости встраивания, а также высокие значения метрики качества стегоизображений, что говорит о повышенной незаметности как для человеческого глаза, так и для многих алгоритмов стегоанализа, так как значения признаков искажаются незначительно.*

Ключевые слова *— стеганография, стегоанализ, цифровые изображения, сжатые JPEG-изображения, операция замены, оптимизация, информативные признаки, демаскирующие признаки, классификация.*

Цитирование: Шумская О. О., Железны М. Адаптивный алгоритм встраивания информации в сжатые JPEG-изображения на основе операции замены. *Информационно-управляющие системы*, 2018, № 5, с. 44–56. doi:10.31799/1684-8853-2018-5-44-56
Citation: Shumskaya O. O., Zelezny M. Adaptive algorithm of replacement-based embedding of data into compressed JPEG images. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 5, pp. 44–56 (In Russian). doi:10.31799/1684-8853-2018-5-44-56

Введение

В настоящее время в различных сферах человеческой деятельности активно применяются информационно-аналитические технологии. Они реализуются в виде информационно-аналитических систем, предназначенных для накопления и комплексного анализа данных различного типа с целью поддержки принятия решений.

Важным составным элементом разработки и эксплуатации информационно-аналитических систем является обеспечение их информационной безопасности, прежде всего в части обеспечения конфиденциальности обрабатываемой и хранимой информации.

Один из подходов к обеспечению конфиденциальности передаваемой и хранимой информации основан на применении методов цифровой стеганографии (науки о способах передачи, хранения информации, обеспечивающих сокрытие факта наличия этой информации в некотором цифровом объекте [1]).

Методы цифровой стеганографии применимы в информационно-аналитических системах, оперирующих мультимедиа-данными, поскольку в качестве стегоконтейнеров в цифровой стеганографии преимущественно применяют цифровые изображения, аудио- и видеоданные. Чаще всего в качестве цифровых объектов-контейнеров используют цифровые изображения. Наиболее часто встречающийся графический формат на сегодня — сжатые JPEG-изображения. Именно в нем сохраняется основная часть всей графики, с которой работают пользователи информационно-аналитических систем.

Эффективность стеганографического встраивания информации в цифровые изображения оценивается с помощью различных показателей. Основным требованием является обеспечение устойчивости перед стегоанализом, что выражается в статистической неразличимости стегоизображений и изображений, не содержащих вложений. Для обеспечения устойчивости перед стегоанализом при встраивании информации в цифровое изображение необходимо, чтобы встраивание не приводило к появлению демаскирующих признаков.

Выявление демаскирующих признаков может быть произведено посредством анализа массивов изображений, обрабатываемых в информационно-аналитической системе (стегоизображений и изображений, не содержащих вложений).

Цель настоящей статьи заключается в разработке алгоритма сокрытия информации в сжатых JPEG-изображениях, позволяющего минимизировать величину вносимых искажений в информативные признаки цифровых изображений.

Стеганографическое встраивание информации в JPEG-изображения

JPEG — это самый распространенный формат изображений. Сжатые JPEG-изображения часто используют в стеганографии в качестве стегоконтейнеров с применением различных алгоритмов встраивания.

Стоит отметить, что большинство работ основаны на обработке изображения, разбитого на непересекающиеся блоки 8×8 пикселей, к каждому из которых применено дискретное косинусное преобразование (ДКП) с последующей квантизацией. Полученные матрицы коэффициентов состоят из DC-коэффициента, который располагается в левом верхнем углу матрицы и содержит основную информацию о блоке изображения, и AC-коэффициентов — остальные 63 коэффициента, в которые производится скрытое встраивание информации.

В статье [2] описывается применение алгоритма встраивания PM1 к JPEG-изображениям, позволяющего получить большую емкость и обеспечить высокую степень безопасности. В качестве области встраивания авторы используют матрицу ДКП-коэффициентов, перемешанных с помощью некоторого ключа. Вложение осуществляется в ненулевые AC-коэффициенты, авторы работы утверждают, что изменение DC-коэффициента или AC-коэффициентов, равных нулю, более заметно. Отрицательные четные и положительные нечетные коэффициенты соответствуют биту со значением «1», отрицательные нечетные и положительные четные — «0», если коэффициент соответствует биту сообщения по этим правилам, то коэффициент сохраняет свое значение, иначе коэффициент увеличивается или уменьшается в произвольном порядке на 1. Если изменение коэффициента приводит к значению «0», то его необходимо заменить на «1» или «–1» в соответствии со значением бита.

Авторы статьи [3] в качестве пространства сокрытия используют последовательности ДКП-коэффициентов равных нулю. Если длина последовательности 3 и более, то коэффициент с самой низкой частотой заменяется согласно правилу: на «0», если бит равен 0, и на «1» или «–1», если бит равен 1. Если последовательность «$x0$», то при $x > 0$ x заменяется на «$x+1$», иначе на «$x-1$». В случае последовательности вида «0x0» при $x > 0$ x заменяется на «$x+1$», иначе на «$x-1$».

В работе [4] пространством сокрытия выступают ненулевые ДКП-коэффициенты, представленные в виде последовательности. Встраивание заключается в увеличении или уменьшении коэффициента в зависимости от его начального значения, а также величины разности между значениями ДКП-коэффициента до округления и после.

Пространство сокрытия в статье [5] — нулевые ДКП-коэффициенты. Встраивание осуществляется путем увеличения или уменьшения коэффициента на 1. Выбор изменяемого коэффициента зависит от самого скрываемого сообщения, т. е. каждый раз высчитывается номер коэффициента в блоке, изменение которого будет свидетельствовать о встраивании именно этой последовательности бит. Данный подход к встраиванию основан на применении кода Хэмминга.

Авторы статьи [6] для встраивания рассматривают ненулевые ДКП-коэффициенты, представленные в виде последовательности. Встраивание осуществляется согласно формуле:

$$C'_i = \begin{cases} C_i + sign(C_i) \times b, \text{ если} |C_i| = 1, \\ C_i + sign(C_i), \text{ если} |C_i| > 1, \end{cases} \quad (1)$$

где C_i — i-й квантованный ДКП-коэффициент, b — встраиваемый бит сообщения,

$$sign(C_i) = \begin{cases} 1, \text{ если } C_i > 0, \\ 0, \text{ если } C_i = 0, \\ -1, \text{ если } C_i < 0. \end{cases}$$

Критерии эффективности стеганографического встраивания информации в JPEG-изображения

Критериями эффективности стеганографического встраивания являются емкость, незаметность (отсутствие видимых искажений цифровых

объектов-стегоконтейнеров) и устойчивость перед стегоанализом [1], представляющего собой науку о способах выявления фактов наличия скрытых сообщений в цифровых объектах. Стегоанализ в большинстве случаев заключается в поиске характеристик (признаков) цифрового объекта, которые изменяются в ходе стеганографического встраивания.

Известны методы стегоанализа, основанные на анализе признаков цифрового изображения в пространственной области, и методы, основанные на анализе признаков в частотной области. В каждом методе стегоанализа значения некоторого набора признаков объединяются в один вектор, с которым уже работает классификатор. Для того чтобы начать классификацию, прежде необходимо провести обучение: классификатор определяет для себя, какие интервалы значений признаков принадлежат «чистым» изображениям и изображениям с вложением. Сопоставляя полученные в ходе обучения данные с рассчитанным вектором признаков исследуемого изображения, классификатор определяет содержание изображения.

В работе [7] описан случай встраивания в ДКП-коэффициенты. Набор признаков для стегоанализа состоит из признаков в частотной области, основанных на соотношениях между энергией, собранной в отдельных частотных коэффициентах ДКП-спектра:

$$F_1 = \frac{E(f_0)}{E(f_{|\eta|=1})}, \qquad (2)$$

где $E(f_0)$ — среднее значение частот нулевых АС-коэффициентов изображения по блокам,

$$F_2 = \frac{\sum_{|\eta|>1} E(f_\eta)}{E(f_{|\eta|=1})}, \qquad (3)$$

где $E(f_{|\eta|=1})$ — среднее значение частот тех АС-коэффициентов изображения, абсолютная величина которых равна 1,

$$F_3 = \frac{En_{|\eta|>1}}{En_{|\eta|\le 1}}, \qquad (4)$$

где $En_{|\eta|>1}$ — энергия тех АС-коэффициентов изображения, абсолютная величина которых > 1.

В исследовании показано, что для аддитивного встраивания информации в квантованные ДКП-коэффициенты JPEG-изображения характерно увеличение значений $E(f_0)$, $\Sigma_{|\eta|>1}E(f_\eta)$, $En_{|\eta|>1}$ и уменьшение значений $E(f_{|\eta|=1})$, $En_{|\eta|\le 1}$.

Авторы объясняют свой выбор именно такого набора признаков тем, что данные величины концентрируют в себе максимальную информацию о внутреннем содержании изображения.

В качестве классификатора авторы выбрали линейный дискриминант Фишера. Такой классификатор достаточно часто встречается в современных работах, он гибкий относительно количества рассматриваемых признаков, так как проецирует весь вектор признаков на прямую. Суть классификации заключается в поиске лучшего направления данной проекции, чтобы величину можно было четко отнести к определенному классу.

Для изображений характерна межблочная корреляция. Во время встраивания вносятся изменения в блоки изображения, что может привести к нарушению связи между блоками

$$F_4 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}}, \qquad (5)$$

где $\bar{x}$, $\bar{y}$ — средние значения АС-коэффициентов соседних блоков, x_i, y_i — i-й АС-коэффициент соседних блоков.

Исследования [8] показали, что добавление параметра межблочной корреляции АС-коэффициентов повышает вероятность обнаружения стеганографических вложений, что говорит об информативности признака в рамках стегоанализа цифровых изображений.

Авторы работ [9, 10] рассматривают в качестве основы стегоанализа 23 признака, как в частотной области, так и в пространственной:

— общая гистограмма ДКП-коэффициентов изображения;

— гистограммы первых пяти АС-коэффициентов, для каждого отдельно (5 признаков);

— двойные гистограммы АС-коэффициентов, значение которых находится в диапазоне [–5, 5], для каждого отдельная гистограмма (11 признаков);

— межблочные зависимости в разных направлениях в пространственной области (6 признаков).

Ранее отмечалось, что стеганографическое встраивание осуществляется преимущественно в АС-коэффициенты со значениями по модулю, близкими к нулю. Поэтому двойные гистограммы для коэффициентов со значениями в диапазоне [–5, 5] концентрируют в себе максимальную информацию о встраивании в частотной области.

Двойная гистограмма представляет собой матрицу, которая отражает, на каком месте сколько раз суммарно по всем блокам встретился коэффициент с определенным значением

$$f_5, \ldots, f_{15} = \frac{\sum_{k=1}^{B}\delta(d, d_k(i, j))}{\left\|\sum_{k=1}^{B}\delta(d, d_k(i, j))\right\|_{L_1}}, \qquad (6)$$

где d — фиксированное значение коэффициента, $d \in [-5, 5]$, B — количество блоков в изображе-

нии, i, j — координаты положения коэффициента в блоке, L_1 норма — максимальная из сумм элементов по столбцам, $\delta(d, d_k(i, j)) = \begin{cases} 1, & d_k(i, j) = d \\ 0, & else \end{cases}$.

Каждую характеристику авторы рассчитывают дважды: для исследуемого изображения (J_1) и для изображения, которое получают путем обрезания исследуемого изображения сверху и слева на 4 пикселя (J_2), как показано на рис. 1. Подобное действие авторы объясняют следующим образом: при обрезании изображения слева и сверху разделение изображения на блоки сдвигается, коэффициенты дискретного косинусного преобразования освобождаются от влияния прошлой квантизации и содержат только статистические данные изображения, которые как раз важны при стегоанализе.

Таким образом, конечным значением признака будет значение функционала:

$$F_5 \ldots F_{15} = \left\| f_5 \ldots f_{15}(J_1) - f_5 \ldots f_{15}(J_2) \right\|_{L_1}. \qquad (7)$$

В работах [9, 10] авторы также используют линейный дискриминант Фишера в качестве классификатора, но отмечают, что применение метода опорных векторов, возможно, повысит надежность обнаружения разработанного алгоритма.

Метод, представленный авторами [11], основан на законе Бенфорда: вероятность появления цифры на первом месте в числе тем выше, чем меньше эта цифра. Основываясь на выводах работы [12], посвященной исследованию справедливости закона Бенфорда в отношении ДКП-коэффициентов JPEG-изображений до и после квантования, авторы предложили частный случай закона Бенфорда, так как квантованные ДКП-коэффициенты не подчиняются строго закону Бенфорда

$$F_{16}, \ldots, F_{24} = N \log_{10}\left(1 + \frac{1}{s + x^q}\right), \qquad (8)$$

где $x = 1, ..., 9$, N, s, q — параметры, зависящие от качества JPEG-сжатия.

Если отклонение реальной величины от ожидаемой превышает некоторый порог, то принимается решение о наличии вложения в данном изображении.

Идея смещения изображения, применяемая в работах [9, 10], показалась довольно интересной. Было принято решение объединить смещение изображения и закон Бенфорда

$$F_{25}, \ldots, F_{33} = \left\| F_{16}, \ldots, F_{24}(J_1) - F_{16}, \ldots, F_{24}(J_2) \right\|_{L_1}. \qquad (9)$$

В приведенных работах идет речь о признаках, формируемых специально для проведения стегоанализа. Однако для решения данной задачи можно использовать и произвольные признаки, представляющие собой некоторые характеристики, рассчитываемые для цифровых изображений. В частности, в таблице представлены текстурные признаки, применяемые в различных задачах распознавания образов [13].

В общем случае стегоанализ цифровых объектов рассматривается как задача двухклассовой классификации, когда для каждого анализируемого объекта выбирается один из двух исходов: нет вложения, или объект содержит скрытые данные.

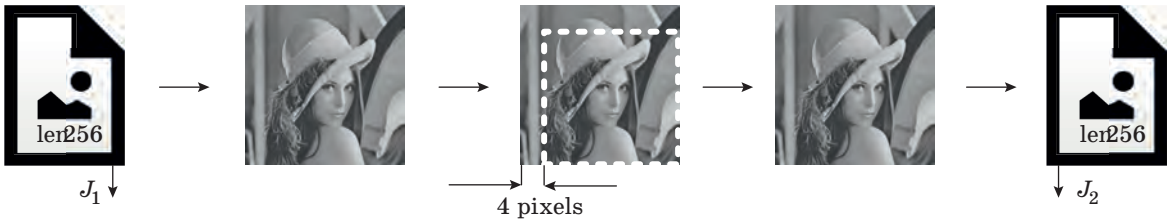
В ходе работы в качестве классификатора в стегоанализаторе применяется наивный байесовский классификатор, так как он показывает хорошие результаты [14, 15], компактен по сравнению с другими методами (в методах, основанных на опорных векторах, нейронных сетях, при увеличении набора признаков объем необходимой памяти возрастает в разы) и достаточно популярен среди ученых, работающих в данной области.

Предположение, что все особенности набора данных независимы, и является причиной такого названия классификатора. Ведь чаще всего характеристики набора данных не независимы.

Упрощенная формула для классификации:

$$P(\text{Класс}A \mid \text{Свойство1, Свойство2}) = \frac{P(\text{Свойство1} \mid \text{Класс}A) \times P(\text{Свойство2} \mid \text{Класс}A)}{P(\text{Свойство1}) \times P(\text{Свойство2})}. \qquad (10)$$

Если видны Свойства 1 и 2, это и есть вероятность того, что данные принадлежат Классу A.



■ ***Рис. 1.*** Исследуемое изображение J_1 и изображение J_2, полученное путем обрезания изображения J_1

■ ***Fig. 1.*** The researched image J_1 and the image J_2 received by trimming of the image J_1

- Признаки в пространственной области изображения
- Features in the spatial domain of image

Признак	Формула	Пояснение
Энергия	$F_{34} = \sum_{i=0}^{N-1}\sum_{j=0}^{N-1} P_{i,j}^2,$ где N — количество градаций яркости стегоизображения, P — матрица смежности	Характеризует однородность изображения и равномерность
Энтропия	$F_{35} = \sum_{i=0}^{N-1}\sum_{j=0}^{N-1} P_{i,j} \log P_{i,j}$	Выражает неравномерность распределения яркостных свойств элементов стегоизображения
Однородность	$F_{36} = \sum_{i=0}^{N-1}\sum_{j=0}^{N-1} \frac{P_{i,j}}{(1+\lvert i+j\rvert)}$	
Контраст	$F_{37} = \sum_{n=0}^{N-1} n^2 \left[\sum_{i=1}^{N}\sum_{j=1}^{N} P_{i,j}\right], \ \lvert i-j\rvert = n$	С увеличением числа локальных вариаций яркости стегоизображения контраст возрастает
Среднее по i	$F_{38} = \sum_{i=1}^{N}\sum_{j=1}^{N} iP_{i,j}$	
Среднее по j	$F_{39} = \sum_{i=1}^{N}\sum_{j=1}^{N} jP_{i,j}$	
Диспресия по i	$F_{40} = \sum_{i=1}^{N}\sum_{j=1}^{N} (i-F_{38})^2 P_{i,j}$	Мера отклонения случайной величины от ее математического ожидания
Диспресия по j	$F_{41} = \sum_{i=1}^{N}\sum_{j=1}^{N} (j-F_{39})^2 P_{i,j}$	
Ковариация	$F_{42} = \sum_{i=1}^{N}\sum_{j=1}^{N} (i-F_{38})(j-F_{39})P_{i,j}$	Мера линейной зависимости двух случайных величин
Корреляция	$F_{43} = \sum_{i=1}^{N}\sum_{j=1}^{N} \frac{(i-F_{38})(j-F_{39})P_{i,j}}{F_{40}}$	Показывает статистическую взаимосвязь двух или более величин

Ранее было отмечено, что важным критерием эффективности стеганографического встраивания является устойчивость перед стегоанализом. Обеспечить незаметность возможно, если организовать встраивание таким образом, чтобы оно не вносило заметных искажений в естественную модель цифрового изображения. С целью максимально возможной эффективности разрабатываемых методов стегоанализа исследователи строят достаточно большие признаковые пространства, включающие десятки и сотни тысяч признаков [16–18]. В совокупности данные признаки дают хорошие результаты, однако вопрос информативности этих признаков во многих исследованиях авторы обходят стороной. Поэтому выбор набора наиболее информативных признаков из большого множества является важной задачей [19–22].

Анализ информативности признаков

Ключевым этапом является выбор информативных признаков, анализ которых позволяет отделять изображения, содержащие встроенную информацию, от чистых изображений. В качестве признаков в задаче стегоанализа используются разнообразные статистические характеристики, рассчитываемые для элементов данных цифровых изображений в пространственной и частотной областях.

Набор признаков может содержать сотни и тысячи элементов, однако некоторые из них могут быть линейно зависимыми, какие-то признаки могут не подходить к решению конкретной задачи. Для того чтобы из огромного множества набранных признаков выделить информативные

для рассматриваемой задачи, необходимо провести эксперименты с применением алгоритма выбора информативных признаков.

Существует много алгоритмов отбора информативных признаков, основанных на выявлении статистических зависимостей (между элементами набора, между элементами набора и выходным значением), сравнительных экспериментах, сложных вычислениях.

Жадные алгоритмы поиска используются часто, так как быстры и дают хороший результат во многих задачах. Группа алгоритмов получила такое название из-за того, что если один из признаков был выбран в поднабор (или исключен), то в дальнейшем он остается в наборе (в случае жадного включения) или навсегда будет отсутствовать (в случае жадного исключения) [23].

Алгоритмы поочередного перебора оценивают важность каждого элемента набора признаков для результата, рассматривая каждый элемент отдельно «в вакууме», т. е. без учета влияния остальных элементов [23]. Однако это не позволяет однозначно отобрать информативные признаки (постановка границы отбора), также в качестве наиболее важных признаков (наиболее информативных) могут оказаться подобные признаки.

Генетический алгоритм — эвристический алгоритм поиска, используемый для решения задач оптимизации и моделирования путем случайного подбора, комбинирования и вариации искомых параметров с использованием механизмов, аналогичных естественному отбору в природе [24].

Для дальнейшей работы принято решение использовать жадный алгоритм с исключением, так как данный метод довольно быстрый и эффективный согласно исследованиям [25, 26].

Эксперименты показали, что набор из 43 признаков в пространственной и частотной областях можно сократить на 17 элементов, при этом повысив общую точность классификации на 19 %.

В результате экспериментов был определен следующий набор из 26 информативных признаков: $\{F_1–F_5, F_8–F_{12}, F_{18}, F_{19}, F_{26}–F_{34}, F_{36}, F_{38}–F_{41}\}$.

Алгоритм встраивания информации в JPEG-изображения

За основу исследования был взят алгоритм, описанный в [27].

Введем общую схему встраивания информации в сжатые JPEG-изображения на основе операции замены квантованных ДКП-коэффициентов. Пусть изображение-контейнер содержит K блоков квантованных ДКП-коэффициентов. Пространство сокрытия представляет собой некоторую подпоследовательность последовательности всех ДКП-коэффициентов $C = c_1c_2...c_L$, $L < 64K$. Часть пространства сокрытия, образованную ДКП-коэффициентами одного блока, назовем областью встраивания данного блока. Будем считать, что нумерация коэффициентов блока осуществляется в порядке его «зигзагообразного» обхода. Секретное сообщение обозначим $M = m_1m_2...m_L$, $m_i \in \{0, 1\}$, $i = \overline{1, L}$. Количество бит, встраиваемых в один блок изображения, обозначим n. Если биты сообщения распределяются по блокам неравномерно, то можем записать $L = \sum_{j=1}^{K} n_j$, $n_j \geq 0$.

Значение, на которое будет заменен ДКП-коэффициент при встраивании в него бита секретного сообщения, назовем величиной замены и обозначим x. Тогда схему встраивания информации в JPEG-изображения на основе операции замены можно представить в виде формулы:

$$c'_i = \begin{cases} x, & m_i = 1, \\ -x, & m_i = 0, \end{cases} \qquad (11)$$

где c'_i — измененное значение ДКП-коэффициента.

Обозначим последовательность ДКП-коэффициентов изображения-контейнера, не входящих в пространство сокрытия, $D = d_1d_2...d_{64K-L}$, и введем дополнительную операцию для изменения данных коэффициентов:

$$d'_i = \begin{cases} x+1, & d_i = x, \\ -x-1, & d_i = -x, \\ d, & \text{иначе}. \end{cases} \qquad (12)$$

Дополнительная операция необходима, чтобы при извлечении опираться только на ДКП-коэффициенты, равные x и $-x$, которые будут соответствовать единичным и нулевым битам секретного сообщения.

Основное преимущество, которое дает представленная схема встраивания по сравнению с другими схемами, заключается в возможности произвольного выбора ДКП-коэффициентов, в которых будут размещены биты встраиваемого сообщения. При этом количество изменяемых коэффициентов для разных блоков изображения-контейнера может быть различным. Это позволит формировать пространство сокрытия для каждого конкретного контейнера наилучшим образом. Устойчивость введенной схемы встраивания перед стегоанализом будет зависеть от статистических характеристик сообщения, объема стеговложения и качества сжатия изображения-контейнера.

Распределение квантованных ДКП-коэффициентов JPEG-изображения близко к обобщен-

ному нормальному распределению [2]. Если распределение нулей и единиц в сообщении будет отлично от равномерного, то столбцы гистограммы ДКП-коэффициентов со значениями x и $-x$ будут иметь вид, не соответствующий нормальному распределению, что послужит демаскирующим признаком. Для предупреждения данной уязвимости сообщение перед встраиванием должно быть сжато или зашифровано. В этом случае столбцы со значениями x и $-x$ на гистограмме ДКП-коэффициентов будут иметь симметричный вид (с некоторой допустимой погрешностью), что соответствует модели исходного изображения.

Другим демаскирующим признаком может послужить изменение высоты данных столбцов. Для предупреждения данной уязвимости длину сообщения следует задавать такой, чтобы она совпадала с количеством ДКП-коэффициентов, по абсолютному значению соответствующих величине замены для данного качества JPEG-сжатия изображения-контейнера. При встраивании сообщения малого объема следует добавить в него поле, хранящее длину, и использовать только часть пространства сокрытия.

Таким образом, устойчивость перед стегоанализом может быть достигнута за счет подстройки параметров встраивания под длину секретного сообщения или характеристики конкретного изображения-контейнера.

Необходимость минимизации искажений естественной модели цифрового изображения в области квантованных ДКП-коэффициентов приводит к появлению задачи оптимизации.

Минимизация искажений естественной модели цифрового изображения в области квантованных ДКП-коэффициентов

В области стеганографии существует большое количество различных алгоритмов встраивания. Авторы работ придумывают новые идеи сокрытия информации в объекте, стараясь повысить незаметность встраивания. Однако тот факт, что вложение не видно человеческому глазу, не означает, что встраивание не внесло искажений в стегоконтейнер. Для оценки визуального качества изображения часто используется метрика PSNR (peak signal to noise ratio) — пиковое отношение сигнал/шум [28]. Чем меньше искажений внесено в стегоизображение, тем больше должна быть эта метрика (ее удобно называть «метрикой сходства»). Число PSNR безразмерно, поскольку единицами измерения и числителя, и знаменателя служат величины пикселей. Тем не менее, из-за использования логарифмов говорится, что число PSNR измеряется в децибелах (дБ).

Формула расчета PSNR:

$$PSNR = 20\log_{10}\left(\frac{255}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(P_i - Q_i)^2}}\right), \quad (13)$$

где 255 — максимальное значение пикселя, n — общее количество пикселей, P_i — пиксели исходного изображения, Q_i — пиксели восстановленного изображения.

Помимо стремления минимизировать различие стегоизображения от исходного, важной является и устойчивость перед стегоанализом. Критерием в данном случае может выступать ошибка классификации стегоизображений.

Таким образом, целевая функция имеет вид:

$$f(x^*) = \max_{n\in N, x\in X}(1 - acc_{n,x}), \quad (14)$$

где $acc_{n,x}$ — точность классификации стегоизображений при встраивании n бит в блок с величиной замены x.

Появление целевой функции требует решения задачи оптимизации. На сегодняшний день существует большое множество алгоритмов оптимизации для различных научных областей, в том числе и для оптимизации стеганографических алгоритмов.

Метод дифференциальной эволюции — один из методов эволюционного моделирования, предназначенный для решения задачи многомерной оптимизации. По классификации оптимизационных методов он относится к классу стохастических методов. Метод дифференциальной эволюции — прямой метод оптимизации, т. е. в ходе его работы требуется только вычисление значения целевой функции, но не ее производных. В общем случае целевые функции, оптимизируемые с помощью данного метода, могут быть не дифференцируемые, нелинейные, с большим количеством переменных [29].

В работах последних лет все чаще применяются биоинспирированные алгоритмы оптимизации, т. е. вдохновленные природой: чаще всего такие методы основаны на действиях различных насекомых, животных. Например, алгоритм пчелиной колонии [30]: на первом шаге пчелы разлетаются и собирают информацию, на втором они собираются и решают, какие направления обладают большим потенциалом (в случае пчел, где больше нектара). Повтор данных шагов позволит выбрать оптимальное решение.

Генетический алгоритм — это эвристический алгоритм поиска, используемый для решения задач оптимизации и моделирования путем случайного подбора, комбинирования и вариации искомых параметров с использованием механиз-

мов, аналогичных естественному отбору в природе [24]. Задача оптимизации решается благодаря применению методов, основанных на естественной эволюции: наследования, мутации, отбора, кроссовера, скрещивания.

Генетический алгоритм уже давно себя зарекомендовал в области стеганографии: многие исследователи выбирают его в качестве алгоритма оптимизации, получая хорошие результаты экспериментов. Преимущество генетического алгоритма в параллельной обработке множества альтернативных решений. Поэтому для дальнейшей работы в качестве алгоритма оптимизации выбран генетический алгоритм.

Ниже представлено описание разработанного алгоритма.

Вход: сообщение $M = m_1 m_2 ... m_L$, $m_i \in \{0, 1\}$, $i = \overline{1, L}$; пустой стегоконтейнер — цифровое изображение, сжатое по методу JPEG; величина замены x; количество бит, встраиваемых в блок, n; параметры генетического алгоритма (число поколений, особей в популяции).

Выход: стегоизображение.

Шаг 1. Изображение-контейнер разбить на K неперекрывающихся блоков размером 8×8 пикселей, вычислить квантованные ДКП-коэффициенты.

Шаг 2. Сгенерировать начальную популяцию из P особей вида $p^k = (p_1^k, p_2^k, \ldots, p_n^k)$, $p_u^k \in [0,63]$.

Шаг 3. Рассчитать значение целевой функции для каждой особи.

Шаг 4. Найти особь с наибольшим значением целевой функции и запомнить ее как p^{best}.

Шаг 5. Для $i = \overline{1, I}$ выполнить:

Шаг 5.1. Сформировать новую популяцию, исключая особей с одинаковыми значениями отдельных элементов в случае их появления.

Шаг 5.2. Рассчитать значение целевой функции для каждой особи.

Шаг 5.3. Обновить p^{best}.

Шаг 6. Осуществить встраивание в ДКП-коэффициенты стегоконтейнера с номерами, полученными в результате применения генетического алгоритма, согласно формуле (11), остальные коэффициенты обработать согласно формуле (12).

Шаг 7. Применить обратное дискретное косинусное преобразование коэффициентов, JPEG-сжатие.

Экспериментальное исследование разработанного алгоритма

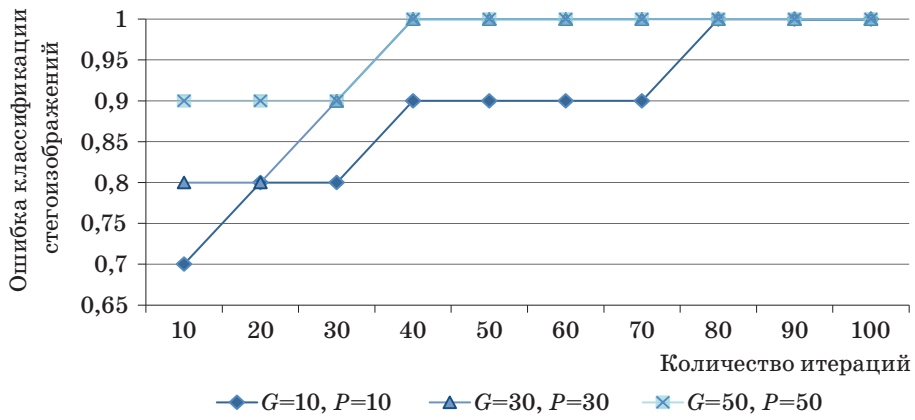
Вычислительные эксперименты проводились на выборке из 10 полутоновых тестовых JPEG-изображений разрешением 256×256 пикселей из базы USC-SIPI ID (University of Southern California Signal and Image Processing Institute Image Database) [31]. Встраиваемые сообщения представляли собой тексты на английском языке.

На рис. 2 показана зависимость ошибки классификации стегоизображений от числа итераций при встраивании сообщения длиной 2604 бита при $x = 3$ и $n = 4$.

Из проведенных экспериментов видно, что при малом числе поколений и особей в популяции алгоритму необходимо достаточно большое количество итераций (~80), чтобы достичь того минимума искажений, которое позволит скрыть от применяемого стегоанализатора факт наличия вложения в стегоизображениях.

Вполне закономерно, что при большом числе популяций и большом их размере необходимо меньше итераций (~30) для достижения поставленной цели. Однако большой объем популяций не оправдывается относительно малым количеством итераций.

Из приведенных результатов приемлем вариант с числом поколений и особей в популяции 30.



■ ***Рис. 2.*** Зависимость величины ошибки классификатора от числа итераций

■ ***Fig. 2.*** Dependence of value of the qualifier error on number of iterations

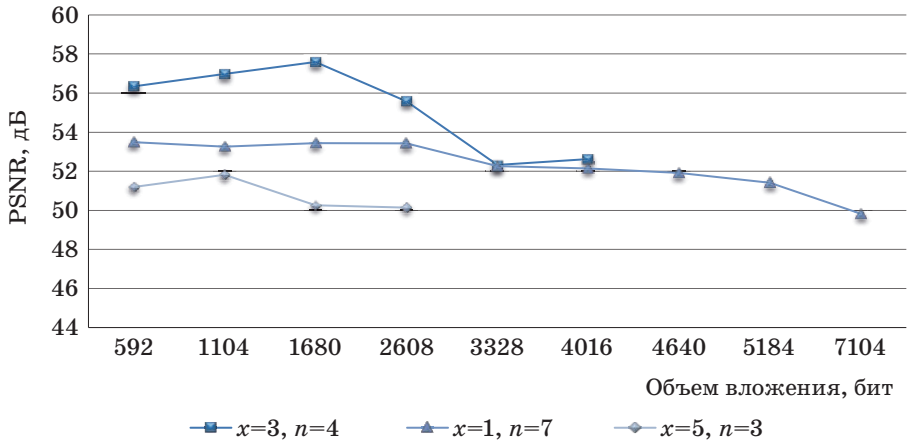
Такие параметры позволяют достичь незаметности встраивания для применяемого стегоанализатора при ~40 итерациях. В дальнейших экспериментах были применены именно эти параметры.

На рис. 3 показана зависимость значения PSNR от объема вложения для трех разных наборов параметров встраивания: величины замены x и количества бит n, встраиваемых в блок. В рассматриваемых изображениях 1024 блока, поэтому емкость стегоконтейнера ограничивается параметром n. Для каждого случая был проведен эксперимент с объемом, близким к максимальной емкости стегоконтейнера.

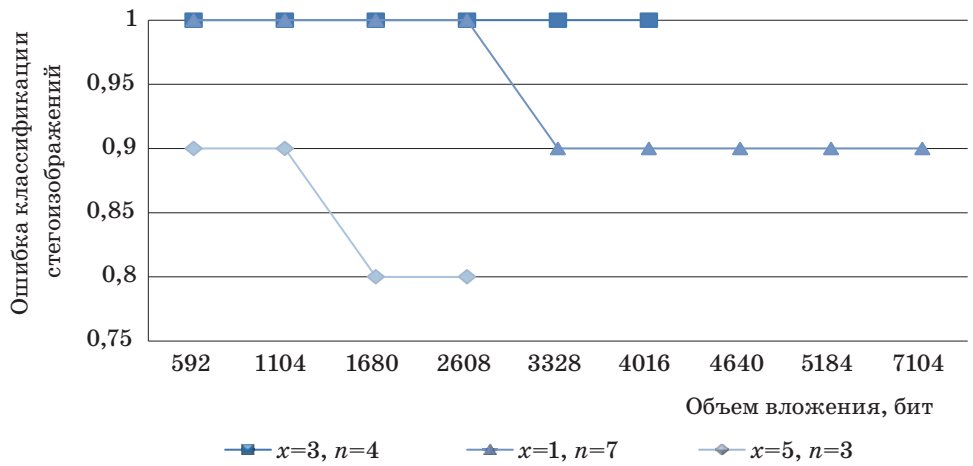
Многие авторы отмечали, что при встраивании желательно изменять коэффициенты, значения которых ближе к 0. В экспериментах с величиной замены, равной 5, значение PSNR всегда ниже, чем у аналогичных экспериментов с другими параметрами. Стоит заметить, что в матрице квантованных ДКП-коэффициентов относительно много коэффициентов со значениями, близкими к 0, и их изменение на величину 5 (не самое часто встречаемое значение в матрице) зачастую может понести большие искажения в конечном результате.

При величине замены, равной 1, значение PSNR ведет себя весьма спокойно, несмотря на увеличение объема вложения. Конечно, при большом вложении значение PSNR опускается ниже 50 дБ, но при этом в маленькое изображение встраивается достаточно большой объем данных.

При параметрах встраивания $x = 3$, $n = 4$ значение PSNR показывает очень хорошие результаты. В методе, который был взят за основу работы, при аналогичных параметрах значение PSNR опускается ниже 50 дБ при вложении порядка 3500 бит, при том, что там изображения в 4 раза больше. При малом объеме вложения адаптивный алгоритм показывает значение PSNR в среднем на 4 дБ ниже по сравнению с описанным в статье [27], но при больших вложениях пока-



■ ***Рис. 3.*** Зависимость значения PSNR от объема вложения
■ ***Fig. 3.*** Dependence of PSNR value on embedding capacity



■ ***Рис. 4.*** Зависимость ошибки классификации от объема вложения
■ ***Fig. 4.*** Dependence of value of the qualifier error on embedding capacity

зывает результаты выше. Однако стоит помнить, что целевая функция в данной работе не связана со значением PSNR, несмотря на то, что это основная метрика качества изображения, его схожести с исходным изображением, а в [27] целевой функцией является PSNR. Именно поэтому в экспериментах адаптивного алгоритма нет такой явной зависимости значения PSNR от объема вложения — «чем больше вложение, тем меньше PSNR».

■ ***Рис. 5.*** Пример работы алгоритма (исходное изображение — слева, стегоизображение — справа)

■ ***Fig. 5.*** Example of the algorithm's work (initial images — at the left, stego-images — on the right)

На рис. 4 показана зависимость ошибки классификации стегоизображений от объема вложения.

В статье [27] авторы проверяют устойчивость алгоритма к стегоанализу для величины замены, равной 3, на основе закона Бенфорда. При объеме встраивания 10 000 бит одно стегоизображение из 8 (разрешением 512×512 пикселей) было классифицировано верно.

При встраивании в изображение разрешением 256×256 пикселей 4016 бит с величиной замены $x = 3$, все стегоизображения были классифицированы неверно.

Эксперимент со встраиванием при $x = 5$ и $n = 3$ показал не только относительно плохой результат по значению PSNR, но и при проверке устойчивости перед стегоанализатором. Даже при малом вложении одно стегоизображение из 10 было классифицировано верно.

На рис. 5 представлены примеры работы алгоритма встраивания 3328 бит при $x = 3$ и $n = 4$.

Заключение

Сформирован набор признаков в пространственной и частотной областях сжатого JPEG-изображения, произведен анализ информативности стегоаналитических признаков, входящих в сформированный набор. Обучен стегоаналитический классификатор на массивах цифровых изображений, содержащих «чистые» изображения, и стегоизображения с использованием сформированного набора информативных признаков. Реализован алгоритм встраивания информации в сжатые JPEG-изображения на основе операции замены. Для минимизации искажений при встраивании была построена целевая функция на основе сформированного набора информативных признаков. Разработан адаптивный алгоритм встраивания информации в сжатые JPEG-изображения на основе операции замены с применением оптимизации. Адаптивность алгоритма заключается в том, что выбор области сокрытия основывается на наборе информативных признаков, который характеризует естественную модель цифрового изображения.

Вычислительные эксперименты показали, что можно встроить более 4 тысяч бит в полутоновое JPEG-изображение разрешением 256×256 пикселей, минимизируя искажения естественной модели цифрового изображения в области квантованных ДКП-коэффициентов настолько, что стегоанализатор на основе наивного байесовского классификатора с набором отобранных с помощью жадного алгоритма с исключением информативных признаков не выявляет факт сокрытия сообщения в изображении в 10 случаях из 10.

Литература

1. **Коханович Г. Ф., Пузыренко А. Ю.** Компьютерная стеганография. Теория и практика. К.: МК-Пресс, 2006. 288 с.
2. **Yu L., Zhao Y., Ni R., Zhu Zh.** PM1 steganography in JPEG images using genetic algorithm // Soft Computing. 2009. vol. 13(4). pp. 393–400.
3. **Yang Ch. N., Kim Ch., Lo Y. H.** Adaptive real-time reversible data hiding for JPEG images // Journal of Real-Time Image Processing. 2018. vol. 14(1). pp. 147–157. doi:10.1007/s11554-015-0555-x
4. **Guo L., Ni J., Shi Y. Q.** Uniform embedding for efficient JPEG steganography // IEEE Transactions on Information Forensics and Security. 2014. vol. 9(5). pp. 814–825.
5. **Sachnev V., Kim H. J., Shi Y., Barni M.** Ternary data hiding technique for JPEG steganography // Digital Watermarking. IWDW 2010. Lecture Notes in Computer Science. 2011. vol. 6525. pp. 202–210.
6. **Huang F., Qu X., Kim H. J., Huang J.** Reversible data hiding in JPEG images // IEEE Transactions on Circuits and Systems for Video Technology. 2015. vol. 26. doi:10.1109/TCSVT.2015.2473235
7. **Jia-Fa M., Xin-Xin M., Gang X., Wei-Guo Sh., Na-Na Zh.** A steganalysis method in the DCT domain // Multimedia Tools and Applications. 2016. № 75. pp. 5999–6019.
8. **Шумская О. О.** Метод стегоанализа JPEG-изображений на основе энергетических признаков в частотной области // Материалы международной научно-технической конференции студентов, аспирантов и молодых ученых «Научная сессия ТУСУР-2017». Томск: «В-Спектр». 2017. Ч. 6. С. 41–44.
9. **Fridrich J.** Feature-Based Steganalysis for JPEG Images and its Implications for Future Design of Steganographic Schemes // Proceedings of the Sixth International Workshop on Information Hiding, Lecture Notes in Computer Science. 2014. vol. 3200. pp. 67–81.
10. **Chen M. C.** Alpha-trimmed Image Estimation for JPEG Steganography Detection // Proceedings of the 2009 IEEE International Conference on Systems, Man, and Cybernetics. San Antonio, Texas, USA. 2009. pp. 4581–4585.
11. **Andriotis P., Oikonomou G., Tryfonas T.** JPEG steganography detection with Benford's Law // Digital Investigation. 2013. vol. 9. pp. 246–257.
12. **Fu D. Shi Y. Q., Su W.** A generalized Benford's law for JPEG coefficients and its applications in image forensics // Proceedings of SPIE 6505, security, steganography, and watermarking of multimedia contents IX. USA, San Jose. 2007. pp. 1L1–1L11.
13. **Мицель А. А., Колодникова Н. В., Протасов К. Т.** Непараметрический алгоритм текстурного анализа аэрокосмических снимков // Известия Томского политехнического университета. 2005. Т. 308(1). С. 65–70.
14. **Berg G., Davidson I., Duan M. Y., Paul G.** Searching For Hidden Messages: Automatic Detection of Steganography // Proceedings of the 15th innovative applications of artificial intelligence conference (IAAI), August 12–14. Acapulco, Mexico. 2003. pp. 51–56.
15. **Maitra S., Paul G., Sarkar S., Lehmann M., Meier W.** New results on generalization of roos-type biases and related keystream of RC // Proceedings of the 6th International conference on cryptology in africa (AFRICACRYPT), June 22–24. Cairo, Egypt. 2013. vol. 7918. pp. 222–239.
16. **Liu Q., Sung A., Qiao M., Chen Z., Ribeiro B.** An improved approach to steganalysis of JPEG images // Inf Sci. 2010. № 180(9). pp. 1643–1655.
17. **Fusheng Y., Gao T.** Novel image splicing forensic algorithm based on generalized DCT coefficient-pair histogram // Proceedings of 10th chinese conference (IGTA 2015). China, Beijing. 2015. pp. 63–71.
18. **Kodovsky J., Fridrich J.** Steganalysis of JPEG images using rich models // Proceedings of SPIE, electronic imaging, media watermarking, security, and forensics XIV. USA, San Francisco. 2012. pp. 7–20.
19. **Орешин А. Н., Саитов И. А., Орешин Н. А.** Стратегия повышения качества услуг видеосвязи на основе фильтрации видеопотока, содержащего кадры-вставки с информационным шумом // Труды СПИИРАН. 2015. Т. 41. С. 57–80.
20. **Карпов А. А., Ронжин А. Л.** Многомодальные интерфейсы в автоматизированных системах управления // Известия высших учебных заведений. Приборостроение. 2005. Т. 48. № 7. С. 9–14.
21. **Ронжин А. Л., Карпов А. А., Леонтьева А. Б., Костюченко Б. Е.** Разработка многомодального информационного киоска // Труды СПИИРАН. 2007. № 5. С. 227–246.
22. **Ронжин А. Л., Будков В. Ю.** Технологии поддержки гибридных Е-совещаний на основе методов аудиовизуальной обработки // Вестник компьютерных и информационных технологий. 2011. № 4. С. 31–35.
23. **Cormen Th. H., Leiserson Ch. E., Rivest R. L., Stein C.** Introduction to algorithms. London: MIT Press. 2013. 1324 p.
24. **Гладков Л. А., Курейчик В. В., Курейчик В. М.** Генетические алгоритмы. — М.: Физматлит, 2006. — 320 с.
25. **Guyon I., Elisseeff A.** An introduction to variable and feature selection // Journal of Machine Learning Research. 2003. vol. 3. pp. 1157–1182.
26. **Molina L. C., Belanche L., Nebot A.** Feature selection algorithms: a survey and experimental evaluation // Proceedings of the 2002 IEEE International conference on data mining. IEEE Computer Society. 2002. pp. 306–313.
27. **Евсютин О. О., Шелупанов А. А., Мещеряков Р. В., Бондаренко Д. О.** Алгоритм встраивания информации в сжатые цифровые изображения на основе операции замены с применением оптимизации // Компьютерная оптика. 2017. Т. 41(3). С. 412–421.

28. **Сэломон Д.** Сжатие данных, изображения и звука. М.: Техносфера, 2004. 368 с.
29. **Storn R., Price K.** Differential evolution — a simple and efficient heuristic for global optimization over continuous spaces // Journal of Global Optimization. 1997. vol. 11. pp. 341–359.
30. **Karaboga D., Basturk B.** On the performance of artificial bee colony (ABC) algorithm // Applied Soft Computing. 2008. vol. 8. pp. 687–697.
31. SIPI Image Database. http://sipi.usc.edu/database/ (дата обращения: 18.04.2018).

UDC 004.056.5
doi:10.31799/1684-8853-2018-5-44-56

Adaptive algorithm of replacement-based embedding of data into compressed JPEG images

O. O. Shumskaya[a], PhD student, orcid.org/0000-0002-8287-5032, shumskaya.oo@gmail.com
M. Zelezny[b], PhD, Deputy Dean, orcid.org/0000-0003-1695-4370
[a]Saint-Petersburg Institute for Informatics and Automation of the RAS, 39, 14 Line, V. O., 199178, Saint-Petersburg, Russian Federation
[b]University of West Bohemia, Pilsen, University Street, 8, no. 2732, 301 00, Plzeň, Czech Republic

Introduction: Digital steganography is an effective approach to ensuring the confidentiality of transferred and stored information. In order to provide stability before steganalysis when data are embedded into digital images, it is important to avoid the appearance of unmasking features caused by the embedding. **Purpose:** Developing an adaptive algorithm of steganographic embedding of data into compressed JPEG images based on a replacement operation, minimizing the distortions introduced to the informative features. **Results:** The paper discusses the importance of unmasking features in steganalysis and their application for adaptability of information concealment algorithms in digital objects. The main features in the spatial and frequency domains of digital images applied in modern steganographic embedding methods are specified. Informative features are selected, excluding linearly dependent features or features without any information about digital object distortion during the embedding. The resulting set has allowed us to increase the accuracy of general classification of images by 19%. On the base of the obtained set of informative features, a replacement-based adaptive modification has been developed for the algorithm of embedding data into compressed JPEG images. This modification minimizes the image container distortions during the embedding due to the use of a criterion function formulated in the article. The algorithm is adaptive because the concealment field is chosen based on the set of informative features which characterize a natural model of a digital image. Computing experiments allowed us to find the best parameter values in order to achieve good embedding capacity and the minimum distortions of the unmasking features. Experiments with the developed adaptive algorithm have demonstrated its increased stability before steganalysis and very good embedding capacity. Also, it has high values of the stego-image quality metrics, making the distortions less noticeable either for human eyes or for numerous steganalysis algorithms, because the values of unmasking features are distorted only slightly.

Keywords – steganography, steganalysis, digital images, compressed jpeg images, replacement operation, optimization, informative features, unmasking features, classification.

Citation: Shumskaya O. O., Zelezny M. Adaptive algorithm of replacement-based embedding of data into compressed JPEG images. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 5, pp. 44–56 (In Russian). doi:10.31799/1684-8853-2018-5-44-56

References

1. Kohanovich G. F., Puzyrenko A. Yu. *Computer steganography. Theory and practice*. K., MK-Press, 2006. 288 p. (In Russian).
2. Yu L., Zhao Y., Ni R., Zhu Zh. PM1 steganography in JPEG images using genetic algorithm. *Soft Computing*, 2009, vol. 13(4), pp. 393–400.
3. Yang Ch. N., Kim Ch., Lo Y. H. Adaptive real-time reversible data hiding for JPEG images. *Journal of Real-Time Image Processing*, 2018, vol. 14(1), pp. 147–157. doi:10.1007/s11554-015-0555-x
4. Guo L., Ni J., Shi Y. Q. Uniform embedding for efficient JPEG steganography. *IEEE Transactions on Information Forensics and Security*, 2014, vol. 9(5), pp. 814–825.
5. Sachnev V., Kim H. J., Shi Y., Barni M. Ternary data hiding technique for JPEG steganography. Digital watermarking. IWDW 2010. *Lecture Notes in Computer Science*, 2011, vol. 6525, pp. 202–210.
6. Huang F., Qu X., Kim H. J., Huang J. Reversible data hiding in JPEG images. *IEEE Transactions on Circuits and Systems for Video Technology*, 2015, vol. 26. doi:10.1109/TCSVT.2015.2473235
7. Jia-Fa M., Xin-Xin M., Gang X., Wei-Guo Sh., Na-Na Zh. A steganalysis method in the DCT domain. *Multimedia Tools and Applications*, 2016, № 75, pp. 5999–6019.
8. Shumskaya O. O. Steganalysis method of JPEG-images on the basis of energy features in frequency domain. *Nauchnaja sessija TUSUR-2017*, 2017, vol. 6, pp. 41–44 (In Russian).
9. Fridrich J. Feature-based steganalysis for JPEG images and its implications for future design of steganographic schemes. Proceedings of the sixth international workshop on information hiding. *Lecture Notes in Computer Science*, 2014, vol. 3200, pp. 67–81.
10. Chen M. C. Alpha-trimmed image estimation for JPEG steganography detection. *Proceedings of the 2009 IEEE international conference on systems, man, and cybernetics*. San Antonio, Texas, USA. 2009, pp. 4581–4585.
11. Andriotis P., Oikonomou G., Tryfonas T. JPEG steganography detection with Benford's law. *Digital Investigation*, 2013, vol. 9, pp. 246–257.
12. Fu D., Shi Y. Q., Su W. A generalized Benford's law for JPEG coefficients and its applications in image forensics. *Proceedings of SPIE 6505, security, steganography, and watermarking of multimedia contents IX*. USA, San Jose. 2007, pp. 1L1–1L11.
13. Mitsel A. A., Kolodnokova N. V., Protasov K. T. Nonparametric algorithm of the textural analysis of space pictures. *News of the Tomsk Polytechnic University*, 2005, vol. 308(1), pp. 65–70 (In Russian).
14. Berg G., Davidson I., Duan M. Y., Paul G. Searching for hidden messages: automatic detection of steganography. *Proceedings of the 15th innovative applications of artificial intelligence conference* (IAAI), August 12–14. Acapulco, Mexico. 2003, pp. 51–56.

15. Maitra S., Paul G., Sarkar S., Lehmann M., Meier W. New results on generalization of roos-type biases and related keystream of RC. *Proceedings of the 6th international conference on cryptology in Africa* (AFRICACRYPT), June 22–24. Cairo, Egypt. 2013, vol. 7918, pp. 222–239.
16. Liu Q., Sung A., Qiao M., Chen Z., Ribeiro B. An improved approach to steganalysis of JPEG images. *Inf. Sci.*, 2010, № 180(9), pp. 1643–1655.
17. Fusheng Y., Gao T. Novel image splicing forensic algorithm based on generalized DCT coefficient-pair histogram. *Proceedings of 10th chinese conference* (IGTA 2015). China, Beijing. 2015, pp. 63–71.
18. Kodovsky J., Fridrich J. Steganalysis of JPEG images using rich models. *Proceedings of SPIE, electronic imaging, media watermarking, security, and forensics XIV*. USA, San Francisco. 2012, pp. 7–20.
19. Oreshin A. N., Saitov I. A., Oreshin N. A. Strategy of the video communication services quality enhancement based on the filtration of a video stream containing snap-insertions with information noise. *SPIIRAS Proceedings*. 2015, vol. 41, pp. 57–80.
20. Karpov A. A., Ronzhin A. L. Multimodal interfaces in automated control systems. *Journal of Instrument Engineering*, 2005, vol. 48, № 7, pp. 9–14 (In Russian).
21. Ronzhin A. L., Karpov A. A., Leontyeva An. B., Kostuchenko B. E. The development of the multimodal information kiosk. *SPIIRAS Proceedings*, 2007, vol. 5, pp. 227–245 (In Russian).
22. Ronzhin A. L., Budkov V. Yu. Support technologies of e-meetings based on methods for audiovisual processing. *Bulletin of Computer and Information Technologies*, 2011, no. 4, pp. 31–35 (In Russian).
23. Cormen Th. H., Leiserson Ch. E., Rivest R. L., Stein C. *Introduction to algorithms*. London, MIT Press, 2013, 1324 p.
24. Gladkov L. A., Kurejchik V. V., Kurejchik V. M. *Genetic algorithms*. 2006, 320 p. (In Russian).
25. Guyon I., Elisseeff A. An introduction to variable and feature selection. *Journal of Machine Learning Research*, 2003, vol. 3, pp. 1157–1182.
26. Molina L. C., Belanche L., Nebot A. Feature selection algorithms: a survey and experimental evaluation. *Proceedings of the 2002 IEEE International conference on data mining. IEEE Computer Society*, 2002, pp. 306–313.
27. Evsutin O. O., Shelupanov A. A., Mescherjakov R. V., Bondarenko D. O. Algorithm of information embedding into compressed digital images on the basis of replacement operation with use of optimization. *Computer optics*, 2017, vol. 41(3), pp. 412–421 (In Russian).
28. Salomon D. *Data compression methods*, 2004. 368 p. (In Russian).
29. Storn R., Price K. Differential evolution — a simple and efficient heuristic for global optimization over continuous spaces. *Journal of Global Optimization*, 1997, vol. 11, pp. 341–359.
30. Karaboga D., Basturk B. On the performance of artificial bee colony (ABC) algorithm. *Applied Soft Computing*, 2008, vol. 8, pp. 687–697.
31. *SIPI Image Database*. Available at: http://sipi.usc.edu/database/ (accessed 18 April 2018).

УВАЖАЕМЫЕ АВТОРЫ!

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

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

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

УДК 621.396.9
doi:10.31799/1684-8853-2018-5-57-65

Научные статьи
Articles

Ограничение несанкционированного доступа в радиотехнических системах с широковещательной передачей информации

С. В. Штанько[a], *канд. техн. наук, доцент, orcid 0000-0002-8391-8911, craft2001@mail.ru*
[a]*Военно-космическая академия им. А. Ф. Можайского, Ждановская наб., 13, Санкт-Петербург, 197198, РФ*

Введение: *использование радиотехнических систем с широковещательной передачей информации связано с возможностью организации сплошных зон радиодоступа при ограниченном числе передающих пунктов и достаточно большом количестве абонентов (потребителей информации). Однако в системах с широковещательной передачей информации имеются особенности реализации ограничения несанкционированного доступа, обусловленные пространственной электромагнитной доступностью радиоканалов и отсутствием обратного канала передачи информации.* ***Цель:*** *разработка принципов распространения ключевой информации и управления селективным доступом абонентов к информации, передаваемой в широковещательных радиотехнических системах различного назначения, с учетом ограничений, накладываемых широковещательным режимом передачи.* ***Метод:*** *модификация используемых в радиотехнических системах с широковещательной передачей информации методов ограничения несанкционированного доступа на основе комбинации симметричных и асимметричных криптоалгоритмов.* ***Результаты:*** *анализ существующих подходов к ограничению несанкционированного доступа в системах с широковещательной передачей информации показал, что ограничения, накладываемые характером функционирования радиотехнических систем с широковещательной передачей информации, в значительной степени затрудняют или делают невозможным использование различных сетевых протоколов защищенного информационного обмена, широко распространенных в компьютерных сетях. Обоснованы принципы (наличие двух или более уровней ключевой информации, возможность замены всей ключевой информации, возможность управления доступом к системе) и способ ограничения несанкционированного доступа к радиоканалам широковещательных систем передачи информации на основе комбинированного использования симметричных и асимметричных криптоалгоритмов. В качестве симметричной части возможно использование блочных или поточных криптоалгоритмов. В качестве криптоалгоритмов асимметричной части рекомендуется использовать математический аппарат эллиптических кривых, как обладающий наилучшими криптографическими и скоростными характеристиками по сравнению с другими типами асимметричных криптоалгоритмов.* ***Практическая значимость:*** *предложенный способ ограничения несанкционированного доступа может быть использован для организации защищенного информационного обмена и управления селективным доступом к информации, передаваемой в широковещательных радиотехнических системах различного назначения.*

Ключевые слова *— широковещательная передача информации, несанкционированный доступ, ключевая информация, симметричные криптоалгоритмы, асимметричные криптоалгоритмы.*

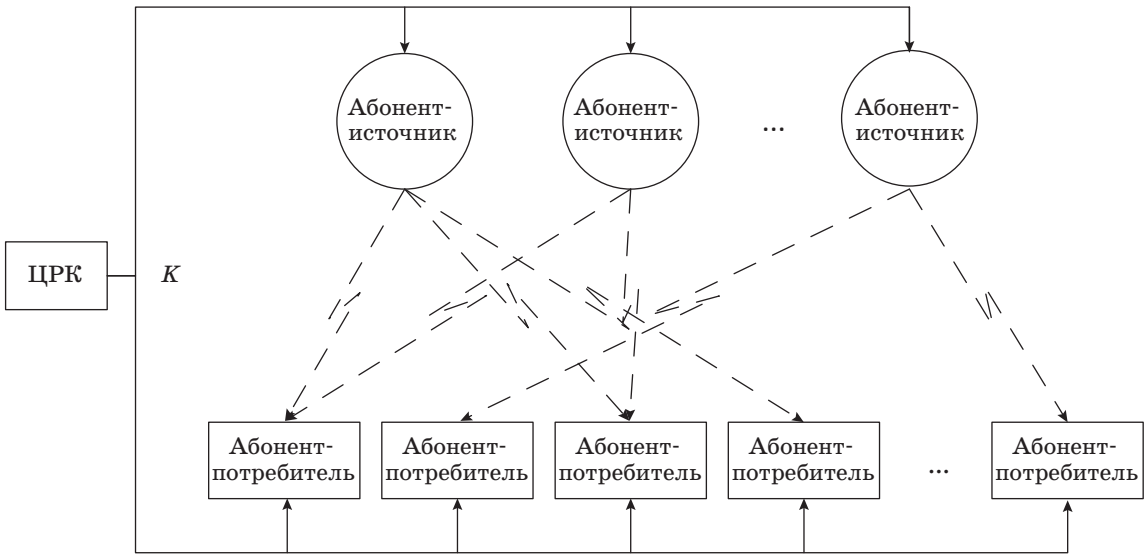
Цитирование: Штанько С. В. Ограничение несанкционированного доступа в радиотехнических системах с широковещательной передачей информации. *Информационно-управляющие системы*, 2018, № 5, с. 57–65. doi:10.31799/1684-8853-2018-5-57-65
Citation: Shtanko S. V. Restriction of unauthorized access in radio systems with broadcast data transmission. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 5, pp. 57–65 (In Russian). doi:10.31799/1684-8853-2018-5-57-65

Введение

Широкое использование радиотехнических систем с широковещательной передачей информации и расширение областей их применения [1–4] приводит к тому, что возникает задача ограничения доступа к информации, передаваемой (транслируемой) такими системами. Пространственная электромагнитная доступность, являющаяся свойством любых радиоканалов, создает условия для несанкционированного доступа (НСД) к информации, передаваемой широковещательными радиотехническими системами [5–8]. При этом особенности таких систем в ряде случаев не позволяют эффективно применять в них существующие криптографические алгоритмы и протоколы защищенного информационного обмена, аутентификации абонентов и распределения ключевой информации [6, 9–13], обеспечивающие селективный доступ и предотвращение НСД к передаваемой информации.

Общий подход к реализации ограничения доступа в системах с широковещательной передачей информации

Рассмотрим типовую радиотехническую систему с широковещательной передачей информации (широковещательную сеть) для большого количества абонентов (рис. 1). Такая система представляет собой ограниченное количество абонентов-источников и достаточно большое количество абонентов-приемников, или потреби-



■ ***Рис. 1.*** Схема радиотехнической системы с широковещательной передачей информации: ЦРК — центр распределения ключей

■ ***Fig. 1.*** Broadcast radio system scheme: ЦРК — rey distribution center

телей информации. При этом накладывается ограничение на отсутствие обратного канала от абонента-приемника (потребителя системы с широковещательной передачей информации) к абоненту-источнику.

При реализации ограничения доступа к информации, передаваемой в широковещательных радиотехнических системах, возникает необходимость защиты таких ресурсов от НСД. Основным техническим методом защиты информации, передаваемой по радиоканалам, от НСД является криптографическая защита информации [14, 15]. Условием ее реализации является распространение ключевой информации между абонентами. Генерацию и распространение ключевой информации K осуществляет ЦРК.

В основу принципа защиты информации от НСД в широковещательных сетях с учетом ограничений, накладываемых широковещательным режимом передачи, может быть положено совместное использование симметричных (одноключевых) и асимметричных (двухключевых) криптоалгоритмов [7, 16]. При этом непосредственно для шифрования передаваемых данных целесообразно использовать симметричные криптоалгоритмы, которые обладают бóльшим быстродействием при аппаратной и программной реализации, чем асимметричные [9, 14, 15], а для процедур распределения ключей, аутентификации абонентов и сообщений и управления доступом целесообразно использовать асимметричные криптоалгоритмы, которые обладают большими возможностями и гибкостью для решения таких задач [17–19].

Опишем систему ограничения НСД к радиотехнической системе с широковещательной передачей информации следующим выражением:

$$S = <\{\mathbf{A}\}, \{KC\}>, \quad (1)$$

где $\mathbf{A}$ — используемые в системе ограничения несанкционированного доступа криптоалгоритма; KC — функция, описывающая процедуру замены ключевой информации криптоалгоритмов.

Теоретико-множественное описание симметричного криптоалгоритма $\mathbf{A}$ определяется четверкой [9, 15]

$$\mathbf{A}_c = <X, K, Y, f>, \quad (2)$$

где $X = \{x_1, x_2, ..., x_n\}$ — множество открытых сообщений; $K = \{k_1, k_2, ..., k_l\}$ — множество ключей; $Y = \{y_1, y_2, ..., y_m\}$ — множество криптограмм; f: $X \times K \to Y$ — функция, однозначно определяющая отображение X на Y; $f(x, k) = y$, $x \in X$, $k \in K$, $y \in Y$.

В выражении (2) функция f определяет семейство отображений f_k: $x \to y$, $k \in K$. Тогда преобразование зашифровывания характеризует функция $E_k(x) = f_k(x) = f(x, k)$, а преобразование расшифровывания характеризует функция $D_k(y) = f^{-1}(y, k) = x$. В симметричном криптоалгоритме для процессов зашифровывания и расшифровывания используется один и тот же ключ k (либо различные ключи, но такие, что один легко вычислить из другого). Для реализации абонентского шифрования посредством симметричных криптоалгоритмов необходимо, чтобы либо

все абоненты обладали одним и тем же ключом k, либо каждая пара абонентов обладала своим ключом k, что достаточно трудно реализовать при большом количестве абонентов.

Процедура замены ключевой информации с использованием симметричных криптоалгоритмов сводится к двум возможным вариантам: либо рекурсивная генерация нового ключа с использованием предыдущего ключа, либо выбор нового ключа из заранее сформированного (сгенерированного) множества [17]. При этом к первому из указанных вариантов замены ключевой информации можно отнести как непосредственно схемы формирования нового ключа на основе некоторой циклической функции от старого ключа, так и передачу нового сгенерированного ключа по открытым каналам, шифруя его старым ключом.

Рекурсивную генерацию нового ключа с использованием предыдущего ключа можно описать следующим образом:

$$k_i = KC(k_{i-1}, r_i), \qquad (3)$$

где r_i — некоторая дополнительная информация, используемая на текущем шаге формирования ключа.

Выбор нового ключа из заранее сформированного множества может быть описан следующим образом:

$$k_i = KC(\{k_l, l = 1..n\}, \varphi), \qquad (4)$$

где φ — правило выбора очередного ключа из множества $\{k_l, l = 1..n\}$; n — мощность ключевого множества.

Следовательно, при использовании только симметричных криптоалгоритмов способа получения нового ключа без передачи дополнительной закрытой информации не существует. Кроме того, симметричные криптоалгоритмы обладают рядом ограничений: они не позволяют реализовывать эффективные процедуры KC замены ключевой информации K и процедуры аутентификации (в случае необходимости подтверждения подлинности абонентов или данных) [17–20].

Данные задачи более эффективно решаются с использованием асимметричных криптоалгоритмов. Теоретико-множественное описание асимметричного криптоалгоритма определяется пятеркой [9, 15]

$$\mathbf{A}_{\text{ac}} = <X, K, Y, E, D>, \qquad (5)$$

где $X = \{x_1, x_2, ..., x_n\}$ — множество открытых сообщений; $K = \{k_1, k_2, ..., k_l\}$ — множество ключей, включающее в себя два подмножества: $K = \{K_E, K_D\}$, K_E — подмножество ключей зашифровывания, K_D — подмножество ключей расшифровывания; $Y = \{y_1, y_2, ..., y_m\}$ — множество криптограмм; E: $X \times K_E \to Y$ — алгоритм зашифровывания; D: $Y \times K_D \to X$ — алгоритм расшифровывания.

В асимметричном алгоритме для процессов зашифровывания и расшифровывания используются различные ключи $k^{\text{о}}$ и $k^{\text{з}}$ такие, что вычисление закрытого ключа $k^{\text{з}}$ по открытому $k^{\text{о}}$ является вычислительно сложной задачей.

Для асимметричных алгоритмов функция замены ключа может быть описана следующим образом:

$$k_i = KC(k_1^{\text{о}}(k_1^{\text{з}}), k_2^{\text{о}}(k_2^{\text{з}}), r_i), \qquad (6)$$

где $k_1^{\text{о}}$, $k_2^{\text{о}}$ — открытые ключи 1-го и 2-го абонентов соответственно; $k_1^{\text{з}}$, $k_2^{\text{з}}$ — закрытые ключи 1-го и 2-го абонентов соответственно.

Выражение (6) в общем случае описывает как непосредственно шифрование информации открытыми ключами абонентов, так и использование открытых ключей для передачи закрытых ключей симметричного криптоалгоритма либо для формирования ключа симметричного криптоалгоритма.

Следовательно, асимметричные криптоалгоритмы позволяют сформировать новый закрытый ключ без передачи закрытой информации. Тем не менее применение асимметричных криптоалгоритмов в системах с широковещательной передачей информации также ограничено, так как отсутствует возможность проведения диалоговых процедур. Абоненты-потребители информации в общем случае являются пассивными участниками информационного обмена и не могут передавать свою ключевую информацию для осуществления процедур формирования сеансовых ключей и аутентификации.

Ограничения, накладываемые характером функционирования радиотехнических систем с широковещательной передачей информации, в значительной степени затрудняют или делают невозможным использование различных сетевых протоколов защищенного информационного обмена, широко распространенных в компьютерных сетях [18, 19]. Существующие в настоящее время подходы к организации защищенного информационного обмена в широковещательных радиотехнических системах в основном сводятся к использованию единственного ключа (фиксированного множества ключей) или к передаче текущих (рабочих) ключей абонентам-потребителям, зашифрованных на долговременных ключах (индивидуальных ключах, мастер-ключах), как это сделано, например, в протоколах защищенного информационного обмена в спутниковом телевидении (протоколы Viaccess, DRECrypt, «Роскрипт» и др.) [3, 4, 7, 8, 21, 22]. Однако всегда

существует ненулевая вероятность компрометации любой ключевой информации: единственного ключа, множества ключей, долговременных ключей, в том числе индивидуальных, — поэтому необходимо иметь возможность замены любой ключевой информации [23].

Для разработки полноценной системы ограничения НСД в спутниковых системах с широковещательной передачей информации необходимо обеспечить реализацию процедуры смены единого ключа, а в лучшем случае — процедуры формирования рабочих ключей, защищаемых долговременным ключом, а также процедуры смены долговременного ключа [24, 25]. Принципиальное значение при предотвращении НСД с использованием симметричных криптоалгоритмов состоит в том, что в соответствии с формулами (3) и (4) они не позволяют осуществлять замену всей ключевой информации без передачи дополнительных закрытых данных. В существующих радиотехнических системах с широковещательной передачей информации с использованием симметричных криптоалгоритмов используются различные многоуровневые схемы, имеющие целью снизить вероятность компрометации всей системы. В таких схемах ключи первого уровня используются только для шифрования ключей второго уровня, а ключи второго уровня используются непосредственно для шифрования передаваемой информации. Рассмотрим существующие способы реализации двухуровневых схем распространения ключевой информации.

Первый способ заключается в использовании единого несменяемого мастер-ключа (долговременного ключа, ключа первого уровня) для шифрования рабочих ключей (ключей второго уровня). Функция замены ключа в этом случае выглядит следующим образом:

$$k_i = KC(k_m, <k_i>), \qquad (7)$$

где k_m — мастер-ключ.

В этом случае новый рабочий ключ не зависит от предыдущего, и схема позволяет в любой момент сформировать новый ключ и распространить его, зашифровав на мастер-ключе. Рабочий ключ заменяется ЦРК с заданной периодичностью либо при необходимости. Малый объем шифруемой на мастер-ключе информации (только периодически распространяемый рабочий ключ) не позволяет нарушителю проводить его эффективный криптоанализ. Тем не менее данный способ ограничения НСД также не позволяет реализовывать механизм замены мастер-ключа.

Второй способ заключается в использовании несменяемых индивидуальных ключей абонентов (ключей первого уровня) и множества рабочих ключей (ключей второго уровня), зашифрованного предварительно на индивидуальных ключах. Расшифровать любой рабочий ключ из данного множества может только тот абонент, на чьем индивидуальном ключе данное множество зашифровано. ЦРК распространяет очередной ключ k_i (периодически либо по необходимости) в зашифрованном виде $<k_i>$ индивидуально для каждого абонента. Каждый абонент расшифровывает очередной рабочий ключ на своем индивидуальном ключе, получая один и тот же ключ. Функция замены ключа в этом случае выглядит следующим образом:

$$k_i = KC(k^{и},\{<k_l>, l = 1..n\}), \qquad (8)$$

где $k^{и}$ — индивидуальный ключ абонента; $\{<k_l>, l = 1..n\}$ — множество зашифрованных на индивидуальном ключе абонента рабочих ключей мощностью n.

Данный способ позволяет заменять рабочие ключи, передавая их абонентам в зашифрованном виде на их индивидуальных ключах, однако так же, как и в первом способе, отсутствует возможность замены долговременных индивидуальных ключей. Кроме того, в случае если индивидуальные ключи известны только самим абонентам, отсутствует возможность изменять сформированное множество ключей.

Способ ограничения НСД в радиотехнических системах с широковещательной передачей информации с комбинированным использованием асимметричных криптоалгоритмов

Существующие способы распространения ключевой информации на основе симметричных криптоалгоритмов не обладают возможностью обновления всей ключевой информации без передачи дополнительных закрытых данных, а все возможные варианты сводятся либо к рекурсивной функции (3), либо к выбору ключа из заранее сформированного множества (4). При использовании двухуровневых ключевых схем (с индивидуальными ключами, мастер-ключами) существует возможность передавать ключи второго уровня (новые рабочие ключи) в соответствии с формулами (7) и (8), однако в этом случае проблема замены ключей просто переносится на ключи первого уровня (долговременные ключи). При этом рассмотренные способы и функции замены ключей (3), (4), (7) и (8) для радиотехнических систем с широковещательной передачей информации охватывают практически все возможные варианты с использованием только симметричных криптоалгоритмов, а другие возможные варианты

являются их комбинацией в том или ином виде. Поэтому для принципиального решения вопроса замены всей ключевой информации необходимо использовать асимметричные криптоалгоритмы.

Таким образом, существуют задача разработки способа ограничения НСД, позволяющего осуществлять распространение ключевой информации в радиотехнических системах с широковещательной передачей информации с учетом их особенностей и ограничений, и задача управления доступом к системе с широковещательной передачей информации, для чего также необходимо осуществлять замену ключевой информации, в том числе индивидуальных ключей с тем, чтобы исключать при необходимости абонентов из числа допущенных к системе. Симметричные криптоалгоритмы обладают рядом ограничений, не позволяющих осуществлять замену всей ключевой информации без передачи дополнительных закрытых данных (таким образом, чтобы текущие ключи не были аргументом функции *КС* для генерации новой ключевой информации). Для решения данной проблемы необходимо модифицировать существующие схемы посредством введения асимметричных криптоалгоритмов и создания комбинированных систем ограничения НСД к радиотехническим системам с широковещательной передачей информации с учетом их особенностей и ограничений.

Для устранения недостатков существующих способов ограничения НСД в радиотехнических системах с широковещательной передачей информации необходимо разработать способ с комбинированным использованием симметричных и асимметричных криптоалгоритмов, позволяющий осуществлять замену всей ключевой информации — ключей первого и второго уровня (рабочих и долговременных), а также осуществлять управление доступом, исключая абонентов в случае их компрометации из числа авторизованных абонентов.

Для реализации такого способа ограничения НСД для шифрования и передачи рабочих ключей $\{k_i\}$ необходимо использовать асимметричный криптоалгоритм. Схема реализации способа ограничения НСД в радиотехнических системах с широковещательной передачей информации с использованием асимметричных криптоалгоритмов представлена на рис. 2. В аппаратуре шифрования *АШ* каждого абонента формируется пара ключей асимметричного криптоалгоритма (открытый и закрытый) $[k_j^{\text{о}}, k_j^{\text{з}}]$, при этом открытые ключи остаются в базе данных (матрице доступности) ЦРК, а закрытые известны только абонентам, для которых они сформированы (рис. 2, *а*). Затем каждый ключ k_i может быть зашифрован открытыми ключами абонентов и передан каждому абоненту индивидуально (рис. 2, *б*):

$$<k_i^j> = E_{Kj}^{\text{о}}(k_i), \qquad (9)$$

где $<k_i^j>$ — зашифрованный новый рабочий ключ на открытом ключе $k_j^{\text{о}}$ абонента j; $E_{Kj}^{\text{о}}$ — процесс зашифровывания на открытом ключе $k_j^{\text{о}}$ абонента j.

Каждый абонент расшифровывает рабочий ключ на своем закрытом ключе, получая один и тот же ключ:

$$k_i = D_{Kj}^{\text{з}}(<k_i^j>), \qquad (10)$$

где $D_{Kj}^{\text{з}}$ — процесс расшифровывания на закрытом ключе $k_j^{\text{з}}$ абонента j.

Процедуру замены ключей можно описать следующей функцией:

$$k_i = KC(k^{\text{о}}, \{<k_l>, l = 1..n\}). \qquad (11)$$

Затем абонент-источник широковещательной передачи информации передает информацию, шифруя ее на рабочем ключе:

$$c = E_{Ki}(m), \qquad (12)$$

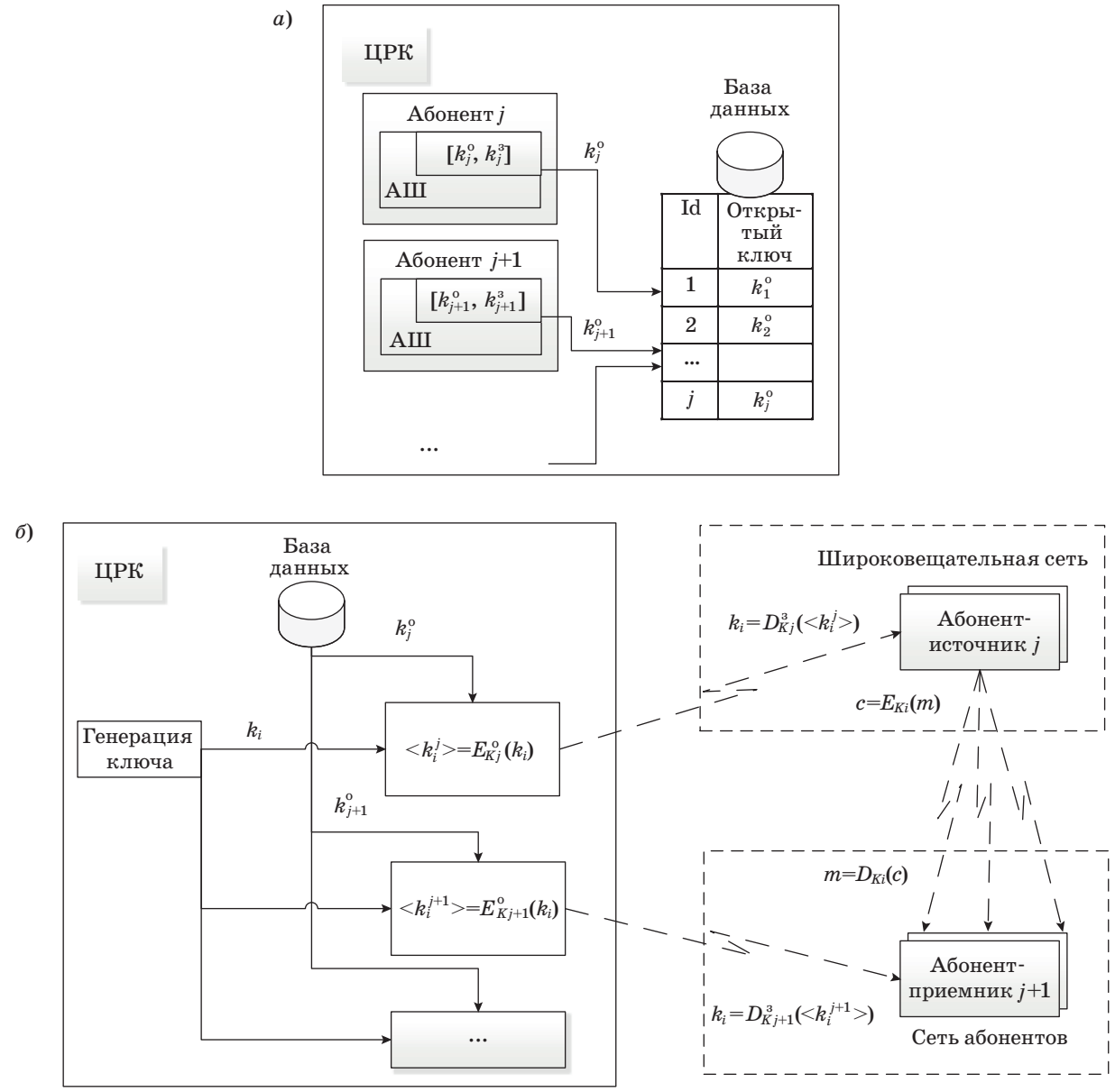
где c — зашифрованные данные; m — исходные открытые данные; E_{Ki} — процедура шифрования на ключе k_i.

Абонент-получатель расшифровывает информацию

$$m = D_{Ki}(c), \qquad (13)$$

где D_{Ki} — процедура расшифровывания на ключе k_i.

При реализации рассмотренного способа, в отличие от способов ограничения НСД на основе симметричных криптоалгоритмов, множество рабочих ключей $\{k_i\}$ не является неизменяемым и может быть в любой момент дополнено или изменено, поскольку в ЦРК имеются открытые ключи всех абонентов, в том числе абонентов-источников, и в любой момент существует возможность сгенерировать новый ключ, зашифровать его на открытых ключах абонентов и передать им. Ключи могут быть переданы абонентам-потребителям как через широковещательную сеть через абонентов-источников, так и по другим каналам. Кроме того, существует возможность смены и самих ключей $[k_j^{\text{о}}, k_j^{\text{з}}]$ абонентов. Ключи могут быть сформированы самими абонентами, и они могут передать открытые ключи $k_j^{\text{о}}$ в ЦРК, оставив закрытые ключи $k_j^{\text{з}}$ у себя. Такая схема позволит управлять доступом абонентов-потребителей информационных ресурсов к сети. В случае необ-



■ ***Рис. 2***. Схема реализации способа ограничения НСД с использованием асимметричных криптоалгоритмов: *а* — первый этап — подготовительный; *б* — второй этап — рабочий

■ ***Fig. 2***. The unauthorized access restriction method scheme with asymmetric crypto algorithms using: *а* — the first stage — the preparatory phase; *б* — the second stage — the worker

ходимости исключения абонента-потребителя из числа авторизованных новый рабочий ключ k_i не шифруется на его открытом ключе k_j^o и не передается по сети, в результате чего абонент-потребитель теряет доступ к ресурсам сети.

При реализации предложенного способа ограничения несанкционированного доступа на основе комбинированного использования симметричных и асимметричных криптоалгоритмов в качестве симметричной части возможно использование блочных или поточных криптоалгоритмов, например, отечественных криптоалгоритмов шифрования, определяемых ГОСТ Р 34.12-2015 (например, шифр «Магма», ранее определяемый ГОСТ 28147-89). Данный шифр имеет три режима шифрования (режим простой замены, режим гаммирования, режим гаммирования с обратной связью) и режим выработки имитовставки. В качестве криптоалгоритмов асимметричной части рекомендуется использовать математический аппарат эллиптических кривых как обладающий наилучшими криптографическими и скоростными характеристиками по сравнению с другими типами асимметричных криптоалгоритмов.

Заключение

Предложенный способ ограничения несанкционированного доступа к радиоканалам широковещательных систем передачи информации на основе комбинированного использования симметричных и асимметричных криптоалгоритмов позволяет решать поставленную задачу замены всей ключевой информации — как рабочих ключей, так и индивидуальных (ключей первого и второго уровня), а также управлять доступом абонентов-получателей к системе. Кроме того, используя асимметричные криптоалгоритмы, можно осуществлять дальнейшее развитие предложенного способа, реализовывать более гибкие схемы и различные сервисы, в том числе сервисы аутентификации. Данный способ может быть применен для организации защищенного информационного обмена и управления доступом к радиотехническим системам с широковещательной передачей информации различного назначения.

Литература

1. **Сакалема Д. Ж.** Подвижная радиосвязь. — М.: Горячая линия–Телеком, 2012. — 512 с.
2. **Карташевский В. Г., Семенов С. Н., Фирстова Т. В.** Сети подвижной связи. — М.: Эко-Трендз, 2001. — 296 с.
3. **Кукк К. И.** Спутниковая связь: прошлое, настоящее, будущее. — М.: Горячая линия–Телеком, 2016. — 256 с.
4. **Костин М. В.** Системы условного доступа//Телеспутник. 2004. № 11(109). С. 62–64.
5. **Нелюб С. А.** Защита речевой информации в радиосетях связи // Инженерный вестник. 2012. № 9. http://ainjournal.ru/doc/501059.html (дата обращения: 15.07.2018).
6. **Мальцев Г. Н., Штанько С. В.** Протоколы аутентификации абонентов и защиты информации на основе асимметричных криптоалгоритмов // Проблемы информационной безопасности. Компьютерные системы. 2003. № 1. С. 51–56.
7. **Штанько С. В., Лесняк Д. А.** Алгоритмы защищенного информационного обмена в радиоканалах космической навигационной системы // Известия высших учебных заведений России. Радиоэлектроника. 2015. № 5. С. 47–51.
8. **Гладченков А.** Спутниковые технологии VSAT и информационная безопасность сети // Журнал сетевых решений/LAN. 2007. № 09. https://www.osp.ru/lan/2007/09/4374549/ (дата обращения: 06.05.2018).
9. **Stallings W.** Cryptography and Network Security Principles and Practices. Fourth Ed. — New Jersey: Prentice Hall, 2005. — 592 p.
10. **Yevpak S. A.** The Special Broadcast Security Scheme based on RM-codes and the Protection from Some Linear Algebraic Attacks// Numerical Algebra with Applications: Proc. of Fourth China-Russia Conf., Rostov-on-Don, SFedU, 2015. P. 183.
11. **Деундяк В. М., Косолапов Ю. В.** Криптосистема на индуцированных групповых кодах // Моделирование и анализ информационных систем. 2016. № 23(2). С. 137–152. https://doi.org/10.18255/1818-1015-2016-2-137-152 (дата обращения: 06.05.2018).
12. **Деундяк В. М., Таран А. А.** О применении кодов Хэмминга в системе распределения ключей для конференций в многопользовательских системах связи // Вестник ВГУ. Серия: Системный анализ и информационные технологии. 2015. № 3. С. 43–50.
13. **Щеглов К. А., Щеглов А. Ю.** Новый подход к защите данных в информационной системе // Изв. вузов. Приборостроение. 2015. Т. 58. № 3. С. 157–166. doi:10.17586/0021-3454-2015-58-3-157-166/issn0021-3454
14. **Шаньгин В. Ф.** Защита информации в компьютерных системах и сетях. — М.: ДМК, 2012. — 592 с.
15. **Schneier B.** Applied Cryptography. Protocols, Algorthms and Source Code in C. — New York: John Wiley & Sons, 2015. — 784 p.
16. **Корниенко А. А., Штанько С. В.** Криптографический протокол защиты информации в радиоканалах сетевых спутниковых систем с использованием асимметричных алгоритмов // Информационно-управляющие системы. 2006. № 5. С. 21–26.
17. **Остроумов О. А., Синюк А. Д.** Протокол открытого формирования трехстороннего ключа // Наукоемкие технологии в космических исследованиях Земли. 2014. Т. 6. Вып. 2. С. 48–52.
18. **Штанько С. В., Жукова Н. А.** Схемы аутентификации данных и пользователей в распределенных информационных системах // Изв. СПбГЭТУ «ЛЭТИ». 2012. № 8. С. 46–51.
19. **Штанько С. В., Лесняк Д. А.** Обеспечение селективного доступа при широковещательной передаче информации // Информационно-управляющие системы. 2016. № 1. С. 74–79. doi:10.15217/issn1684-8853.2016.1.74
20. **Osipov D. S., Frolov A. A., Zyablov V. V.** Multiple Access System for a Vector Disjunctive Channel//Problems of Information Transmission. 2012. Vol. 48. N 3. P. 243–249.
21. **Frolov A. A., Zyablov V. V.** A New Coding Method for a Multiple-Access System with a Large Number of Active users// Proc. IEEE Information Theory Workshop (ITW), 2015, April 26 –May 1, Jerusalem, Israel. 2015. P. 1–5.
22. **Анисимов Д. В., Дмитриев С. В.** Управление доступом к среде передачи данных в беспроводных сетях стандарта IEEE 802.11 с учетом ненасыщенного со-

стояния канала // Информационные системы и технологии. 2018. № 4(108). С. 99–107. issn2072-8964

23. **Мальцев Г. Н., Панкратов А. В., Лесняк Д. А.** Исследование вероятностных характеристик изменения защищенности информационной системы от несанкционированного доступа нарушителей // Информационно-управляющие системы. 2015. № 1. С. 50–59. doi:10.15217/issn1684-8853.2015.1.50
24. **Gentry C., Waters B.** Adaptive Security in Broadcast Encryption Systems// Advances in Cryptology — EUROCRYPT 2009. Springer, 2009. P. 171–188.
25. **Blundo C., Mattos L. A. F., Stinson D. R.** Trade-offs between Communication and Storage in Unconditionally Secure Schemes for Broadcast Encryption and Interactive Key Distribution//Advances in Cryptology. LNCS. 1996. Vol. 1109. P. 387–400.

UDC 621.396.9
doi:10.31799/1684-8853-2018-5-57-65

Restriction of unauthorized access in radio systems with broadcast data transmission

S. V. Shtanko[a], PhD, Tech., Associate Professor, orcid 0000-0002-8391-8911, craft2001@mail.ru
[a]A. F. Mozhaiskii Military Space Academy, 13, Zhdanovskaia Emb., 197198, Saint-Petersburg, Russian Federation

Introduction: The usage of broadcast radio systems is associated with the possibility of organizing continuous radio access zones with a limited number of transmission points and a considerable number of subscribers (information consumers). However, radio broadcasting systems have specific features of unauthorized access restriction related to the spatial electromagnetic accessibility of radio channels and the absence of a reverse information channel. **Purpose:** Developing principles for key information distribution and for the control over selective access of subscribers to the data transmitted in broadcast radio systems of various purposes, taking into account the restrictions imposed by the broadcasting mode of transmission. **Method:** Modification of the unauthorized access restriction methods used in radio broadcast systems, based on a combination of symmetric and asymmetric crypto algorithms. **Results:** The analysis of the existing approaches to unauthorized access restriction in broadcast radio systems has shown that the restrictions imposed by the peculiarities of these systems often make it difficult or impossible to apply various network protocols of protected data exchange commonly used in computer networks. We substantiate the principles (presence of two or more levels of key information; possibility to replace all the key information; ways to control the access to the system) and a way to restrict unauthorized access to the radio channels of broadcast systems, based on the combined use of symmetric and asymmetric crypto algorithms. As the symmetric part, you can use block or stream ciphers. As algorithms for the asymmetric part, elliptic curves are recommended, as they have the best cryptographic and speed characteristics as compared to other types of asymmetric cryptographic algorithms. **Practical relevance:** The proposed unauthorized access restriction method can be used to organize secure data exchange and selective control over the access to data transmitted in broadcast radio systems of various purposes.

Keywords — broadcast data transmission, unauthorized access, key information, symmetric crypto algorithms, asymmetric crypto algorithms.

Citation: Shtanko S. V. Restriction of unauthorized access in radio systems with broadcast data transmission. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 5, pp. 57–65 (In Russian). doi:10.31799/1684-8853-2018-5-57-65

References

1. Sakalema D. Zh. *Podvizhnaia radiosviaz'* [Mobile Radio Communication]. Moscow, Goriachaia liniia–Telecom Publ., 2012. 512 p. (In Russian).
2. Kartashevskii V. G., Semenov S. N., Firstova T. V. *Seti podvizhnoi sviazi* [Mobile Networks]. Moscow, Eko-Trendz Publ., 2001. 296 p. (In Russian).
3. Kukk K. I. *Sputnikovaia sviaz': proshloe, nastoiashchee, budushchee* [Satellite Communication: Past, Present, Future]. Goriachaia liniia–Telecom Publ., 2016. 256 p. (In Russian).
4. Kostin M. V. Conditional Access Systems. *Telesputnik*, 2004, no. 11(109), pp. 62–64 (In Russian).
5. Nelyub S. A. Speech Information Protection in Radio Networks. *Inzhenernyj vestnik*, 2012, no. 9. Available at: http://ainjournal.ru/doc/501059.html (accessed 15 July 2018) (In Russian).
6. Mal'tsev G. N., Shtan'ko S. V. Protocols of Authentication of Subscribers and Information Security on the Basis of Asymmetric Cryptoalgorithms. *Problemy informatsionnoi bezopasnosti. Komp'iuternye sistemy* [Information Security Problems. Computer Systems], 2003, no. 1, pp. 51–56 (In Russian).
7. Shtan'ko S. V., Lesniak D. A. Algorithms for Secure Information Exchange in Space Radio Navigation Systems. *Izvestiia vysshikh uchebnykh zavedenii Rossii. Radioelektronika*, 2015, no. 5, pp. 47–51 (In Russian).
8. Gladchenkov A. Satellite VSAT Technology and Information Security Network. *Zhurnal setevykh reshenii/LAN*, 2007, no. 9. Available at: https://www.osp.ru/lan/2007/09/4374549/ (accessed 06 May 2018) (In Russian).
9. Stallings W. *Cryptography and Network Security Principles and Practices*. Fourth Ed. New Jersey, Prentice Hall, 2005. 592 p.
10. Yevpak S. A. The Special Broadcast Security Scheme based on RM-codes and the Protection from Some Linear Algebraic Attacks. *Proc. of Fourth China-Russia Conference "Numerical Algebra with Applications"*, Rostov-on-Don, SFedU, 2015, pp. 183.
11. Deundyak V. M., Kosolapov Y. V. Cryptosystem Based on Induced Group Codes. *Modelirovanie i analiz informatsionnykh sistem*, 2016, no. 23(2), pp. 137–152. Available at: https://doi.org/10.18255/1818-1015-2016-2-137-152 (accessed 06 May 2018) (In Russian).
12. Deundyak V. M., Taran A. A. On the Hamming Codes Application in the Key Distribution System for Conferencing in a Multi-user Communication Systems. *Vestnik VGU. Seriya: Sistemnyj analiz i informacionnye tekhnologii*, 2015, no. 3, pp. 43–50. issn1995-5499 (In Russian).
13. Shcheglov K. A., Shcheglov A. Yu. New Approach to Data Securing in Information System. *Izvestiya vysshikh uchebnykh zavedeniy. Priborostroenie*, 2015, vol. 58, no. 3, pp. 157–166 (In Russian). doi:10.17586/0021-3454-2015-58-3-157-166/issn0021-3454
14. Shan'gin V. F. *Zashchita informacii v komp'yuternyh sistemah i setyah* [Information Security in Computer Systems and Networks]. Moscow, DMK Publ., 2012. 592 p. (In Russian).
15. Schneier B. *Applied Cryptography. Protocols, Algorthms and Source Code in C*. New York, John Wiley & Sons, 2015. 784 p.

16. Kornienko A. A., Shtanko S. V. Information Security Protocol for Network Satellite Systems using Asymmetric Algorithms. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2006, no. 5, pp. 21–26 (In Russian).
17. Ostroumov O. A., Sinyuk A. D. Protocol of Open Formation of a Tripartite Key. *Naukoemkie tekhnologii v kosmicheskikh issledovaniyakh Zemli* [High Tech in Earth Space Research], 2014, no. 2(82), pp. 48–52 (In Russian).
18. Shtanko S. V., Zhukova N. A. Schemes of Authentication of Data and users in the Distributed Information Systems. *Izvestiia SPbGETU «LETI»*, 2012, no. 8, pp. 46–51 (In Russian).
19. Shtanko S. V., Lesniak D. A. Providing Selective Access for Broadcasting Information Transfer. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2016, no. 1, pp. 74–79 (In Russian). doi:10.15217/issn1684-8853.2016.1.74
20. Osipov D. S., Frolov A. A., Zyablov V. V. Multiple Access System for a Vector Disjunctive Channel. *Problems of Information Transmission*, 2012, vol. 48, no. 3, pp. 243–249.
21. Frolov A. A., Zyablov V. V. A New Coding Method for a Multiple-Access System with a Large Number of Active users. *Proc. IEEE Information Theory Workshop (ITW)*, 2015, April 26–May 1, Jerusalem, Israel, 2015, pp. 1–5.
22. Anisimov D. V., Dmitriev S. V. Managing Access to the Transmission Medium in Wireless Networks IEEE 802.11 Standard Taking into Account the Unsaturated State of the Channel. *Informatsionnye sistemy i tekhnologii* [Information Systems and Technologies], 2018, no. 4(108), pp. 99–107 (In Russian). issn2072-8964
23. Maltsev G. N., Pankratov A. V., Lesniak D. A. Probabilistic Characteristics of Information System Security Changes under Unauthorized Access. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2015, no. 1, pp. 50–59 (In Russian). doi:10.15217/issn1684-8853.2015.1.50
24. Gentry C., Waters B. Adaptive Security in Broadcast Encryption Systems. *Advances in Cryptology — EUROCRYPT 2009*, Springer, 2009, pp. 171–188.
25. Blundo C., Mattos L. A. F., Stinson D. R. Trade-offs between Communication and Storage in Unconditionally Secure Schemes for Broadcast Encryption and Interactive Key Distribution. *Advances in Cryptology, LNCS*, 1996, vol. 1109, pp. 387–400.

Научный журнал
«ИНФОРМАЦИОННО-УПРАВЛЯЮЩИЕ СИСТЕМЫ»
выходит каждые два месяца.

Стоимость годовой подписки (6 номеров) для подписчиков России — 6000 рублей,
для подписчиков стран СНГ — 6600 рублей, включая НДС 18%, таможенные и почтовые расходы.

Подписку на печатную версию журнала можно оформить в любом отделении связи по каталогу:
«Роспечать»: № 15385 — полугодовой индекс,
а также через посредство подписных агентств:
«Северо-Западное агентство „Прессинформ“»
Санкт-Петербург, тел.: (812) 335-97-51, 337-23-05,
эл. почта: press@crp.spb.ru, zajavka@crp.spb.ru,
сайт: http://www.pinform.spb.ru
«МК-Периодика» (РФ + 90 стран)
Москва, тел.: (495) 681-91-37, 681-87-47,
эл. почта: export@periodicals.ru, сайт: http://www.periodicals.ru
«Деловая пресса»
Москва, тел.: (495) 962-11-11, эл. почта: podpiska@delpress.ru,
сайт: http://delpress.ru/contacts.html
«Коммерсант-Курьер»
Казань, тел.: (843) 291-09-99, 291-09-47, эл. почта: kazan@komcur.ru,
сайт: http://www.komcur.ru/contacts/kazan/
«Урал-Пресс» (филиалы в 40 городах РФ)
Сайт: http://www.ural-press.ru
«Идея» (Украина)
Сайт: http://idea.com.ua
«BTL» (Узбекистан)
Сайт: http://btl.sk.uz/ru/cat17.html и др.

На электронную версию нашего журнала (все выпуски, годовая подписка, один выпуск, одна статья)
вы можете подписаться на сайтах НЭБ: http://elibrary.ru; РУКОНТ: http://www.rucont.ru;
ИВИС: http://www.ivis.ru; Некс-Медиа: http://biblioclub.ru/index.php?page=news&id=11196

Полнотекстовые версии журнала за 2002–2017 гг.
в свободном доступе на сайте журнала (http://www.i-us.ru),
НЭБ (http://www.elibrary.ru)
и Киберленинки (http://cyberleninka.ru/journal/n/informatsionno-upravlyayuschiesistemy).

УДК 621.391
doi:10.31799/1684-8853-2018-5-66-78

Научные статьи
Articles

Решетчатые сигнально-кодовые конструкции для каналов с линейными искажениями

Ф. А. Таубин[a], *доктор техн. наук, профессор, orcid.org/0000-0002-8781-9531, ftaubin@yahoo.com*
[a]*Санкт-Петербургский государственный университет аэрокосмического приборостроения, Б. Морская ул., 67, Санкт-Петербург, 190000, РФ*

Введение: *реальные каналы связи при высокоскоростной передаче данных характеризуются заметными линейными искажениями, в результате чего эффективность известных сигнально-кодовых конструкций, ориентированных на каналы без искажений, существенно снижается. Это означает, что структура сигнально-кодовых конструкций для каналов с линейными искажениями должна быть согласована с особенностями искажений, вносимых каналом.* ***Цель:*** *построение на основе многомерных сигнальных множеств многоуровневых сигнально-кодовых конструкций с многоэтапным декодированием, ориентированных на каналы с линейными искажениями, и исследование основных характеристик этих конструкций.* ***Результаты:*** *разработан новый подход к построению сигнально-кодовых конструкций для каналов с линейными искажениями, базирующийся на комбинировании двух процедур: преобразовании исходного канала в совокупность независимых подканалов без памяти и частотно-временного многоуровневого кодирования. Высокая степень гибкости предложенной кодовой конструкции позволяет эффективно учитывать и компенсировать влияние канальных искажений за счет рационального выбора числа подканалов и оптимизации распределения энергии между подканалами, а также обеспечивает широкий диапазон обменных соотношений между помехоустойчивостью и сложностью декодирования. В явном виде получены соотношения, связывающие между собой основные параметры предложенного класса сигнально-кодовых конструкций. Полученные соотношения позволяют установить обменные соотношения между скоростью передачи, минимальным расстоянием и сложностью декодирования. Приведены примеры конструкций, построенных с использованием многомерных решеток Барнса — Уолла. Сравнительный анализ построенных решетчатых сигнально-кодовых конструкций показал, что с увеличением значения допустимой сложности декодирования более предпочтительными оказываются конструкции с большими значениями как количества подканалов, так и мощности сигнального множества.* ***Практическая значимость:*** *представленный класс сигнально-кодовых конструкций позволяет эффективно учитывать и компенсировать влияние линейных искажений, характерных для реальных каналов передачи данных, обеспечивая при этом сравнительно небольшое значение пик-фактора передаваемых сигналов и тем самым улучшение коэффициента полезного действия передатчика.*

Ключевые слова *— кодированная модуляция, сигнально-кодовые конструкции, каналы с линейными искажениями, многомерные сигнальные множества, решетки Барнса — Уолла, многоуровневые коды, многоэтапное декодирование.*

Цитирование: Таубин Ф. А. Решетчатые сигнально-кодовые конструкции для каналов с линейными искажениями. *Информационно-управляющие системы*, 2018, № 5, с. 66–78. doi:10.31799/1684-8853-2018-5-66-78
Citation: Taubin F. A. Trellis-coded modulation for linear distortion channels. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 5, pp. 66–78 (In Russian). doi:10.31799/1684-8853-2018-5-66-78

Введение

Кодированная модуляция, возникшая в результате объединения процедур модуляции и кодирования [1–3], представляет собой весьма эффективную технику надежной передачи данных, позволяющую обеспечить одновременно как значительный кодовый выигрыш, так и высокую спектральную эффективность. Доминирующей тенденцией при исследовании кодированной модуляции и разработке конкретных методов передачи (сигнально-кодовых конструкций согласно установившейся в отечественной литературе терминологии) являлась и является поныне [4, 5] ориентация на простейшую модель передающей среды — канал с аддитивным белым гауссовым шумом (АБГШ).

Вместе с тем, как хорошо известно, отличительной особенностью реальных каналов передачи данных является жесткое ограничение рабочей полосы частот фильтрами минимально-фазового типа, вызывающими при передаче дискретных сообщений межсимвольную интерференцию (МСИ). Уровень МСИ резко возрастает с увеличением скорости передачи, в результате чего непосредственное использование в высокоскоростных системах сигнально-кодовых конструкций, ориентированных на канал с АБГШ, связано с существенным снижением величины реализуемого энергетического выигрыша. Это следует из результатов ряда работ, связанных с анализом помехоустойчивости известных методов кодированной модуляции в каналах с МСИ [6–8].

Достаточно общий подход к организации передачи по каналу с искажениями базируется на том факте, что кодер, модулятор и канал с МСИ можно описать в совокупности с помощью единой (общей) решетчатой диаграммы [9]. Известные методы конструирования решетчатых кодов с хорошими дистанционными характеристиками могут быть

адаптированы применительно к этой ситуации. Однако чрезмерно большое число состояний единой решетчатой диаграммы порождает по меньшей мере две проблемы. Во-первых, усложняется вычисление даже основной дистанционной характеристики — минимального евклидова расстояния. Во-вторых, процедуры декодирования, оперирующие с сокращенным набором состояний, трудно поддаются анализу [10, 11]. Тем самым существенно ограничивается возможность использования этого подхода для практически интересных каналов.

В качестве альтернативного подхода, позволяющего устранить МСИ (путем введения защитных интервалов), в этой ситуации часто рассматривается передача на основе ортогонального частотного разделения с мультиплексированием (Orthogonal Frequency Division Multiplexing — OFDM). Существенным недостатком этой технологии является высокое значение пик-фактора, так как передаваемый OFDM-сигнал представляет собой сумму большого числа (как правило, более 100) модулированных гармоник; в результате для усиления OFDM-сигналов требуется высокая линейность амплитудной и равномерность фазоамплитудной характеристик усилительного тракта, что приводит к снижению коэффициента полезного действия передатчика. Кроме того, известные подходы к введению кодированной модуляции в схему с OFDM-передачей ограничены, как правило, рассмотрением традиционных методов с двумерными сигнальными множествами [12, 13], что заметно ограничивает ее эффективность.

Рассматриваемый в данной работе новый метод кодированной модуляции для каналов с линейными искажениями позволяет преодолеть трудности, возникающие при использовании известных подходов. Он базируется на комбинировании двух процедур: преобразования исходного канала в совокупность независимых подканалов без памяти и частотно-временного многоуровневого кодирования на основе многомерных сигнальных множеств. Основная идея первоначально была сформулирована автором в работах [14, 15]. Близкий по смыслу подход (но для каналов с дискретным временем) рассматривался в работах [16, 17].

Модель передачи

Двоичная информационная последовательность $\mathbf{a}$ разбивается на k-блоки $a_0, a_1, a_2, ..., a_i = (a_{i1}, ..., a_{ik})$ и кодируется линейным кодом C над полем $\mathbb{F}_2^m$ со скоростью $R = k/n$ бит/символ. Последовательность кодовых символов $\mathbf{c} = (\mathbf{c}_0, \mathbf{c}_1, \mathbf{c}_2, ...)$, $\mathbf{c}_j = (c_{j1}, ..., c_{jm})$ с помощью отображения $\varphi: \mathbb{F}_2^m \to \mathbb{C}^N$ преобразуется в последовательность канальных символов $\mathbf{b}_0$, $\mathbf{b}_1$, $\mathbf{b}_2$, ...; канальные символы представляют собой комплексные N-мерные векторы $\mathbf{b}_j = (b_{j1}, ..., b_{jN})$, $b_{jl} \in \mathbb{C}$, принимающие значения из 2^m-точечного множества B, $B \subset \mathbb{C}^N$. Обычно множество значений канальных символов называют сигнальным множеством, и хотя в общем случае употребление прилагательного «сигнальное» не всегда уместно, будем придерживаться традиционного названия. Совокупность всех последовательностей канальных символов будем называть модуляционным кодом U.

Последовательность канальных символов $\mathbf{b} = (\mathbf{b}_0, \mathbf{b}_1, \mathbf{b}_2, ...)$ с тактовой частотой $1/T$ передается по каналу посредством абсолютной модуляции сигналов с полным откликом, т. е. комплексная огибающая $u(\cdot)$ суммарного передаваемого сигнала имеет вид

$$u(t) = \sqrt{2RE_b / \sigma^2} \sum_{j \geq 0} \mathbf{b}_j \mathbf{a}(t - jT),\ t \geq 0,$$

где E_b — энергия, затрачиваемая на передачу одного бита, σ^2 — среднее значение квадрата евклидовой нормы канальных символов $\mathbf{b}_j$,

$$\sigma^2 = \overline{\|\mathbf{b}_j\|^2}; \qquad (1)$$

черта сверху в (1) означает усреднение по распределению на сигнальном множестве B, индуцированному распределением информационных k-блоков, $\mathbf{a}(\cdot) = \mathrm{col}\ (a_1(\cdot), ..., a_N(\cdot))$ — вещественная вектор-функция, компоненты которой образуют ортонормированную систему в пространстве $L^2(0, T)$.

Пару $(U, \mathbf{a}(\cdot))$ будем называть входной сигнально-кодовой конструкцией (СКК).

Основными количественными характеристиками СКК являются: а) коэффициент использования полосы частот $\gamma = \upsilon/W$, где $\upsilon = R/T$ — скорость передачи, W — полоса частот, занимаемая модулирующей вектор-функцией $\mathbf{a}(\cdot)$, б) минимальное евклидово расстояние

$$\min_{\mathbf{b}, \hat{\mathbf{b}} \in U, \mathbf{b} \neq \hat{\mathbf{b}}} \left(\frac{RE_b}{\sigma^2} \sum_{j \geq 0} \left\| \mathbf{b}_j - \hat{\mathbf{b}}_j \right\|^2 \right)^{1/2}. \qquad (2)$$

Будем полагать, что передача ведется по линейному стационарному каналу с симметричной относительно несущей частоты передаточной характеристикой; низкочастотный эквивалент (комплексная огибающая) $h(\cdot)$ импульсной характеристики такого канала является вещественной функцией. Комплексная огибающая $r(\cdot)$ сигнала на выходе канала имеет вид

$$r(t) = \sqrt{2RE_b / \sigma^2} \sum_{j \geq 0} \mathbf{b}_j \mathbf{g}(t - jT) + n(t),\ t \geq 0,$$

где $\mathbf{g}(\cdot) = \mathrm{col}(g_1(\cdot), \ldots, g_N(\cdot))$, $g_l(\cdot)$ — свертка функций $a_l(\cdot)$ и $h(\cdot)$, $n(\cdot)$ — комплексный АБГШ с двусторонней спектральной плотностью N_0. В терминах СКК влияние линейных искажений, вносимых каналом, выражается в замене модулирующей вектор-функции $\mathbf{a}(\cdot)$ ее реакцией $g(\cdot)$ на выходе канала. Иначе говоря, исходная СКК $(U, \mathbf{a}(\cdot))$ трансформируется в выходную СКК $(U, \mathbf{g}(\cdot))$ с минимальным евклидовым расстоянием

$$\delta = \min_{\mathbf{b},\hat{\mathbf{b}}\in U,\mathbf{b}\neq\hat{\mathbf{b}}} \left(\frac{RE_b}{\sigma^2} \int_0^\infty \left| \sum_{j\geq 0} \left(\mathbf{b}_j - \hat{\mathbf{b}}_j\right) \mathbf{g}\left(t - jT\right) \right|^2 \mathrm{d}t \right)^{1/2}. \quad (3)$$

Принятый сигнал обрабатывается когерентным демодулятором, который формирует последовательность $\mathbf{r} = (\mathbf{r}_0, \mathbf{r}_1, \mathbf{r}_2, \ldots)$, $\mathbf{r}_j = (r_{j1}, \ldots, r_{jN})$ N-мерных комплексных векторов, следующих с тактовой частотой $1/T$ и имеющих вид

$$\mathbf{r}_j = \int_0^\infty r(t)\mathbf{p}(t - jT)\mathrm{d}t = 0,\ j \geq 0,$$

где $\mathbf{p}(\cdot) = (p_1(\cdot), \ldots, p_N(\cdot))$ — демодулирующая вектор-функция такая, что совокупность всех ее временных сдвигов, кратных T, образует базис в линейном пространстве, натянутом на совокупность всех временных сдвигов вектор-функции $\mathbf{g}(\cdot)$. Последовательность $\mathbf{r} = (\mathbf{r}_0, \mathbf{r}_1, \mathbf{r}_2, \ldots)$ «мягких решений» поступает в декодер модуляционного кода U, который формирует оценку $\hat{\mathbf{a}}$ переданной информационной последовательности.

При декодировании по максимуму правдоподобия (МП) главный член вероятности ошибки декодирования есть, очевидно, $O(\exp(-\delta^2/4N_0))$, и в этом смысле МП-декодер реализует расстояние δ. В общем случае анализ обменных соотношений между сложностью декодера и реализуемым расстоянием и последующая оптимизация компонентов СКК представляют собой весьма сложную задачу. Одно из главных препятствий — сложная зависимость между характеристиками компонентов СКК, в частности минимальным расстоянием (2) и «реализуемым» расстоянием δ (3) на выходе канала. Ситуация заметно упрощается при исследовании многоуровневых кодов и сигнально-кодовых конструкций, основанных на совместном использовании модулирующих вектор-функций $\mathbf{a}(\cdot)$, позволяющих исключить МСИ. Именно этот класс СКК и рассматривается в дальнейшем.

Модулирующая вектор-функция

Выбор модулирующей функции в рассматриваемой СКК определяется видом комплексной огибающей $h(\cdot)$ импульсной характеристики канала. Введем семейство функций $\{h_t(\cdot) | t \geq T\}$ вида

$$h_t(\tau) = \begin{cases} h(t - \tau),\ \tau \in [0, T], \\ 0,\ \tau \notin [0,\ T]; \end{cases}$$

и пусть X есть линейное подпространство, натянутое на совокупность функций $\{h_t(\cdot) | t \geq T\}$. Ортогональное дополнение подпространства X в пространстве $L^2(0, T)$ обозначим через F. Очевидно, для всякого элемента $f(\cdot)$, принадлежащего подпространству F,

$$\int_0^T h(t - \tau) f(\tau)\mathrm{d}\tau = 0,\ t \geq T,$$

т. е. при использовании элементов из F в качестве компонентов модулирующей функции $\mathbf{a}(\cdot)$ МСИ в канале исключается. Для описания структуры подпространства F в явном виде введем следующие обозначения: обозначим через H интегральный оператор в $L^2(0, T)$ с ядром $h(\cdot, \cdot)$ вида $h(t, \tau) = h(t - \tau)$, $0 \leq \tau \leq t \leq T$; пусть H_F есть сужение оператора H на подпространство F и пусть E представляет собой H-образ подпространства F.

Утверждение 1. Пусть подпространство E плотно в $L^2(0, T)$. Тогда оператор H_F^{-1}, обратный к оператору H_F, может быть представлен в виде

$$H_F^{-1} = \sum_{i\geq 0} \lambda_i \varphi_i \psi_i,$$

где $\{\lambda_i\}$ — сингулярные числа оператора H_F^{-1}, $\{\varphi_i\}$ и $\{\psi_i\}$ — ортонормированные полные (в подпространстве F и пространстве $L^2(0, T)$ соответственно) системы функций, такие, что ψ_i — собственная функция оператора $\left(H_F^{-1}\right)^* H_F^{-1}$, соответствующая собственному числу λ_i^2, и

$$\varphi_i = \lambda_i^{-1} H_F^{-1} \psi_i. \quad (4)$$

Доказательство 1. В силу плотности E в $L^2(0, T)$ оператор $\left(H_F^{-1}\right)^* H_F^{-1}$ самосопряжен и поэтому обладает полной в $L^2(0, T)$ системой собственных функций $\{\psi_i\}$ [18]. Ортонормированность совокупности $\{\varphi_i\}$ следует из (4), а ее полнота в F — из полноты $\{\psi_i\}$ в E и того, что F есть область значений оператора H_F^{-1}.

Исчерпывающее описание класса импульсных характеристик $h(\cdot)$ таких, что подпространство F не пусто и подпространство E плотно в $L^2(0, T)$, неизвестно. Вместе с тем этот класс, как видно из следующего ниже утверждения, достаточно широк.

Утверждение 2. Пусть фильтр в канале описывается линейным дифференциальным уравнением порядка l с постоянными коэффициентами, т. е. оператор H^{-1} представляет собой дифференциальный оператор порядка l. Тогда:

1) коразмерность подпространства F относительно $L^2(0, T)$ равна l; 2) подпространство E есть совокупность всех функций y из $L^2(0, T)$, имеющих абсолютно непрерывные производные вплоть до порядка $l-1$ и удовлетворяющих граничным условиям

$$\left.\frac{\mathrm{d}^i y}{\mathrm{d}t^i}\right|_{t=0} = 0, \quad \left.\frac{\mathrm{d}^i y}{\mathrm{d}t^i}\right|_{t=T} = 0, \; 1 \le i \le l-1.$$

Доказательство 2. Пусть $x_1, ..., x_m$ — корни характеристического многочлена дифференциального оператора H^{-1} и пусть l_j — кратность корня x_j, $l_1 + l_2 + ... + l_m = l$. Тогда $h(\cdot)$ можно представить в виде

$$h(t) = \sum_{j=1}^{m} \sum_{i=1}^{l_j} k_{ij} t^{i-1} e^{x_j t}, \; t \ge 0, \qquad (5)$$

где $\{k_{ij} \mid 1 \le i \le l_j, 1 \le j \le m\}$ — некоторые коэффициенты [19, 20]. Из представления (5) следует, что функция $h_t(\cdot)$ при $t \ge T$ есть линейная комбинация l линейно независимых функций $q_1(\cdot), ..., q_l(\cdot)$, заданных на $[0, T]$ и имеющих вид

$$q_i(\tau) = \tau^{\alpha_i} \exp(x_{\beta_i} \tau),$$

где

$$\beta_i = \max\left\{ j \ge 1 \,\middle|\, \sum_{k=1}^{j-1} l_k < i \right\}, \; \alpha_i = i - \sum_{k-1}^{\beta_i - 1} l_k - 1.$$

Поэтому в подпространство F входят те и только те функции, которые ортогональны каждой функции из совокупности $\{q_i(\cdot) \mid 1 \le i \le l\}$.

2. Существование абсолютно непрерывных производных вплоть до порядка $l-1$ для функций, входящих в E, следует из того, что H^{-1} — дифференциальный оператор порядка l. Граничные условия в точке $t = 0$ вытекают из представления (5) для $h(\cdot)$. Справедливость граничных условий в точке $t = T$ следует из представления (5) и того, что любая функция $x(\cdot)$ из подпространства F должна быть ортогональна каждой функции из совокупности $\{q_i(\cdot) \mid 1 \le i \le l\}$.

Из п. 2 утверждения 2 следует, что $\left(H_F^{-1}\right)^* H_F^{-1}$ есть линейный дифференциальный оператор порядка $2l$; его область определения $E_0 \subset E$ и состоит из функций, имеющих абсолютно непрерывные производные вплоть до порядка $2l - 1$ и удовлетворяющих дополнительным граничным условиям, которые определяются конкретным видом $h(\cdot)$. Это в свою очередь означает, что система функций $\{\varphi_i\}$ представляет собой совокупность гармонических функций, параметры которых определяются путем решения стандартным образом формируемой системы уравнений.

Замечание. Ограничение на вид канального фильтра, указанное в утверждении 2, эквивалентно тому, что его передаточная характеристика есть дробно-рациональная функция. Отметим, что передаточные характеристики, не являющиеся дробно-рациональными функциями, например передаточная характеристика UTP кабеля, могут быть достаточно точно аппроксимированы подходящей дробно-рациональной функцией.

Пусть сингулярные числа λ_i, $i = 1, 2, ...$, упорядочены в неубывающем порядке. Выберем компоненты модулирующей $\mathbf{a}(\cdot)$ и демодулирующей $\mathbf{p}(\cdot)$ функций следующим образом:

$$a_i(\cdot) = \varphi_i(\cdot), \; p_i(\cdot) = \psi_i(\cdot), \; 1 \le i \le N.$$

Тогда N-мерные комплексные векторы на выходе демодулятора имеют вид

$$\mathbf{r}_j = \sqrt{2RE_b/\sigma^2}\,\mathbf{b}_j\mathbf{D} + \mathbf{n}_j, \; j \ge 0, \qquad (6)$$

где $\mathbf{D} = \mathrm{diag}\left\{\lambda_1^{-1}, ..., \lambda_N^{-1}\right\}$, $\mathbf{n}_j$ — комплексный гауссов N-мерный вектор с независимыми компонентами, имеющими нулевое среднее и дисперсию N_0. Представление (6) можно интерпретировать как разложение исходного канала на N параллельных независимых подканалов с гауссовым шумом и коэффициентами передачи $\lambda_1^{-1}, ..., \lambda_N^{-1}$. Другими словами, исходный модуляционный код U трансформируется в канале (для выбранной модулирующей функции $\mathbf{a}(\cdot)$) в «выходной» код $V = \{\mathbf{bD} \mid \mathbf{b} \in U\}$.

Многоуровневый модуляционный код U

Модуляционный код U удобно рассматривать как каскадную конструкцию: внешняя ступень кодирования осуществляется линейным кодом C над полем $\mathbb{F}_2^m$, а роль внутреннего кода выполняет сигнальное множество B. Многоуровневая или обобщенная каскадная конструкция включает, как известно, последовательность вложенных внутренних подкодов и совокупность внешних (уровневых) кодов.

Внутренний код B и его вложенные подкоды конструируются следующим образом. Вначале строится прообраз внутреннего кода, обозначим его B^p, как конечное подмножество N-мерной целочисленной комплексной решетки G^N. Пусть

Λ_0 — подрешетка комплексной N-мерной двоичной решетки G^N, $\Lambda_0 \subset G^N$. Подрешетка Λ_0 определяет разбиение G^N/Λ_0 решетки G^N на подрешетку Λ_0 и ее смежные классы; пусть порядок этого разбиения $|G^N/\Lambda_0| = \exp_2(m_0)$. Будем полагать, что решетка Λ_0 является mod–2 μ-depth решеткой, т. е. $\Lambda_0 \supset \varphi^\mu G^N$, где $\varphi = 1 + i$, $i^2 = -1$, μ — целое положительное число. С точки зрения приложений основной интерес представляют 1-depth и 2-depth решетки. Смежные классы разбиения $\Lambda_0/\varphi^\mu G^N$ могут быть представлены с помощью генераторов (порождающих элементов) фактор-группы $[\Lambda_0/\varphi^\mu G^N]$, которые в свою очередь можно интерпретировать как порождающие матрицы некоторых μ линейных двоичных кодов $C^{(\mu-1)}$, ..., $C^{(0)}$ [21]. В результате 1-depth и 2-depth решетки Λ_0 представляются схематично в виде комплексных кодовых формул $\Lambda_0 = \varphi G^N + C^{(0)}$ и $\Lambda_0 = \varphi^2 G^N + \varphi C^{(1)} + C^{(0)}$ соответственно.

Пусть, далее, Λ_1, Λ_2, ..., Λ_L — последовательность вложенных подрешеток решетки Λ_0, т. е. $\Lambda_0 \supset \Lambda_1 \supset \cdots \supset \Lambda_{L-1} \supset \Lambda_L$, и $\Lambda_0/\Lambda_1/\cdots/\Lambda_L$ есть L-уровневое разбиение подрешетки Λ_0. Будем полагать, что для всех l разбиение Λ_{l-1}/Λ_l имеет порядок $\exp_2(m_l)$, т. е. $|\Lambda_{l-1}/\Lambda_l| = \exp_2(m_l)$, $1 \le l \le L$, $m_l \ge 1$ и $\sum_{l=1}^{L} m_l = m$. Пусть $\mathbf{g}_{l1}, \ldots, \mathbf{g}_{lm_l}$ — генераторы (порождающие элементы) фактор-группы $[\Lambda_{l-1}/\Lambda_l]$; $\mathbf{g}_{ls} \in (\Lambda_{l-1} \setminus \Lambda_l)$, $1 \le s \le m_l$. Обозначим через $\mathbf{G}_l$ матрицу, строками которой являются генераторы $\mathbf{g}_{l1}, \ldots, \mathbf{g}_{lm_l}$, и пусть $\mathbf{G}$ — матрица размера $m \times N$, составленная из подматриц $\mathbf{G}_l$, $1 \le l \le L$:

$$\mathbf{G} = \begin{pmatrix} \mathbf{G}_1 \\ \vdots \\ \mathbf{G}_L \end{pmatrix}.$$

Введем вектор $\mathbf{x}_0$, сдвигающий решетку относительно начала координат и определяемый как

$$\mathbf{x}_0 = \arg\min_{\mathbf{x} \in \Lambda_L} \overline{\left\| (\mathbf{c}\mathbf{G} + \mathbf{x}) \bmod \Lambda_L - \mathbf{z} \right\|^2},$$

где $\mathbf{c}$ — m-мерный двоичный вектор, $\mathbf{c} \in \{0, 1\}^m$, $\mathbf{z} = \frac{1}{2}(1 + i, \ldots, 1 + i)$, $i^2 = -1$, черта сверху означает усреднение по всем равнораспределенным векторам $\mathbf{c}$. Совокупность точек

$$B^{(p)} = \left\{ (\mathbf{c}\mathbf{G} + \mathbf{x}_0) \bmod \Lambda_L - \mathbf{z} \,\middle|\, \mathbf{c} \in \{0,1\}^m \right\} \qquad (7)$$

представляет собой прообраз внутреннего кода (сигнального множества) B, мощность которого $q = 2^m$. Введем далее положительную диагональную матрицу $\Theta = \mathrm{diag}\{\theta_1, \ldots, \theta_N\}$, задающую (определяющую) распределение энергии между подканалами. Внутренний код (сигнальное множество) B определим как

$$B = \left\{ b \,\middle|\, b = \mathbf{x}\Theta,\ \mathbf{x} \in B^{(p)} \right\}. \qquad (8)$$

Отметим, что матрицу $\mathbf{G}$ в (7) можно интерпретировать как порождающую матрицу внутреннего кода (сигнального множества) B. Среднее значение квадрата евклидовой нормы слов внутреннего кода (канальных символов)

$$\sigma^2 = \overline{\|\mathbf{x}\Theta\|^2},$$

где усреднение проводится по всем элементам множества $B^{(p)}$. Полагая канальные символы равнораспределенными, получаем

$$\sigma^2 = 2^{-m} \sum_{\mathbf{x}_i \in B^{(p)}} \overline{\|\mathbf{x}\Theta\|^2}.$$

Разбиение $\Lambda_0/\Lambda_1/\cdots/\Lambda_L$ однозначно определяет L-уровневое разбиение $B/B_1/\cdots/B_L$ внутреннего кода (сигнального множества) B на вложенные подкоды такие, что подкод B_l, $1 \le l \le L-1$, порождается подматрицами $\mathbf{G}_{l+1}, \ldots, \mathbf{G}_L$ матрицы $\mathbf{G}$ и состоит из $\exp_2(m_{l+1} + \ldots + m_L)$ слов (точек), а одноточечное множество B_L однозначно определяется векторами $\mathbf{x}_0$ и $\mathbf{z}$. Смежному классу Λ_{li} в разбиении Λ_{l-1}/Λ_l, $0 \le i \le \exp_2 m_l - 1$, определяемому как

$$\Lambda_{li} = \Lambda_l + \sum_{s=1}^{m_l} \chi_s(i)\mathbf{g}_{ls}, \qquad (9)$$

где $\chi_s(i)$ — s-й компонент двоичного представления числа i, соответствует «смежный класс» B_{li} в разбиении B_{l-1}/B_l, $B_0 = B$, определяемый тем же вектором сдвига, что и в (9): $B_{li} = B_l + \boldsymbol{\tau}_{li}$,

$$\boldsymbol{\tau}_{li} = \left(\left(\sum_{s=1}^{m_l} \chi_s(i)\mathbf{g}_{ls} \right) \bmod \Lambda_L \right) \Theta.$$

Далее, пусть код C есть прямое произведение L уровневых кодов C_l, $1 \le l \le L$, таких, что l-й уровневый код C_l в свою очередь суть прямое произведение двоичных кодов C_{ls}, называемых компонентными кодами l-го уровня. Скорость $R = k/n$ кода C равна сумме скоростей компонентных кодов $R_{ls} = k_{ls}/n_{ls}$:

$$R = \sum_{l=1}^{L} \sum_{s=1}^{m_l} R_{ls};$$

при этом размер кодового блока n кода C равен наименьшему общему кратному величин n_{ls}, $1 \le l \le L$, $1 \le s \le m_l$, т. е. из каждого входного информационного k-блока, где $k = nR$, кодом C_{ls} кодируются n/n_{ls}-подблоков длины k_{ls}. Кодовый символ $\mathbf{c}_j$ кода

C представляет собой вектор длины m, компонентами которого являются j-е символы кодовых последовательностей $\mathbf{c}_{ls} \in \mathrm{C}_{ls}$. Порождающая матрица кода C очевидным образом определяется через порождающие матрицы компонентных кодов C_{ls}. Преобразование символов $\mathbf{c}_j$ кода C в кодовые символы $\mathbf{b}_j$ модуляционного кода U определяется фактически выражениями (7) и (8):

$$\mathbf{b}_j = ((\mathbf{c}_j\mathbf{G} + \mathbf{x}_0) \bmod \Lambda_L - \mathbf{z})\boldsymbol{\Theta},\ j \geq 0.$$

На выходе канала символы $\mathbf{b}_j$ модуляционного кода U трансформируются в символы $\mathbf{b}_j\mathbf{D}$ «выходного» кода $V = \{\mathbf{bD} \mid \mathbf{b} \in U\}$. Минимальное расстояние «выходного» кода V легко оценивается снизу с помощью стандартных рассуждений. Обозначим через Y_l «выходное» подмножество кода V, соответствующее подмножеству B_l

$$Y_l = \{\mathbf{bD} \mid \mathbf{b} \in B_l\},$$

и пусть

$$K_{ls}(c) = \bigcup_{i:\chi_s(i)=c} \left(Y_l + \boldsymbol{\tau}_{li}\mathbf{D}\right),\ c \in \{0,1\}. \qquad (10)$$

Тогда максимизированное по $\boldsymbol{\Theta}$ минимальное евклидово расстояние на выходе канала δ можно оценить снизу как

$$\delta \geq \max_{\boldsymbol{\Theta}} \min_{1\leq l\leq L} \min_{1\leq s\leq m_l} \rho_{ls}\sqrt{d_{ls}RE_b/\sigma^2}, \qquad (11)$$

где ρ_{ls} — евклидово расстояние между подмножествами $K_{ls}(0)$ и $K_{ls}(1)$, d_{ls} — минимальное расстояние компонентного кода C_{ls}. Очевидно, что $\min\limits_s \rho_{ls}$ совпадает с минимальным расстоянием Δ_{l-1} подмножества Y_{l-1}:

$$\Delta_{l-1} = \min_{\mathbf{e}} \overline{\left\|\mathbf{eD}\boldsymbol{\Theta}\right\|^2}, \qquad (12)$$

где $\mathbf{e}$ — линейная комбинация генераторов $\mathbf{g}_{l1}, \ldots, \mathbf{g}_{lm_l}$. Таким образом, справедливо следующее утверждение.

Утверждение 3. При использовании идентичных (для каждого из L уровней) компонентных кодов с минимальными расстояниями $d_1, \ldots, d_L$ граница (11) приобретает вид

$$\delta \geq \max_{\boldsymbol{\Theta}} \min_{1\leq l\leq L} \Delta_{l-1}\sqrt{d_lRE_b/\sigma^2}. \qquad (13)$$

Нижняя граница в (13) достигается, т. е. минимальное расстояние δ совпадает с правой частью (13), если на каждом уровне l для всех пар (s, t), $1 \leq s, t \leq m_l$, $\rho_{ls} = \rho_{lt}$.

Величины Δ_{l-1}/σ, $1 \leq l \leq L$ в правой части (13) определяются конкретным видом внутреннего кода (сигнального множества) B и его вложенными подкодами, тогда как внешняя ступень кодирования — код C — определяет величины d_lR, $1 \leq l \leq L$.

Многоступенчатое декодирование модуляционного кода U

Декодирование многоуровневого кода в целом — по максимуму правдоподобия, имеет, как правило, неприемлемо высокую сложность. Поэтому в качестве альтернативы была предложена процедура многоступенчатого декодирования, характеризующаяся существенно меньшей сложностью реализации. Идея многоступенчатого декодирования (multistage decoding) многоуровневого кода основана, как известно, на последовательном декодировании уровневых кодов с использованием результатов декодирования кодов предыдущих уровней [23, 24]. Для рассматриваемой конструкции декодирование кода U сводится, очевидно, к независимому декодированию кодов C_l, $1 \leq l \leq L$. Будем полагать, что декодирование l-го уровневого кода C_l выполняется путем независимого декодирования m_l компонентных кодов C_{ls}, $1 \leq s \leq m_l$. В качестве мягких решений используются евклидовы расстояния между N-мерными комплексными векторами $\mathbf{r}_j$, $j \geq 0$ на выходе демодулятора и соответствующими объединениями множеств вида (10). Формальная запись многоступенчатого алгоритма декодирования, представленная в [24], применительно к декодированию кода U имеет следующий вид.

1. Для $s = 1, 2, \ldots, m_l$ определить

$$\hat{\mathbf{c}}_{1s} = \arg\min_{\mathbf{c}\in C_{1s}} \Gamma_{1s}(\mathbf{c}),$$

где

$$\Gamma_{1s}(\mathbf{c}) = \sum_{j\geq 0} \min_{\mathbf{v}\in K_{1s}(c_j)} \left\|\mathbf{r}_j - \mathbf{v}\right\|^2. \qquad (14)$$

2. Сформировать m_1-мерные векторы $\hat{\mathbf{c}}_{1j}$, $j = 0, 1, 2, \ldots$, состоящие из компонентов $\hat{c}_{1sj}$, $1 \leq s \leq m_1$.

3. Для $l = 2, 3, \ldots, L$ выполнять:

— для $s = 1, 2, \ldots, m_l$ определить $\hat{\mathbf{c}}_{ls} = \arg\min_{\mathbf{c}\in C_{ls}} \Gamma_{ls}(\mathbf{c})$, где $\Gamma_{ls}(\mathbf{c})$ имеет вид, аналогичный (14), с тем лишь отличием, что минимум в каждом слагаемом отыскивается на таком сдвиге множества $K_{ls}(c_j)$, который определяется ранее вынесенными решениями $\hat{\mathbf{c}}_{1j}, \ldots, \hat{\mathbf{c}}_{l-1,j}$;

— сформировать m_l-мерные векторы $\hat{\mathbf{c}}_{lj}$, $j = 0, 1, 2, \ldots$, состоящие из компонентов $\hat{c}_{lsj}$, $1 \leq s \leq m_l$.

Последовательность векторов $\hat{\mathbf{c}}_j = (\hat{\mathbf{c}}_{1j}, \ldots, \hat{\mathbf{c}}_{Lj})$, $j = 0, 1, 2, \ldots$, образует результат декодирования.

Сложность реализации многоступенчатого алгоритма можно оценить, суммируя сложности декодирования уровневых кодов. Число операций,

затрачиваемых на декодирование одного бита l-го уровневого кода, определяется следующими компонентами:

— числом операций $N(Y_{l-1}/Y_l)$, необходимых для отыскания в каждом смежном классе разбиения Y_{l-1}/Y_l элемента множества, ближайшего к произвольной точке $\mathbf{r} \in \mathbb{C}^N$;

— числом операций Q_l, необходимых для формирования по полученным $\exp_2 m_l$ расстояниям $2m_l$ составляющих (по две на каждый компонентный код) приращений метрик $\Gamma_{ls}(\cdot)$, $1 \le s \le m_l$;

— числом операций M_{ls}, необходимых для декодирования кодового символа s-го компонентного кода.

В результате число операций типа сложение/сравнение, приходящееся на один бит,

$$\varpi = (n/k)\sum_{l=1}^{L}\left(N(Y_{l-1}/Y_l) + Q_l + \sum_{s=1}^{m_l} M_{ls}\right). \quad (15)$$

В силу изоморфности разбиений Y_{l-1}/Y_l и Λ_{l-1}/Λ_l величина $N(Y_{l-1}/Y_l)$ совпадает с аналогичным образом определяемой величиной $N(\Lambda_{l-1}/\Lambda_l)$, значения которой для ряда разбиений приведены в [21]. Число операций Q_l для формирования приращений метрик определяется как $Q_l = m_l(\exp_2 m_l - 2)$. Число операций M_{ls} для s-го компонентного сверточного кода с кодовым ограничением ν_{ls} при использовании алгоритма Витерби (АВ) равно $(\exp_2 \nu_{ls})(\exp_2 k_{ls} - 1)/n_{ls}$; для выколотого кода со скоростью $R_{ls} = (n_{ls} - 1)/n_{ls}$ АВ может быть реализован с числом операций $M_{ls} \le R_{ls}\exp_2(\nu_{ls} + 1)$.

Верхняя граница вероятности ошибки на бит P_b при многоступенчатом декодировании кода U может быть получена с использованием техники, развитой в работе [24]. Эта граница имеет довольно громоздкий вид, поэтому ограничимся здесь лишь ее главным членом. Обозначим через $\tilde{\delta}$ правую часть границы (11) и введем величину

$$N_{\tilde{\delta}} = \sum_{(l,s)\in\Omega} A(d_{ls})(\eta_{ls})^{d_{ls}},$$

где Ω — совокупность пар (l, s) таких, что $\rho_{ls}\sqrt{d_{ls}RE_b/\sigma^2} = \tilde{\delta}$, $A(d_{ls})$ — суммарный информационный вес петель веса d_{ls} в решетчатой диаграмме кода C_{ls}, η_{ls} — ошибочный коэффициент пары множеств $\{K_{ls}(0), K_{ls}(1)\}$, т. е. максимальное число соседей произвольного элемента из $K_{ls}(0)$, находящихся от него на расстоянии ρ_{ls} в решающей области множества $K_{ls}(1)$.

Утверждение 4. Пусть код U декодируется с использованием многоступенчатой процедуры. Тогда при $N_0 \to 0$ вероятность ошибки на бит

$$P_b \le N_{\tilde{\delta}}\, e^{-\tilde{\delta}^2/4N_0}(1 + o(1)). \quad (16)$$

Величину $N_{\tilde{\delta}}$, как следует из (16), можно интерпретировать как число ближайших соседей, учитываемых при многоступенчатом декодировании, или, другими словами, это есть наибольшее число точек, в которых шар радиуса $\tilde{\delta}/2$ с центром в точке, соответствующей произвольной кодовой последовательности $\mathbf{v} \in V$, касается границы решающей области. Очевидно, величина $N_{\tilde{\delta}}$ не меньше, чем истинное число ближайших соседей (kissing number). Учет ложных соседей находит свое отражение в тенденции к размножению ошибок, возникающих при декодировании уровневых кодов.

Будем полагать далее, что обменное соотношение между нижней границей $\tilde{\delta}$ и сложностью ϖ рассматривается в качестве критерия при совместном выборе параметров СКК.

Конструкции на основе решеток Барнса — Уолла

С точки зрения приложений основной интерес при выборе конкретных вариантов исходной решетки Λ_0 (и ее разбиений) представляют решетки Барнса — Уолла (Barnes — Wall). Совокупность N-мерных комплексных решеток Барнса — Уолла $\Lambda(r, s)$, $s \ge r \ge 0$, $s = \log_2 N$ (r — порядок решетки), образует последовательность вложенных разбиений: $\Lambda(s, s) \supset \Lambda(s-1, s) \supset \cdots \supset \Lambda(1, s) \supset \Lambda(0, s)$; при этом решетка $\Lambda(r, s)$ может быть представлена в виде комплексной кодовой формулы [22]

$$\Lambda(r, s) = \varphi^{s-r} G^N + \sum_{r \le h < s} \varphi^{h-r} RM(h, s),$$

где $RM(h, s)$ — код Рида — Маллера длины 2^s и порядка h.

В табл. 1 приведены параметры ряда вариантов сигнальных множеств и их разбиений, построенных на основе решеток Барнса — Уолла и их главных (principal) подрешеток.

Коды Рида — Маллера RM (0, 1) и RM (0, 2) в табл. 1 представляют собой коды с повторением, RM (1, 3) — код Хэмминга (8, 4), коды RM (1, 2) и RM (2, 3) — коды с проверкой на четность.

В табл. 2 приведены средние значения квадрата евклидовой нормы канальных символов для сигнальных множеств, перечисленных в табл. 1.

Будем полагать, что каждый из уровневых кодов C_l, $1 \le l \le L$, многоуровневого кода U состоит из идентичных двоичных компонентных кодов с минимальными расстояниями $d_1, \ldots, d_L$ соответственно. Тогда, подставляя в правую часть границы (13) значения квадратов минимальных расстояний $\Delta_0^2, \ldots, \Delta_{L-1}^2$ (табл. 1) и среднее значение квадрата евклидовой нормы σ^2 (табл. 2) и выполняя максимизацию по $\boldsymbol{\Theta}$, задающей распределение энергии между подканалами, получаем мак-

■ ***Таблица 1.*** Параметры сигнальных множеств и их разбиений

■ ***Table 1.*** Parameters of signal sets and their partitions

#	N	Исходная решетка Λ_0	L	$m_0, m_1, ..., m_L$	$q = 2^m$	Квадраты минимальных расстояний (на выходе канала) $\Delta_0^2, ..., \Delta_{L-1}^2$
1	1	G	2	$m_0 = 0; m_1 = 1; m_2 = 1$	4	$\Delta_0^2 = \theta_1^2/\lambda_1^2; \Delta_1^2 = 2\Delta_0^2$
2	1	G	3	$m_0 = 0; m_1 = 1;$ $m_2 = 1; m_3 = 1$	8	$\Delta_0^2 = \theta_1^2/\lambda_1^2; \Delta_1^2 = 2\Delta_0^2; \Delta_2^2 = 4\Delta_0^2$
3	1	G	4	$m_0 = 0; m_1 = 1;$ $m_2 = 1; m_3 = 1; m_4 = 1$	16	$\Delta_0^2 = \theta_1^2/\lambda_1^2; \Delta_1^2 = 2\Delta_0^2; \Delta_2^2 = 4\Delta_0^2; \Delta_3^2 = 8\Delta_0^2$
4	2	$\varphi G^2 + RM\ (0, 1)$	2	$m_0 = 1; m_1 = 2; m_2 = 1$	8	$\Delta_0^2 = \min\limits_{1 \le i \le 2} 2\theta_i^2/\lambda_i^2; \Delta_1^2 = 2\sum\limits_{i=1}^{2} \theta_i^2/\lambda_i^2$
5	2	$\varphi G^2 + RM\ (0, 1)$	3	$m_0 = 1; m_1 = 2;$ $m_2 = 2; m_3 = 1$	32	$\Delta_0^2 = \min\limits_{1 \le i \le 2} 2\theta_i^2/\lambda_i^2; \Delta_1^2 = 2\sum\limits_{i=1}^{2} \theta_i^2/\lambda_i^2;$ $\Delta_2^2 = 8\theta_1^2/\lambda_1^2$
6	2	G^2	4	$m_0 = 0; m_1 = 1;$ $m_2 = 2; m_3 = 3; m_4 = 1$	64	$\Delta_0^2 = \min\limits_{1 \le i \le 2} \theta_i^2/\lambda_i^2; \Delta_1^2 = 2\Delta_0^2;$ $\Delta_2^2 = 2\sum\limits_{i=1}^{2} \theta_i^2/\lambda_i^2; \Delta_3^2 = 8\theta_1^2/\lambda_1^2$
7	4	$\varphi^2 G^4 + \varphi\, RM\ (1, 2) +$ $+ RM\ (0, 2)$	2	$m_0 = 4; m_1 = 3;$ $m_2 = 1$	16	$\Delta_0^2 = \min\limits_{\substack{1 \le i,j \le 4; \\ i \ne j}} 2(\theta_i^2/\lambda_i^2 + \theta_j^2/\lambda_j^2)$ $\Delta_1^2 = 2\sum\limits_{i=1}^{4} \theta_i^2/\lambda_i^2$
8	4	$\varphi G^4 + RM\ (1, 2)$	3	$m_0 = 1; m_1 = 3;$ $m_2 = 3; m_3 = 1$	128	$\Delta_0^2 = \min\limits_{1 \le i,j \le 4} (\theta_i^2/\lambda_i^2 + \theta_j^2/\lambda_j^2); \Delta_1^2 = 2\Delta_0^2;$ $\Delta_2^2 = 2\sum\limits_{i=1}^{4} \theta_i^2/\lambda_i^2$
9	4	G^4	4	$m_0 = 0; m_1 = 1;$ $m_2 = 3; m_3 = 3;$ $m_4 = 1$	256	$\Delta_0^2 = \min\limits_{1 \le i \le 4} \theta_i^2/\lambda_i^2; \Delta_1^2 = \min\limits_{1 \le i,j \le 4} (\theta_i^2/\lambda_i^2 + \theta_j^2/\lambda_j^2);$ $\Delta_2^2 = 2\Delta_1^2; \Delta_3^2 = 2\sum\limits_{i=1}^{4} \theta_i^2/\lambda_i^2$
10	8	$\varphi^2 G^8 + \varphi RM\ (2, 3) +$ $+ RM\ (1, 3)$	2	$m_0 = 5; m_1 = 7;$ $m_2 = 4$	2^{11}	$\Delta_0^2 = \min\limits_{\substack{1 \le i,j,l,m \le 8; \\ i \ne j \ne l \ne m}} (\theta_i^2/\lambda_i^2 + \theta_j^2/\lambda_j^2 + \theta_l^2/\lambda_l^2 + \theta_m^2/\lambda_m^2);$ $\Delta_1^2 = 2\Delta_0^2$

■ ***Таблица 2***. Средние значения квадрата евклидовой нормы сигнальных множеств
■ ***Table 2***. Mean values of the squared Euclidean norm of signal sets

Число подканалов N	Номер варианта в табл. 1	Средние значения квадрата евклидовой нормы σ^2
1	1	$\theta_1^2/2$
1	2	$3\theta_1^2/2$
1	3	$5\theta_1^2/2$
2	4	$(\theta_1^2+\theta_2^2)/2$
2	5, 6	$5\theta_1^2/2+\theta_2^2/2$
4	7, 8, 9	$\sum_{i=1}^{4}\theta_i^2/2$
8	10	$\sum_{i=1}^{8}\theta_i^2/2$

симизированное по Θ евклидово расстояние $\tilde{\delta}$, определяющее при большом отношении сигнал-шум, как отмечалось выше, помехоустойчивость кода U при многоступенчатом декодировании. Результаты максимизации нижней границы $\tilde{\delta}$ (в ряде случаев совпадающей с минимальным расстоянием δ) для сигнальных множеств (внутренних кодов), перечисленных в табл. 1, представлены в табл. 3.

Отыскание параметров предпочтительных СКК в рамках обменного соотношения «расстояние $\tilde{\delta}$ — сложность ϖ» осуществляется следующим образом. При заданном ограничении на сложность декодирования ϖ (15) многоуровневого кода U выполняется максимизация евклидова расстояния $\tilde{\delta}$ (для каждого из вариантов сигнальных множеств в табл. 3 посредством перебора по всем возможным наборам компонентных кодов и оптимизации длительности тактового интервала T при фиксированном значении скорости передачи $\upsilon = R/T$ (или коэффициента использования полосы частот $\gamma = \upsilon/W$)).

Рассмотрим в качестве примера поиск параметров предпочтительных СКК для канала, низкочастотный эквивалент передаточной характеристики которого описывается фильтром Баттерворта первого порядка с частотой среза $W/2$ Гц. В этом случае:

квадраты сингулярных чисел —

$$\lambda_i^2 = 1+\left(i/WT\right)^2 = 1+\left(i\gamma/R\right)^2,\ 1\le i\le N;$$

компоненты модулирующей функции $\mathbf{a}(\cdot)$ —

$$a_i(t) = \sqrt{2/T}\,\frac{WT}{\sqrt{i^2+W^2T^2}}\left(-\frac{k}{WT}\cos\frac{k\pi t}{T}+\sin\frac{k\pi t}{T}\right),$$
$$t\in[0,\ T];$$

компоненты демодулирующей функции $\mathbf{p}(\cdot)$ —

$$p_i(t) = \sqrt{2/T}\sin\frac{k\pi t}{T},\ t\in[0,T].$$

Подставляя значения λ_i^2 в выражения для δ^2/RE_b (правая колонка табл. 3), получаем соотношения для δ^2/E_b, зависящие от коэффициента использования полосы частот γ, скорости кода R и минимальных расстояний компонентных кодов. Так, для первого варианта сигнального множества $\delta^2/RE_b = 2\min\left(d_1,\ 2d_2\right)/\lambda_1^2$, поэтому $\delta^2/E_b = 2\min(d_1,\ 2d_2)R/(1+(\gamma/R)^2)$. Перебирая затем для каждого варианта сигнального множества по всем допустимым значениям скорости кода R и по наборам допустимых компонентных кодов с фиксированным значением скорости R многоуровневого кода U, получаем совокупность возможных сигнально-кодовых конструкций. Каждая конструкция характеризуется нормированным значением квадрата минимального расстояния δ^2/E_b (или его нижней границей), зависящим от коэффициента использования полосы частот γ, и сложностью многоэтапного декодирования ϖ, измеряемой числом операций (типа сложение/сравнение) на бит. Так, для конструкции с первым вариантом (из табл. 1) сигнального множества и двухуровневым кодом, включающим на первом уровне сверточный код со скоростью 1/3 и минимальным расстоянием 8 и на втором уровне — сверточный код со скоростью 2/3 и минимальным расстоянием 2, получаем $\delta^2/E_b = 8/(\gamma^2+1)$, $\varpi = 6{,}3$ операций/бит.

Сравнительный анализ построенных таким образом сигнально-кодовых конструкций показывает, что с увеличением значения допустимой сложности декодирования ϖ возрастают как размерность N, так и мощность q сигнального мно-

■ ***Таблица 3.*** Нормированные значения квадрата минимального расстояния для конструкций на основе сигнальных множеств из табл. 1

■ ***Table 3.*** Normalized values of the square minimum distance for construction based on signal sets from Table 1

#	N	$q = 2^m$	Нормированное значение квадрата минимального расстояния δ^2/RE_b
1	1	4	$2\min(d_1,\ 2d_2)\big/\lambda_1^2$
2	1	8	$2\min(d_1,\ 2d_2,\ 4d_3)\big/3\lambda_1^2$
3	1	16	$2\min(d_1,\ 2d_2,\ 4d_3,\ 8d_4)\big/5\lambda_1^2$
4	2	8	$4d_1/(\lambda_1^2+\lambda_2^2)$, если $d_1 \le 2d_2$; $4d_1/(\lambda_1^2(d_1/d_2-1)+\lambda_2^2)$, если $d_1 > 2d_2$
5	2	32	$4d_1/(5\lambda_1^2+\lambda_2^2)$, если $d_1 \le 4d_3$ и $d_1 \le 2d_2$; $\ge 16d_1d_3/(5d_1\lambda_1^2+4d_3\lambda_2^2)$, если $d_1 > 4d_3$ и $d_2 \ge 4d_1d_3/(d_1+4d_3)$; $\ge 4\min(d_1,\ 2d_2,\ 4d_3)\big/(5\lambda_1^2+\lambda_2^2)$ в остальных случаях
6	2	64	$\ge 2\min(d_1,\ 2d_2)\big/(5\lambda_1^2+\lambda_2^2)$, если $\min(d_1,\ 2d_2) \le 8d_4$ и $\min(d_1,\ 2d_2) \le 4d_3$; $\ge \dfrac{16\min(d_1,\ 2d_2)d_4}{5\min(d_1,2d_2)\lambda_1^2+8d_4\lambda_2^2}$, если $\min(d_1,\ 2d_2) > 8d_4$ и $d_3 > \dfrac{4\min(d_1,\ 2d_2)}{\min(d_1,\ 2d_2)+8d_4}$; $\ge 2\min(d_1,\ 2d_2,\ 4d_3,\ 8d_4)\big/(5\lambda_1^2+\lambda_2^2)$ в остальных случаях
7	4	16	$8d_1\big/\sum_{i=1}^4\lambda_i^2$, если $d_1 \le 2d_2$; $16d_2\big/\sum_{i=1}^4\lambda_i^2$, если $d_1 > 2d_2$
8	4	128	$4\min(d_1,\ 2d_2)\big/\sum_{i=1}^4\lambda_i^2$, если $\min(d_1,\ 2d_2) \le 4d_3$; $16d_3\big/\sum_{i=1}^4\lambda_i^2$, если $\min(d_1,\ 2d_2) > 4d_3$
9	4	256	$2\min(d_1,\ 2d_2,\ 4d_3)\big/\sum_{i=1}^4\lambda_i^2$, если $\min(d_1,\ 2d_2,\ 4d_3) \le 8d_4$; $16d_4\big/\sum_{i=1}^4\lambda_i^2$, если $\min(d_1,\ 2d_2,\ 4d_3) > 8d_4$
10	8	2^{11}	$8d_1\big/\sum_{i=1}^8\lambda_i^2$, если $d_1 \le 2d_2$; $16d_2\big/\sum_{i=1}^8\lambda_i^2$, если $d_1 > 2d_2$

жества наиболее предпочтительной (в смысле δ^2/E_b) сигнально-кодовой конструкции. Так, при ограничении на допустимую сложность декодирования $\varpi \le 100$ операций/бит и для всех значений коэффициента использования полосы частот γ наиболее предпочтительной (в рамках рассматриваемого подхода) оказывается следующая сигнально-кодовая конструкция:

— сигнальное множество: количество подканалов $N = 4$, исходная решетка G^4, мощность сигнального множества $q = 256$, число уровней разбиения $L = 4$;

— многоуровневый код U: скорость кода $R = 6{,}25$, на первом уровне используется один сверточный код со скоростью 1/2 и минимальным расстоянием 10, на втором уровне — три сверточных кода со скоростью 3/4 и минимальным расстоянием 5, на третьем уровне — три сверточных кода со скоростью 7/8 и минимальным расстоянием 3, на четвертом уровне — безызбыточный код.

Для этой конструкции нормированное значение квадрата минимального расстояния

$$\delta^2/E_b = 130{,}2/\left(\gamma^2 + 5{,}2\right). \qquad (17)$$

Сопоставляя правую часть (17) с нормированным значением квадрата минимального расстояния δ^2_{OFDM}/E_b при кодированной (решетчатым кодом Унгербоека над КАМ-созвездием) OFDM-передаче со 128 поднесущими, длительностью сигнала $T = 128/W$ и ограничением на сложность декодирования $\varpi \le 100$, получаем, что при $\gamma \ge 1$, даже без учета потери в скорости передачи, вызванной введением защитного интервала, $\delta^2_{OFDM}/E_b < 130{,}2/\left(\gamma^2 + 5{,}2\right)$. Асимптотический энергетический выигрыш (относительно коди-

рованной OFDM-передачи) быстро растет с увеличением коэффициента использования полосы частот γ. Так, при $\gamma = 1$ бит/с·Гц асимптотический энергетический выигрыш составляет 0,21 дБ, тогда как при $\gamma = 6$ бит/с·Гц выигрыш достигает 5,04 дБ. Отметим, что энергетический выигрыш представленной СКК достигается за счет использования в качестве сигнального множества многомерной решетки, согласованного (в определенном смысле) с каналом распределения энергии между подканалами и гибкой структурой многоуровневого модуляционного кода, обеспечивающей широкий диапазон обменного соотношения между помехоустойчивостью и сложностью декодирования.

Заключение

В работе в явном виде получены соотношения, связывающие между собой основные параметры предложенного класса сигнально-кодовых конструкций. Полученные соотношения позволяют установить обменные соотношения между скоростью передачи, минимальным расстоянием и сложностью декодирования. Приведены примеры конструкций, построенных с использованием многомерных решеток Барнса — Уолла. Сравнительный анализ построенных решетчатых сигнально-кодовых конструкций показал, что с увеличением значения допустимой сложности декодирования более предпочтительными оказываются конструкции с большими значениями как размерности N (количества подканалов), так и мощности q сигнального множества. В качестве примера для канала, низкочастотный эквивалент передаточной характеристики которого описывается фильтром Баттерворта первого порядка, приведены параметры лучшей (в смысле минимального расстояния) конструкции при ограничении на допустимую сложность декодирования $\varpi \le 100$ операций/бит. Сопоставление этой конструкции с возможным конкурирующим вариантом — кодированной OFDM-передачей — показало, что асимптотический энергетический выигрыш (относительно кодированной OFDM-передачи) быстро растет с увеличением коэффициента использования полосы частот γ, достигая 5,04 дБ при $\gamma = 6$ бит/с·Гц. При этом за рамками сравнения остались дополнительные факторы, увеличивающие выигрыш: отсутствие защитных интервалов, снижающих скорость передачи, и небольшое значение пик-фактора передаваемых сигналов, позволяющее повысить коэффициент полезного действия передатчика.

Литература

1. **Benedetto S., Marsan M. A., Allegtendo G., Csachin E.** Combined coding and modulation: theory and applications. *IEEE Transactions on Information Theory*, 1988, vol. 34, no. 5, pp. 1121–1151. doi:10.1109/18.2631
2. **Biglieri E., Divsalar D., McLane P. J., Simon M. K.** *Introduction to trellis coded modulation with application*. New York, McMillan Publishing Company, 1991. 576 p.
3. **Anderson J. B., Svensson A., Helstrom W.** *Coded modulation systems*. New York, Kluwer Academic / Plenum Publishers, 2003. 486 p.
4. **Seidl M., Schenk A., Stierstorfer C., Huber J. B.** Polar-coded modulation. *IEEE Transactions on Communications*, 2013, vol. 61, no. 10, pp. 4108–4119. doi:10.1109/ TCOMM.2013.090513.130433
5. **Xiao X., Hong Y., Viterbo E., Gupta A.** Trellis coded modulation for informed receivers. *Proceedings of 2017 IEEE International Conference on Communications Workshops* (ICC Workshops), Paris, France, 2017, pp. 955–960. doi:10.1109/ICCW.2017.7962782
6. **Carlisle C. J., Taylor D. P., Shafi M., Kennedy W. K.** Performance bounds for trellis-coded modulation on time-dispersive channels. *IEEE Transactions on Communications*, 1994, vol. 42, no. 8, pp. 2534–2542. doi:10.1109/26.310613
7. **Forney G. D., Ungerboeck G.** Modulation and coding for linear gaussian channels. *IEEE Transactions on Information Theory*, 1998, vol. 44, no. 6, pp. 2384–2415. doi:10.1109/18.720542
8. **Xia X.-G.** *Modulated coding for inter-symbol interference channels*. New York, McGraw-Hill, 2001. 512 p.
9. **Biglieri E.** High level modulation and coding for nonlinear satellite channels. *IEEE Transactions on Communications*, 1984, vol. 32, no. 5, pp. 616–626.
10. **Euboglu M. V., Qureshi S. U.** Reduced-state sequence estimation for coded modulation on inter-symbol interference channels. *IEEE Journal on Selected Areas in Communications*, 1989, vol. 7, no. 6, pp. 989–995. doi:10.1109/49.29621
11. **Taubin F.** Reduced-state decoding for trellis coded modulation on nonlinear inter-symbol interference channels. *Lecture Notes in Computer Science*, 1994, vol. 829, pp. 88–96.
12. **Misra A., Sarma K. K.** TCM-coded OFDM assisted by ANN in wireless channels. *International Journal of Smart Sensors and Ad Hoc Networks* (IJSSAN), 2012, vol. 1, no. 3, pp. 50–55.
13. **Chide N., Deshmukh S., Borole P. B., Chore N.** An overview of OFDM variants and their applications. *International Journal of Electronics Communication and Computer Engineering*, 2013, vol. 4, pp. 47–51.
14. **Таубин Ф. А.** Сигнально-кодовые системы для линейного канала. *Девятая Всесоюзная конференция*

по теории кодирования и передачи информации. Тезисы докладов. Одесса, 1998, ч. 3. С. 125–128.
15. **Taubin F.** Trellis coded multitone modulation for linear distortion channels. *Proceedings of the IEEE International Workshop on Information Theory*, 1994, Moscow, pp. 102–106.
16. **Зяблов В. В., Коробков Д. Л., Портной С. Л.** Высокоскоростная передача сообщений в реальных каналах. М.: Радио и связь, 1991. 288 с.
17. **Kasturia S., Aslanis J. T., Cioffi J. M.** Vector Coding for Partial Response Channels. *IEEE Transactions on Information Theory*, 1990, vol. 36, no 4, pp. 741–762. doi:10.1109/18.53735
18. **Riesz F., Nagy B. S.** Functional analysis. London and Glasgow, 1956. 478 p.
19. **Камке Э.** Справочник по обыкновенным дифференциальным уравнениям. М.: Наука, 1971. 576 с.
20. **Наймарк М. А.** Линейные дифференциальные операторы. М.: Наука, 1969. 528 с.
21. **Forney G. D.** Coset codes — Part I: Introduction and geometrical classification. *IEEE Transactions on Information Theory*, 1988, vol. 34, no. 5, pp. 1121–1151. doi:10.1109/18.21245
22. **Forney G. D.** Coset codes — Part II: Binary lattices and related codes. *IEEE Transactions on Information Theory*, 1988, vol. 34, no. 5, pp. 1152–1187. doi:10.1109/18.21246
23. **Calderbank A.** Multilevel codes and multistage decoding. *IEEE Transactions on Communications*, 1989, vol. 37, no. 3, pp. 222–229. doi:10.1109/26.20095
24. **Taubin F. A., Trofimov A. N.** Pipeline decoding of embedded trellis codes: error-tolerance analysis. *Problems Inform. Transmission*, 1990, vol. 26, no. 4, pp. 332–343.

UDC 621.391
doi:10.31799/1684-8853-2018-5-66-78

Trellis-coded modulation for linear distortion channels

F. A. Taubin[a], Dr. Sc., Tech., orcid.org/0000-0002-8781-9531, ftaubin@yahoo.com
[a]Saint-Petersburg State University of Aerospace Instrumentation, 67, B. Morskaia St., 190000, Saint-Petersburg, Russian Federation

Introduction: Real communication channels used for high-speed data transmission are characterized by substantial linear distortions leading to a significant decrease in the effectiveness of the known coded modulation techniques developed for distortion-free channels. This means that a modified coded modulation technique should be created, which would take into account the channel distortions. **Purpose:** Using multidimensional signal sets to build up multilevel trellis-coded modulation constructions with multidimensional constellations and multi-stage decoding, matched to the channels with linear distortions, and study of their main characteristics. **Results:** A new approach was developed to the design of trellis-coded modulation constructions for channels with linear distortions, based on the combination of two techniques: decomposition of the channel into a set of independent memoryless subchannels and time-frequency multilevel coding. The high level of flexibility of the proposed trellis-coded modulation constructions makes it possible to effectively take into account linear channel distortions and compensate for them, due to the rational choice of the number of the subchannels and optimization of the energy distribution between the subchannels. It also provides a wide-range tradeoff between performance and decoding complexity. The relationships between the basic parameters of the proposed class of trellis-coded modulation constructions were obtained in an explicit form. These relationships allow you to provide tradeoffs between the transmission rate, the minimum distance, and the decoding complexity. Examples of multilevel trellis-coded modulation constructions based on multidimensional Barnes — Wall lattices are given. A comparative study showed that as the value of the permissible decoding complexity increases, constructions with larger values of both the number of subchannels and the cardinality of the signal constellation become more preferable. **Practical relevance:** The proposed class of trellis-coded modulation constructions allows you to effectively take into account and to overcome the influence of linear channel distortions common in communication channels used for high-speed data transmission. An additional advantage of these constructions is a relatively small value of the crest factor of transmitted signals and thereby more efficient use of the transmitter's power.

Keywords — coded modulation, trellis coding, linear distortion channels, multidimensional signal sets, Barnes — Wall lattices, multilevel codes, multistage decoding.

Citation: Taubin F. A. Trellis-coded modulation for linear distortion channels. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 5, pp. 66–78 (In Russian). doi:10.31799/1684-8853-2018-5-66-78

References

1. Benedetto S., Marsan M. A., Allegtendo G., Csachin E. Combined coding and modulation: theory and applications. *IEEE Transactions on Information Theory*, 1988, vol. 34, no. 5, pp. 1121–1151. doi:10.1109/18.2631
2. Biglieri E., Divsalar D., McLane P. J., Simon M. K. *Introduction to trellis coded modulation with application*. New York, McMillan Publishing Company, 1991. 576 p.
3. Anderson J. B., Svensson A., Helstrom W. *Coded modulation systems*. New York, Kluwer Academic / Plenum Publishers, 2003. 486 p.
4. Seidl M., Schenk A., Stierstorfer C., Huber J. B. Polar-coded modulation. *IEEE Transactions on Communications*, 2013, vol. 61, no. 10, pp. 4108–4119. doi:10.1109/ TCOMM.2013.090513.130433
5. Xiao X., Hong Y., Viterbo E., Gupta A. Trellis coded modulation for informed receivers. *Proceedings of 2017 IEEE International Conference on Communications Workshops* (ICC Workshops), Paris, France, 2017, pp. 955–960. doi:10.1109/ICCW.2017.7962782
6. Carlisle C. J., Taylor D. P., Shafi M., Kennedy W. K. Performance bounds for trellis-coded modulation on time-dispersive channels. *IEEE Transactions on Communications*, 1994, vol. 42, no. 8, pp. 2534–2542. doi:10.1109/26.310613

7. Forney G. D., Ungerboeck G. Modulation and coding for linear gaussian channels. *IEEE Transactions on Information Theory*, 1998, vol. 44, no. 6, pp. 2384–2415. doi:10.1109/18.720542
8. Xia X.-G. *Modulated coding for inter-symbol interference channels*. New York, McGraw-Hill, 2001. 512 p.
9. Biglieri E. High level modulation and coding for nonlinear satellite channels. *IEEE Transactions on Communications*, 1984, vol. 32, no. 5, pp. 616–626.
10. Euboglu M. V., Qureshi S. U. Reduced-state sequence estimation for coded modulation on inter-symbol interference channels. *IEEE Journal on Selected Areas in Communications*, 1989, vol. 7, no. 6, pp. 989–995. doi:10.1109/49.29621
11. Taubin F. Reduced-state decoding for trellis coded modulation on nonlinear inter-symbol interference channels. *Lecture Notes in Computer Science*, 1994, vol. 829, pp. 88–96.
12. Misra A., Sarma K. K. TCM-coded OFDM assisted by ANN in wireless channels. *International Journal of Smart Sensors and Ad Hoc Networks* (IJSSAN), 2012, vol. 1, no. 3, pp. 50–55.
13. Chide N., Deshmukh S., Borole P. B., Chore N. An overview of OFDM variants and their applications. *International Journal of Electronics Communication and Computer Engineering*, 2013, vol. 4, pp. 47–51.
14. Taubin F. A. Coded modulation for linear channels. *Proceedings of the 9th All-union conference on coding theory and information transmission*, 1998, Odessa, part 3, pp. 125–128 (In Russian).
15. Taubin F. Trellis coded multitone modulation for linear distortion channels. *Proceedings of the IEEE International Workshop on Information Theory*, 1994, Moscow, pp. 102–106.
16. Zyablov V. V., Korobkov D. L., Portnoy S. L. *High-rate information transmission in physical channels*. Moscow, Radio i Svyaz Publ., 1991. 288 p. (In Russian).
17. Kasturia S., Aslanis J. T., Cioffi J. M. Vector coding for partial response channels. *IEEE Transactions on Information Theory*, 1990, vol. 36, no. 4, pp. 741–762. doi:10.1109/18.53735
18. Riesz F., Nagy B. S. *Functional analysis*. London and Glasgow, 1956. 478 p.
19. Kamke E. *Handbook of ordinary differential equations*. Moscow, Nauka Publ., 1971. 576 p. (In Russian).
20. Naimark M. A. *Linear differential operators*. Moscow, Nauka Publ., 1969. 528 p. (In Russian).
21. Forney G. D. Coset codes — Part I: Introduction and geometrical classification. *IEEE Transactions on Information Theory*, 1988, vol. 34, no. 5, pp. 1121–1151. doi:10.1109/18.21245
22. Forney G. D. Coset codes — Part II: Binary lattices and related codes. *IEEE Transactions on Information Theory*, 1988, vol. 34, no. 5, pp. 1152–1187. doi:10.1109/18.21246
23. Calderbank A. Multilevel codes and multistage decoding. *IEEE Transactions on Communications*, 1989, vol. 37, no. 3, pp. 222–229. doi:10.1109/26.20095
24. Taubin F. A., Trofimov A. N. Pipeline decoding of embedded trellis codes: error-tolerance analysis. *Problems Inform. Transmission*, 1990, vol. 26, no. 4, pp. 332–343.

УВАЖАЕМЫЕ АВТОРЫ!

Научная электронная библиотека (НЭБ) продолжает работу по реализации проекта SCIENCE INDEX. После того как Вы зарегистрируетесь на сайте НЭБ (http://elibrary.ru/defaultx.asp), будет создана Ваша личная страничка, содержание которой составят не только Ваши персональные данные, но и перечень всех Ваших печатных трудов, имеющихся в базе данных НЭБ, включая диссертации, патенты и тезисы к конференциям, а также сравнительные индексы цитирования: РИНЦ (Российский индекс научного цитирования), h (индекс Хирша) от Web of Science и h от Scopus. После создания базового варианта Вашей персональной страницы Вы получите код доступа, который позволит Вам редактировать информацию, помогая создавать максимально объективную картину Вашей научной активности и цитирования Ваших трудов.

УДК 621. 391.3
doi:10.31799/1684-8853-2018-5-79-87

Articles

Theorem about key capacity of a communication network

A. D. Sinjuk[a], *PhD, Associate Professor, orcid.org/0000-0003-0608-4359*
O. A. Ostroumov[a], *PhD, Post-Graduate Student, orcid.org/0000-0003-1674-6248, oleg-26stav@mail.ru*
[a]*S. M. Budyonny Military Academy of Telecommunications, 3, Tikhoretskii Ave., 194064, Saint-Petersburg, Russian Federation*

Introduction: *An important aspect of cryptographic telecommunication systems is the encryption key control problem. The most complicated stages in its solution are the safe generation of the keys, their distribution, and their delivery to legitimate subscribers via protected communication channels, which is fairly expensive, sometimes slow and not always possible. As an alternative, the keys can be generated by transferring information via telecommunication channels, being possibly exposed to a violator. The known estimations of information efficiency look like a solution of sophisticated information theory problems for certain ways of open key coordination between two legitimate subscribers. Efficiency estimations for the conditions of network key generation are not known.* ***Purpose:*** *A strictly conclusive search for potential estimates of information efficiency of open network key generation.* ***Results:*** *Within the formulated statement of the problem, we have proposed a violator model and a network channel connectivity model which is a combination of a broadcast channel connecting three legitimate subscribers and an intercept channel at the output of which the violator controls the transferred information. The information exchange is based on the proposed models of a random coder and deterministic decoder with a specially developed asymptotic method of key generation. In order to assess the process, we introduce a system of quality indicators and requirements which differs from the known ones by its definition of "information" speed of network key generation. We also introduce a term of key network capacity which determines the asymptotic information efficiency of the key generation. We have formulated and rigorously proved a theorem about key capacity. The boundary values have been substantiated.* ***Practical relevance:*** *The obtained results develop the known scientific achievements in the field of open key coordination theory and can be used by specialists in the design and development of key control subsystems in modern cryptographic information security systems which provide closed network information exchange.*

Keywords *— communication network, legitimate subscribers, violator, network key, broadcast communication channel, intercept channel, random coder and deterministic decoder, joint information, asymptotic method of network secret key agreement by public discussion, quality indicators and requirements to a network key, network key generation speed, theorem about key capacity of a communication network of minimum size, key capacity assessment.*

Citation: Sinjuk A. D., Ostroumov O. A. Theorem about key capacity of a communication network. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 5, pp. 79–87. doi:10.31799/1684-8853-2018-5-79-87

Introduction

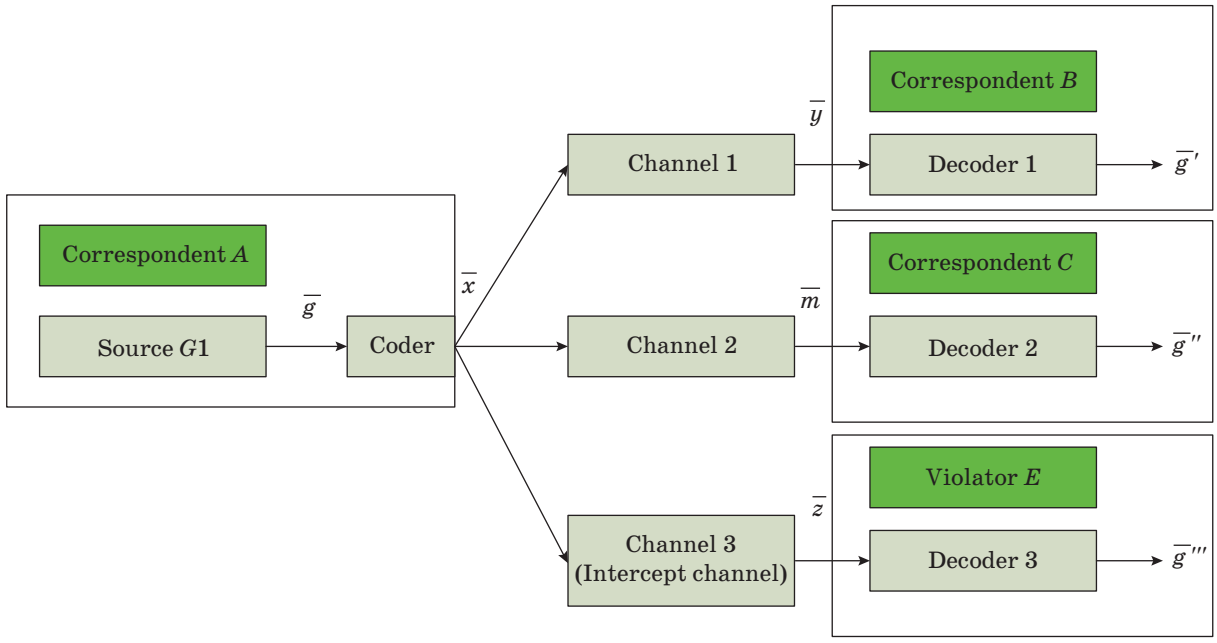
When ensuring information security of telecommunication systems which use cryptographic methods to protect information, it is most important to find an effective solution for the problem of forming, distributing and delivering keys to the subscribers. At the moment, this problem is solved by using protected communication channels, which is expensive, slow and not always possible. Therefore, it is of great practical interest to develop methods for generating these keys with open telecommunication channels. In these circumstances, we need to estimate the information efficiency of generating keys via open communication channels in order to optimize the secret key agreement by public discussion methods which are now under development. The obtained results develop the known scientific achievements in the network key sharing by public discussion [1–5].

Statement of the problem

We have to evaluate the information efficiency of generating a common key for legitimate subscribers (LS) of a communication network, called A, B and C, while exchanging data between them via channels available for a violator E. We need to ensure that a network key is promptly and reliably generated for the LS and the information about this key has a low level for E.

Let us consider the following generalized description of information transfer via the network [6, 7]. The LS A has a transmitter (coder). Two other LS B and C, and the violator E have three independently functioning receivers (decoders) whose inputs are fed with output signals of various channels. The transmitter gets messages $\overline{g}$ from a source $G1$ which LS A should simultaneously pass to the receivers 1, 2 and 3 (i. e. to the LS B and C, and the violator E). The channel connectivity model (CCM) of the network is shown in Fig.

Let the source $G1$ be described by a model of a discrete stationary source without memory [8–10]. The alphabet of the source is specified by a set G, consisting of t letters $G = \{g1, ..., g_t\}$. In every time unit, it independently chooses the i^{th} letter from the alphabet with a preset probability $p(g_i)$ [11], being specified by the assembly $\{G, p(g)\}$ [11]. Let the source generate a message $\overline{g}$, which is a sequence of



■ Channel connectivity model

k letters, and $\bar{g} \in G^k$, where G^k is a Cartesian k^{th} degree of the set G. The total number of the sequences from the source M_0 is

$$M_0 = t^k. \quad (1)$$

Generation probability $\bar{g}$ is equal to

$$p(\bar{g}) = \prod_{i=1}^{k} p\left(g^{(i)}\right), \quad (2)$$

where $g^{(i)}$ is the i^{th} element of the sequence $\bar{g}$.

Information parameter of the source is the entropy [8–10] which is equal to

$$H_S = H(G), \quad (3)$$

where $H(G)$ is the entropy of the assembly G.

Let us assume that all the communication channels in the CCM are described by models of discrete symmetric communication channels without memory (DSC) [10–12]. The complex of two channels with a common input (at the output of the coder LS A) and outputs (at the inputs of receivers 1 and 2 which are LS B and C respectively) can be described by a model of a broadcast channel (BC) [13–15] which binds the three LS in a framework of a loose network with the smallest number of subscribers [7, 8]. Signal transmission is determined by two channels with an input alphabet X, output alphabets Y and M, and matrices of transition probabilities $P_1 = \{p(y/x)\}$, $P_2 = \{p(m/x)\}$, $x \in X$, $y \in Y$, $m \in M$. Alphabets X, Y and M are finite. Let us denote the BC as $\{X, Y, M; p(y/x), p(m/x)\}$; the channels $\{X, Y; p(y/x)\}$ and $\{X, M; p(m/x)\}$ are components of the BC [13]. The DSC from the coder output of LS A to the input of the receiver 3 of violator E will be called an intercept channel (IC). Signal transmission via the IC $\{X, Z; p(z/x)\}$ is determined by the input alphabet X, output alphabet Z and a matrix of transition probabilities $P_3 = \{p(z/x)\}$, $x \in X$, $z \in Z$. The components of the BC and IC are independent channels:

$$p(y, m, z/x) = p(y/x)\, p(m/x)\, p(z/x).$$

Let us assume that the alphabets of $G1$, BC and IC coincide, and $|G| = |X| = |Y| = |M| = |Z| = t$.

Models of random coder and deterministic decoder

Let us build a random coder and a deterministic decoder which determine the asymptotic transmission of information for generating a network key in the model shown in Fig. Let a large n be specified. We need to build up a code (some approaches are discussed in [9, 16–18]) with which information will be transferred via the BC. We will consider only such codes for which the coder is an identity mapping on a set of code words. In this case, codes will be defined by sets of code words and decoding mappings, not by coders and decoders as it was in [9]. For information transmission, let us use a random coder. Out of a set X^n, we will choose a set of code words V; this set is a highly probable set of typical sequences [9, 11]. For building up V, the test channel method

can be used, discussed in details in [9]. Let the cardinality J of the set V be no more than

$$J = |V| < 2^{n(F(X;\,Y;\,M)\,-\,\tau)},$$

where $F(X; Y; M)$ is the average joint information (JI) of the BC defined in [19, 20], and τ is a certain positive number ($\tau > 0$).

A set of code words V specifies a certain (n, ε_1) code on X^n with the speed

$$R < F(X; Y; M) - \tau. \qquad (4)$$

In accordance with the direct BC coding theorem [6] for the (n, ε_1) code we provide an average probability of erroneous decoding no more than ε_1, $\varepsilon_1 > 0$. Let us split V into M_0 disjoint subsets C_i of the same cardinality. The cardinality D of any i^{th} subset C_i, where $i = 1, 2, ..., M_0$, is equal to $D = J / M_0$.

In accordance with the theorem of highly probable sets [9, 11], when n is large, probabilities of elements V are close to each other (i. e. the distribution of probabilities on the elements of the set V is close to uniform) and $P(V)$ is the total probability of all elements of the set V (the probability of (n, ε_1) code) is close to 1, i. e. for an arbitrarily small $\rho > 0$

$$P(V) > 1 - \rho. \qquad (5)$$

Then the probabilities of the elements C_i are also close to each other. The expression (5) means that for any sequence $\bar{x}, \bar{x} \in X^n$, with probability $p(\bar{x} \in C_i) \geq \tilde{n}$, we can find C_i to which $\bar{x}$ belongs. For each i^{th} sequence $\bar{g}_i$ of the source $G1$, a certain C_i is chosen. Let us consider random coding. Let $\bar{g}_i$ appear at the output of $G1$. The coder chooses C_i. Then randomly, with probability $1/D$, it chooses from the sequences C_i a code word and sends it to the input of the BC (and the IC). Such coding by (n, ε_1) code determines the probability distribution on X^n

$$p(\bar{x}) = \begin{cases} \dfrac{1}{D} p(\bar{g}_i), \text{ для всех } \bar{x} \in C_i, \text{ где } i = 1, \ldots, M_0 \\ 0, \text{ для всех остальных } \bar{x} \notin \bigcup\limits_{i=1}^{M_0} C_i. \end{cases} \qquad (6)$$

The next statement (proved in [6]) determines that JI does not grow when a $G1$ sequence is transferred by a code word as long as n characters with the use of random coding, as compared to the JI of the BC $F(X^n; Y^n; M^n)$.

Statement 1. Let $F(X^n; Y^n; M^n)$ be defined for (6) and $F(G^k; Y^n; M^n)$ be the average JI between the $G1$ output and the outputs of the BC. Then

$$F(G^k; Y^n; M^n) \leq F(X^n; Y^n; M^n), \qquad (7)$$

where G^k is Cartesian k^{th} degree of the set, and the equality in (7) is satisfied if C_i consists of just one code word.

Let us discuss the deterministic decoder model [6] and find the probability of correct decoding. LS B and C choose C_i as solving areas corresponding to $\bar{g}_i$, and $S^i \subseteq Y^n$, $Q^i \subseteq M^n$, where $i = 1, ..., M_0$. For each code word $u_i = 1, ..., M_0$, let us define a joint solving area L which would unite the solving areas $S^i \subseteq Y^n$, $Q^i \subseteq M^n$. Let $J1$ be a joint assemble of solutions which is a result of mapping of the assemble $Y^n M^n$ onto a set of solutions. Each couple of sequences $(\bar{y}, \bar{m}) \in Y^n M^n$ determines the solution $j \in J1$ according to the following rule:

$$j1 = \begin{cases} j1_i, \text{ если } (\bar{y}, \bar{m}) \in L_i, \; i = 1, \ldots, M_0, \\ j1_{M_0+1}, \text{ если } (\bar{y}, \bar{m}) \notin \bigcup\limits_{i=1}^{M_0} L_i. \end{cases} \qquad (8)$$

In [6] it is proved that after transmitting a $G1$ sequence by a code word of n characters with random coding and decoding by the rule (8), the uncertainty of the source with known outputs of the BC $HF(G^k/Y^n, M^n)$ does not exceed the uncertainty of the BC $HF(X^n/Y^n, M^n)$ specified in [6].

After transmitting a code word with probability distribution (6) and decoding by the rule (8), the full group of events is defined as

$$\Pr\left((\bar{y}, \bar{m}) \notin \bigcup_{i=1}^{M_0} L_i\right) + \Pr\left((\bar{y}, \bar{m}) \in \bigcup_{i=1}^{M_0} L_i\right) = 1, \qquad (9)$$

where $\Pr\left((\bar{y}, \bar{m}) \notin \bigcup\limits_{i=1}^{M_0} L_i\right)$ is the probability of making a decision j_{M0+1} (probability of giving up the decoding) by rule (8) or the probability of combination $(\bar{y}, \bar{m}) \in Y^n M^n$, which both stay "outside the (n, ε_1) code". Then, according to (5)

$$\Pr\left((\bar{y}, \bar{m}) \notin \bigcup_{i=1}^{M_0} L_i\right) < 1 - (1 - \rho)^2.$$

For $\Pr\left((\bar{y}, \bar{m}) \notin \bigcup\limits_{i=1}^{M_0} L_i\right)$, "inside the (n, ε_1) code", according to (8), let us write

$$\Pr\left((\bar{y}, \bar{m}) \in \bigcup_{i=1}^{M_0} L_i\right) = P_{ed} + P_{cd}, \qquad (10)$$

where P_{cd}, P_{ed} are the probabilities of correct and erroneous decoding of the (n, ε_1) code respectively, while $P_{ed} < \varepsilon_1$. Then the sought probability

$$P_{cd} \geq (1 - \rho^2) - \varepsilon_1. \qquad (11)$$

The model of a violator. Asymptotic method of forming a network key

The IC output for the violator E, if a $G1$ sequence is chosen and a random coder is used, is a random choice of a $G1$ message and uniform distribution of the input sequence at the IC input. It is assumed that the violator uses a *passive* strategy, observing the information exchange between the LS at the IC output. Other cases are discussed in [21–24]. It is assumed that E knows full descriptions of LS actions in forming a key, the (n, ε_1) code and the source $G1$. The violator's ignorance degree can be measured by the uncertainty rate ω [25]

$$\omega = H(G^k / Z^n). \qquad (12)$$

Here is a general description of the key generation method for a large n:

1. The LS A, using $G1$, randomly chooses $\overline{g}_i$, $\overline{g}_i \in G^k$, where $i = 1, ..., M_0$.

2. The LS A, using random coding, associates $\overline{g}_i$, $\overline{g}_i \in G^k$, with a code word $\overline{x}$, $\overline{x} \in V$ of n characters.

3. The LS A transmits $\overline{x}$, $\overline{x} \in V$ via the BC (and the IC), LS B, C (and E).

4. The LS B receives a sequence $\overline{y}$, $\overline{y} \in Y^n$ at the output of the first component BC. The LS C receives $\overline{m}$, $\overline{m} \in M^n$ at the output of the second component channel (the violator E receives $\overline{z}$, $\overline{z} \in Z^n$ at the IC output).

5. The LS B, with a probability higher than $1 - \rho$, according to the decoding rule which states that the receiver 1 makes a decision about the transferred code word, does the following: if $\overline{y} \in S_i$, where $i = 1, ..., M_0$, it makes a decision about the message $\overline{g}_i$ and gets $\overline{g}'$, $\overline{g}' \in G^k$, if $\overline{y} \notin \bigcup S_i$, the decision is to give up the decoding. The LS C, with a probability higher than $1 - \rho$, according to the decoding rule which states that the receiver 2 makes a decision about the transferred code word, does the following: if $\overline{m} \in Q_i$, where $i = 1, ..., M_0$, it makes a decision about the message $\overline{g}_i$ and gets $\overline{g}''$, $\overline{g}'' \in G^k$, if $\overline{m} \notin \bigcup Q_i$ (with a probability lower than ρ), the decision is to give up the decoding.

6. With a probability higher than $(1 - \rho)^2$, sequences $\overline{g}'$, $\overline{g}''$, $\overline{g}$ are chosen as a key for the LS B, C and A respectively.

The system of quality score and requirements to a key

The main quality indicators of a generated key can be reduced to reliable transmission of a large number of bits of a "good" key and a small leakage of information to the violator E. The term "good" means that the quality indicators of a key fit certain requirements adequate to the conditions of its generation and usage. The expression (9) shows that with a certain non-zero probability the keys may be not generated altogether. Furthermore, the keys (see expression (10)) may not coincide. Let us take into account all the cases in our estimation of probability P_ε. Apparently, P_ε is an addition to the correct decoding probability P_{cd} defined in (11). Then

$$P_\varepsilon = 1 - P_{cd} < 1 - (1 - \rho)^2 + \varepsilon_1.$$

As the next indicator, it would be expedient to choose the speed I_E with which the violator receives information about the key [25], i. e. the average amount of mutual information [9, 10, 12] between the assemblies of the source G^k and sequences at the IC output Z^n, assigned to the length n of a code word (CW) $I_E = I(G^k; Z^n)/n$.

The third indicator is the quality of the generated key Ω treated as closeness of its probability distribution to the uniform distribution of the probabilities of the key characters [11, 25]

$$\Omega = k(\log_2 t - H_s) / n, \qquad (13)$$

where k is the length of $G1$ messages, t is the volume of $G1$ alphabet, H_s is the entropy of $G1$ from (3).

Let us introduce the "information" speed of key generation.

Definition 1. A number H_3 is called speed of BC key generation if for a sufficiently large n and arbitrarily small $\varepsilon_2 > 0$, $\varepsilon_3 > 0$, $\varepsilon_4 > 0$ и $\varepsilon_5 > 0$, when the asymptotic method of key generation is used, a key is generated which meets the following requirements:

$$P_\varepsilon < \varepsilon_2; \qquad (14)$$
$$I_E < \varepsilon_3; \qquad (15)$$
$$\Omega < \varepsilon_4; \qquad (16)$$
$$H_3 < (H(G^k) - \varepsilon_5)/n. \qquad (17)$$

The introduced indicator determines the amount of information about the key generated, which includes one channel character of a code word transferred via the BC. From [8–10, 12] it is known that the code speed is equal to

$$R = (\log M)/n, \qquad (18)$$

where M is the code volume, and n is the code length.

It follows from definition 1 that H_3 is the "information" speed of the key generation, and from (18) it follows that R is the "information" speed of the transmission. Analysis of the asymptotic method of key generation shows that in the discussed method, a key is also transferred with random coding. The code speed in (18) is the maximum amount of information which can be transferred in a single channel character. Analysis of (17) and (18) shows that the maximum achievable H_3 does not exceed R, i. e.

$H_3 \le R$, and $H_3 = R$, provided that the violator's IC is "cut". We have to find the maximum H_3, as it determines the information efficiency of open generation of a network key, which will be represented as the key capacity. Therefore, we have to specify the bounds of R as a (n, ε_1) code parameter for any $\varepsilon > 0$ within

$$\left|\frac{1}{n}\log_2(M_0) - R\right| < \varepsilon, \qquad (19)$$

where M_0 is the volume of the set of $G1$ messages.

Then the research goal is reduced to finding the maximum achievable value of H_3 with a maximum possible R.

Key capacity theorem

In order to prove the theorem, let us prove the following statement:

Statement 2. Let $I(X^k; Z^n)$ be an average mutual information between the input and output of an IC by distribution (6), and let $I(G^k; Z^n)$ be an average mutual information between the output of the source $G1$ and output of the IC obtained with the use of a random coder. Then

$$I(X^k; Z^n) \ge I(G^k; Z^n) \qquad (20)$$

and

$$H(X^n / Z^n) \ge H(G^k / Z^n). \qquad (21)$$

The equality in (20) and (21) is achieved if each subset C_i of a random coder, where $i = 1, 2, ..., M_0$ consists of just one code word.

Proof

Let us consider mutual information $I(X^n, G^k); Z^n)$:

$$I((X^n, G^k); Z^n) = I(X^n, Z^n) + I(G^k; Z^n / X^n) = \\ = I(G^k; Z^n) + I(X^n; Z^n / G^k).$$

Taking into account that the output of the IC is determined only by its input, we have $H(Z^n / G^k, X^n) = H(Z^n / X^n)$, and hence $I(G^k; Z^n / X^n) = 0$. Then

$$I(X^n, Z^n) = I(G^k; Z^n) + I(X^n; Z^n / G^k). \qquad (22)$$

The expression (22) proves that the inequality (20) is true.

The analysis of the joint probability distribution law at $G1$ output in (2) and at the BC input in (6) shows that

$$H(X^n, G^k) = H(X^n) + H(G^k / X^n) = \\ = H(G^k) + H(X^n / G^k).$$

Taking into account (6), the BC input can uniquely determine the source output, then $H(G^k / X^n) = 0$. Therefore, we can write

$$H(X^n) = H(G^k) + H(X^n / G^k) \ge H(G^k). \qquad (23)$$

The equality sign in (23) is achieved if each C_i consists of just one code word. Let us unwind (22).

$$H(X^n) - H(X^n / Z^n) = H(G^k) - H(G^k / Z^n) + \\ + H(X^n / G^k) - H(X^n / Z^n, G^k).$$

Taking into account (23), we have

$$H(G^k / Z^n) = H(X^n / Z^n) - H(X^n / Z^n, G^k).$$

The last expression proves that the inequality (21) is true. The statement is proved.

Definition 2. The key capacity C_3 of a BC is the maximum achievable H_3 with which a key is generated meeting the requirements (14) – (17) for arbitrarily small $\varepsilon_2 > 0$, $\varepsilon_3 > 0$, $\varepsilon_4 > 0$ and $\varepsilon_5 > 0$:

$$C_3 = \max H_3.$$

Theorem. Let, in the above-described channel connectivity conditions, the subscribers of a connection network (BC) use the asymptotic method of key generation with uniform distribution of character probabilities at $G1$ output, a random coder of (n, ε_1) code for transmission and a deterministic decoder for reception in order to generate a network key. Let the violator follow a passive strategy of information intercept. Let the value of BC capacity C exceed the value of IC capacity C_w. Then

$$C_3 = C - C_w, \qquad (24)$$

where C is the capacity of the BC [8, 9, 15], and C_w is the capacity of the IC

$$C_w = \log_2 L + \sum_{j=1}^{L} p_j \log_2 p_j,$$

where L is the volume of the IC output alphabet, $p_1, ..., p_L$ are the elements in the first line of the transient probability matrix of the IC [9].

Proof

In accordance with the definition of BC uncertainty $HF(X^n / Y^n, M^n)$ from [6], using Fano's inequality [9, 12], let us write:

$$HF(X^n / Y^n, M^n) \le h(\lambda) + \lambda \log(M_0),$$

where λ is the average probability of a BC decoding error [6]. Then for any χ, $\chi > 0$, depending on ε_1, we can write

$$HF(X^n / Y^n, M^n) / n < \chi. \quad (25)$$

Let us find out how H_3 is interconnected with the violator's uncertainty speed ω from (12) when the requirements (14) – (17) are met, for arbitrarily small $\varepsilon_2 > 0$, $\varepsilon_3 > 0$, $\varepsilon_4 > 0$ and $\varepsilon_5 > 0$. Add up the left and right parts of the inequalities which determine the demands to the key (15) and (17), move I_E to the right-hand side, represent the latter as average mutual information [8, 9, 12] and finally get

$$H_3 < H(G^k / Z^n) / n + \varepsilon_3 - \varepsilon_5. \quad (26)$$

Now, taking into account the expression (21) and statement 1, we can rewrite (26) as

$$H_3 < H(X^n / Z^n) / n + \varepsilon_3 - \varepsilon_5. \quad (27)$$

If we add up the left and right parts of the inequalities (25) and (27) and move $HF(X^n / Y^n, M^n) / n$ to the right-hand side, we will get

$$H_3 < H(X^n / Z^n) / n - HF(X^n / Y^n, M^n) / n + \varepsilon_3 - \varepsilon_5 + \chi. \quad (28)$$

Adding and subtracting $H(X^n) / n$ in the right-hand side of (28), according to the definitions of JI and mutual information from [9, 22], will give us the following:

$$H_3 < F(X^n; Y^n; M^n) / n - I(X^n; Z^n) / n + \varepsilon_3 - \varepsilon_5 + \chi. \quad (29)$$

Using the results of the theorems about BC information capacity [6] and DSC information capacity [9], we can rewrite (29) as

$$H_3 < \max_{\{p(x)\}} F(X; Y; M) - \max_{\{p(x)\}} I(X; Z) + \varepsilon_3 - \varepsilon_5 + \chi. \quad (30)$$

Results of the theorems about average JI maximization [19, 20] and average mutual information maximization [9] $\max\limits_{\{p(x)\}} F(X; Y; M)$ and $\max\limits_{\{p(x)\}} I(X; Z)$ can be achieved with independent and uniform distribution of BC and IC input character probabilities. For (n, ε_1) code speed defined in (4), the following is true:

$$R < \max_{\{p(x)\}} F(X; Y; M) - \tau. \quad (31)$$

Let a sufficiently large n ($n \to \infty$) be chosen. Then, according to the theorem of highly probable sets [9, 11], the probability of (n, ε_1) code $P(V)$ will tend to 1, and $\rho \to 0$. Then at the input of the BC (and IC), uniformly distributed (with nearly equal probability) code words will appear. Due to this, the distribution of BC and IC input character probabilities is very close to uniform. The results of the theorem about average JI maximization and the direct theorem of BC coding [6] allow us to rewrite (31) (note that $\tau \to 0$ when $n \to \infty$):

$$R < C. \quad (32)$$

This means that (n, ε_1) code speed can be maximized as closely to the BC capacity value as we want, but will never exceed it. Consequently, $R \to \max$. Then it follows from the direct theorem of BC coding [6] that $\varepsilon_1 \to 0$, too. Now from $\rho \to 0$ it follows that $\varepsilon_2 \to 0$, too. Then the fulfillment of requirement (14) determines that $P_\varepsilon \to 0$. In this conditions, $HF(X^n / Y^n, M^n) \to 0$. From the inequality (25), we can see that $\chi \to 0$, too. Using the results of the theorems about information capacity and the direct theorem of DSC coding [9] with uniform distribution of code words at the IC input

$$\max_{\{p(x)\}} I(X; Z) = C_w, \quad (33)$$

Let us assume that $C > C_w$. When $n \to \infty$, we can find a code whose value R will satisfy the inequality

$$C_w < R < C. \quad (34)$$

Using the results of the information capacity theorem and the converse theorem about DSC coding from [9] for the left-hand side of (34), i. e. the condition $C_w < R$, we can claim that the average probability of erroneous decoding in the IC will be higher than a certain preset positive number. Consequently, $I(X^n; Z^n) \to 0$. Then, taking into account (20) the statement 2, we can say that

$$I(G^k; Z^n) \to 0. \quad (35)$$

From the analysis (35), we can conclude that $\varepsilon_3 \to 0$ in the requirement (15). Let us estimate how close the key assembly distribution is to uniform probability distribution using the parameter Ω defined in the requirement (16). If $n \to \infty$, then, according to the definition of R in (4) and the restriction (34), k (the length measured in characters of the generated sequence $\bar{g}$ of the source $G1$) will also grow, i. e. $k \to \infty$. Let a random probability distribution be specified at the source output. According to the highly probable set theorem [9], when $k \to \infty$, a sequence at the $G1$ output is split into a highly probable subset of typical sequences U and a subset of non-typical sequences $\bar{U}$. The probabilities of elements U are close to each other, and $P(U)$, the to-

tal probability of the elements U, will be close to 1. Taking this into account, we can write

$$\log_2(M_0) \geq H(G^k), \quad (36)$$

where M_0 is the volume of the set of $G1$ sequences defined in (1), and $H(G^k)$ is the entropy of its assembly. The equality sign in (36) is achieved when the probability distribution of the messages at the source output is uniform. For that, the discrete stationary memoryless source $G1$ must have a uniform distribution law for the character probabilities in its messages [8, 9, 11]. Therefore, let us choose this distribution law at $G1$ output. In this case, $\Omega \to 0$ in (13), and hence $\varepsilon_4 \to 0$ in the requirement (16). Taking into account the above-mentioned conditions and considering the joint functioning of the source and a random coder, we can see that when $n \to \infty$, the highly probable subset of $G1$ output sequences expands to a highly probable subset of code words at the output of a random coder of (n, ε_1) code (this takes place at the input of BC and IC). Further analysis of using the asymptotic method of key generation shows that the method can be reduced to transferring coded sequences via the BC so that the violator who can observe only the output of his/her IC (and who knows the code in use) cannot restore a message you send. Analysis of (36) shows that when $n \to \infty$, the speed R of (n, ε_1) code will tend to its maximum in (18) and hence $\varepsilon \to 0$ in (19). Since the information about a "higher quality" key is transferred via the BC, H_3 will tend to the maximum R and, consequently, in the requirement (17) $\varepsilon_5 \to 0$. Furthermore, taking into account the definition 2 and inequality (30), let us write

$$C_3 = \max_{\{p(x)\}} \left[F(X; Y; M) - I(X; Z)\right] + \varepsilon_3 - \varepsilon_5 + \chi. \quad (37)$$

Under the condition described above, the key requirement parameters (14) – (17) $\varepsilon_2 \to 0$, $\varepsilon_3 \to 0$, $\varepsilon_4 \to 0$, $\varepsilon_5 \to 0$, and $\chi \to 0$ in (25). Then, taking into account the fulfillment of (31) – (33), we can rewrite (37) as

$$C_3 = C - C_w. \quad (38)$$

Expression (38) coincides with expression (24) from the statement of the theorem. Thus the theorem is proved.

Estimating the boundary values of the key capacity

If $C \leq C_w$, then $C_3 = 0$. Consequently, the lower boundary of C_3

$$0 \leq C_3.$$

In order to find the upper boundary of C_3, let us use the following statement whose proof is given in [6].

Statement 3. C_3 is bounded above

$$C_3 \leq \max F(X; Y; M / Z),$$

where the equality sign is set in the case of statistical independence of the assembly X at the input and assembly Z at the output of the IC (the IC is "cut" [7, 10, 12]).

Conclusion

The work contains new scientific results about open generation of keys for a communication network which includes three subscribers of connected BCs. We propose models of a random coder, deterministic decoder, asymptotic generation method and a system of quality indicators and requirements to a network key. We introduce a term of key capacity of a BC. We have formulated and proved a theorem about the key capacity of a BC, which provides the ways to estimate the asymptotic information efficiency of the studied process. We also have shown its boundary values.

The problems for further research should be assessing the key capacity for information exchange within a binary BC, search for regularities which determine the ways of forming a "good" key, and establishing integral links with earlier scientific results in the filed of secret key agreement by public discussion. On the base of the latter, a comprehensive comparative analysis should be performed, with the development of a methodology for estimating the gain of the proposed key generation method. Another goal could be determining the conditions and developing new methods for providing higher key capacity of a BC.

The obtained results can be helpful for the specialists in the design and optimization of key control subsystems in modern cryptographic information security systems which provide uninterrupted closed network information exchange.

References

1. **Siavoshani M. J., Mishra S., Diggavi S. N., Fragouli Ch.** Group secret key agreement over state-dependent wireless Broadcast channels, *IEEE ISIT 2011*, 2011, pp. 1960–1964.
2. **Shimizu T., Iwai H., Sasaoka H.** Group secret key agreement based on radio propagation characteristics in wireless relaying systems. *IEICE Transactions* (IEICET), 2012, 95-B(7), pp. 2266–2277.
3. **Naito M., Watanabe Sh., Matsumoto R., Uyematsu T.** Secret key agreement by soft-decision of signals in

gaussian maurer's model. *IEICE Transactions* (IEICET), 2009, 92-A(2), pp. 525–534.
4. **Li Zh., Wang H., Fang H.** Group-based cooperation on symmetric key generation for wireless body area networks. *IEEE Internet of Things Journal*, 2017, no. 4(6), pp. 1955–1963.
5. **Halford T. R., Courtade T. A., Chugg K. M., Li X., Thatte G.** Energy-efficient group key agreement for wireless networks. *IEEE Transactions on Wireless Communications*, 2015, no. 14 (10). pp. 5552–5564.
6. **Sinjuk A. D.** *Formirovanie trekhstoronnego shifrkliucha po otkrytym kanalam sviazi s oshibkami* [Forming a three-way encryption key through open communication channels with errors], SPb, VAS, 2009. 360 p. (In Russian).
7. **Shangin V. F.** *Informacionnaja bezopasnost' komp'juternyh sistem i setej* [Information security of computer systems and networks]. Moscow, Forum Publ., 2013. 416 p. (In Russian).
8. **Sklar B.** *Digital communications: fundamentals and applications*. Los Angeles, University of California, 2007. 1104 p.
9. **Kolesnik V. D., Poltyrev G. Sh.** *Kurs teorii informacii* [The course of information theory]. Moscow, Nauka Publ., 1982. 416 p. (In Russian).
10. **Sannikov V. G.** *Teorija informacii i kodirovanija* [Theory of information and coding]. Moscow, MTUSI Publ., 2015. 96 p. (In Russian).
11. **Bure V. M., Parilina E. M.** *Teorija verojatnostej i matematicheskaja statistika* [Theory of Probability and Mathematical Statistics]. SPb, Lan' Publ., 2013. 416 p. (In Russian).
12. **Bikkenin R. R., Chesnokov M. N.** *Teorija jelektricheskoj svjazi* [The theory of electrical communication]. Moscow, Akademija Publ., 2010. 329 p. (In Russian).
13. **Nair Ch., Gamal. A.** The capacity region of a class of three-receiver Broadcast channels with degraded message sets. *IEEE Transactions on Information Theory* (TIT), 2009, no. 55(10), pp. 4479–4493.
14. **Gohary R. H., Davidson T. N.** The capacity region of a product of two unmatched physically degraded gaussian broadcast channels with three individual messages and a common message. *IEEE Transactions on Information Theory* (TIT). 2013, no. 59(1), pp. 76–103.
15. **Kim H., Gamal A.** Capacity theorems for broadcast channels with two channel state components known at the receivers. *IEEE Transactions on Information Theory*. 2016, no. 62 (12), pp. 6917–6930.
16. **Muramatsu J., Miyake Sh.** construction of codes for the wiretap channel and the secret key agreement from correlated source outputs based on the hash property. *IEEE Transactions on Information Theory* (TIT). 2012, no. 58(2), pp. 671–692.
17. **Gao Y., Tuncel E.** Wyner-Ziv coding over broadcast channels: hybrid digital/analog schemes. *IEEE Transactions on Information Theory* (TIT). 2011, no. 57(9), pp. 5660–5672.
18. **Liang Y., Kramer G.** rate regions for relay broadcast channels. *IEEE Transactions on Information Theory* (TIT). 2007, no. 53(10), pp. 3517–3535.
19. **Ostroumov O. A., Sinjuk A. D.** Study of joint information. *Informacija i kosmos*. 2017, no. 3, pp. 55 — 58 (In Russian).
20. **Ostroumov O. A., Sinjuk A. D.** The information-information model of transmission of the general information of the broadcasting channel. *Vestnik komp'juternyh i informacionnyh tehnologij*. 2017, no. 11, pp. 29–36 (In Russian).
21. **Yakovlev V., Korzhik V., Bakaev M.** Protocols of key formation based on communication channels with noise under conditions of active interception using ex-tractors. *Problems of Information Security. Computer systems*. 2006, no. 1, pp. 60 — 81(In Russian).
22. **Dodis Y., Kanukurthi B., Katz J., Reyzin L., Smith A.** Robust fuzzy extractors and authenticated key agreement from close secrets. *IEEE Transactions on Information Theory* (TIT). 2012, no) 58(9), pp. 6207–6222.
23. **Alezabi K. A., Hashim F., Hashim Sh. J., Ali B. M.** An efficient authentication and key agreement protocol for 4G (LTE) networks. *IEEE region 10 symposium*. 2014. pp. 502 – 507.
24. **Wazid M., Das A. K., Kumar N., Odelu V., Reddy A. G., Park K., Park Y.** Design of lightweight authentication and key agreement protocol for vehicular ad hoc networks. *IEEE Access*. 2017, no. 5, pp. 14966 – 14980.
25. **Maurer U.** Secret key agreement by public discussion based on common information. *IEEE Trans. on IT*. 1993, no. 39, pp. 733–742.

УДК 621.391.3
doi:10.31799/1684-8853-2018-5-79-87

Теорема о ключевой пропускной способности сети связи

А. Д. Синюк[a], доктор техн. наук, доцент, orcid.org/0000-0003-0608-4359
О. А. Остроумов[a], канд. техн. наук, адъюнкт, orcid.org/0000-0003-1674-6248, oleg-26stav@mail.ru
[a]Военная академия связи им. Маршала Советского Союза С. М. Буденного. Тихорецкий проспект, 3, Санкт-Петербург, 194064, РФ

Введение: важнейшим аспектом функционирования криптографических телекоммуникационных систем представляется решение задачи управления ключами шифрования. Сложными этапами ее решения являются безопасное формирование, распреде-

ление и доставка ключей законным корреспондентам с использованием защищенных каналов связи, что достаточно дорого, не всегда оперативно и возможно. Альтернативой выступают способы формирования ключей посредством передачи информации по каналам электросвязи, которая, возможно, становится известной нарушителю. Поиск известных оценок информационной эффективности для некоторых способов открытого ключевого согласования двух законных корреспондентов представлял решение сложных теоретико-информационных задач. Результаты же решения подобных задач для более сложных условий, связанных с введением третьего корреспондента и открытым формированием теперь уже сетевого ключа, до настоящего времени не известны. **Цель исследования:** строго доказательный поиск потенциальных оценок информационной эффективности открытого сетевого формирования ключей. **Результаты:** произведена постановка задачи оценки информационной эффективности открытого ключевого согласования сети и определены условия, обеспечивающие ее решение. В рамках первого результата предложены модели нарушителя и сетевой канальной связности трех законных корреспондентов. Последняя модель представляет совокупность широковещательного канала, связывающего корреспондентов, и канала перехвата нарушителя. Обмен информацией основан на предложенных моделях случайного кодера и детерминированного декодера посредством разработанного асимптотического метода формирования ключа. В целях достижения второго результата представлена система показателей качества и требований, отличающаяся от известных определением «информационной» скорости формирования сетевого ключа. Введен термин ключевой пропускной способности сети, определяющий асимптотическую информационную эффективность формирования ключа. Представленные строгое доказательство теоремы о ключевой пропускной способности и обоснование ее граничных значений определяют формирование окончательных условий, обеспечивающих решение теоретико-информационной задачи оценки эффективности открытого формирования сетевого ключа.

Ключевые слова — сеть связи, законные корреспонденты, нарушитель, сетевой ключ, широковещательный канал связи, канал перехвата, случайный кодер и детерминированный декодер, совместная информация, асимптотический метод формирования сетевого ключа, показатели качества и требования к сетевому ключу, скорость формирования сетевого ключа, теорема о ключевой пропускной способности сети связи минимального объема, оценка ключевой пропускной способности.

Цитирование: Синюк А. Д., Остроумов О. А. Теорема о ключевой пропускной способности сети связи. *Информационно-управляющие системы*, 2018, № 5, с. 79–87. doi:10.31799/1684-8853-2018-5-79-87

Citation: Sinjuk A. D., Ostroumov O. A. Theorem about key capacity of a communication network. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 5, pp. 79–87. doi:10.31799/1684-8853-2018-5-79-87

ПАМЯТКА ДЛЯ АВТОРОВ

Поступающие в редакцию статьи проходят обязательное рецензирование.

При наличии положительной рецензии статья рассматривается редакционной коллегией. Принятая в печать статья направляется автору для согласования редакторских правок. После согласования автор представляет в редакцию окончательный вариант текста статьи.

Процедуры согласования текста статьи могут осуществляться как непосредственно в редакции, так и по e-mail (ius.spb@gmail.com).

При отклонении статьи редакция представляет автору мотивированное заключение и рецензию, при необходимости доработать статью — рецензию. Рукописи не возвращаются.

Редакция журнала напоминает, что ответственность за достоверность и точность рекламных материалов несут рекламодатели.

UDC 621.396
doi:10.31799/1684-8853-2018-5-88-93

Articles

Efficiency evaluation of scheduling algorithms for delay-sensitive traffic in OFDM downlink

I. A. Pastushok[a]***,*** *PhD, Tech., Senior Lecture, orcid.org/0000-0002-3296-562X, i.pastushok@k36.org*
[a]*Saint-Petersburg State University of Aerospace Instrumentation, 67, B. Morskaia St., 190000, Saint-Petersburg, Russian Federation*

Introduction: *Common introduction of industrial control systems based on wireless networks leads to the emergence of a new way of using wireless networks, called the Industrial Internet of Things (IIoT). This framework assumes a large but finite number of users for which you have to provide the transmission of short messages with low delay and limited delay variation.* ***Purpose:*** *Efficiency evaluation of resource allocation algorithms for delay-sensitive traffic in a downlink OFDM channel.* ***Methods:*** *An OFDM system is presented as a set of* ***M/G/1*** *queuing systems. A convex optimization problem is introduced, whose solution determines the lower bound of the average delay of message transmission in delay-sensitive traffic over the set of scheduling algorithms under study.* ***Results:*** *An algorithm is proposed for calculating the lower bound of the average delay. This algorithm is fed by system parameters: the average message size, the input flow intensity, and the average maximum throughput for each device connected to the base station. Based on these input parameters, the optimization problem is introduced and solved by a numerical method. The solution of the optimization problem determines the lower bound of the average message transmission delay.* ***Practical relevance:*** *The obtained result can be used by system developers for planning and deploying OFDM networks. It can also serve as a reference when developing Ultra-Reliable and Low Latency Communication technology.*

Keywords *— industrial internet of things, OFDM, average delay, queuing systems, mathematical optimization, URLLC.*

Citation: Pastushok I. A. Efficiency evaluation of scheduling algorithms for delay-sensitive traffic in OFDM downlink. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 5, pp. 88–93. doi:10.31799/1684-8853-2018-5-88-93

Introduction

Wireless technologies are now widely introduced into industry, leading to the emergence of a new way of using wireless networks, called the Industrial Internet of Things (IIoT). This way assumes that there are multiple devices which establish a wireless connection to a remote server. The server uses a sequence of control messages transmitted over a downlink to remotely control and monitor a particular technological process or the enterprise as a whole [1–3].

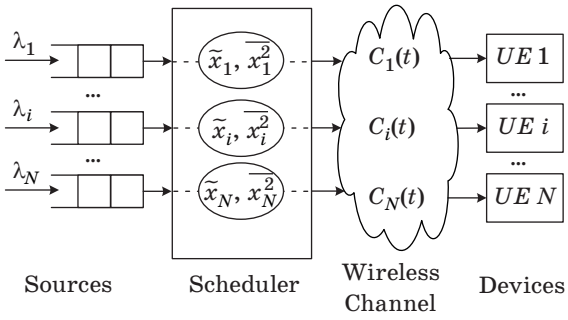
Within the IIoT concept framework, the International Telecommunication Committee [4] and 3GPP consortium [5] are standardizing the network scenario of Ultra-Reliable and Low Latency Communication (URLLC) designed to satisfy the key performance indicators for the industry. However, the 3GPP consortium has just started this standardization procedure.

It is worth noting that a number of studies are known, tasked with reviewing [6] or theoretical analysis [7] of URLLC scenario for 5G networks. In [6] and [7], the authors discuss possible usage of non-orthogonal multiple access technology to achieve key performance indicators of a wireless network required for the operation of an information system. Unlike these works, this article aims to obtain efficiency estimates for the existing orthogonal communication systems when a delay-sensitive traffic is transmitted via a downlink channel. This will give the engineers and network developers a tool for assessing the efficiency of using orthogonal networks when planning and deploying them.

The article is organized as follows. First, we give a description for a model of the system under study, along with a system of assumptions. We analyze the system as a queuing system and formulate an optimization problem whose solution will allow us to find the lower bound of an average message transmission delay in an OFDM system. Then we propose an algorithm for calculating the lower bound of an average delay of downlink message transmission in an OFDM system. Finally, we present some numerical results and conclusions.

System model

In this article, we discuss a development of the model proposed in [8], applied to delay-sensitive traffic transmission. The networks contains a remote server that coordinates the functioning of a finite number N of devices connected to the same base station. The functioning of each device i is coordinated by forming a sequence of messages (each message has a unique serial number j) as long as $V_{i,j}$, $j=\overline{1,\infty}$ bits, and the time intervals between the arrival of messages are distributed according to an exponential law with the parameter λ_i [9]. Via an absolutely reliable channel, the messages are transmitted to the base station (BS) where they get into



■ ***Fig. 1.*** System model

queues (each queue corresponds to a unique device) to be transmitted via a wireless channel (see Fig. 1).

General statements

We discuss the operation of a downlink in OFDM mode: all the operation time is split into slots of an equal duration; within a slot, a limited frequency zone can be distributed between several devices. Because of flat fadings in the wireless channel, its state changes in time for each device: in each time moment t, the channel state can be represented as a vector $C(t) = \left[C_i(t), i = \overline{1,N} \right]$, where $C_i(t)$ is the speed of transmitting data for the device i if it is given all the frequency zone for transmission at the time moment t.

The BS scheduler whose task is distributing the fractions of the wireless channel between the devices makes a decisive contribution to the system efficiency. The scheduler task solution can be described with a vector $\alpha(t) = \left[\alpha_i(t), i = \overline{1,N} \right]$, where $\alpha_i(t)$ is the frequency zone fraction given to the device i at the time moment t. An obvious limitation is that the available frequency band for data transmission is finite:

$$\forall t : \sum_{i=1}^{N} \alpha_i(t) \le 1. \qquad (1)$$

The main efficiency indicators for delay-sensitive traffic are [5, 4]:

— average delay of message transmission for the device i, see the expression (2);

— delay dispersion for message transmission via the wireless channel.

The delay dispersion can be large because the resources for data transmission are given irregularly. This is considered to be a negative effect. Therefore, the delay dispersion is limited from above by a certain value.

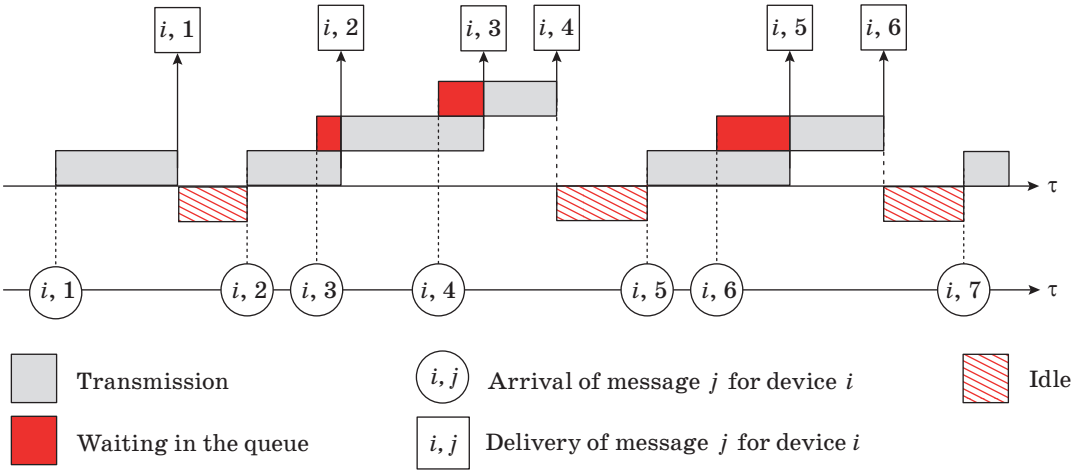
The average delay of message transmission is determined by the following expression:

$$\overline{T_i} = \lim_{\tau \to \infty} \frac{1}{H_i(\tau)} \sum_{j=1}^{H_i(\tau)} T_{i,j}, \qquad (2)$$

where $H_i(\tau)$ is the number of fully delivered messages for the device i during the time interval $[0, \tau]$, and $T_{i,j}$ is the time interval from the moment when a message j for the device i gets into a BS queue, to the moment when the last bit of the message is received by the device i. In its turn, the value $T_{i,j}$ consists of the duration of waiting for a message in the BS queue ($w_{i,j}$) and the duration of transmitting the message via the wireless channel ($x_{i,j}$):

$$T_{i,j} = w_{i,j} + x_{i,j}. \qquad (3)$$

Note that as a result of the scheduler's operation, during the transmission of a message j for the device i, the value of $x_{i,j}$, is determined which directly affects the waiting time in the queue for the subsequent messages. This process is demonstrated in Fig. 2.



■ ***Fig. 2.*** Graphical representation of delivering messages for device i

The generalized result of the scheduler's operation for each user i are pairs of values $\widetilde{x_i}$ (average delay in message transmission) and $\overline{x_i^2}$ (coefficient of variation in message transmission).

Assumptions

In this article, we use the following assumptions:

— $\forall i: V_{i,j}, j=\overline{1,\infty}$ is an ergodic random process with finite mathematical expectation $\widetilde{V}_i$.

— $\forall i: x_{i,j}, j=\overline{1,\infty}$ is a sequence of identically distributed random values independent from each other and from the input flow; their mathematical expectation is $\widetilde{x_i}$ and coefficient of variation is $\overline{x_i^2}$.

— The wireless channel state changes in such a way that the following equality holds:

$$\int_{t_{i,j}}^{t_{i,j}+x_{i,j}} \alpha_i(t)C_{i,j}(t)dt = C_{i,j}\int_{t_{i,j}}^{t_{i,j}+x_{i,j}} \alpha_i(t)dt = V_{i,j}, \quad (4)$$

where $t_{i,j}$ is the time when a message j starts to be transmitted to the device i.

— $\forall i: C_{i,j}^{-1}, j=\overline{1,\infty}$ is an ergodic random process with finite mathematical expectation $\widetilde{C}_i^{-1}$.

The presented description of the system and the assumptions introduced above describe an infinite and uncountable set of hypothetical planning algorithms $\mathcal{A}$ with possible analytical estimates.

Analysis of the system

In this section, we analyze the system under study as a queuing system, and formulate an optimization problem whose solution will characterize the efficiency of OFDM systems in delay-sensitive message transmission.

The system model as a queuing system

With the assumptions made above, the system can be represented as a combination of N queues **M/G/1**, where each queue corresponds to a unique device. An average delay of message transmission for device i (see expressions (2) and (3)) can be found by the Pollaczek — Khinchin formula [10]:

$$\overline{T_i} = \widetilde{x_i} + \frac{\lambda_i \overline{x_i^2}}{2(1-\lambda_i \widetilde{x_i})}. \quad (5)$$

Thus we obtained an expression for an average delay of message transmission for device i ($\overline{T_i}$) as a function of an average duration of message transmission via the wireless channel for the user i ($\widetilde{x_i}$).

An important parameter of the system is the average idle time for the **M/G/1** queue per message, which is determined by the following expression [11]:

$$\overline{I_i} = \frac{1}{\lambda_i}. \quad (6)$$

Formulation of the optimization problem

One of the central results of [8] is the interrelation between the characteristics of a wireless centralized communication network and the quality of user service when transmitting video data by the HTTP protocol. With reasoning similar to that in [8], we can find a real interrelation in terms of the system under study.

Lemma. For a set of scheduling algorithms $\mathcal{A}$ satisfying the introduced assumptions, the following inequality is true:

$$\sum_{i=1}^{N} \frac{\widetilde{V}_i \widetilde{C}_i^{-1}}{\widetilde{x_i} + \overline{I_i}} \le 1. \quad (7)$$

The proof of this Lemma is similar to the proof given in [8]. Therefore, it is not given in this article.

Let us define the efficiency index of an OFDM system transmitting delay-sensitive traffic as an average value of message transmission delay over the entire set of devices with a given coefficient of variation for each particular device: $T(A) = \sum_{i=1}^{N} \overline{T_i}(A)$.

It is important to note that the value of the message transmission delay is determined by a specific scheduling algorithm A installed on the BS.

Let us introduce the problem of estimating the efficiency of an OFDM system as a problem of finding the lower bound of an average message transmission delay over the set of scheduling algorithms $\mathcal{A}$:

$$\widetilde{T} = \inf_{A\in\mathcal{A}} T(A). \quad (8)$$

The value of the lower bound of an average message transmission delay can be obtained as a solution to a nonlinear optimization problem (9).

$$\text{Minimize: } T^* = \frac{1}{N}\sum_{i=1}^{N}\left(\widetilde{x_i} + \frac{\lambda_i \overline{x_i^2}}{2(1-\lambda_i \widetilde{x_i})}\right)$$

$$\text{Subject to:} \quad , \quad (9)$$

$$\begin{cases} \sum_{i=1}^{N} \dfrac{\widetilde{V}_i \widetilde{C}_i^{-1}}{\widetilde{x_i} + (\lambda_i)^{-1}} - 1 \le 0 \\ \widetilde{x_i} \in \left[0,\ (\lambda_i)^{-1}\right],\ i=\overline{1,N} \end{cases}$$

where $\widetilde{x_i}$, $i=\overline{1,N}$ are the optimized parameters, λ_i and $\widetilde{V}_i$ are characteristics of the input data flow

for device i, $\widetilde{C}_i^{-1}$ is a characteristic of the state of the wireless channel for the user i, and $\overline{x_i^2}$ is the coefficient of variation specified for the duration of message transmission via the wireless channel for the device i.

Lower bound of average message transmission delay in an OFDM system

The optimization problem (9) belongs to the class of convex nonlinear programming problems. However, due to the sophisticated structure of the objective function, it is difficult to obtain a closed expression for the lower bound value. Therefore, in this work, we will use numerical methods to solve the optimization problems. In convex optimization problems, they allow you to find the global extremum point of an objective function with limitations [12, 13].

However, the use of a numerical procedure for finding an extremum can be computationally expensive. Therefore, below we analyze the cases when the lower bound value can be obtained without a numerical procedure. Here is a consecutive analysis of the optimization problem (9).

Consider the case when the inequality $\sum_{i=1}^{N}\widetilde{V}_i\widetilde{C}_i^{-1}\lambda_i - 1 \le 0$ holds. Then an apparent solution of the optimization problem (9) is $\widetilde{x_i} = 0,\ i = \overline{1,\ N}$ and $T^* = 0$.

Next, let us use the solution method based on the dual Lagrange function [12]. To this end, we introduce an auxiliary function f^*:

$$f^*\left(\widetilde{x}\right) = \begin{cases} \dfrac{1}{N}\sum_{i=1}^{N}\left(\widetilde{x_i} + \dfrac{\lambda_i\overline{x_i^2}}{2(1-\lambda_i\widetilde{x_i})}\right), \ \widetilde{x_i} \in \widetilde{X} \\ \infty, \text{ overwise} \end{cases}, \quad (10)$$

where $\widetilde{X}$ is a set of admissible values of the parameter $\widetilde{x_i} \in \left[0,\ (\lambda_i)^{-1}\right],\ i = \overline{1,\ N}$.

Using the expression (10), we can represent the optimization problem as follows:

$$\begin{aligned} &\text{Minimize: } f^*\left(\widetilde{x}\right) \\ &\text{Subject to:} \\ &\begin{cases} \sum_{i=1}^{N}\dfrac{K_i}{\widetilde{x_i} + (\lambda_i)^{-1}} - 1 \le 0 \end{cases} \end{aligned}, \quad (11)$$

where $K_i = \widetilde{V}_i\widetilde{C}_i^{-1},\ i = \overline{1,\ N}$.

The Lagrange function for the optimization problem (11) takes the following form:

$$L\left(\widetilde{x},\ \mu\right) = \frac{1}{N}\sum_{i=1}^{N}\left(\widetilde{x_i} + \frac{\lambda_i\overline{x_i^2}}{2(1-\lambda_i\widetilde{x_i})}\right) + \mu\left[\sum_{i=1}^{N}\frac{K_i}{\widetilde{x_i} + (\lambda_i)^{-1}} - 1\right]. \quad (12)$$

The dual Lagrange function for the expression (12) has the following form:

$$g(\mu) = -\mu + \inf_{x\in\widetilde{X}}\frac{1}{N}\sum_{i=1}^{N}\left(\widetilde{x_i} + \frac{\lambda_i\overline{x_i^2}}{2(1-\lambda_i\widetilde{x_i})} + \frac{\mu K_i N}{\widetilde{x_i} + (\lambda_i)^{-1}}\right) = -\mu + \sum_{i=1}^{N}\frac{\mu K_i\lambda_i}{2} = \mu\left(\sum_{i=1}^{N}\frac{K_i\lambda_i}{2} - 1\right). \quad (13)$$

On the base of expression (13), the problem of finding the lower bound can be represented as a solution of the optimization (8) problem (14) [12].

$$\begin{aligned} &\text{Maximize: } \mu\left(\sum_{i=1}^{N}\frac{K_i\lambda_i}{2} - 1\right) \\ &\text{Subject to:} \\ &\{\mu \ge 0 \end{aligned}. \quad (14)$$

The solution of the optimization problem (14) is the following expression.

$$\widetilde{x^*} = \begin{cases} 0, \sum_{i=1}^{N}\dfrac{K_i\lambda_i}{2} \le 1 \\ \infty, \text{ overwise} \end{cases}. \quad (15)$$

The expression (15) determines the bound of the non-existence of scheduling algorithms which satisfy the introduced assumptions.

Based on the reasoning presented in this section, we introduce the main theorem.

Theorem. The lower bound value for an average message transmission delay in an ODFM system over all possible scheduling algorithms which satisfy the assumptions can be calculated as follows:

$$\widetilde{T} = \begin{cases} 0, \ \sum_{i=1}^{N}K_i\lambda_i \le 1 \\ \infty, \ \sum_{i=1}^{N}\dfrac{K_i\lambda_i}{2} \ge 1 \\ \chi\left(K_i,\ \lambda_i,\ \overline{x_i^2},\ i = \overline{1,\ N}\right), \text{ overwise} \end{cases},$$

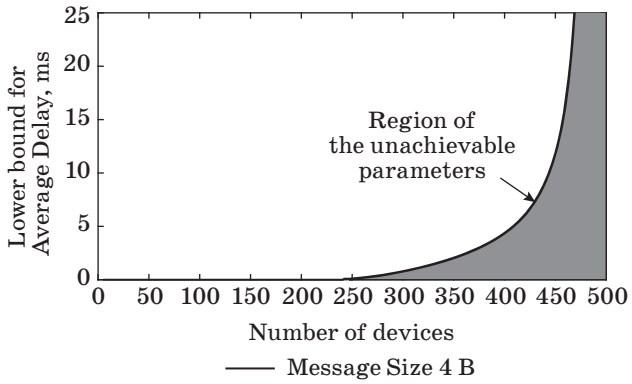
where $\chi(\cdot)$ is a solution of the optimization problem by a numerical method with the given accuracy.

Numerical example

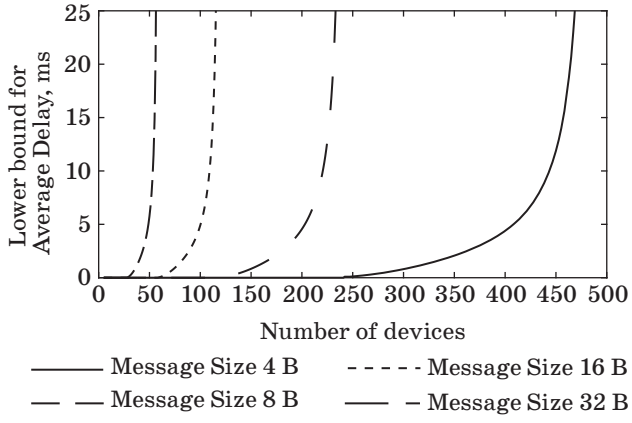
The obtained results are demonstrated on the base of LTE system parameters. We consider the functioning of one cell in which users are placed with an equal step on their way from the base station. The maximum achievable user's data transmission rate is in the area from 0.5 MB/s to 5 MB/s, which corresponds to 5 MHz bandwidth [14].

Fig. 3 demonstrates how the lower bound value depends on the number of users in a cell when the size of the transmitted messages is 4 B and the arrival rate is 250 per second for each user in the cell. At the qualitative level, the following statement is true: for given network parameters, there is no scheduling algorithm which provides an average message transmission delay of 5ms when there are 450 users in a cell.

Fig. 4 demonstrates the values of the obtained lower boundaries with fixed input flow intensities for various message sizes. From the calculation for the network scenario described above it follows that with a fixed value of the average message transmission delay, the maximum number of subscribers with a possible scheduling algorithm from the set $\mathcal{A}$ decreases linearly with the message size.



■ ***Fig. 3.*** Dependence of the lower bound value on the number of devices when $V_i = 4$ B



■ ***Fig. 4.*** Dependence of the lower bound values on the number of devices for various message sizes

Conclusion

An algorithm is proposed for calculating the lower bound value for an average message transmission delay in OFDM systems. The obtained result allows Internet communication system developers to plan OFDM systems deployed in industry. It can also serve as a reference for communication systems that provide URLLC.

The author is supported by scientific project No 8.8540.2017 «Development of data transmission algorithms in IoT systems with limitations on the devices complexity».

References

1. **Wan J., Tang S., Shu Z., Li D., Wang S., Imran M., Vasilakos A. V.** Software-defined industrial internet of things in the context of industry 4.0. *IEEE Sensors Journal*, 2016, no 20, pp. 7373–7380. doi:10.1109/JSEN.2016.2565621
2. **Shrouf F., Ordieres J., Miragliotta G.** Smart factories in Industry 4.0: A review of the concept and of energy management approached in production based on the Internet of Things paradigm. *2014 IEEE International Conference on Industrial Engineering and Engineering Management*, 2014, pp. 697–701. doi:10.1109/IEEM.2014.7058728
3. **Chen B., Wan J., Shu L., Li P., Mukherjee M., Yin B.** Smart factory of industry 4.0: key technologies, application case, and challenges. *IEEE Access*, 2018, vol. 6, pp. 6505–6519. doi:10.1109/ACCESS.2017.2783682
4. ITU-R M.2083-0. IMT Vision — framework and overall objectives of the future development of IMT for 2020 and beyond. *ITU-R*, 2015. 21 p.
5. 3GPP TS 22.261 16.4.0. Service requirements for next generation new services and markets. *3GPP*, 2017. 55 p.
6. **Bennis M., Debbah M., Poor V.** *Ultra-reliable and low-latency wireless communication: tail, risk and scale*. 2018. Available at: https://arxiv.org/pdf/1801.01270.pdf (accessed 12 September 2018).
7. **Popovski P., Trillingsgaard K., Simeone O., Durisi G.** *5G wireless network slicing for eMBB, URLLC, and mMTC: A communication-theoretic view*. 2018. Available at: https://arxiv.org/pdf/1804.05057.pdf (accessed 12 September 2018).
8. **Pastushok I., Bakin E.** On interrelation of video streaming characteristics in centralized wireless networks. *IEEE wireless communications letters*, 2018,

vol. 7, no. 2, pp. 226–229. doi:10.1109/LWC.2017.2767597
9. 3GPP TR 36.872 V12.1.0. Small cell enhancements for E-UTRA and E-UTRAN — Physical layer aspects. *3GPP*, 2013. 100 p.
10. **Khinchin A.** Matematicheskaia teoriya stacionarnoi ocheredi. *Matematicheskii sbornik*, 1932, vol. 39, no. 4, pp. 73–84 (In Russian).
11. **Kleinrock L.** *Theory. Volume 1. Queueing systems.* Wiley-Interscience, 1975. 417 p.
12. **Boyd S., Vandenberghe L.** *Convex optimization*. Cambridge University Press, 2004. 716 p.
13. **Byrd R. H., Gilbert J. C., Nocedal J.** A trust region method based on interior point techniques for nonlinear programming. *Mathematical programming*. 2000, vol 89, no 1, pp. 149–185. doi:10.1007/s101070000189
14. **Sesia S., Toufik I., Baker M.** *Lte, the UMTS long term evolution: from theory to practice*. Wiley Publishing, 2009. 752 p.

УДК 621.396
doi:10.31799/1684-8853-2018-5-88-93

Оценка эффективности алгоритмов планирования для передачи чувствительного к задержке трафика по нисходящему каналу в режиме OFDM

И. А. Пастушок[a], канд. техн. наук, старший преподаватель, orcid.org/0000-0002-3296-562X, i.pastushok@k36.org
[a]Санкт-Петербургский государственный университет аэрокосмического приборостроения, Б. Морская ул., 67, Санкт-Петербург, 190000, РФ

Введение: повсеместное внедрение систем управления, основанных на беспроводной связи, в промышленных объектах формирует специфичный сценарий использования беспроводных систем, который получил название Индустриальный интернет вещей. В рамках данного сценария предполагается наличие большого, но конечного числа пользователей, для которых должна быть обеспечена передача коротких сообщений с низкой задержкой и ограничением на дисперсию задержки. **Цель исследования:** оценка эффективности алгоритмов планирования распределения частотно-временных ресурсов для систем связи с ортогональным частотным мультиплексированием каналов (OFDM) при передаче чувствительного к задержке трафика в нисходящем канале связи. **Методы:** система OFDM представлена в виде множества систем массового обслуживания M/G/1, произведена постановка оптимизационной задачи, решение которой определяет нижнюю границу средней задержки передачи сообщений чувствительного к задержке трафика по множеству исследуемых алгоритмов планирования. **Результаты:** предложен алгоритм вычисления значения нижней границы средней задержки передачи сообщений чувствительного к задержкам трафика в системе OFDM. Предложенный алгоритм принимает на вход параметры системы: средний размер сообщения, интенсивность входного потока и среднее значение максимально достижимой скорости передачи данных для каждого устройства, подключенного в базовой станции. На основе введенных параметров системы производится постановка оптимизационной задачи и ее решение численным алгоритмом. Решение оптимизационной задачи определяет значение нижней границы средней задержки передачи сообщений по множеству исследуемых алгоритмов планирования в нисходящем канале связи систем OFDM. **Практическая значимость:** полученный результат может быть использован разработчиками систем при планировании и развертывании OFDM сетей, а также является опорным результатом для разрабатываемой технологии сверхнадежной связи с низкими задержками (URLLC).

Ключевые слова — индустриальный интернет вещей, OFDM, средняя задержка, теория массового обслуживания, математическая оптимизация, URLLC.

Цитирование: Pastushok I. A. Efficiency evaluation of scheduling algorithms for delay-sensitive traffic in OFDM downlink. *Информационно-управляющие системы*, 2018, № 5, с. 88–93. doi:10.31799/1684-8853-2018-5-88-93
Citation: Pastushok I. A. Efficiency evaluation of scheduling algorithms for delay-sensitive traffic in OFDM downlink. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 5, pp. 88–93. doi:10.31799/1684-8853-2018-5-88-93

UDC 621.371
doi:10.31799/1684-8853-2018-5-94-103

Surveys

Evolution of multiple-access networks — cellular and non-cellular — in historical perspective. Part 2

A. M. Sergeev[a], *Senior Lecturer, orcid.org/0000-0002-4788-9869*
N. Sh. Blaunstein[b, c], *Dr. Sc., Phys.-Math., Professor, nathan.blaunstein@hotmail.com*
[a]*Saint-Petersburg State University of Aerospace Instrumentation, 67, B. Morskaia St., 190000, Saint-Petersburg, Russian Federation*
[b]*Ben-Gurion University of the Negev, POB 653, 1, Ben-Gurion St., Beer-Sheva, 74105, Israel*
[c]*Jerusalem College of Technology — Lev Academic Center, 21 Havaad Haleumi, POB 16031, Jerusalem, 91160, Israel*

Introduction: *The goal of this issue is the analysis of evolution of the current and novel wireless networks, from second generation (2D) to fifth generation (5G), as well as changes in technologies and their corresponding theoretical background and protocols — from Bluetooth, WLAN, WiFi and WiMAX to LTE, OFDM/OFDMA, MIMO and LTE/MIMO advanced technologies with new hierarchy of cellular maps design — femto/pico/micro/macro.* ***Methods:*** *We use new theoretical frameworks for description of the advanced technologies, such as multicarrier diversity technique, OFDM and OFDM novel approach, MIMO aspects description based on multi-beam antennas approach, various cellular maps design based on a new algorithms of femto/pico/micro/macrocell deployment, and a new methodology of a new MIMO/LTE system integration based on multi-beam antennas.* ***Results:*** *We have created a new methodology of multi-carrier diversity description for novel multiple-access networks, of usage of OFDM/OFDMA modulation to obey inter-user and inter-symbol interference in multiple-access networks, of how to obey the multiplicative noises occurring in the multiple-access wireless networks, caused by multi-ray phenomena, and finally, of how to overcome propagation effects occurring in the terrestrial communication links by use combination of MIMO and LTE technologies based on multi-beam antennas. For these purposes we present new stochastic approach that accounts for the terrain features, such as buildings' overlay profile, buildings' density around the base station and each user antennas, and so forth. These parameters allow us to estimate for each situation occurs at the built-up terrain area the effects of fading, as a source of multiplicative noise.* ***Practical relevance:*** *New methodology of how to estimate effects of multiplicative noise, inter-user and inter-symbol interference, occurring in the terrestrial wireless networks, allows us to predict a-priory practical aspects of the current and new multiple-access wireless communication networks, such as: the users' capacity and user's links spectral efficiency for various configurations of cells deployment — femto, pico, micro, and macro, as well as the novel MIMO/LTE system configuration for future networks of 4^{th} and 5^{th} generation deployment.*

Keywords *— capacity, multiple-input-multiple-output (MIMO), MIMO channel, spectral efficiency, K-factor, space-time diversity, spatial multiplexing urban environment, dense layout of buildings.*

Citation: Sergeev A. M., Blaunstein N. Sh. Evolution of multiple-access networks — cellular and non-cellular — in historical perspective. Part 2. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 5, pp. 94–103. doi:10.31799/1684-8853-2018-5-94-103

Continuation.
Start in *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 4, pp. 86–104.
doi:10.31799/1684-8853-2018-4-86-104

MIMO modern networks design in the space and time domains

We will now present some advanced technology concepts based on adaptive multi-beam or phased-array antenna applications through the prism of the physical layer description accounting for the "reaction" of the multipath outdoor channel with fading on radio propagation within such a channel [22, 49–62]. These techniques are fully described in references [63–82].

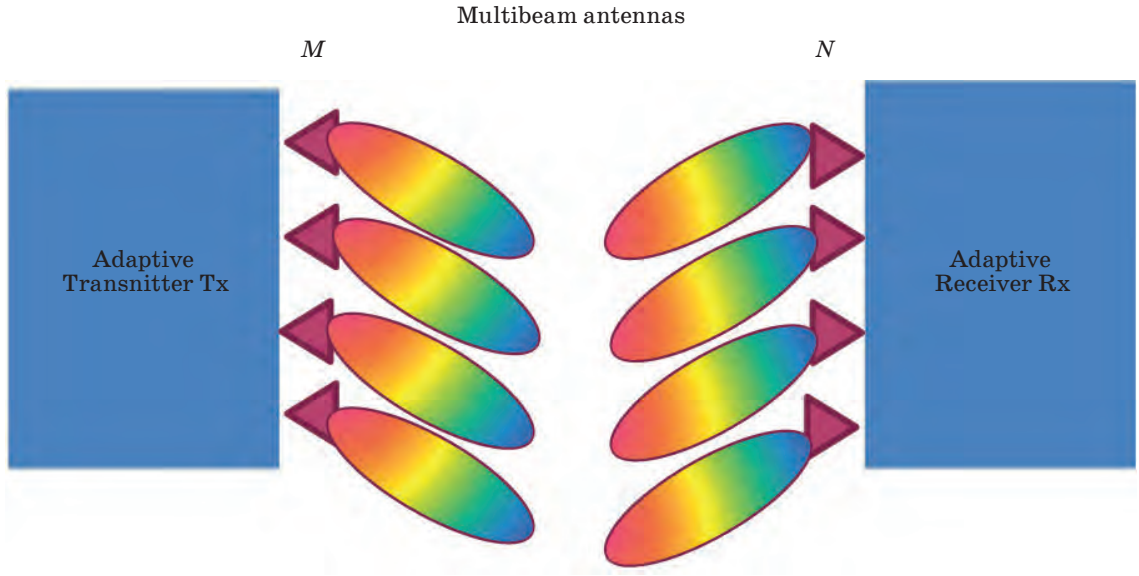
Multiple-input-multiple-output (MIMO) communication systems with multi-beam or multiple-element antennas arranged at both ends of the communication link have been introduced during the last decade to increase of spectral efficiency and communication link reliability that can be achieved via spatial and time diversity techniques. In Fig. 11 is shown an example of how to arrange multibeam antennas (4×4 beams, i. e., 16 beams) that manage and control many clients servicing via the mobile/stationary broadband internet or sensors' networking.

In modern MIMO systems, the two basic techniques usually used [63–82] to mitigate multipath fading phenomena and increase efficiency of such networks, are:

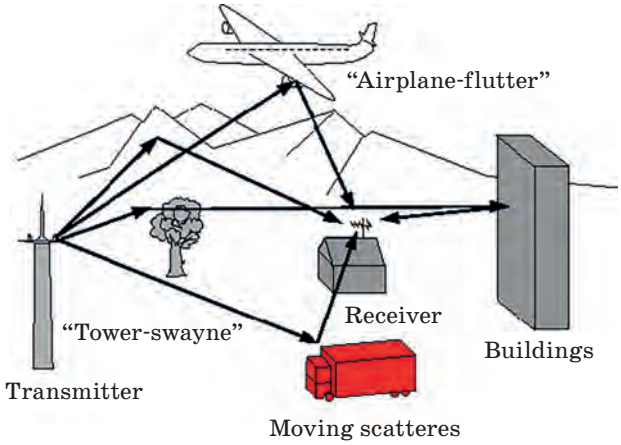
— spatial multiplexing as a space-time modulation techniques;

— diversity modulation technique as a special case of a space-time modulation technique.

According to the first procedure, each transmitting antenna element sends to the receiver *independent* (e. g. non-correlated) streams of signal data accounting for a strong multipath phenomenon occurring in each channel with the Rayleigh fading [63, 64] caused by multiple reflection and diffrac-



■ ***Fig. 11.*** MIMO system based on $M \times N$ multibeam antennas



■ ***Fig. 12.*** Schematically presented effects of multiray phenomena occurring in the real wireless MIMO channel that causes strong fading on the input/output signals

tion occurring in the real wireless environment (Fig. 12). This idea of spatial multiplexing was first proposed in Reference [47], which then was adapted in practice of MIMO systems deployment in [66, 67]. Initially spatial multiplexing systems used narrowband channel for each antenna element of the MIMO system with a small delay spread (i. e., the large bandwidth of coherency B_{cn}, see previous section). In modern MIMO systems spatial multiplexing was adapted for wideband channels too in conjunction with OFDM modulation technique [68–72].

In contrast to spatial multiplexing, the diversity modulation technique deals with *dependent* (e. g. correlated) streams of data from each transmit antenna element of desired MIMO system. There are several diversity modulation techniques, which can be found in references [73–77]. In this book, will be using the term MIMO diversity technique that combines space-time diversity and spatial-time multiplexing according to references [78–82].

Following the results obtained in [22], we will describe the MIMO system capacity, as an example of space-diversity and time-diversity techniques adapted for the use of the multibeam (e. g., multielement) antennas in various built-up environments. For this case, the desired formulas according to the unified multiparametric stochastic approach described in [21, 22, 95–102], will be rearranged to obtain simple relations between the MIMO antenna element number and the parameters of the terrain.

The spectral efficiency of the MIMO technique strongly depends on the diversity among multiple channels which is determined by the spatial statistical behavior of the MIMO fading channel, partly characterized by the special spatial fading correlation function (or coefficients) described in [22]. It was shown in [22, 49–55] that a sufficiently high diversity in the received multipath replicas of the transmitted signal can be achieved within a "rich" scattering environment where the communication channel capacity linearly increases with the number of the transmitter and the receiver antennas. Moreover, [49] was shown that the spatial fading correlation coefficient for *dependent* data streams (or de-correlation coefficient for *independent* data streams) can be determined by the system parameters, such as antenna elements spacing and their number. At the same time, this coefficient can be determined by the propagation conditions, such as the received energy spread in the AOA (angle-of-arrival), TOA (time-of-arrival) and Doppler domains, as it is described in [21, 22].

In practice, scattering environment is a scenario-dependent and as a result, spatial de-correlation

characteristics are also scenario dependent [22, 49–52, 55]. Therefore, an accurate modeling of the spatial de-correlation characteristics of the MIMO channels is crucial for investigation of the scenario-specific spectral efficiency, which serves as the main criterion for communication systems design and wireless networks planning. Thus, we will derive the MIMO channel capacity for various propagation conditions as a function of the spatial correlation and the received distribution in the AOA-TOA domain. The proposed spatial fading correlation will be introduced via the stochastic multiparametric model of the urban propagation conditions in a joint AOA-TOA domain described in [21, 22, 50–52]. As a result of the proposed stochastic approach, it can be shown that the spatial fading correlation parameters depend on the propagation phenomena, such as multiple scattering, reflection, diffraction, as well as on the waveguide propagation along streets, which characterize urban environment propagation conditions.

Modeling of MIMO channel capacity

Usually, as follows from Fig. 11, there are several output antennas and input antennas assembled at the transmitter and the receiver, which we will denote them by M and N respectively.

For the uncorrelated antennas (e. g. working as separate independent antenna elements) arranged at the MIMO channel, the spectrally normalized to bandwidth B_w [in Hz] capacity C [e. g., *spectral efficiency* measured in bit/s/Hz] was defined in [50] as

$$\tilde{C}_{\text{uncorr}} = N\log_2\left(\left(K_m\cdot\left(\frac{P_m}{N}\right)_{\text{add}}\right)\Big/\left(K_m+\left(\frac{P_m}{N}\right)_{\text{add}}\right)\right), \quad (26a)$$

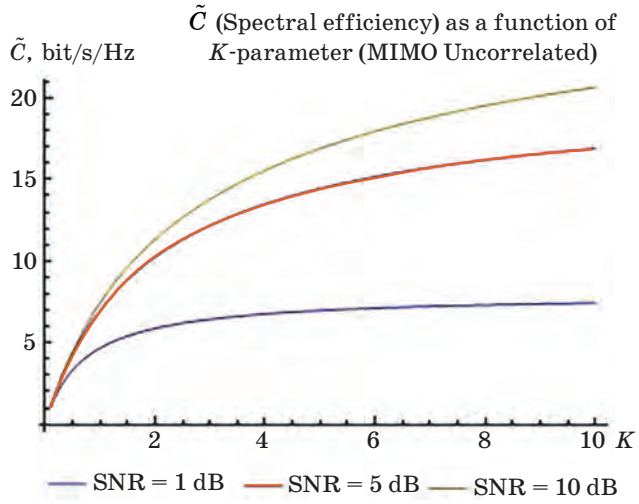
where $(P_m/N)_{\text{add}}$ is the signal to additive Gaussian noise ratio, which usually is taken into account in the literature. We also accounted the multiplicative noise caused by fading multipath phenomena, occurring in each of N input channels, where Km = (LOS-component)/(Multipath-component) — is the ratio of the coherent (e. g., deterministic) component of the signal and incoherent (stochastic) component of the signal caused the multiplicative noise.

At the same time, for the case of correlated antennas (i. e., working as unified whole transmitter and receiver antenna) the spectral efficiency [measured in bit/s/Hz] can be presented according to [50] as

$$\tilde{C}_{\text{corr}} = \log_2\left(1+MN\left(K_m\times\left(\frac{P_m}{N}\right)_{\text{add}}\right)\Big/\left(K_m+\left(\frac{P_m}{N}\right)_{\text{add}}\right)\right). \quad (26b)$$

Thus, mentioned above can be proved by a special numerical simulations carried out for different values of SNR = $(P_m/N)_{\text{add}}$ [in dB], and for various amounts of elements M and N at the transmitter (output) and receiver (input) antennas, respectively. We should notice before presenting some results of numerical computations that the uncorrelated arrangement of the MIMO antenna elements at the both terminal sides allows to obtain much higher capacity, e. g., higher spectral efficiency of each M and N channels with respect to the case of the correlated arrangement of the antenna elements into the unique antenna. This can be clearly seen from formulas (26a) and (26b), where in (26a) the number of elements at the receiver N is outside the logarithmic function and C increases linearly with increase of elements N at the input of the receiver. At the same time, in (26b) N and M are inside the logarithm, that is, capacity and the spectral efficiency increases logarithmically (i. e., very slow) with increase of number of elements N and M. In [50, 51] are presented several variants of MIMO system arrangements to proof mentioned above. Therefore, we will present here only the case of the uncorrelated arrangement of the antenna elements, M and N. Thus, Fig. 13 presents a spectral efficiency vs. the fading factor K, as a ratio of the coherent and incoherent (multipath) components of the signal at the input of the multi-element ($M = 2$) transmitter antenna and multi-element ($N = 4$) receiving antenna, according to scenario shown in Fig. 12.

As clearly seen from the presented illustration, with increase of the coherent component (e. g. line-of-sight component) of the income signals at the input of the multi-element receiver antenna with respect to the multipath component (caused by multi-diffraction and multi-scattering from obstruc-



■ ***Fig. 13.*** Spectral efficiency vs. the K-parameter of fading for various values of signal to additive noise ratio for $M = 2$ and $N = 4$

tions located in area of users' service), that is, with increase of K-factor of fading, the spectral efficiency increases sharply till K~4÷6, and then a saturation of the process becomes evident.

This effect depends on SNR and becomes for smaller (K~2) with increase of SNR from 1 to 10 dB. In other words, increasing SNR inside the MIMO system, it can be easier to obey effect of fading caused by real conditions of each land-to-land communication channel.

Considering now, according to [51], the uplink scenario where the multiple receiving antennas in the BS are spatially separated, the correlation among the received replicas of the transmitted signal is determined by the propagation conditions and the spatial separation distance, that is,

$$\rho = \int_0^{2\pi} e^{ikr\sin(\varphi)} f(\varphi)\mathrm{d}\varphi, \qquad (27)$$

where $k = 2\pi/\lambda$ is a wavenumber, λ is a wavelength; r is a spatial spacing between the transmitter or the receiver antennas, which in practice is limited by the physical dimensions of the platform or installation constraints; φ is AOA; $f(\varphi)$ is an angular spectrum. Conventionally it is assumed that $M > N$.

The influence of the MIMO channel correlation, represented by the transmitter and the receiver spatial correlation matrices on the channel capacity is intensively studied in the literature (see [22, 49–51, 53]). Following [50, 51], we will represent the influence of the spatial fading correlation in (27) on the MIMO channel capacity in (26) in the more convenient and simple manner to understand the matter:

$$\tilde{C} = N\log_2\left(1 + (1-\rho)\frac{M}{N}\frac{SNR}{B_w}\right). \qquad (28)$$

Notice that the spatial fading correlation in (28) decreases the MIMO channel capacity. Moreover, as was shown in addition, note that the received energy spread in the AOA-TOA domain is inversely proportional to the correlation and therefore directly proportional to the achievable channel capacity. Thus, the higher received signal spread is represented by a smaller correlation, $\rho \to 0$, which results in a higher MIMO channel capacity according to equation (28).

Fading correlation in space-time doman in urban environment with complicated building layout

considering the below-the-rooftops propagation conditions [51], and using the corresponding formulas presented in [22], the fading correlation in (26) can be rewritten as follows:

$$\rho_{\text{below}} = \int_0^{2\pi} f_{\text{below}}(\varphi) e^{jkd\sin\varphi}\mathrm{d}\varphi, \qquad (29)$$

where

$$f_{\text{below}}(\varphi) = \int_{\tau=0}^{t}\left(\frac{\bar{L}\nu^2\tau d^2(\tau^2-1)}{2\pi(\tau-\cos\varphi)}\beta(\varphi)e^{-\frac{2\bar{L}\nu\tau d}{\pi}}\right)\times$$
$$\times e^{\left(-\frac{d(\tau^2-1)|\ln\chi|}{(\tau-\cos\varphi)a'(\varphi)}\right)}\mathrm{d}\tau, \qquad (30)$$

where d is the straight-line distance between the MS (mobile subscriber) and the BS, km; τ is the ratio between the actual distance that signal travels from the MS to the BS and d; φ is the AOA; ν is the building density per square km [i. e., in km^{-2}];

$$\beta(\varphi) = \sin^2\left(\frac{1}{2}\arcsin\left(\frac{d\sin(\varphi)}{\tilde{r}}\right)\right),$$

$\chi = \bar{L}/(\bar{L}+\bar{l})$ is the street "discontinuity" parameter, where $\bar{L}$ is the average lengths of the buildings, km; $\bar{l}$ is the average lengths of the slits (gaps) between the buildings, km; $a' = \sqrt{\frac{4a^4}{\lambda^2 n^2} + a^2}$, where a is the average width of streets, km, and n is the street waveguide mode number.

Considering wave propagation above the rooftops, and using formulas derived in [22], the correlation coefficient in (26) can be obtained as follows [51]:

$$\rho_{\text{above}} = \int_0^{2\pi} f_{\text{above}}(\varphi) e^{jkd\sin\varphi}\mathrm{d}\varphi, \qquad (31)$$

where

$$f_{\text{above}}(\varphi) = \int_{\tau=0}^{t}\left[\frac{\bar{L}\nu^2\tau d^2(\tau^2-1)}{2\pi(\tau-\cos\varphi)}\times\right.$$
$$\times\beta(\varphi)\frac{\bar{h}}{h_R}e^{-\frac{2\bar{L}\nu}{\pi}\left(\tilde{r}(\tau,\varphi)+\frac{\bar{h}}{h_R}\frac{d(\tau^2-1)}{2(\tau-\cos\varphi)}\right)}+$$
$$\left.+\frac{\nu}{2}\beta(\varphi)\frac{(h_R-\bar{h})\tilde{r}(\tau,\varphi)}{\bar{h}}e^{-\frac{2\bar{L}\nu\tilde{r}(\tau,\varphi)}{\pi}}\right]\times$$
$$\times e^{\left(-\frac{d(\tau^2-1)|\ln\chi|}{(\tau-\cos\varphi)a'(\varphi)}\right)}\mathrm{d}\tau, \qquad (32)$$

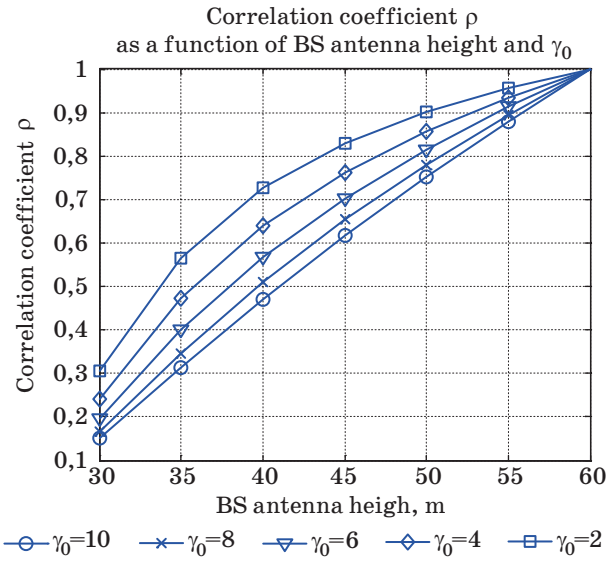
here, assuming a uniform distribution of building heights; $\bar{h} = \frac{h_1+h_2}{2}$ is the average height of building profiles, where h_1 and h_2 are the minimum

and maximum building heights; h_R is the BS antenna height, and $\tilde{r}(\tau, \varphi) = \frac{d(\tau^2 - 2\tau\cos\varphi + 1)}{2(\tau - \cos\varphi)}$. Below, we will analyze MIMO channel capacity in specific urban scenarios.

Correlation coefficient analysis in urban scene

The spatial fading correlation is analyzed in various urban propagation conditions simulated by the proposed f_{below} and f_{above} models according to (30) and (32), respectively. Following [51], we analyse an urban environment with the following parameters of the experiment described there: the two BS receiving antennas with a separation distance of $\frac{r}{\lambda} = 10$ located in the urban scene with the following parameters: $\gamma_0 = 8$ km^{-1}, $\chi = 0.5$, $d = 0.3$ km, pseudo-LOS is at $\varphi = 0°$, $\bar{h} = \frac{h_1 + h_2}{2} = 15$ m, $\nu = 250$ km^{-2}. Figure 14 shows the correlation coefficient ρ from (32) as a function of the BS antenna height h_R and the buildings' density parameter, $\gamma_0 = \frac{2\bar{L}\nu}{\pi}$, which describes the clutter density in urban scenario.

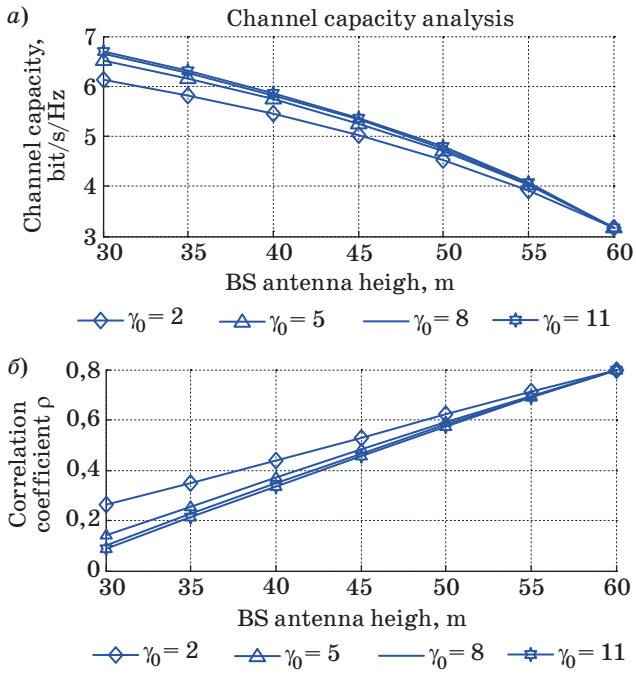
As follows from results presented in Fig. 14, the spatial fading correlation is directly proportional to the BS antenna height. We can outline that the results shown in Fig. 14 agree with those obtained in [52, 55]. In addition, Fig. 14 shows that the spatial fading correlation is inversely proportional to the buildings' density parameter, γ_0 (e. g., *clutter* density), which determines the received signal diversity in the AOA-TOA domain. Therefore, the denser urban environment (higher γ_0) results in lower correlation among the received replica of the transmitted signal. Finally, we notice that the influence of the parameter γ_0 on the spatial fading correlation decreases with increasing BS antenna height. This observation can be explained by the fact that the significance of the built-up environment structure for a particular urban scene (distribution of the scatterers) decreases with increasing BS antenna height. In practice, the result shown in Fig. 14 can be used as a guideline for the BS antenna height selection required to achieve the predefined correlation (and as a result the predefined MIMO channel capacity) in a specific urban scenario.



■ ***Fig. 14.*** Correlation coefficient vs. the BS antenna height h_R for various $\gamma_0 = 2, 4, 6, 8, 10$ km^{-1} in the urban environment with $r = 10\lambda$, $\chi = 0.5$, $d = 0.3$ km, pseudo-LOS is at 0°, $\bar{h} = 15$ m, $\nu = 250$ km^{-2}

MIMO channel capacity estimation

Now, we will analyze the effect of the spatial fading correlation on the MIMO channel capacity by using (29) and (31) in (32). In Fig. 15, *a* shows the MIMO channel capacity as a function of the BS antenna heights for a variety of the buildings' density γ_0 in the simulated urban environment with the following parameters: SNR = 10 dB, separation distance between two BS receiver antennas of $r = 10\lambda$, $\chi = 0.5$, $d = 0.5$ km, pseudo LOS is at $\varphi = 0°$, and $\bar{h} = 25$ m.

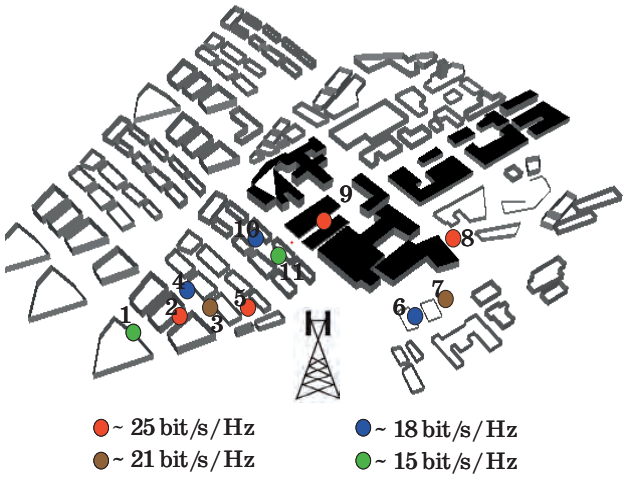


■ ***Fig. 15.*** MIMO channel capacity vs. the correlation coefficient and γ_0 in the urban scenario with SNR = = 10 dB, $r/\lambda = 10$, $\chi = 0.5$, $d = 0.5$ km, pseudo LOS is at $\varphi = 0°$, and $\bar{h} = 25$ m

We show that the MIMO channel capacity in Fig. 15, *a*, to be inversely proportional to the spatial fading correlation in Fig. 15, *b*. Thus, the increase in the spatial fading correlation ρ from 0.2 to 0.8 results in the normalized channel capacity degradation from approximately 6.5 bit/s/Hz to approximately 3.2 bit/s/Hz.

Notice that the MIMO channel capacity is directly proportional to the buildings' density γ_0 in the urban environment. Thus, the increase in the density γ_0 from 2 to 11 km^{-1}, results in the channel capacity improvement.

This simulations show that in additional to the conventional dependence of the MIMO channel capacity on SNR, it strongly depends on the location-specific parameters of the urban environment (via the statistical model of the urban propagation conditions), such as BS antenna height, buildings' density, average buildings' height, and so forth.



■ ***Fig. 16***. The MIMO channel capacity is simulated urban scenario with the BS antenna height of 15 m, MIMO system is with 4 × 4 antenna elements, and SNR = 20 dB; the numbers show places where numerical experiment was carried out

MIMO channel capacity analysis in predefined urban scenario

Now we will evaluate the MIMO channel capacity in a simulated "virtual" urban scenario, using an urban topographic plan taken from [22, 51]. The MIMO channel capacity was evaluated in 11 representative locations characterized by the SNR of 20 dB. The SNR was obtained using the ray-tracing software tool [54] as a ratio between the received signal strength and the noise with an equivalent bandwidth of 5 MHz. A BS antenna height of 15 m was simulated. The separation distance of $r = 10\lambda$ was simulated between 4 receiving antennas. The urban environment with the following parameters was simulated using (30) and (32): $\chi = 0.8$, $\bar{h} = 25$ m, $\gamma_0 = 6$ km^{-1}. The inter-user interference was neglected in this simulation.

Figure 16 shows the simulated urban scene. Colored dots in Fig. 16 show the tested MS locations with the SNR of 20 dB. The color pattern represents the achievable 4 × 4 MIMO channel capacity. Numbers near each color circle, from 1 to 9, indicate position of the MS antenna with respect to the BS antenna, shown in the picture, during the computer experiment. Notice that different locations (with equal SNR) are characterized by different spatial fading correlation and therefore the resulting MIMO channel capacity varies among these locations. We also notice that results shown in Fig. 16 agree with the results obtained during the measurement reported in references [45, 55].

The above obtained results allow us to summarize that the MIMO channel capacity in an urban environment has the BS antenna elevation and location specific factor and strongly depends on the spatial fading correlation determined by the scattering, reflection, diffraction and waveguide propagation phenomena.

To be continued.

References

1. **Jakes W. C.** *Microwave Mobile Communications*. Wiley, New York, 1974.
2. **Lee S. C. Y.** *Mobile Cellular Telecommunication Systems*. McGraw-Hill, New York, 1989.
3. **Steele R.** *Mobile Radio Communication*. IEEE Press, 1992.
4. **Proakis J. G.** *Digital Communications*. 3d ed. McGraw-Hill, New York, 1995.
5. **Stuber G. L.** *Principles of Mobile Communications*. Kluwer Academic Publishers, Boston, 1996.
6. **Peterson R. L., Ziemer R. E., and Borth D. E.** *Introduction to Spread Spectrum Communications*. Prentice Hall PTR, New Jersey, 1995.
7. **Rappaport T. S.** *Wireless Communications: Principles and Practice*. Prentice Hall PTR, New Jersey, 1996; (2nd ed.) in 2001.
8. **Steele R., and Hanzo L.** *Mobile Communications*. 2nd ed. John Wiley & Sons, Chickester, 1999.
9. **Li J. S., and Miller L. E.** *CDMA Systems Engineering Handbook*. Artech House, Boston-London, 1998.
10. **Saunders S. R.** *Antennas and Propagation for Wireless Communication Systems*. John Wiley& Sons, Chickester, 2001.
11. **Burr A.** *Modulation and Coding for Wireless Communications*. Prentice Hall PTR, New Jersey, 2001.
12. **Molisch A. F.** (Ed.). *Wideband Wireless Digital Communications*. Prentice Hall PTR, New Jersey, 2000.

13. **Paetzold M.** *Mobile Fading Channels: Modeling, Analysis, and Simulation*. John Wiley & Sons, Chickester, 2002.
14. **Simon M. K., Omura J. K., Scholtz R. A., and Levitt B. K.** *Spread Spectrum Communications Handbook*. McGraw-Hill, New York, 1994.
15. **Glisic S. and Vucetic B.** *Spread Spectrum CDMA Systems for Wireless Communications*. Artech House, Boston-London, 1997.
16. **Dixon R. C.** *Spread Spectrum Systems with Commercial Applications*. John Wiley & Sons, Chickester, 1994.
17. **Viterbi A. J.** *CDMA: Principles of Spread Spectrum Communication*. Addison-Wesley Wireless Communications Series, 1995.
18. **Goodman D. J.** *Wireless Personal Communication Systems*. Addison-Wesley, Reading, Massachusetts, 1997.
19. **Schiller J.** *Mobile Communications*. 2nd ed. Addison-Wesley Wireless Communications Series, 2003.
20. **Molisch A. F.** *Wireless Communications*. John Wiley & Sons, Chickester, 2007.
21. **Blaunstein N. and Christodoulou C.** *Radio Propagation and Adaptive Antennas for Wireless Communication Links*. 1st ed. Wiley & Sons, New Jersey, 2007.
22. **Blaunstein N. and Christodoulou C.** *Radio Propagation and Adaptive Antennas for Wireless Communication Networks — Terrestrial, Atmospheric and Ionospheric*. 2nd ed. Wiley & Sons, New Jersey, 2014.
23. **Hadar O., Bronfman I., and Blaunstein N.** Optimization of Error Concealment Based on Analysis of Fading Types. Part 1. Statistical Description of the Wireless Video Channel, Models of BER Determination and Error Concealment of Video Signals. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2017, no. 1, pp. 72–82. doi:10.15217/issn1684-8853.2017.1.72
24. **Hadar O., Bronfman I., and Blaunstein N.** Optimization of Error Concealment Based on Analysis of Fading Types. Part 2. Modified and New Models of Video Signal Error Concealment. Practical Simulations and their Results. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2017, no. 2, pp. 67–76. doi:10.15217/issn1684-8853.2017.2.67
25. **Vostrikov A., Kurtyanik D., Sergeev A.** Choosing Embedded WI-FI Module for Mobile Optic-Information Systems. *Vestnik*, 2018, no. 4, pp. 26–29 (In Russian).
26. **Vostrikov A., Balonin Yu., Kurtyanik D., Sergeev A., Sinitsyna O.** On Hybrid Method of Video Data Protection in IP-networks. *Telekommunikatsii* [Telecommunications], 2018, no. 2, pp. 34–39 (In Russian).
27. **Erosh I., Sergeev A., Filatov G.** Protection of Images During Transfer via Communication Channels. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2007, no. 5, pp. 20–22 (In Russian).
28. **Sergeev A.** Generalized Mersenne Matrices and Balonin's Conjecture. *Automatic Control and Computer Sciences*, 2014, vol. 48, no. 4, pp. 214–220.
29. **Sergeev A. M., Blaunstein N. S.** Orthogonal Matrices with Symmetrical Structures for Image Processing. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2017, no. 6, pp. 2–8 (In Russian). doi:10.15217/issn1684-8853.2017.6.2
30. *Modulation and Coding Techniques in Wireless Communications*. Ed. by Krouk E., and Semenov S. John Wiley & Sons, Chichester, United Kindom, 2011.
31. *Specification of the Bluetooth System*. Dec. 1, 1999. Available at: www.bluetooth.com. (accessed 15 August 2017).
32. **Junaid M., Mufti M., and Ilyas M. U.** Vulnerabilities of IEEE 802.11i Wireless LAN. *Trans. Eng., Comput. and Technol.*, Feb. 2006, vol. 11, pp. 228–233.
33. *IEEE 802.11 Working Group*. Available at: http://grouper.ieee.org/groups/802/11/index.html (accessed 15 August 2017).
34. *Wireless Ethernet Compatibility Alliance*. Available at: http://www.wirelessethernet.org/index.html (accessed 15 August 2017).
35. **Sharon O., and Altman E.** An Efficient Polling MAC for Wireless LANs. *IEEE/ACM Trans. on WiMAX Systems Evaluation Methodology V2.1etworking*, 2001, vol. 9, no. 4, pp. 439–451.
36. IEEE std. 802.11-1999: *Wireless LAN Medium Access Control* (*MAC*) *and Physical Layer* (*PHL*) *Specifications*, 1999.
37. **Qainkhani I. A., and Hossain E.** A Novel QoS-aware MAC Protocol for Voice Services over IEEE 802.11-based WLANs. *J. Wireless Communic. and Mobile Comput.*, 2009, vol. 9, pp. 71–84.
38. *Wireless LAN Medium Access Control and Physical Layer Specification*. IEEE Press, 1999, Jan. 14.
39. **Zyren J.** *Reliability of IEEE 802.11 High Rate DSSS WLANs in a High Density Bluetooth Environment*; 802.11 section, 8–6, 1999.
40. **Perahia E.** IEEE 802.11n Development: History, Process, and Technology. *IEEE Communic. Magazine*, 2008, vol. 46, pp. 46–55.
41. **Ni Q., Romshani L., and Turletti T.** A Survey of QoS Enhancements for IEEE 802.11 Wireless LAN. *J. Wireless Communic. and Mobile Comput.*, 2004, vol. 4, no. 5, pp. 547–566.
42. **Wang W., Liew S. C., and Li V. O. K.** Solutions to Performance Problems in VoIP over 802.11 Wireless LAN. *IEEE Trans. Veh. Tech.*, 2005, vol. 54, no. 1, pp. 366–384.
43. **Robinson J. W., and Randhawa T. S.** Saturation Throughput Analysis of IEEE 802.11e Enhanced Distributed Coordination Function. *IEEE J. Select. Areas of Communic.*, 2004, vol. 22, no. 5, pp. 917–928.
44. **Wang P., Jiang H., and Zhuang W.** 802.11e Enhancement for Voice Service. *IEEE Wireless Communic.*, 2006, vol. 13, no. 1, pp. 30–35.
45. **Perez-Costa X., and Camps-Mur D.** IEEE 802.11e QoS and Power Saving Features Overview and Analysis of Combined Performances. *IEEE Wireless Communic.*, 2010, vol. 17, no. 2, pp. 88–96.

46. **Kopsel A., and Wolisz A.** Voice Ptransmission in an IEEE 802.11 WLAN Based Access Network. *Proc. of 4th ACM Int. Workshop on Wireless Mobile Multimedia (WoWMoM)*, Rome, Italy, 2001, pp. 24–33.
47. **Veeraraghavan M., Chocker N., and Moors T.** Support of Voice Services in IEEE 802.11 Wireles LANs. *Proc. of IEEE INFOCOM'01*, 2001, vol. 1, pp. 488–497.
48. **Kim Y-J., and Suh Y-J.** Adaptive Polling MAC Schemes for IEEE 802.11 Wireless LANs Supporting Voice-over-IP (VoIP) Services. *J. Wireless Communic. and Mob. Comput.*, 2004, vol. 4, pp. 903–916.
49. **Andersen J. B.** Array Gain and Capacity of Known Random Channels with Multiple Element Arrays at Both Ends. *IEEE J. Selected Areas in Coomun.*, 2000, vol. 18, pp. 2172–2178.
50. **Blaunstein N., and Yarkoni N.** Capacity and Spectral Efficiency of MIMO Wireless Systems in Multipath Urban Environment with Fading. *Proc. of the European Conf. on Antennas and Propagation*, *EuCAP-2006*, Nice, France, 2006, pp. 111–115.
51. **Tsalolihin E., Bilik I., and Blaunstein N.** MIMO Capacity in Space and Time Domain for Various Urban Environments. *Proc. of 5th European Conf. on Antennas and Propagation, EuCAP*, Rome, Italy, 11–15 April, 2011, pp. 2321–2325.
52. **Chizhik D., Farrokhi F., Ling J., and Lozano A.** Effect of Antenna Separation on Capacity of BLAST in Correlated Channels. *IEEE Commun. Letters*, 2000, vol. 4, no. 11.
53. **Gesbert D., Shafi M., Shiu D., Smith P., and Naguib A.** From Theory to Practice: An Overview of MIMO Space-Time Coded Wireless Systems. *IEEE Journal on Selected Areas in Comm.*, 2003, vol. 21, no. 3, pp. 281–302.
54. *Radioplan*. RPS user Manual 5.4. Available at: http://www.actix.com (accessed 15 August 2017).
55. **Philippe J., Schumacher L., Pedersen K., Mogensen P., and Frederiksen F.** A Stochastic MIMO Radio Channel Model with Experimental Validation. *IEEE J. Selected Areas in Commun.*, 2002, vol. 20, no. 6, pp. 1211–1226.
56. **Gesbert D., Boleskei H., Gore D. A., and Paulraj A. J.** Outdoor MIMO Wireless Channels: Models and Performance Prediction. *IEEE Trans. Commun.*, 2002, vol. 50, no. 6, pp. 1926–1934.
57. **Boleskei H., Borgmann M., and Paulraj A. J.** On the Capacity of OFDM-based Spatial Multiplexing Systems. *IEEE Trans. Commun.*, 2002, vol. 50, no. 1, pp. 225–234.
58. **Boleskei H., Borgmann M., and Paulraj A. J.** Impact of the Propagation Environment on the Performance of Space-Frequency Coded MIMO-OFDM. *IEEE J. Select. Areas Commun.*, 2003, vol. 21, no. 2, pp. 427–439.
59. **Chizik D., Ling J., Wolniansky P. W., Valenzuela R. A., Costa N., and Huber K.** Multiple-imput-multiple-output Measurements and Modeling in Manhattan. *IEEE J. on Selected Areas in Comm.*, 2003, vol. 23, no. 2, pp. 321–331.
60. **Oyman O., Nabar R. U., Boleskei H., and Paulraj A. J.** Characterizing the Statistical Properties of Mutual Information in MIMO Channels. *IEEE Trans. Signal Processing*, 2003, vol. 51, pp. 2784–2795.
61. **Paulraj A. J., Gore D. A., Nabar R. U., and Boleskei H.** An Overview of MIMO Communications — A Key to Gigabit Wireless. *Proc. of IEEE*, 2004, vol. 92, no. 2, pp. 198–218.
62. **Forenza A.,** et al. Adaptive MIMO Transmission for Exploiting the Capacity of Spatially Correlated Channels. *IEEE Trans. Vehic. Technol.*, 2007, vol. 56, no. 2, pp. 619–630.
63. **Foschini G. J., and Gans M. J.** On Limits of Wireless Communications in a Fading Environment when using Multiple Antennas. *Wireless Person. Communic.*, 1998, vol. 6, no. 3, pp. 311–335.
64. **Proakis J. G.** *Digital Communications*. 4th ed. McGraw-Hill, New York, 2001.
65. **Paulraj A. J., and Kailath T.** *Increasing Capacity in Wireless Broadcast Systems using Distributed Transmission/Directional Reception (DTDR).* US patent 5,345,599, Sept. 6, 1994.
66. **Foschini G. J.** Layered Space-time Architecture for Wireless Communication in a Fading Environment when using Multiple Antennas. *Bell Labs. Tech. J.*, 1996, vol. 1, no. 2, pp. 41–59.
67. **Golden G. D., Foschini G. J., Valenzula R. A., and Wolniansky P. W.** Direction Algorithm and Initial Laboratory Results using the V-BLAST Space-time Communication Architecture. *Electron. Lett.*, 1999, vol. 35, no. 1, pp. 14–15.
68. **Nabar R. U., Bolcskei H., Erceg V., Gesbert D., and Paulraj A. J.** Performance of Multiantenna Signaling Techniques in the Presence of Polarization Diversity. *IEEE Trans. Signal Process.*, 2002, vol. 50, no. 10, pp. 2553–2562.
69. **Zheng L., and Tse D.** Diversity and Multiplexing: A Fundamental Tradeoff in Multiple Antenna Channels. *IEEE Trans. Inform. Theory*, 2003, vol. 49, no. 5, pp. 1073–1096.
70. **Varadarajan B., and Barry J. R.** The Rate-diversity Trade-off for Linear Space-time Codes. *Proc. IEEE Vehicular Tech. Conf.*, 2002, vol. 1, pp. 67–71.
71. **Godovarti M, and Nero A. O.** Diversity and Degrees of Freedom in Wireless Communications. *Proc. ICASSP*, May 2002, vol. 3, pp. 2861–2864.
72. **Raleigh G. G., and Cioffi J. M.** Spatio-temporal Coding for Wireless Communication. *IEEE Trans. Communic.*, 1998, vol. 46, no. 3, pp. 357–366.
73. **Wittniben A.** Base Station Modulation Diversity for Digital Simulcast. *Proc. IEEE Vehicular Tech. Conf.*, May 1991, pp. 848–853.
74. **Seshadri N., and Winters J. H.** Two Signaling Schemes for Improving the Error Performance of Frequency-Division-Duplex (FDD) Transmission Systems using Transmitter Antenna Diversity. *Int. J.*

Wireless Inform. Networks, 1994, vol. 1, no. 1, pp. 49–60.
75. **Alamouti S. M.** A Simple Transmit Diversity Technique for Wireless Communications. *IEEE J. Select. Areas Communic.*, 1998, vol. 16, no. 8, pp. 1451–1458.
76. **Tarokh V., Seshandri N., and Calderbank A. R.** Space-time Codes for High Data Rate Wireless Communication: Performance Criterion and Code Construction. *IEEE Trans. Inform. Theory*, 1999, vol. 45, no. 5, pp. 1456–1467.
77. **Ganesan G., and Stoica P.** Space-time Block Codes: A Maximum SNR Approach. *IEEE Trans. Inform. Theory*, 2001, vol. 47, no. 4, pp. 1650–1656.
78. **Hassibi B., and Hochwald B. M.** High-rate Codes that are Linear in Space and Time. *IEEE Trans. Inform. Theory*, 2002, vol. 48, no. 7, pp. 1804–1824.
79. **Health Jr., R. W., and Paulraj A. J.** Linear Dispersion Codes for MIMO Systems based on Frame Theory. *IEEE Trans. Signal Process.*, 2002, vol. 50, no. 10, pp. 2429–2441.
80. **Winters J. H.** The Diversity Gain of Transmit Diversity in Wireless Systems with Rayleigh Fading. *IEEE Trans. Veh. Technol.*, 1998, vol. 47, no. 1, pp. 119–123.
81. **Bjerke B. A., and Proakis J. G.** Multiple-antenna Diversity Techniques for Transmission over Fading Channels. *Proc. Wireless Communic. and Networking Conf.*, Sept. 1999, vol. 3, pp. 1038–1042.
82. **Heath Jr., R. W., and Paulraj A. J.** Switching between Diversity and Multiplexing in MIMO Systems. *IEEE Trans. Communic.*, 2005, vol. 53, no. 6, pp. 962–968.
83. **Chandrasekhar V., Andrews J. G., and Gatherer A.** Femtocell Networks: A Survey. *IEEE Commun. Magazine*, 2003, vol. 46, no. 9, pp. 59–67.
84. **Shannon C. E.** A Mathematical Theory of Communication. *Bell System Tech. J.*, July and October 1948, vol. 27, pp. 379–423 and pp. 623–656.
85. **Yeh S.-P., Talwar S., Lee S.-C., and Kim H.** WiMAX Femtocells: A Perpective on Network Architecture, Capacity, and Coverage. *IEEE Commun. Magazine*, 2008, vol. 46, no. 10, pp. 58–65.
86. **Knisely D. N., Yoshizawa T., and Favichia F.** Standardization of Femtocells in 3GPP. *IEEE Commun. Magazine*, 2009, vol. 47, no. 9, pp. 68–75.
87. **Knisely D. N., and Favichia F.** Standardization of Femtocells in 3GPP2. *IEEE Commun. Magazine*, 2009, vol. 47, no. 9, pp. 76–82.
88. **Chandrasekhar V., and Andrews J. G.** Uplink Capacity and Interference Avoidance for Two-tier Femtocell Networks. *IEEE Trans. Wireless Commun.*, 2009, vol. 8, no. 7, pp. 3498–3509.
89. **Calin D., Claussen H., and Uzunalioglu H.** On Femto Deployment Architectures and Macrocell Offloading Benefits in Joint Macro-femto Deployments. *IEEE Commun. Magazine*, 2010, vol. 48, no. 1, pp. 26–32.
90. **Kim R. Y., Kwak J. S., and Etemad K.** WiMAX Femtocel: Requirements, Challenges, and Solutions. *IEEE Commun. Magazine*, 2009, vol. 47, no. 9, pp. 84–91.
91. **Lopez-Perez D., Valcarce A., de la Roche G., and Zhang J.** OFDMA Femtocells: A Roadmap on Interference Avoidance. *IEEE Commun. Magazine*, 2009, vol. 47, no. 9, pp. 41–48.
92. **Chandrasekhar V., Andrews J. G., Muharemovic T., Shen Z., and Gatherer A.** Power Control in Two-tier Femtocell Networks. *IEEE Trans. Wireless Commun.*, 2009, vol. 8, no. 8, pp. 4316–4328.
93. **Yavuz M., Meshkati F., Nanda S.,** et al. Interference Management and Performance Analysis of UMTS/HSPA+Femtocells. *IEEE Commun. Magazine*, 2009, vol. 47, no. 9, pp. 102–109.
94. *Femto Forum*. Available at: http://www.femtoforum.org/femto/ (accessed 15 August 2017).
95. **Blaunstein N. S., and Sergeev M. B.** Channel Capacity Prediction for Femtocell-Macrocell Deployment Strategies in the Urban Environments with Congested Layout of Users. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2012, no. 3, pp. 54–62 (In Russian).
96. **Tsalolihin E., Bilik I., Blaunstein N., and Babich Y.** Channel Capacity in Mobile Broadband Heterogeneous Networks based on Femto Cells. *Proc. of EuCAP-2012 Int. Conf.*, Prague, Czech Republic, March 26–30, 2012, pp. 1–5.
97. **Blaunstein N., and Levin M.** VHF/UHF Wave Attenuation in a City with Regularly Spaced Buildings. *Radio Science*, 1996, vol. 31, no. 2, pp. 313–323.
98. **Blaunstein N.** Prediction of Cellular Characteristics for Various Urban Environments. *J. Anten. and Propagatat. Magazine*, 1999, vol. 41, no. 6, pp. 135–145.
99. **Blaunstein N.** Average Field Attenuation in the Nonregular Impedance Street Waveguide. *IEEE Trans. on Anten. and Propagat.*, 1998, vol. 46, no. 12, pp. 1782–1789.
100. **Blaunstein N., Katz D., Censor D.,** et al. Prediction of Loss Characteristics in Built-up Areas with Various Buildings' Overlay Profiles. *J. Anten. and Propagat. Magazine*, 2002, vol. 44, no. 1, pp. 181–192.
101. **Yarkoni N., Blaunstein N., and Katz D.** Link Budget and Radio Coverage Design for Various Multipath Urban Communication Links. *Radio Science*, 2007, vol. 42, no. 2, pp. 412–427.
102. **Katz D., Blaunstein N., Hayakawa M., and Kishiki Y. S.** Radio Maps Design in Tokyo City based on Stochastic Multi-parametric and Deterministic Ray Tracing Approaches. *J. Anten. and Propag. Magazine*, 2009, vol. 51, no. 5, pp. 200–208.

УДК 621.371
doi:10.31799/1684-8853-2018-5-94-103

Эволюция многопроцессорных систем связи — сотовых и несотовых — в исторической перспективе. Часть 2

А. М. Сергеев[а], старший преподаватель, orcid.org/0000-0002-4788-9869,
Н. Ш. Блаунштейн[б, в], доктор физ.-мат. наук, профессор, nathan.blaunstein@hotmail.com
[а]Санкт-Петербургский государственный университет аэрокосмического приборостроения, Санкт-Петербург, РФ
[б]Негевский университет им. Бен-Гуриона, ПОБ 653, Бен-Гуриона ул., 1, г. Беэр-Шева, 74105, Израиль
[в]Иерусалимский технологический колледж, Хавад Халейми, 21, ПОБ 16031, Иерусалим, 91160, Израиль

Постановка проблемы: целью данного обзора является анализ эволюции систем беспроводной связи от второй генерации (2D) до пятой генерации (5G), а также изменения технологий и их существующих теоретических основ и протоколов — от Bluetooth, WLAN, WiFi и WiMAX до LTE, OFDM/OFDMA, MIMO и LTE/MIMO — продвинутых технологий с новой иерархической структурой дизайна сотовых карт femto/pico/micro/macro. **Методы:** использованы новые теоретические подходы для описания продвинутых технологий, таких как многопользовательская техника разделения пользователей, OFDM и OFDM-новейший подход, новые аспекты описания MIMO-систем на базе использования многолучевых антенн, дизайн различных сотовых карт на основе новых алгоритмов построения фемто/пико/микро/макро сот, а также новой методологии интегрирования новой MIMO/LTE-системы с помощью многолучевых антенн. **Результаты:** создана новая методология описания многопользовательского разделения, использования комбинированной OFDM/OFDMA-модуляции для обхождения интерференции между пользователями и между символами в новых многопроцессорных системах, мультипликативных шумов, имеющих место в беспроводных многопроцессорных системах связи, вызванных явлениями многолучевости. В итоге предложено, как обойти эффекты распространения, имеющие место в наземных каналах связи, используя комбинацию MIMO- и LTE-технологий, основанных на применении многолучевых антенн. Для этих целей разработан новый стохастический подход к проблеме, учитывающий особенности застройки земной поверхности, такие как профиль застройки домов, плотность застройки домов вокруг антенн базовой станции и пользователей и т. д. Эти характеристики позволяют в итоге оценить эффекты федингa как источника мультипликативного шума. **Практическая значимость:** новая методология оценки эффектов, созданных мультипликативным шумом, интерференцией между пользователями и между символами, имеющими место в наземных системах беспроводной связи, позволяет прогнозировать практические аспекты существующих и новых многопроцессорных беспроводных систем связи, такие как емкость (количество) пользователей и спектральная эффективность каналов пользователей для различных конфигураций построения сот — фемто/пико/микро/макро, а также новейших конфигураций систем MIMO/LTE для построения будущих систем 4-го и 5-го поколений.

Ключевые слова — пропускная способность, многочисленный вход-многочисленный выход (MIMO), канал MIMO, пространственное мультиплексирование, спектральная эффективность, фактор K, городская среда, плотная застройка домов.

Цитирование: Sergeev A. M., Blaunstein N. Sh. Evolution of multiple-access networks — cellular and non-cellular — in historical perspective. Part 2. *Информационно-управляющие системы*, 2018, № 5, с. 94–103. doi:10.31799/1684-8853-2018-5-94-103
Citation: Sergeev A. M., Blaunstein N. Sh. Evolution of multiple-access networks — cellular and non-cellular — in historical perspective. Part 2. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 5, pp. 94–103. doi:10.31799/1684-8853-2018-5-94-103

К статье
Галининой О. С., Андреева С. Д., Тюрликова А. М.
«Учет специфики доступа большого числа устройств при межмашинном взаимодействии в современных сотовых сетях» («Информационно-управляющие системы», 2018, № 4, с. 105–114.).
На странице 112, правый столбец, последним абзацем вводится дополнение текста:
Исследования Тюрликова Андрея Михайловича выполнены при поддержке гранта РФФИ 17-07-00142 «Разработка моделей и методов анализа надежности соединений в гетерогенных сетях 5G для концепции Интернета надежных вещей».

К статье
Матвеева Н. В., Тюрликова А. М.
«Слотовый ALOHA с итерационной процедурой разрешения коллизий. Стабильность и нестабильность» («Информационно-управляющие системы», 2018, № 3, с. 89–97.).
На странице 96, правый столбец, последним абзацем вводится дополнение текста:
Работа выполнена в рамках инициативного научного проекта № 8.8540.2017/БЧ «Разработка алгоритмов передачи данных в системах IoT с учетом ограничений на сложность устройств».

УДК 612.8 + 57.089
doi:10.31799/1684-8853-2018-5-104-111

Articles

Algorithm for automatic estimation of human brain activity features during mental task evaluation

V. A. Maksimenko[a], *PhD, Associate Professor, orcid.org/0000-0002-4632-6896*
A. E. Runnova[a], *PhD, Associate Professor, orcid.org/0000-0002-2102-164X*
R. A. Kulanin[a], *Junior Researcher, orcid.org/0000-0001-6810-8024*
P. A. Protasov[a], *Junior Researcher, orcid.org/0000-0003-1451-4582*
M. O. Zhuravlev[a, b], *PhD, Associate Professor, orcid.org/ 0000-0002-8620-1609*
P. Chholak[c], *Junior Researcher, orcid.org/0000-0002-6437-7750*
A. Pisarchik[c], *PhD, Professor, orcid.org/0000-0003-2471-2507*
A. E. Hramov[a, b], *Ph. D, Professor, orcid.org/0000-0003-2787-2530, hramovae@sstu.ru*
[a]*Yuri Gagarin State Technical University of Saratov, 77, Politechnicheskaia St., 410054, Saratov, Russian Federation*
[b]*Saratov State University named after N. G. Chernyshevsky, 83, Astrakhanskaia St., 410012, Saratov, Russian Federation*
[c]*Technical University of Madrid, Calle Ramiro de Maeztu, 7, 28040, Madrid, Spain*

Introduction: *It is known that many types of human activity involve a generation of particular patterns in electroencephalographic recordings with common properties for different subjects. Among them one can highlight the brain response to the visual stimuli in occipital lobe or motor-related activity in motor cortex. At the same time, more complex human activity can induce different scenarios of the neural dynamics in brain, which depends on the human personal features. Personality is more pronounced during mental task processing. In particular, it is shown that human personality causes individual scenarios during decision-making and affects learning performance. We suppose that individual features of human personality, when we wish to define the ways of how a human processes mental tasks, affect neural network dynamics and therefore can be seen in electroencephalographic recordings.* ***Purpose:*** *Development of the algorithm for estimating the personal spatio-temporal and time-frequency features of electrical brain activity during mental task evaluation.* ***Results:*** *We propose algorithm, which allows to reveal individual features of the brain activity during completion of mental tasks based on multichannel electroencephalogram analysis. Algorithm is implemented in brain-computer system and tested during experimental session for the subjects who perform the Schulte table test. We show, that revealed individual features of brain activity can predict the properties of human attention.* ***Practical relevance:*** *We believe that the results are of the great interest for testing and diagnostics. It can be the starting point for development of automatic intelligent systems for estimation and control of human mental abilities.*

Keywords *— continuous wavelet transformation, electroencephalography, mental task evaluation.*

Citation: Maksimenko V. A., Runnova A. E., Kulanin R. A., Protasov P. A., Zhuravlev M. O., Chholak P., Pisarchik A., Hramov A. E. Algorithm for automatic estimation of human brain activity features during mental task evaluation. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 5, pp. 104–111. doi:10.31799/1684-8853-2018-5-104-111

Introduction

It is known, that many types of human activity involve a generation of particular patterns in electroencephalographic (EEG) recordings with common properties for different subjects. For instance, the perception of visual stimuli is known to induce an event-related response of the neuronal brain network, in particular, a decrease in alpha-wave (8–12 Hz) and an increase in beta-wave (15–30 Hz) activities [1–3]. Such a behavior reflects different cognitive functions, namely, the alpha-wave suppression is associated with visual [4] or auditory [5] attention, while the beta-wave activation relates to information processing [6] and an alerted state [7].

Different physiological and psychological states (e. g., sleep stages, arousal, etc.) are known to possess specific properties of neural activity. For instance, motor-related brain activity is manifested in the brain as a specific scenario of neural activity with well-defined frequency and spatial localization. Particularly, it is characterized by event related desynchronization (ERD) in alpha/mu- and beta-bands [8]. The same features are observed during motor imagery in specially trained subjects [9, 10]. However, different scenarios occur in untrained subjects, where EEG patterns can vary from subject to subject [11]. Such a variation is caused by the task complexity when each subject chooses his own strategy to process the task, that results in individual time-frequency and spatio-temporal EEG structures. Along with motor imagery, the personality is more pronounced during mental task processing. It was also shown that human personality causes individual scenarios during decision-making [12] and affects learning performance [13].

We suppose that individual features of human personality, when we wish to define the ways of how a human processes mental tasks, affect neural network dynamics and therefore can be seen in EEG recordings. Correlation between EEG and personal features provides possibility to estimate such human features as personality traits and mental abilities. It should be noted that this problem was attacked yet in 1973. By analyzing resting states, Edwards and Abbott [14] tried to reveal personality traits in EEG signals. However, their attempt was unsuccessful because personality is not manifested when a person is at rest. Until now, this problem remains open [15, 16]

In the present work, we propose algorithm, which allows to reveal individual features of the brain activity during completion of mental tasks based on multichannel EEG analysis. Algorithm is implemented in brain-computer system and tested during experimental session for the subjects who perform the Schulte table test.

Algorithm for EEG analysis

The proposed algorithm is schematically illustrated in Fig. 1. via a flowchart. One can see that it is evaluated in seven steps.

Step I: Acquisition of multichannel EEG with the help of non-invasive electrodes located on the surface of the head, according to the arrangement of 10–20. Electrical brain activity signals are recorded with a sampling frequency of 250 Hz. The recorded signals are processed by a bandpass filter.

Step II: EEG signals recorded in different parts of the cortex are divided into two equal groups. The first group contains the channels located in the left hemisphere (Fp1, F7, F3, T3, C3, P3, T5, O1), the second group contains the channels located in the right hemisphere (Fp2, F8, F4, T4, C4, P4, T6, O2). The channels located in the interhemispheric region (Fz, Cz, Pz) are excluded from consideration.

Step III: For each channel $X_n(t)$ (in the first and second group), wavelet transformation is performed. The wavelet energy spectrum $E^n(f,\ t) = \sqrt{W_n(f,\ t)^2}$ is calculated for each EEG channel in the frequency range 10–40 Hz. Here, $W_n(f,\ t)$ is the complex-valued wavelet coefficients calculated as [17]

$$W_n(f,\ t) = \sqrt{f}\int_{t-4/f}^{t+4/f} X_n(t)\varphi^*(f,\ t)dt, \qquad (1)$$

where $n = 1,\ ...,\ N$ is the EEG channel number ($N = 19$) being the total number of channels used for the analysis) and "*" defines the complex conjugation. The mother wavelet function $\varphi(f,\ t)$ is the Morlet wavelet often used for the analysis of neurophysiological data, defined as

$$\varphi(f,\ t) = \sqrt{f}\pi^{1/4} e^{j\omega_0 f(t-t_0)} e^{f(t-t_0)^2/2}, \qquad (2)$$

where $w_0 = 2\pi$ is the central frequency of the mother Morlet wavelet.

Step IV: For each channel, the obtained wavelet energy spectrum is analyzed in several frequency ranges (in accordance with Table 1.)

For these bands the values of wavelet energy $E^n_\delta(t), E^n_\theta(t), E^n_\alpha(t), E^n_{\beta_1}, E^n_{\beta_1}, E^n_\gamma$ for each -th EEG channel are calculated as

$$E^n_{\delta,\theta,\alpha,\beta_1,\beta_2,\gamma}(t) = \frac{1}{\Delta f}\int_{f\in\delta,\theta,\alpha,\beta_1,\beta_2,\gamma} E^n(f,\ t)df. \qquad (3)$$

Based on Eq. (3) the percentages of the spectral energy distributed in the considered bands are estimated as

$$e^n_{\delta,\theta,\alpha,\beta_1,\beta_2,\gamma}(t) = E^n_{\delta,\theta,\alpha,\beta_1,\beta_2,\gamma}(t) / E^n_0(t)\ (\times 100\%), \qquad (4)$$

where is defined as the whole energy and calculated as

$$E^n_0(t) = \frac{1}{\Delta f}\int_{1\text{Hz}}^{40\text{Hz}} E^n(f,\ t)df. \qquad (5)$$

Step V: In order to describe the ratio between high frequency and low frequency brain activity for each channel the coefficient ε^n is calculated via equation

$$\varepsilon^n = E^n_{\text{HF}} / E^n_{\text{LF}}, \qquad (6)$$

where

$$E^n_{\text{HF}}(t) = \frac{1}{\Delta f}\int_{f>10\text{Hz}} E^n(f,\ t)df, \qquad (7)$$

■ ***Table 1***. Frequency bands of EEG signals

Name	Definition	Frequency range, Hz
Delta-band	δ	1–4
Theta-band	θ	4–8
Alpha-band	α	8–13
Beta1-band	b_1	13–23
Beta2-band	b_2	32–34
Gamma-band	γ	34–40

$$E^n_{\mathrm{LF}}(t) = \frac{1}{\Delta f} \int\limits_{f<10\mathrm{Hz}} E^n(f,\, t)df. \qquad (8)$$

Step VI: The coefficients ε^n are calculated for each EEG channel for both during the task evaluation (active phase) and during resting state (passive phase). The obtained values of ε^n are averaged over the channels belonging to the first and secon groups (see Step II for groups definition)

$$\varepsilon_{\mathrm{LH}} = \frac{1}{N_{\mathrm{LH}}} \sum_n \frac{E^n_{\mathrm{HF}}}{E^n_{\mathrm{LF}}},$$
$$n = \{\mathrm{Fp1},\ \mathrm{F3},\ \mathrm{F7},\ \mathrm{C3},\ T3,\ P3,\ T5,\ O1\}, \qquad (9)$$
$$N_{\mathrm{LH}} = 8.$$

$$\varepsilon_{\mathrm{RH}} = \frac{1}{N_{\mathrm{RH}}} \sum_n \frac{E^n_{\mathrm{HF}}}{E^n_{\mathrm{LF}}},$$
$$n = \{\mathrm{Fp2},\ \mathrm{F4},\ \mathrm{F8},\ \mathrm{C4},\ T4,\ P4,\ T6,\ O2\}, \qquad (10)$$
$$N_{\mathrm{RH}} = 8.$$

As the result, the values $\varepsilon^{active}_{\mathrm{LH}}$, $\varepsilon^{active}_{\mathrm{RH}}$, $\varepsilon^{passive}_{\mathrm{LH}}$, $\varepsilon^{passive}_{\mathrm{RH}}$ are obtained.

Step VII: Based on the obtained coefficients $\varepsilon^{active}_{\mathrm{LH}}$, $\varepsilon^{active}_{\mathrm{RH}}$, $\varepsilon^{passive}_{\mathrm{LH}}$, $\varepsilon^{passive}_{\mathrm{RH}}$ which characterize the activity of the left and right hemispheres, the lateralization coefficient for the active and passive phases $k^{active} = \varepsilon^{active}_{\mathrm{RH}} / \varepsilon^{active}_{\mathrm{LH}}$, $k^{passive} = \varepsilon^{passive}_{\mathrm{LH}} / \varepsilon^{passive}_{\mathrm{RH}}$ are calculated.

Mental ability evaluation

In order to compare the results of EEG analysis with the human mental abilities we used Schulte tables. Such method is frequently used as a psychodiagnostic test for studying properties of human attention. It allows to determine working effectiveness and ability, as well as resistance to external interference. It is known, that the time of the -th table completion can be used to evaluate personal criteria:

1. Work efficiency *WE* (the arithmetic mean of the values of table completion times)

$$WE = \frac{\tau_1 + \tau_2 + \ldots + \tau_R}{R}, \qquad (11)$$

2. Warming-up work indicator *WU* (the ratio of the working time which subject spend for the first table to the value of work efficiency)

$$WU = \frac{\tau_1}{WE}, \qquad (12)$$

3. Psychological stability *PS* (the human ability to sustain the operational activity for a long period of time).

$$PS = \frac{\tau_R - 1}{WE}. \qquad (13)$$

The work efficiency is known to illustrate the attention consistency and performance. The resulted *WU* close to or lower than 1 indicates good warming-up, while 1 and higher means that the subject needs longer preparation time (warm-up) for the main work. The *PS* results close to 1 and less indicate a good psychological stability.

Data processing and main results

The results of the proposed algorithm evaluation are shown in Fig. 2. on the single subject example.

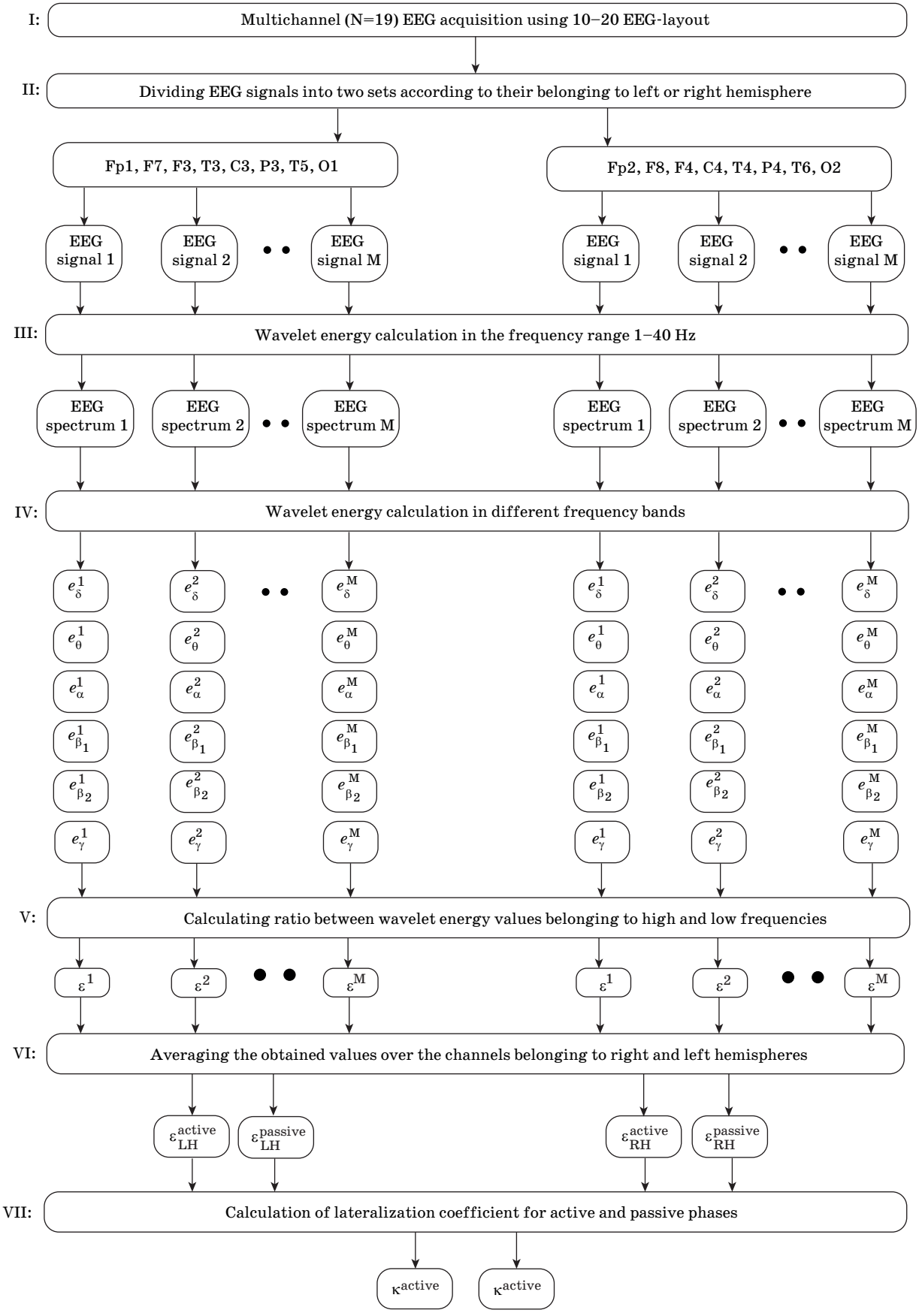
Fig. 2 (a) demonstrates the typical EEG recordings obtained in left and right hemispheres during the step I of the algorithm.

Fig. 2 (b) shows the values of $e^{\mathrm{F4}}_{\delta,\theta,\alpha,\beta_1,\beta_2,\gamma}(t)$ calculated during for a single EEG trial recorded from the frontal lobe, specifically, from the F4 electrode. One can see, that when the active phase is replaced by the passive phase, the values of $e^{\mathrm{F4}}_{\delta,\theta}(t)$ calculated for low frequencies (namely, δ, and θ frequency bands) rapidly increase, while the values of $e^{\mathrm{F4}}_{\alpha,\beta_1,\beta_2,\gamma}(t)$, calculated for α, b_1, b_2, and γ frequency bands, pronouncedly decrease. Such a dynamical behavior repeats itself during subsequent completions of the Schulte tables.

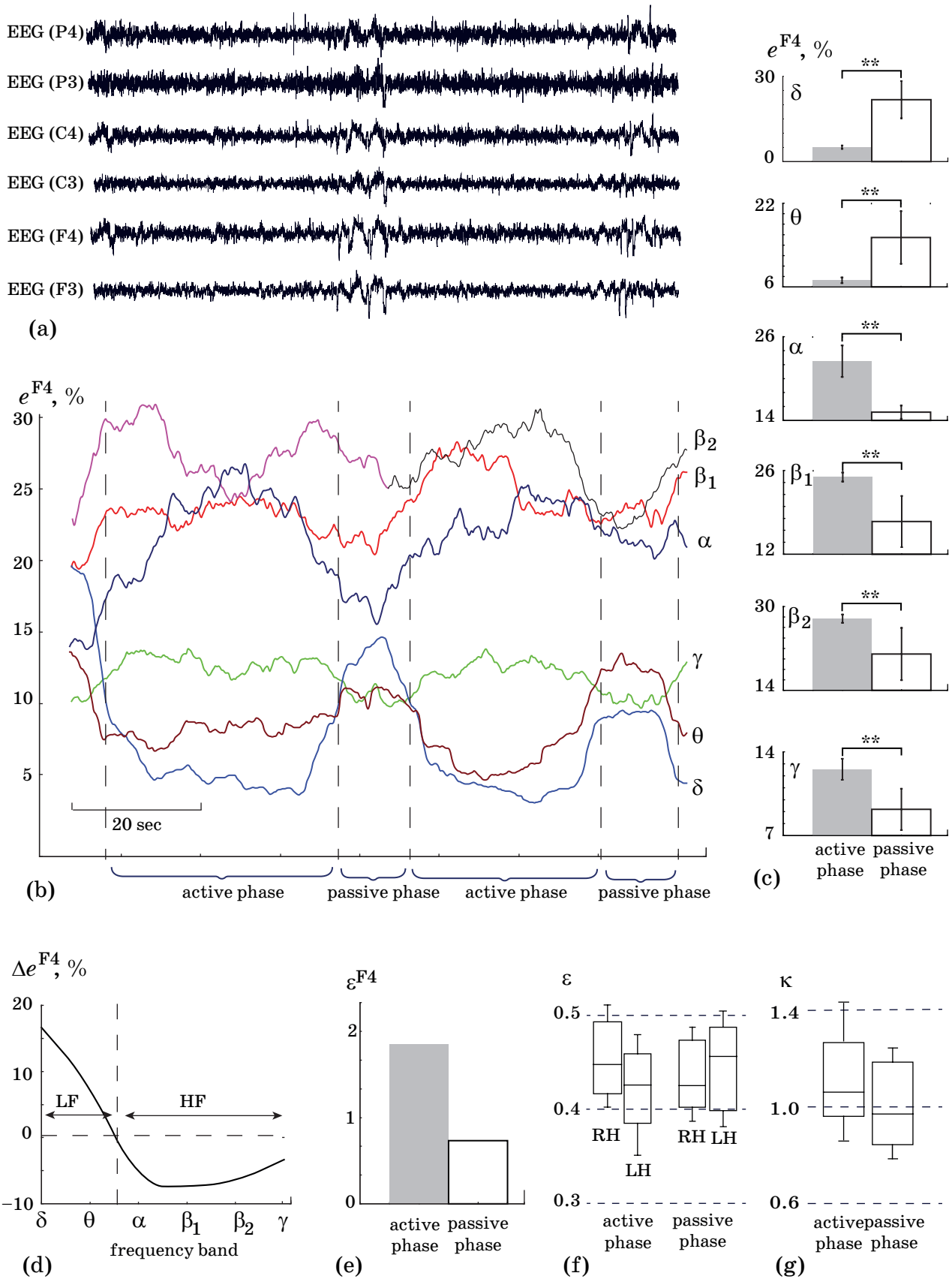
Fig. 2 (c) shows the results of statistical analysis of the values $e^{\mathrm{F4}}_{\delta,\theta,\alpha,\beta_1,\beta_2,\gamma}$ calculated for the time intervals corresponding to $N = 5$ consecutive active and passive sessions. Data are shown as mean SD. Obtained results demonstrate the significant increase of $e^{\mathrm{F4}}_{\delta,\theta}$ and significant decrease of $e^{\mathrm{F4}}_{\alpha,\beta_1,\beta_2,\gamma}$ during the transition from passive to active phase (**$p < 0.01$ via nonparametric Whitney U-test).

Fig. 2 (d) demonstrates the distinctive features between the mean values $e^{\mathrm{F4}}_{\delta,\theta,\alpha,\beta_1,\beta_2,\gamma}$ obtained for the active and passive phases for each frequency band. One can see that in the low frequency range, which includes δ, and θ frequency bands, such difference is positive ($De^{\mathrm{F4}} > 0$), while in the high frequency range (α, b_1, b_2, and γ frequency bands) it is negative ($De^{\mathrm{F4}} < 0$).

According to this result, one can easily distinguish active and passive phases, based on the consideration of EEG properties, i. e., by comparing the energy of the spectral components belonging either to high (HF) or low (LF) frequency bands. For this purpose, it is convenient to use coefficient ε^n



■ ***Fig. 1.*** Flowchart of the algorithm for EEG analysis. Each step is marked as I...VII on the left-hand side



■ ***Fig. 2.*** The results of the proposed algorithm evaluation

(Eq. 6), which reflects the ratio between the values of spectral energy in the high and low frequency ranges. In particular, for the considered F4 electrode, the values of ε^{F4}, shown in Fig. 2 (e) are significantly lower during the passive phase than during the active phase.

Thus, the time frequency analysis performed for a single EEG recording demonstrates a pronounced change in the ratio between the energy of high and low spectral components. At the same time, along with the features of time-frequency structure revealed in a single EEG, the spatio-temporal features of electrical brain activity also play an important role. This is mostly reflected in hemispheric differences commonly observed in electrical activity of the brain associated with the completion of mental tasks [18–22].

The spatio-temporal features are taken into account in our algorithm by consideration of the values $\varepsilon_{LH}^{active}$, $\varepsilon_{RH}^{active}$, $\varepsilon_{LH}^{passive}$, $\varepsilon_{RH}^{passive}$ calculated during step VI by averaging over the channels, belonging to left and right hemispheres.

Fig. 2 (f) demonstrates the values $\varepsilon_{LH}^{active}$, $\varepsilon_{RH}^{active}$, $\varepsilon_{LH}^{passive}$, $\varepsilon_{RH}^{passive}$ calculated for group of 8 subjects during active and passive phases. Data are shown as median and 25–75 percentiles (box) and outlines (whiskers). One can see that there are differences in the electrical activity in the hemispheres during active and passive phases. Namely, during active phase subjects of the considered group exhibit increase of high-freauency activity in right hemisphere, while during the passive phase such increase is observed in left hemisphere. As the result, median value of lateralization coefficient becomes >1 for active pahse and <1 for passive phase. At the same time, such deviations in median lateralization coefficient are insignificant. It evidences the variability of this coefficient between subjects in the group.

It can be supposed, that such variability is connected with personal differences which affect the process of mental task accomplishing. According to this we have applied algorithm, described above for the group of 20 subjects and compared the behaviour of lateralization coefficient with the results of psychodiagnostic test. We have shown that the subjects, for which the lateralization coefficient is close to unity for both active and passive phases demonstrate the lowest value of work efficiency ($WE > 40$ seconds) and the lowest degree of psychological stability (PS~1.0). On the contrary, subjects, for which the lateralization coefficient $\kappa > 1.0$ for active phases and $\kappa < 1.0$ for passive phases demonstrate the value of work efficiency much higher (WE~30 seconds) as well as the higher degree of psychological stability ($PS < 0.9$).

Materials and methods

Twenty healthy men (33±7 years), participated at the experiment. All participants provided informed written consent before participating in the experiment. The experimental procedure was performed in accordance to the Helsinki's Declaration and approved by the local Ethics Committee of the Yuri Gagarin State Technical University of Saratov.

Experiments was carried out during the first half of the day. All participants performed a series of simple psycho-diagnostic tests using the Schulte tables to study their attention features. Shulte table is a simplified version of Zahlen-Verbindungs-Test (ZVT) [23, 24], widely used in Russia [25]. The Schulte table is a 5×5 matrix of random numbers from 1 to 25. The psychological task was to find all numbers in a reverse order. During these *active* experimental phases, each person had to complete $R = 5$ tables. For every i-th testing series, the completion time T_i was registered. Between the active phases, each volunteer had a short resting interval referred to as a *passive* experimental phase. Length of active phases was varied from 30 to 50 second depending on the speed of task completion. Length of passive phases was set as 10 seconds.

Electrical brain activity was recorded with multi-channel EEG-acquisition system — electroencephalograph-reorder Encephalan-EEGR-19/26 (Russia) with multiple EEG channels and the two-button input device. To study EEGs the monopolar registration method and the classical ten-twenty electrode system were used.

Conclusion

We propose the algorithm for the estimation of the spatio-temporal and time-frequency features of electrical brain activity during the mental task evaluation. Time-frequency features of the brain activity are estimated by analyzing EEG spectral energy in high- and low- frequency bands. Spatio-temporal features are estimated with the help of lateralization coefficient. Proposed algorithm is implemented in brain-computer interface and tested and tested during experimental session for the subjects who perform the Schulte table test. We demonstrate, that the dynamics of lateralization coefficient reflects the personal features of the brain activity, which correlates with the properties of human attention. In particular we show that that the subjects, for which the lateralization coefficient is close to unity for both active and passive phases demonstrate the lowest value of work efficiency and the lowest degree of psychological stability. On the contrary, subjects, for which the lateralization coefficient $\kappa > 1.0$ for active phases and $\kappa < 1.0$ for

passive phases demonstrate the value of work efficiency much higher as well as the higher degree of psychological stability.

We believe that the results are of the great interest for testing and diagnostics. It can be the starting point for development of automatic intelligent systems for estimation and control of human mental abilities.

This work has been supported by Russian Science Foundation (grant No. 16-12-10100).

References

1. **Maksimenko V. A., Runnova A. E., Frolov N. S., Makarov V. V., Nedaivozov V. O., Koronovskii A. A., Pisarchik A. N., Hramov A. E.** Multiscale neural connectivity during human sensory processing in the brain. *Phys. Rev. E*, 2018, vol. 97, 052405, pp. 1–9.
2. **Maksimenko V. A., Runnova A. E., Zhuravlev M. O., Makarov V. V., Nedayvozov V. O., Grubov V. V., Pchelintceva S. V., Hramov A. E., Pisarchik A. N.** Visual perception affected by motivation and alertness controlled by a noninvasive brain-computer interface. *PloS One*, 2017, vol. 12(12), p. e0188700.
3. **Foxe J. J., Snyder A. C.** The role of alpha-band brain oscillations as a sensory suppression mechanism during selective attention *Frontiers in Psychology*, 2011, vol. 2, p. 154.
4. **Sauseng P., Klimesch W., Stadler W., Schabus M., Doppelmayr M., Hanslmayr S., Gruber W. R., Birbaumer N.** A shift of visual spatial attention is selectively associated with human EEG alpha activity. *European Journal of Neuroscience*, 2005, vol. 22(11), pp. 2917–2926.
5. **Basar E., Güntekin B.** A short review of alpha activity in cognitive processes and in cognitive impairment. *International Journal of Psychophysiology*, 2012, vol. 86(1), pp. 25–38.
6. **Sehatpour P., Molholm S., Schwartz T. H., Mahoney J. R., Mehta A. D., Javitt D. C., Stanton P. K., Foxe J. J.** A human intracranial study of long-range oscillatory coherence across a frontal–occipital–hippocampal brain network during visual object processing. *Proceedings of the National Academy of Sciences*, 2008, vol. 105(11), pp. 4399–4404.
7. **Gola M., Magnuski M., Szumska I., Wróbel A.** EEG beta band activity is related to attention and attentional deficits in the visual performance of elderly subjects. *International Journal of Psychophysiology*, 2013, vol. 89(3), pp. 334–341.
8. **Duann J. R., Chiou J. C.** A comparison of independent event-related desynchronization responses in motor-related brain areas to movement execution, movement imagery, and movement observation. *PloS One*, 2016, vol. 11(9), e0162546.
9. **Guillet A., Di Rienzo F., Maclntyre T., Moran A., Collet C.** Imagining is not doing but involves specific motor commands: a review of experimental data related to motor inhibition. *Frontiers in Human Neuroscience*, 2012, vol. 6, p. 247.
10. **Wolpaw J. R., McFarland D. J.** Control of a two-dimensional movement signal by a noninvasive brain-computer interface in humans. *Proceedings of the National Academy of Sciences of the United States of America*, 2004, vol. 101(51), pp. 17849–17854.
11. **Maksimenko V. A., Pavlov A. N., Runnova A. E., Nedaivozov V. O., Grubov V. V., Koronovskii A. A., Pchelintseva S. V., Pitsik E., Pisarchik A. N., Hramov A. E.** Nonlinear analysis of brain activity, associated with motor action and motor imaginary in untrained subjects. *Nonlinear Dynamics*, 2018, vol. 91(4), pp. 2803–2817.
12. **Franken I. H. A., Muris P.** Individual differences in decision-making. *Personality and Individual Differences*, 2005, vol. 39(5), pp. 991–998.
13. **Chamorro-Premuzic T., Furnham A.** Personality, intelligence and approaches to learning as predictors of academic performance. *Personality and Individual Differences*, 2008, vol. 44(7), pp. 1596–1603.
14. **Edwards A. L., Abbott R. D.** Measurement of personality traits: theory and technique. *Annual Review of Psychology*, 1973, vol. 24(1), pp. 241–278.
15. **Roslan N. S., Izhar L. I., FayeI., Saad M. N. M., Sivapalan S., Rahman M. A.** Review of EEG and ERP studies of extraversion personality for baseline and cognitive tasks. *Personality and Individual Differences*, 2017, vol. 119, pp. 323–332.
16. **Korjus K., Uusberg A., Uusberg H., Kuldkepp N., Kreegipuu K., Allik J., Vicente R., Aru J.** Personality cannot be predicted from the power of resting state EEG. *Frontiers in Human Neuroscience*, 2015, vol. 9(63), pp. 1–7.
17. **Pavlov A. N., Hramov A. E., Koronovskii A. A., Sitnikova Yu. E., Makarov V. A., Ovchinnikov A. A.** Wavelet analysis in neurodynamics. *Physics-Uspekhi*, 2012, vol. 55(9), pp. 845–875.
18. **Park C., Looney D., Kidmose P., Ungstrup M., Mandic D. P.** Time-frequency analysis of EEG asymmetry using bivariate empirical mode decomposition. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2011, vol. 19(4), pp. 366–373.
19. **Rogers L. J, Zucca P, Vallortigara G.** Advantages of having a lateralized brain. *Proceedings of the Royal Society of London B: Biological Sciences*, 2004, 271, (Suppl 6), pp. S420-S422.
20. **Barry R. J, Clarke A. R, Johnstone S. J.** A review of electrophysiology in attention-deficit/hyperactivity disorder: I. Qualitative and quantitative electroencephalography. *Clinical neurophysiology*, 2003, vol. 114(2), pp. 171–183.
21. **Luschekina E, Khaerdinova O. Y, Luschekin V., Strelets V.** Interhemispheric differences in the spectral power and coherence of EEG rhythms in children

with autism spectrum disorders. *Human Physiology*, 2017, vol.43(3), pp. 265–273.
22. **Santarnecchi E, Tatti E, Rossi S, Serino V, Rossi A.** Intelligence-related differences in the asymmetry of spontaneous cerebral activity. *Human brain mapping*, 2015, vol. 36(9), pp. 3586–3602.
23. **Oswald W. D., Roth E.** *Der zahlen-verbindungs-test: (ZVT), ein sprachfreier intelligenz-test zur messung der "kognitiven leistungsgeschwindigkeit", Handanweisung*. Verlag für Psychologie Hogrefe, 1987, 62 p.
24. **Neubauer A. C, Knorr E.** Three paper-and-pencil tests for speed of information processing: Psychometric properties and correlations with intelligence. *Intelligence*, 1998, vol. 26(2), pp.123–151.
25. **Pavlenko V., Lutsyuk N., Borisova M.** Correlation of the characteristics of evoked EEG potentials with individual peculiarities of attention in children. *Neurophysiology*, 2004, vol. 36(4), pp. 276–284.

УДК 612.8 + 57.089
doi:10.31799/1684-8853-2018-5-104-111

Алгоритм для автоматического детектирования особенностей активности мозга во время выполнения когнитивных задач

В. А. Максименко[а], канд. физ.-мат. наук, доцент, orcid.org/0000-0002-4632-6896
А. Е. Руннова[а], канд. физ.-мат. наук, доцент, orcid.org/0000-0002-2102-164X
Р. А. Куланин[а], младший научный сотрудник, orcid.org/0000-0001-6810-8024
П. А. Протасов[а], младший научный сотрудник, orcid.org/0000-0003-1451-4582
М. О. Журавлев[а, б], канд. физ.-мат. наук, доцент, orcid.org/ 0000-0002-8620-1609
П. Чолак[с], младший научный сотрудник, orcid.org/0000-0002-6437-7750
А. Писарчик[с], PhD, профессор, orcid.org/0000-0003-2471-2507
А. Е. Храмов[а, б], доктор физ.-мат. наук, профессор, orcid.org/0000-0003-2787-2530, hramovae@sstu.ru
[а] Саратовский государственный технический университет имени Гагарина Ю. А., Политехническая ул., 77, Саратов, 410054, РФ
[б]Саратовский государственный университет им. Н. Г. Чернышевского, Астраханская ул., 83, Саратов, 410012, РФ
[с]Мадридский политехнический университет, ул. Рамиро де Маэцту, 7, 28040, Мадрид, Испания

Введение: многие виды человеческой деятельности ассоциируются с возникновением характерных паттернов на электроэнцефалографических записях, которые обладают общими свойствами для разных испытуемых. Среди них можно выделить отклик мозга на визуальные стимулы, регистрируемый в затылочной области, или нейронную активность, связанную с двигательными функциями, регистрируемую в моторной коре. В то же время, более сложная деятельность человека может вызывать различные сценарии нейронной динамики в зависимости от индивидуальных особенностей человека. Наиболее значительно данный эффект проявляется при выполнении человеком когнитивных задач. В частности, показано, что индивидуальные особенности определяют сценарии нейронной активности при принятии решений и влияют на эффективность обучения. Можно предположить, что индивидуальные особенности человеческой личности определяют стратегию, которую человек использует при решении когнитивных задачи, что, в свою очередь, отражается на динамике нейронной сети мозга и может быть детектировано на электроэнцефалографических записях. **Цель:** разработка алгоритма оценки индивидуальных пространственно-временных и частотно-временных характеристик электрической активности головного мозга при решении когнитивных задачи. **Результаты:** предложен алгоритм, позволяющий выявить индивидуальные особенности функционирования нейронной сети мозга при выполнении когнитивных задач на основе анализа многоканальных электроэнцефалограмм. Алгоритм реализован в виде интерфейса мозг-компьютер и протестирован на группе испытуемых, которые выполняют тест Шульте. Показано, что выявленные индивидуальные особенности активности мозга могут ассоциироваться со свойствами человеческого внимания в процессе решения задач. **Практическая значимость:** полученные результаты представляют большой интерес для тестирования и диагностики. Они являются основой для разработки автоматических интеллектуальных систем для оценки и контроля умственных способностей человека.

Ключевые слова — вейвлетное преобразование, электроэнцефалография, решение когнитивной задачи.

Цитирование: Maksimenko V. A., Runnova A. E., Protasov P. A., Zhuravlev M. O., Chholak P., Pisarchik A., Hramov A. E. Algorithm for automatic estimation of human brain activity features during mental task evaluation. *Информационно-управляющие системы*, 2018, № 5, с. 104–111. doi:10.31799/1684-8853-2018-5-104-111
Citation: Maksimenko V. A., Runnova A. E., Protasov P. A., Zhuravlev M. O., Chholak P., Pisarchik A., Hramov A. E. Algorithm for automatic estimation of human brain activity features during mental task evaluation. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 5, pp. 104–111. doi:10.31799/1684-8853-2018-5-104-111

УДК 338.984
doi:10.31799/1684-8853-2018-5-112-119

Научные статьи
Articles

Мультиагентный метод построения сменно-суточных заданий для задачи планирования производственных ресурсов в реальном времени

А. Н. Лада[а, б], *руководитель проекта, orcid.org/0000-0002-1022-7212, lada@kg.ru*
И. В. Майоров[а, б], *аналитик, orcid.org/0000-0002-3130-8142*
[а]*ООО «Научно-производственная компания «Интеллектуальные Транспортные Системы», Московское шоссе, 17, Самара, Самара, 443013, РФ*
[б]*Институт проблем управления сложными системами РАН, Россия, Садовая ул., 61, Самара, 443020, РФ*

Постановка проблемы: *задача согласованного управления производственными ресурсами предприятий, производящих электротехническую продукцию, — комплексная проблема, обладающая высоким уровнем сложности ввиду разнообразия типов используемых ресурсов, зависимости производственных процессов от множества факторов и условий. Традиционные методы планирования для решения этой проблемы оказываются недостаточно эффективными.* ***Цель исследования:*** *на основе мультиагентного подхода разработать метод оперативного планирования производства электротехнического оборудования.* ***Результаты:*** *разработан мультиагентный метод построения сменно-суточных заданий для планирования производственных ресурсов предприятия, производящего электротехническую продукцию, на основе онтологии предметной области, заданных критериев, предпочтений и ограничений. Заказы состоят из связанных операций, описываемых технологическими картами. В итерационном процессе обмена сообщениями агенты планировщика улучшают текущие значения критериев равномерности загрузки и минимизации времени выполнения для построения сменно-суточных заданий. Разработанный метод планирования производственных ресурсов позволяет строить расписания выполнения связанных операций в системе ресурсов по событиям реального времени. При выполнении пула заказов на предприятии ООО «ПК «Электрум» (г. Самара) была обеспечена равномерность загрузки оборудования и исполнителей, а также снижено на 10 % количество задержек выполнения заказов.* ***Практическая значимость:*** *система, разработанная на основе предлагаемого метода, может работать автономно или совместно с имеющейся системой складского учета материалов и готовой продукции. Подход не ограничен рамками описанной предметной области и применим в других отраслях, требующих решения аналогичных производственных задач. Ожидается получение экономического эффекта за счет снижения простоев производственных ресурсов и повышения показателей их эффективности.*

Ключевые слова *— предметная область, электротехническая продукция, мультиагентные методы, управление производственными ресурсами, онтология производственного предприятия, планирование в реальном времени.*

Цитирование: Лада А. Н., Майоров И. В. Мультиагентный метод построения сменно-суточных заданий для задачи планирования производственных ресурсов в реальном времени. *Информационно-управляющие системы*, 2018, № 5, с. 112–119. doi:10.31799/1684-8853-2018-5-112-119
Citation: Lada A. N., Mayorov I. V. Multi-agent method of constructing daily-shift schedule for real-time industrial resource management. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 5, pp. 112–119 (In Russian). doi:10.31799/1684-8853-2018-5-112-119

Введение

Современные системы управления производственными процессами должны обеспечивать не только планирование материальных, трудовых, финансовых ресурсов, но также учитывать комплексный характер взаимодействия различных подразделений и взаимное согласование планов с учетом окружающей обстановки [1–4]. Внешние события, заключающиеся в приходе новых заказов, изменения требований и характеристик уже запланированных, пересмотр сроков поставок материалов, а также разница, возникающая при фактическом выполнении, приводят к необходимости перерасчета существующих и утвержденных планов. В связи с этим традиционные методы планирования, использующие многочисленные методы дискретной оптимизации, теории расписаний и эвристики [5–10], не вполне подходят к современным задачам, для которых необходим постоянный пересчет планов по внешним и внутренним событиям, поскольку на практике приходится прибегать к упрощениям в постановках задач, не все ограничения и факторы могут быть учтены и т. п. В таких случаях более соответствует реальности мультиагентный подход, с помощью которого создаются системы производственного планирования, управления проектами, транспортной логистики и во многих других областях [11–19].

Постановка задачи начального планирования производственных ресурсов

Пусть имеем набор заказов на производство изделий O_n, $n = 1$, s и набор производственных

ресурсов (станки и другое оборудование) R_j, $j = 1, m$. Каждый заказ характеризуется технологической картой, с описанием всех деталей D^i_k, $k = 1, p$, которые также могут состоять из других деталей (возникает многоуровневая вложенность). Каждая деталь D^i_k описывается материалами M^{ik}_z, $z = 1, q$, из которых она изготавливается, и упорядоченным набором технологических операций TO^{ik}_j, $j = 1, m$, которые требуется произвести с материалом и/или другой деталью на ресурсе R_j за известное время обработки детали TD^{ik}_j. В общем виде технологическая карта описывается следующей структурой:

$[O_n]$
  $\{[D^n{}_1]$
    $[M^{n,1}{}_1]$
    $[M^{n,1}{}_q]$
    $\{[D^{n,1}{}_1]$
      $[M^{n,1,1}{}_1]$
      $[M^{n,1,1}{}_q]$
      $\{[D^{n,1,1,\ldots}{}_p]\}$
      $[TO^{n,1,1}{}_1]\,[R_1]\,[TD^{n,1,1}{}_1]$
      $[TO^{n,1,1}{}_m]\,[R_m]\,[TD^{n,1,1}{}_m]\}$
    $\{[D^{n,1}{}_p]\}$
    $[TO^{n,1}{}_1]\,[R_1]\,[TD^{n,1}{}_1]$
    $[TO^{n,1}{}_m]\,[R_m]\,[TD^{n,1}{}_m]\}$
$\{[D^n{}_p]\}$

Для каждого ресурса R_j задано ежесуточное временное окно $[TRs_j; TRf_j]$ доступности этого ресурса для работы (рабочая смена станка), с учетом режима труда и отдыха рабочих, которые на них работают. Одна и та же деталь не может обрабатываться на нескольких ресурсах одновременно, т. е. пока она обрабатывается на одном ресурсе, второй будет свободен и может быть использован для обработки другой детали. Требуется составить сменно-суточный план работы для каждого ресурса R_j по производству всех заказов O_i согласно их технологическим картам, с минимальным простоем ресурсов R_j. При планировании, кроме производственных мощностей (оснастки, комплектующих и материалов), требуется учет календарей доступности исполнителей в подразделениях, их квалификации и специализации. Предприятию в целом необходимо поддерживать равномерность загрузки оборудования, подразделений и исполнителей, не допуская простоев. В отличие от существующих MES систем, вся информация об оборудовании, исполнителях, календарях, структурах ресурсов хранится в базе знаний на основе онтологии производства [20]. Таким образом, необходимо сквозное многокритериальное планирование заказов на ресурсах в вертикальных и горизонтальных структурах по входящим внешним и внутренним событиям.

Метод решения задачи построения начального плана

Для решения задачи построения первоначального расписания предлагается использовать «жадный» итерационный метод, где детали всех заказов распределяются по производственным ресурсам последовательно согласно следующему алгоритму: заказы O_i обрабатываются последовательно от 1 до s. Из всех деталей D^i_k i-го заказа вначале выбираются те, что лежат на самом низком уровне технологической карты, затем уровнем выше и так далее, до самого верхнего уровня. Детали, лежащие на одном уровне, обрабатываются последовательно, согласно их порядковому номеру в уровне, с учетом последовательного выполнения технологических операций TO^{ik}_j на ресурсах R_j. При планировании операций для каждой детали проверяется наличие необходимых материалов M^{ik}_z, требующихся для ее производства, в случае их недостатка деталь пропускается. Анализируются свободные места в расписании нужного ресурса с учетом окна доступности его текущей смены работы $[TRs_j; TRf_j]$ и в первую очередь заполняются свободные места, образовавшиеся за счет планирования предыдущих операций, начиная с самой ранней. Если длительность свободного места недостаточна для выполнения операции, то анализируется следующее свободное место. В худшем случае, если не удалось встроиться ни в одно доступное свободное место, операция встает в самый конец плана. Если с учетом ее постановки происходит выход за пределы окна работы текущей смены ресурса $[TRs_j; TRf_j]$, она становится первой на следующей доступной смене данного ресурса. Все детали последующих заказов обрабатываются аналогичным образом. Алгоритм повторяется до тех пор, пока все операции деталей всех заказов не будут распределены по всем ресурсам с учетом их смен работы.

Пример решения задачи построения начального плана

Для решения задачи построения начального плана рассмотрим пример планирования двух заказов O_1 и O_2 на изготовление узла фиксации трансформатора для трансформаторной подстанции КТПН-УХЛ1 и рычага включения к ней. Технологическая карта для данных заказов имеет вид:

$[O_1]$ Узел фиксации трансформатора КТПН-УХЛ1

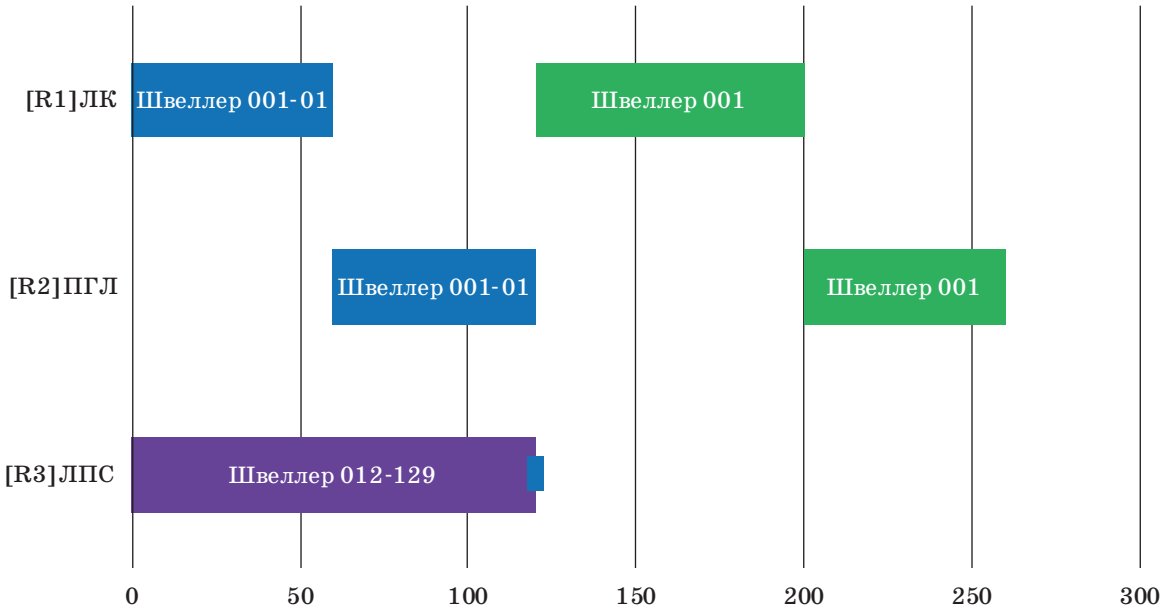
  $\{[D^1{}_1]$ Швеллер 001
    $[M^{1,1}{}_1]$ Лист 4,0 ГОСТ 1050-13 1,5 кг

{[$D^{1,1}_1$] Швеллер 001-01
[$M^{1,1,1}_1$] Лист 4,0 ГОСТ 1050-13 1 кг
[$TO^{1,1,1}_1$] Раскрой металла [R_1] Лазерный комплекс [$TD^{1,1,1}_1$] 60 с
[$TO^{1,1,1}_2$] Гибка [R_2] Пресс гидравлический листогиб [$TD^{1,1,1}_2$] 60 с
[$TO^{1,1}_1$] Раскрой металла [R_1] Лазерный комплекс [$TD^{1,1}_1$] 80 с
[$TO^{1,1}_2$] Гибка [R_2] Пресс гидравлический листогиб [$TD^{1,1}_2$] 60 с}}
{[D^1_2] Швеллер 012-129
[$M^{1,2}_1$] Лист 3,0 ГОСТ 1050-13 2,5 кг
[$TO^{1,2}_3$] Резка [R_3] Ленточнопильный станок [$TD^{1,2}_3$] 120 с}
[O_2] Рычаг включения трансформатора КТПН-УХЛ1
{[D^2_1] Фланец 019
[$M^{2,1}_1$] Лист 6,0 ГОСТ 14637-89 0,2 кг
[$TO^{2,1}_1$] Раскрой металла [R_1] Лазерный комплекс [$TD^{2,1}_1$] 50 с
[$TO^{2,1}_2$] Гибка [R_2] Пресс гидравлический листогиб [$TD^{2,1}_2$] 70 с}
{[D^2_2] Втулка 005-01
[$M^{2,2}_1$] Труба 30х6 ГОСТ 8734-75 0,2 кг
[$TO^{2,2}_3$] Резка [R_3] Ленточнопильный станок [$TD^{2,2}_3$] 160 с}

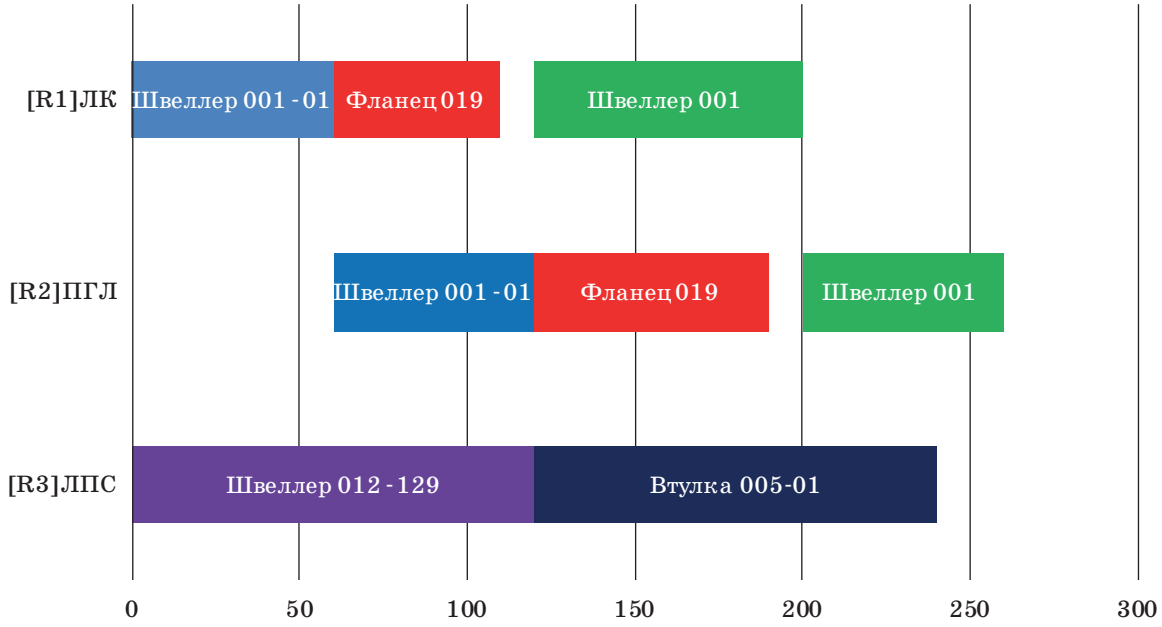
Для простоты будем полагать, что все производственные ресурсы (в нашем примере их 3) доступны для работы круглосуточно (временные окна [TRs_j; TRf_j] не ограничены), а также есть в наличии и в нужном количестве все производственные материалы, требуемые в технологических картах.

Распределение обработки деталей по станкам начинается с самого глубокого уровня технологической карты, в нашем примере это деталь *Швеллер* 001-01, она раскраивается на лазерном комплексе за 60 с, после чего переходит на пресс гидравлический листогиб, где гнется также 60 с. Далее переходим на уровень выше и планируем родительскую деталь *Швеллер* 001, она в свою очередь раскраивается и гнётся, но только после предыдущей дочерней детали, поставим ее в план выполнения по станкам и запланируем последнюю деталь *Швеллер* 012-129. Она обрабатывается на другом отдельном ресурсе и никак не зависит от предыдущих двух деталей, поэтому поставим ее в план первой, полученное расписание представим диаграммой загрузки ресурсов (рис. 1).

Перейдем к планированию заказа O_2, который мы получили одновременно с первым, мы должны максимально использовать свободные области в уже имеющемся расписании, построенном после пла-



■ ***Рис. 1.*** Диаграмма первоначального распределения операций по ресурсам
■ ***Fig. 1.*** The diagram of the initial operations on resources allocation



■ ***Рис. 2.*** Диаграмма планирования заказа O_2
■ ***Fig. 2.*** The diagram of scheduling order O_2

нирования первого заказа. Заказ O_2 состоит из двух деталей, которые изготавливаются на одном уровне, поэтому порядок их обработки не важен. Начинаем с детали *Фланец* 019, она раскраивается на лазерном комплексе за 50 с, ищем первое свободное место на ресурсе R_1, есть свободное место после обработки детали *Швеллер* 001-01 продолжительностью 60 с, времени хватает, поэтому ставим деталь в эту область. Далее *Фланец* 019 требуется обработать на R_2 за время 70 с, ищем первое свободное место на ресурсе R_2, после момента окончания обработки на R_1. Пустой участок в самом начале длиною 60 с нам не подходит, идем дальше, находим место после детали *Швеллер* 001-01 длиной 80 с, ставим *Фланец* 019 туда. Обрабатываем последнюю деталь *Втулка* 005-01, она обрабатывается только на R_3 за время 160 с, находим ближайшее свободное место нужной длины после *Швеллер* 012-129 и ставим деталь туда. Новое расписание загрузки ресурсов показано на рис. 2.

Динамическое перепланирование по фактическим событиям

Задача построения динамического плана по фактическим событиям в реальном времени является более сложной, чем статическая задача построения начального плана. В такой задаче подразумевается, что ее условия могут произвольно динамически меняться с течением времени, могут быть добавлены новые заказы, отменены или частично изменены уже известные заказы, стать недоступными ресурсы, но чаще всего могут возникать задержки при выполнении уже существующего плана, что требует его адаптивной перестройки.

Разрабатываемый нами подход основан на сопоставлении заказам и ресурсам программных агентов с их локальными и зачастую противоположными интересами, способных реагировать на изменения состава заказов и ресурсов, выявлять конфликты в расписании, принимать решения и взаимодействовать между собой для разрешения конфликтов и поиска компромиссов путём переговоров (взаимных уступок) [17, 19, 20]. Это позволяет находить согласованные решения и поддерживать баланс интересов агентов и всей системы, в общем случае представляющий многокритериальную целевую функцию. С каждым ресурсом R_j связан агент ресурса, с каждой деталью D_k^i — агент детали. Агенты могут отправлять и получать сообщения и принимать решения согласно своей логике и текущей ситуации, которая определяется состоянием каждого агента. Текущие состояния агентов изменяются в моменты поступления заказов и фиксации внешних фактических событий.

При поступлении нового заказа создаются агенты деталей согласно технологической карте этого заказа. Они рассылают запрос на свое размещение на агентах ресурсов, которые, в свою очередь, анализируют свое текущее состояние, наличие свободных окон, оценивают свою загрузку, предлагают свободные места агентам деталей

для размещения. Агент детали стремится запланировать себя на нужных ресурсах как можно раньше. Агент ресурса R_j, в свою очередь, стремится быть постоянно загруженным и минимизировать образующиеся простои внутри рабочей смены $[TRs_j; TRf_j]$, которые рассчитываются по формуле:

$$Dtime^j = TRf^j - TRs^j - \sum_{k=1}^{p} TD_k^j,$$

где $k = (1, ..., p)$ — индексы размещенных деталей заказов на ресурсе R_j, $TD^j{}_k$ — продолжительность обработки этих деталей. Глобальная целевая функция планирования F определяется как суммарное время простоя всех ресурсов:

$$F = \left\{ P \to \max, \sum_{j=1}^{m} Dtime^j \to \min \right\},$$

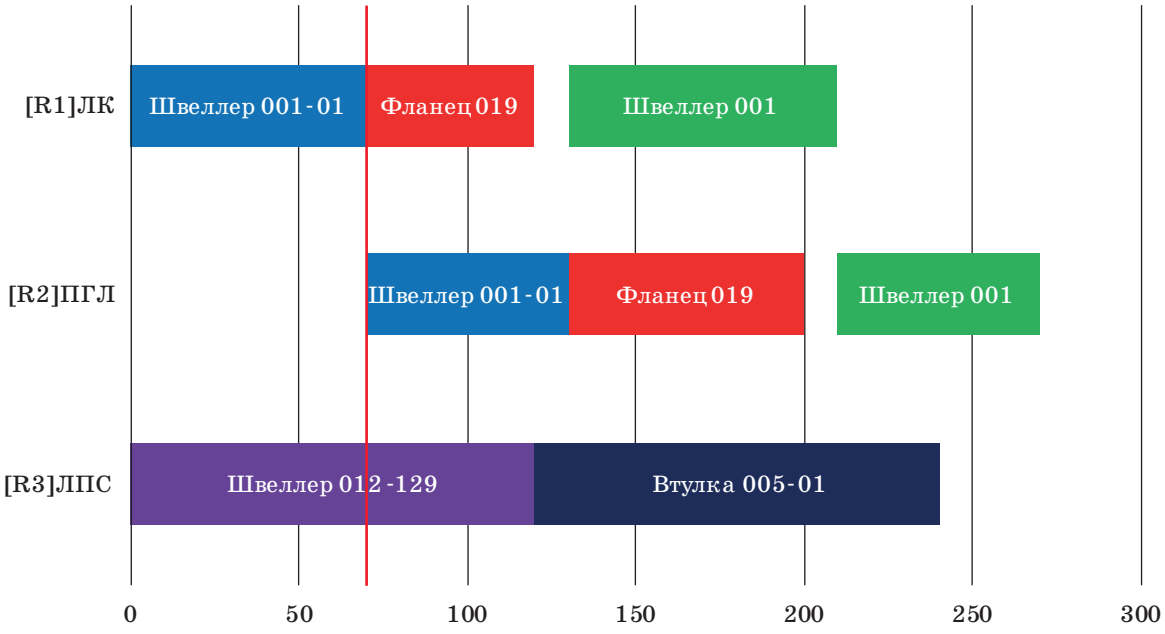
где P — общее число запланированных деталей на всех ресурсах. При улучшении глобальной целевой функции (F — число распределенных деталей и суммарное время простоя ресурсов) текущий вариант распределения деталей по ресурсам принимается в качестве рабочей версии плана, после чего агенты не размещенных и плохо размещенных деталей пробуют улучшить свое положение на ресурсах за счет переговоров с другими деталями с просьбой о «подвижках». Если в результате этих переговоров глобальная целевая функция улучшилась, новая версия плана принимается в качестве рабочей и процесс повторяется до тех пор, пока не перестанут поступать новые фактические события или новые переговоры не будут приводить к глобальному улучшению.

Пример динамического перепланирования по фактическим событиям

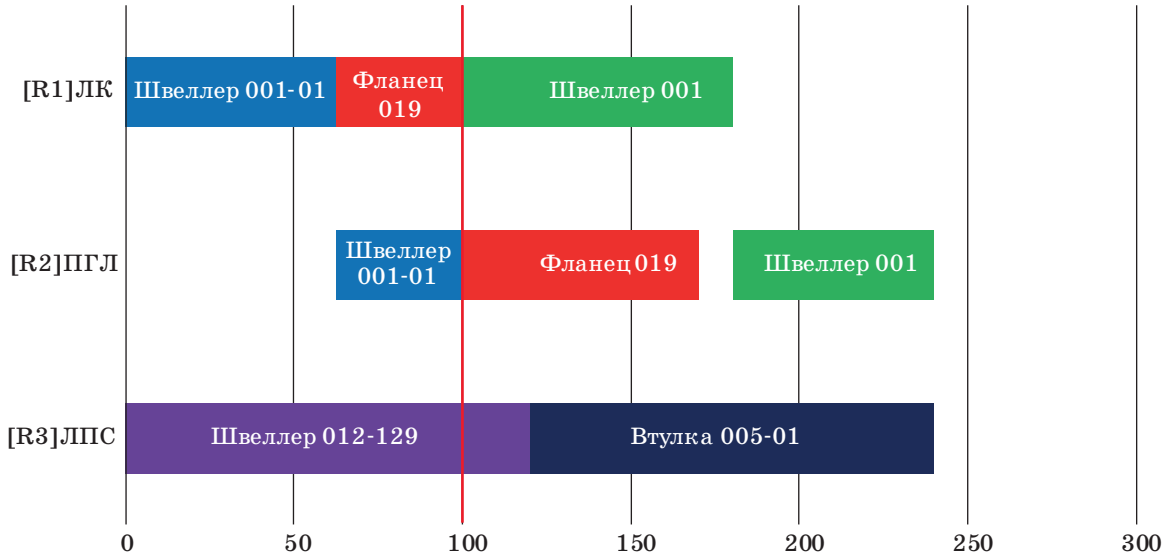
Рассмотрим построенный ранее начальный план и предположим, что теперь мы будем каким-то образом (например, с помощью терминалов рабочих) фиксировать фактические события о завершении обработки каждой детали на каждом ресурсе, после чего последующий план нужно будет адаптивно перестраивать. Для простоты будем отсчитывать время от 0.

Пусть в момент времени T_{70} наступило событие завершения изготовления детали *Швеллер* 001-01 на ресурсе R_1 (ожидалось, что изготовление завершится в T_{60}). В результате реакции на данное событие все расписание на ресурсе R_1 сдвигается на 10 с вправо. Поскольку деталь *Швеллер* 001-01 обрабатывается также на ресурсе R_2 после R_1, расписание на R_2 также сдвинется вправо на 10 с. С учетом всех изменений расписание примет вид, показанный на рис 3.

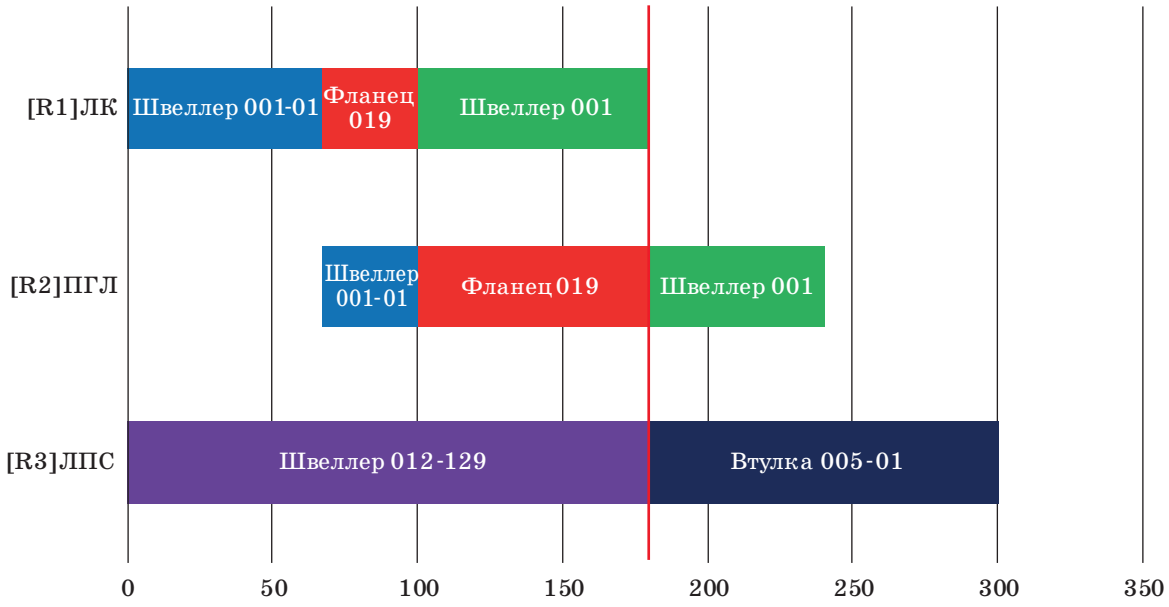
Пусть в момент времени T_{100} наступило событие завершения изготовления детали *Фланец* 019 на ресурсе R_1 и *Швеллер* 001-01 на ресурсе R_2. В результате более раннего завершения этих операций можно начать последующие операции раньше. Новая версия плана приведена на рис. 4.



■ ***Рис. 3.*** Реакция системы планирования на событие завершения изготовления T_{70}
■ ***Fig. 3.*** The diagram of scheduling system reaction for the completion event of T_{70}



■ ***Рис. 4.*** Расписание после события перепланирования T_{100}
■ ***Fig. 4.*** The diagram of the schedule after the rescheduling event of T_{100}



■ ***Рис. 5.*** Стабилизированный план после учета событий перепланирования
■ ***Fig. 5.*** The diagram of the stabilized schedule after considering all events

Пусть в момент времени T_{180} наступило событие завершения изготовления детали *Швеллер* 001 на ресурсе R_1 и детали *Фланец* 019 на ресурсе R_2, но обработка детали *Швеллер* 012-129 на ресурсе R_3 еще не завершена. В результате перестроим расписание на всех трех ресурсах и получим новую версию плана, представленную на рис. 5.

При сравнении версий плана, полученных после фазы начального планирования (в момент времени T_0) и по ходу его выполнения, можно сделать вывод, что операции распределены по ресурсам с обеспечением минимального простоя ресурсов по событиям реального времени.

Ожидаемые результаты

Применение рассмотренного метода позволит создавать интеллектуальные системы управления процессами производства в реальном времени. Ожидается получение экономического эф-

фекта за счет снижения простоев ресурсов в результате адаптивной перестройки сменно-суточных расписаний.

Систему, разработанную на основе предложенного подхода, используют на Самарском предприятии ООО «ПК «Электрум», где в результате внедрения была обеспечена равномерность загрузки оборудования и исполнителей, а также снижено на 10 % количество задержек выполнения производственных заказов. Ввиду отсутствия в предлагаемом подходе непосредственной зависимости от предметной области производства электротехнической продукции, описанный метод можно применять в других отраслях, где используются технологические операции, требующие решения аналогичных задач планирования.

Статья подготовлена на основе материалов научных исследований в рамках Госбюджетной темы ИПУСС РАН № АААА-А16-116040410059-7 «Разработка и исследование моделей, методов и алгоритмов построения планов сменно-суточных заданий при производстве продукции в условиях неопределенности и высокой динамики изменений производственной обстановки».

Литература

1. **Framinan J. M., Leisten R., García R. R.** Manufacturing scheduling systems: an integrated view on models, methods and tools. London, Springer. 2014. 400 p.
2. **Leung J. Y. T.** (ed.). Handbook of scheduling: algorithms, models, and performance analysis. CRC Press. 2004, 1216 p.
3. **Pinedo M., Zacharias C., Zhu N.** Scheduling in the service industries: an overview //Journal of systems science and systems engineering. 2015. Vol. 24. № 1. P. 1–48.
4. **Sule D. R.** Production planning and industrial scheduling: examples, case studies and applications. CRC press. 2007, 560 p.
5. **Pinedo M.** Scheduling. Theory, algorithms, and systems. Springer. 2016, 676 p.
6. **Chapman S. N.** The fundamentals of production planning and control. Prentice Hall. 2006, 272 p.
7. **Driessel R., Mönch L.** Variable neighborhood search approaches for scheduling jobs on parallel machines with sequence-dependent setup times, precedence constraints, and ready times // Computers & industrial engineering. 2011, vol. 61, № 2, pp. 336–345.
8. **Mehrabi M. G., Ulsoy A. G., Koren Y.** Reconfigurable manufacturing systems: key to future manufacturing // Journal of intelligent manufacturing. 2000, vol. 11, № 4, pp. 403–419.
9. **Лазарев А. А., Гафаров Е. Р.** Теория расписаний. Задачи и алгоритмы. М.: Физический факультет МГУ, 2011. 222 с.
10. **Wooldridge M.** An introduction to multiagent systems. John Wiley & Sons. 2009, 484 p.
11. **Jennings N. R., Wooldridge M. J.** (ed.). Agent technology: foundations, applications, and markets. Springer Science & Business Media. 2012, 325 p.
12. **Meisels A.** Distributed search by constrained agents: algorithms, performance, communication. Springer Science & Business Media. 2008, 216 p.
13. **Виттих В. А., Моисеева Т. В., Скобелев П. О.** Принятие решений на основе консенсуса с применением мультиагентных технологий // Онтология проектирования. 2013. № 2 (8). С. 20–25.
14. **Скобелев П. О.** Интеллектуальные системы управления ресурсами в реальном времени: принципы разработки, опыт промышленных внедрений и перспективы развития // Приложение к теоретическому и прикладному научно-техническому журналу «Информационные технологии. 2013. № 1. С. 1–32.
15. **Skobelev P.,** et al. Practical approach and multi-agent platform for designing real time adaptive scheduling systems // Proceedings of the XII International Conference on Practical Applications of Agents and Multi-Agent Systems (PAAMS 2014), 4–6 June, 2014, Salamanca, Spain. CCIS 0430. Spinger. 2014, pp. 1–12.
16. **Skobelev P.** Systems for real time adaptive resource management. In industrial agents: emerging applications of software agents in industry. Paulo Leitão, Stamatis Karnouskos (ed.). Elsevier. 2015, pp. 207–230.
17. **Skobelev P. O., Lakhin O. I., Polnikov A. S., Simonova E. V.** Approach to the solution of aerospace product lifecycle management problem based on network-centric principles // V. Marik, et al. (eds). Proceedings of the 7th International Conference on Industrial Applications of Holonic and Multi-Agent Systems (HoloMAS'2015), September 2–4, 2015, Valencia, Spain. LNAI 9266. Springer. 2015, pp. 169–178. doi:10.1007/978-3-319-22867-9_15
18. **Скобелев П. О.** и др. Оперативное управление ресурсами цехов предприятий на основе мультиагентного подхода // Труды XIX Международной конференции «Проблемы управления и моделирования в сложных системах», Самара, 12–15 сентября 2017 г. Самара: ОФОРТ, 2017. С. 474–485.
19. **Rzevski G., Skobelev P.** Managing complexity. Wit Press, 2014. 216 p.
20. **Baclawski K.,** et al. Коммюнике Онтологического Саммита 2017 — Искусственный интеллект, машинное обучение, логический вывод и онтологии // Онтология проектирования. 2017. Т. 7. № 2(24). С. 227–238.

UDC 338.984
doi:10.31799/1684-8853-2018-5-112-119

Multi-agent method of constructing daily-shift schedule for real-time industrial resource management

A. N. Lada[a, b], Project Manager, orcid.org/0000-0002-1022-7212, lada@kg.ru
I. V. Mayorov[a, b], Analyst, orcid.org/0000-0002-3130-8142
[a]Software Engineering Company «Smart Transport Systems», 17, Moscovskoe Sh., 443013, Samara, Russian Federation
[b]Institute for the Control of Complex Systems of Russian Academy of Sciences, 61, Sadovaya St., 443020, Samara, Russian Federation

Introduction: Coordinated production resource management of enterprises manufacturing electrical products is a complex problem with a high level of complexity due to the variety of types of resources used, and the dependence of production processes on a variety of factors and conditions. The traditional scheduling methods are not efficient enough for this problem. **Purpose:** Developing a method for rapid scheduling of electrical equipment production, based on the multi-agent approach. **Results:** A multi-agent method is developed for constructing daily-shift schedule for managing production resources of an enterprise manufacturing electrical products, based on the subject area ontology, given criteria, preferences and limitations. The orders consist of related operations described by technological cards. In the iterative messaging process, the scheduler agents improve the current values of load uniformity criteria and minimize the execution time for building daily-shift jobs. The developed method of managing the production resources allows you to build schedules for performing related operations in a system of resources by real-time events. When executing an order pool at LLC «PC» Electrum» (Samara city), the uniformity of equipment load and performers was well maintained, and the number of delays in fulfilling the orders was reduced by 10 %. **Practical relevance:** The system developed based on the proposed method can work either autonomously or along with an existing system of warehouse accounting of materials and finished products. The approach is not limited to the described subject area, being applicable in other industries which require similar production tasks. The economic effect is expected to be obtained by reducing idle production resources and improving their efficiency.

Keywords — knowledge domain, electrical products, multi-agent methods, production resource management, production enterprise ontology, real-time scheduling.

Citation: Lada A. N., Mayorov I. V. Multi-agent method of constructing daily-shift schedule for real-time industrial resource management. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 5, pp. 112–119 (In Russian). doi:10.31799/1684-8853-2018-5-112-119

References

1. Framinan J. M., Leisten R., García R. R. *Manufacturing scheduling systems: an integrated view on models, methods and tools*. London, Springer, 2014, 400 p.
2. Leung J. Y. T. (ed.). *Handbook of scheduling: algorithms, models, and performance analysis*. CRC Press, 2004, 1216 p.
3. Pinedo M., Zacharias C., Zhu N. Scheduling in the service industries: an overview. *Journal of Systems Science and Systems Engineering*, 2015, vol. 24, no. 1, pp. 1–48.
4. Sule D. R. *Production planning and industrial scheduling: examples, case studies and applications*. CRC press, 2007, 560 p.
5. Pinedo M. *Scheduling. Theory, algorithms, and systems*. Springer, 2016, 676 p.
6. Chapman S. N. *The fundamentals of production planning and control*. Prentice Hall, 2006, 272 p.
7. Driessel R., Mönch L. Variable neighborhood search approaches for scheduling jobs on parallel machines with sequence-dependent setup times, precedence constraints, and ready times. *Computers & Industrial Engineering*, 2011, vol. 61, no. 2, pp. 336–345.
8. Mehrabi M. G., Ulsoy A. G., Koren Y. Reconfigurable manufacturing systems: key to future manufacturing. *Journal of intelligent manufacturing*, 2000, vol. 11, no. 4, pp. 403–419.
9. Lazarev A. A., Gafarov E. R. Theory of schedules. *Problemy i algoritmy* [Problems and algorithms]. Moscow, Physics Faculty of Moscow State University, 2011, 222 p. (In Russian).
10. Wooldridge M. *An introduction to multiagent systems*. John Wiley & Sons, 2009, 484 p.
11. Jennings N. R., Wooldridge M. J. (ed.). *Agent Technology: Foundations, Applications, and Markets*. Springer Science & Business Media, 2012, 325 p.
12. Meisels A. *Distributed Search by Constrained Agents: algorithms, performance, communication*. Springer Science & Business Media, 2008, 216 p.
13. Vittikh V. A., Moiseeva T. V., Skobelev P. O. Making decisions on the basis of consensus using multi-agent technologies. *Ontologiya proektirovaniya* [Ontology of Designing], 2013, no. 2(8), pp. 20–25 (In Russian).
14. Skobelev P. Intellectual resource management systems in real time: the principles of development, the experience of industrial applications and development prospects. *Prilozhenie k teoreticheskomy i prikladnomy naychno-tekhicheskomy zhyrnaly "Informacionnye tekhologii"* [An appendix to the theoretical and applied scientific and technical journal Information Technologies], 2013, no. 1, pp. 1–32 (In Russian).
15. Skobelev P., et al. Practical approach and multi-agent platform for designing real time adaptive scheduling systems [*Proc. of the XII Int. Conf. on practical applications of agents and multi-agent systems* (PAAMS 2014)], CCIS 0430, Spinger, 2014, pp. 1–12.
16. Skobelev P. Multi-agent systems for real time adaptive resource management. In: Industrial Agents: Leitão, P., Karnouskos, S. (ed.). *Emerging Applications of Software Agents in Industry*. Elsevier, 2015, pp. 207–230.
17. Skobelev P. O., Lakhin O. I., Polnikov A. S., Simonova E. V. Approach to the solution of aerospace product lifecycle management problem based on network-centric principles. V. Marik, et al. (eds). [*Proc. of the 7th Int. Conf. on industrial applications of holonic and multi-agent systems* (HoloMAS'2015)], LNAI 9266. Springer, 2015, pp. 169–178. doi:10.1007/978-3-319-22867-9_15
18. Skobelev P. O., et al. Operational management of the resources of the enterprises' shops on the basis of the multi-agent approach. *Trudy XIX mezdynarodnoj konferencii "Problemy upravlenija I modelirovanija v slozhnykh sistemakh* [Proc. of the XIX Int. Conf. "Problems of control and modeling in complex systems" (PCMCS 2017)]. Samara, OFORT, 2017, pp. 474–485 (In Russian).
19. Rzevski G., Skobelev P. *Managing complexity*. Wit Press, 2014, 216 p.
20. Baclawski K., et al. Communique of the Ontological Summit 2017 — Artificial Intelligence, machine learning, inference and ontology. *Ontologiya proektirovaniya* [Ontology of designing], 2017, vol. 7, no. 2(24), pp. 227–238 (In Russian).

СВЕДЕНИЯ ОБ АВТОРАХ

БЛАУНШТЕЙН Натан Шаевич

Профессор Иерусалимского технологического института, профессор-эмиритус кафедры систем связи инженерного факультета Негевского университета им. Бен-Гуриона, Беэр-Шева, Израиль.
В 1972 году окончил Томский государственный университет по специальности «Радиофизика и электроника, включая квантовую».
В 1991 году защитил диссертацию на соискание ученой степени доктора физико-математических наук.
Является автором около 200 научных публикаций, в том числе 12 монографий, пяти патентов и трех изобретений.
Область научных интересов — радиофизика, системы проводной и беспроводной связи, радары, оптика и лидары.
Эл. адрес: nathan.blaunstein@hotmail.com

ЖЕЛЕЗНЫ Милош

Доцент, заместитель декана по концепции обучения и педагогической деятельности Западночешского университета, Пльзень, Чехия.
В 1994 году окончил магистратуру Западночешского университета, Пльзень, Чехия, по специальности «Кибернетика и техника управления».
В 2002 году защитил диссертацию на соискание доктора наук (PhD).
Является автором 70 научных публикаций.
Область научных интересов — человеко-машинное взаимодействие, обработка аудиовизуальных сигналов, распознавание образов.
Эл. адрес: zelezny@kky.zcu.cz

КОНОВАЛОВ Александр Сергеевич

Профессор кафедры управления и информатики в технических системах Санкт-Петербургского государственного университета аэрокосмического приборостроения, почетный работник высшего профессионального образования РФ.
В 1968 году окончил Ленинградский институт авиационного приборостроения по специальности «Электрооборудование летательных аппаратов».
В 1998 году защитил диссертацию на соискание ученой степени доктора технических наук.
Является автором более 200 научных публикаций.
Область научных интересов — автоматическое и интеллектуальное управление.
Эл. адрес: Konovalov@c4t.com

БУРАКОВ Михаил Владимирович

Доцент кафедры управления и информатики в технических системах Санкт-Петербургского государственного университета аэрокосмического приборостроения.
В 1984 году окончил Ленинградский институт авиационного приборостроения по специальности «Автоматизированные системы управления».
В 1994 году защитил диссертацию на соискание ученой степени кандидата технических наук.
Является автором более 150 научных публикаций.
Область научных интересов — системы интеллектуального управления, нечеткие регуляторы, нейронные сети, эволюционные алгоритмы.
Эл. адрес: bmv@sknt.ru

ЖУРАВЛЕВ Максим Олегович

Доцент кафедры физики открытых систем Саратовского государственного университета им. Н. Г. Чернышевского и кафедры геоэкологии и инженерной геологии СГТУ им. Гагарина Ю. А.
В 2011 году окончил Саратовский государственный университет им. Н. Г. Чернышевского по специальности «Физик».
В 2014 году защитил диссертацию на соискание ученой степени кандидата физико-математических наук.
Является автором более 20 научных публикаций.
Область научных интересов — приложение вейвлетного анализа к задачам нелинейной динамики, классическая и хаотическая синхронизация колебаний в распределенных системах электронно-волновой природы и др.
Эл. адрес: zhuravlevmo@gmail.com

КУЛАНИН Роман Анатольевич

Научный сотрудник научно-образовательного центра «Нелинейная динамика сложных систем» Саратовского государственного технического университета им. Гагарина Ю. А.
В 2010 году окончил Саратовский государственный социально-экономический университет по специальности «Производственный менеджмент».
Является автором пяти научных публикаций.
Область научных интересов — экспериментальное исследование активности головного мозга при выполнении когнитивных задач, обработка, моделирование и статистический анализ данных биологической природы.
Эл. адрес: rkulanin2010@yandex.ru

ЛАДА
Александр
Николаевич

Генеральный директор ООО «НПК «Интеллектуальные Транспортные Системы», Самара.
В 2006 году окончил Самарский государственный технический университет по специальности «Прикладная математика и информатика».
Является автором 30 научных публикаций.
Область научных интересов — мультиагентные технологии для создания интеллектуальных систем управления ресурсами в реальном времени, использующих принципы самоорганизации и эволюции.
Эл. адрес: lada@kg.ru

ЛЕБЕДЕВА
Наталья
Валерьевна

Ассистент кафедры динамики и управления полетом ракет и космических аппаратов МГТУ имени Н. Э. Баумана, Москва, инженер ПАО РКК «Энергия» им. С. П. Королёва, Королёв.
В 2008 году окончила Московский государственный технический университет им. Н. Э. Баумана.
Является автором более 20 научных публикаций.
Область научных интересов — управление летательными аппаратами, методы обработки телеметрической информации.
Эл. адрес: trigonella@mail.ru

МАКСИМЕНКО
Владимир
Александрович

Доцент Саратовского государственного технического университета им. Гагарина Ю. А., старший научный сотрудник научно-образовательного центра «Нелинейная динамика сложных систем».
В 2012 году окончил Саратовский государственный университет им. Н. Г. Чернышевского по специальности «Физика открытых нелинейных систем».
В 2015 году защитил диссертацию на соискание ученой степени кандидата физ.-мат. наук.
Является автором более 50 научных публикаций.
Область научных интересов — анализ устойчивости динамических режимов, реализующихся в пространственно-распределенных системах различной природы и др.
Эл. адрес:
maximenkovl@gmail.com

ЛАШКОВ
Игорь
Борисович

Младший научный сотрудник лаборатории интегрированных систем автоматизации Санкт-Петербургского института информатики и автоматизации РАН.
В 2014 году окончил с отличием Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики по специальности «Информационные системы и технологии».
Является автором 12 научных публикаций.
Область научных интересов — управление знаниями, профилирование, базы данных, мобильные технологии, облачные технологии.
Эл. адрес: igor-lashkov@ya.ru

МАЙОРОВ
Игорь
Владимирович

Руководитель отдела разработки математических моделей, методов и алгоритмов ООО «НПК «Разумные решения», Самара.
В 1989 году окончил физический факультет Куйбышевского государственного университета по специальности «Физика».
В 2017 году защитил диссертацию на соискание ученой степени кандидата технических наук.
Является автором более 40 научных публикаций.
Область интересов — системы искусственного интеллекта, мультиагентные системы, теория оптимизации, дискретная математика.
Эл. адрес: imayorov@kg.ru

МАЛЬЦЕВ
Георгий
Николаевич

Профессор кафедры космических радиотехнических систем Военно-космической академии им. А. Ф. Можайского, Санкт-Петербург, заслуженный деятель науки РФ, действительный член Академии космонавтики им. К. Э. Циолковского.
В 1980 году окончил Военный инженерный Краснознаменный институт им. А. Ф. Можайского по специальности «Радиотехнические системы комплексов».
В 1994 году защитил диссертацию на соискание ученой степени доктора технических наук.
Является автором более 300 научных публикаций и 27 патентов на изобретения.
Область научных интересов — обработка сигналов в радиотехнических и оптико-электронных информационных системах и др.
Эл. адрес:
georgy_maltsev@mail.ru

ОСТРОУМОВ Олег Александрович

Адъюнкт кафедры общепрофессиональных дисциплин Военной академии связи им. Маршала Советского Союза С. М. Буденного, Санкт-Петербург.
В 2009 году окончил Ставропольский военный институт связи РВ по специальности «Радиосвязь, радиовещание и телевидение».
Является автором трех научных публикаций.
Область научных интересов — математическое моделирование каналов связи, помехоустойчивость сигналов, многомерные сигнальные конструкции, разнесенный прием.
Эл. адрес: oleg-26@mail.ru

ПИСАРЧИК Александр Николаевич

Профессор, ведущий научный сотрудник Центра биомедицинских технологий Технического университета Мадрида, академический редактор журнала PLoS One и др.
В 1976 году окончил Белорусский государственный университет по специальности «Физика».
В 1990 году защитил диссертацию на соискание ученой степени кандидата физико-математических наук.
Является автором более 300 научных публикаций и пяти патентов на изобретения.
Область научных интересов — нелинейная динамика, хаос, синхронизация, нейронная динамика, мультистабильность, стохастические эффекты, анализ нейрофизиологических данных и интерфейсы мозг-компьютер.
Эл. адрес:
alexander.Pisarchik@ctb.upm.es

РУННОВА Анастасия Евгеньевна

Докторант, доцент кафедры автоматизации, управления, мехатроники Саратовского государственного технического университета им. Гагарина Ю. А., старший научный сотрудник Научно-образовательного центра «Нелинейная динамика сложных систем».
В 2005 году окончила Саратовский государственный университет им. Н. Г. Чернышевского по специальности «Физика».
В 2008 году защитила диссертацию на соискание ученой степени кандидата физико-математических наук.
Является автором более 40 научных публикаций.
Область научных интересов — теория динамических систем, нейрофизиология, методы обработки данных.
Эл. адрес: anefila@gmail.com

ПАСТУШОК Игорь Анатольевич

Старший преподаватель кафедры инфокоммуникационных систем Санкт-Петербургского государственного университета аэрокосмического приборостроения.
В 2014 году окончил Санкт-Петербургский государственный университет аэрокосмического приборостроения по специальности «Комплексная защита объектов информатизации».
Является автором 12 научных публикаций.
Область научных интересов — математическая оптимизация, беспроводные сети, алгоритмы распределения ресурсов, теория вероятности, имитационное моделирование.
Эл. адрес: i.pastushok@vu.spb.ru

ПРОТАСОВ Павел Александрович

Младший научный сотрудник Саратовского государственного технического университета имени Гагарина Ю. А.
В 2013 году окончил Саратовский государственный университет им. В. И. Разумовского по специальности «Клиническая психология».
Является автором пяти научных публикаций.
Область научных интересов — теория разрядных вычислений, методы проектирования спецпроцессоров для систем контроля и управления, оптико-информационные системы.
Эл. адрес:
protasovpa213@gmail.com

СЕРГЕЕВ Александр Михайлович

Старший преподаватель кафедры вычислительных систем и сетей Санкт-Петербургского государственного университета аэрокосмического приборостроения.
В 2004 году окончил Санкт-Петербургский государственный университет аэрокосмического приборостроения по специальности «Вычислительные машины, комплексы, системы и сети».
Является автором 32 научных публикаций.
Область научных интересов — численные методы, теория вычислительных процессов, проектирование специализированных процессоров.
Эл. адрес: asklab@mail.ru

СИНЮК
Александр
Демьянович

Доцент, профессор кафедры общепрофессиональных дисциплин Военной академии связи им. Маршала Советского Союза С. М. Буденного, Санкт-Петербург.
В 1989 году окончил Полтавское высшее военное командное училище связи им. К. С. Москаленко по специальности «Командная тактическая войск связи».
В 2014 году защитил диссертацию на соискание ученой степени доктора технических наук.
Является автором более 170 научных публикаций и десяти патентов на изобретения.
Область научных интересов — общая теория связи, теория информации, методы защиты информации, подсистемы управления ключами криптографических систем.
Эл. адрес: eentrop@rambler.ru

ТАТАРНИКОВА
Татьяна
Михайловна

Профессор кафедры безопасности информационных систем Санкт-Петербургского государственного университета аэрокосмического приборостроения.
В 1993 году окончила Восточно-Сибирский технологический институт по специальности «Электронно-вычислительные машины, комплексы, системы и сети».
В 2007 году защитила диссертацию на соискание ученой степени доктора технических наук.
Является автором более 100 научных публикаций.
Область научных интересов — инфокоммуникации, взаимодействие неоднородных сетей.
Эл. адрес: tm-tatarn@yandex.ru

ХРАМОВ
Александр
Евгеньевич

Заведующий кафедрой автоматизации, управления, мехатроники Саратовского государственного технического университета им. Гагарина Ю. А., профессор кафедры электроники, колебаний и волн Саратовского государственного университета им. Н. Г. Чернышевского.
В 1996 году окончил Саратовский государственный университет им. Н. Г. Чернышевского.
В 2005 году защитил диссертацию на соискание ученой степени доктора физико-математических наук.
Является автором более 200 научных публикаций.
Область научных интересов — нейронаука, теория сложных сетей, вейвлет-анализ и его приложения в нелинейной динамике и нейронауке, нейроинтерфейсы.
Эл. адрес: hramovae@gmail.com

СОЛОВЬЕВ
Сергей
Владимирович

Доцент кафедры динамики и управления полетом ракет и космических аппаратов Московского государственного технического университета имени Н. Э. Баумана, Москва, ведущий конструктор ПАО РКК «Энергия» имени С. П. Королёва, Королев.
В 1993 году окончил Московский государственный технический университет им. Н. Э. Баумана по специальности «Ракетные двигатели».
В 1996 году защитил диссертацию на соискание ученой степени кандидата технических наук.
Является автором более 30 научных публикаций.
Область научных интересов — проектирование автоматических космических аппаратов, управление космическим полетом, интеллектуальный анализ данных.
Эл. адрес:
sergey.soloviev@scsc.ru

ТАУБИН
Феликс
Александрович

Профессор кафедры информационных систем Санкт-Петербургского государственного университета аэрокосмического приборостроения.
В 1968 году окончил Ленинградский институт авиационного приборостроения по специальности «Радиоэлектронные устройства систем управления».
В 1992 году защитил диссертацию на соискание ученой степени доктора технических наук.
Является автором более 80 научных публикаций, девяти патентов и двух изобретений.
Область научных интересов — цифровые системы связи, методы помехоустойчивого кодирования, широкополосные системы, беспроводные сети.
Эл. адрес: ftaubin@yahoo.com

ЧОЛАК
Парт

Аспирант Технического университета Мадрида.
В 2017 году получил диплом бакалавра Индийского института технологии по специальности «Механическая инженерия», в 2018 году — степень магистра по специальности «Биоинженерные науки» в Техническом университете Мадрида.
Является автором пяти научных публикаций.
Область научных интересов — анализ нейрофизиологических сигналов и выявление биомаркеров, связанных с памятью, болезнью Альцгеймера, воображаемыми движения, принятием решений.
Эл. адрес:
parthchholak@gmail.com

ШТАНЬКО
Сергей
Владимирович

Докторант кафедры космических радиотехнических систем Военно-космической академии имени А. Ф. Можайского, Санкт-Петербург.
В 1999 году окончил Военный инженерно-космический университет им. А. Ф. Можайского по специальности «Радиоэлектронные системы».
В 2004 году защитил диссертацию на соискание ученой степени кандидата технических наук.
Является автором 46 научных публикаций и восьми учебных изданий.
Область научных интересов — защита информации в радиотехнических комплексах управления, сбора и передачи информации.
Эл. адрес: craft2001@mail.ru

ЯКИМОВ
Виктор
Леонидович

Докторант кафедры космических радиотехнических систем Военно-космической академии им. А. Ф. Можайского, Санкт-Петербург.
В 2000 году окончил Военный инженерно-космический институт им. А. Ф. Можайского по специальности «Радиоэлектронные системы космических аппаратов».
В 2005 году защитил диссертацию на соискание ученой степени кандидата технических наук.
Является автором более 50 научных публикаций.
Область научных интересов — моделирование сложных систем, методы обработки телеметрической информации, техническая диагностика, методы обработки изображений.
Эл. адрес: yakim78@yandex.ru

ШУМСКАЯ
Ольга
Олеговна

Аспирант, младший научный сотрудник лаборатории автономных робототехнических систем Санкт-Петербургского института информатики и автоматизации РАН.
В 2018 году окончила Томский государственный университет систем управления и радиоэлектроники по специальности «Информационно-аналитические системы безопасности».
Является автором 15 научных публикаций, двух патентов на программы для ЭВМ и одного патента на базу данных.
Область научных интересов — стеганография, стегоанализ, цифровые объекты, методы классификации.
Эл. адрес:
shumskaya.oo@gmail.com