

ИНФОРМАЦИОННО- УПРАВЛЯЮЩИЕ СИСТЕМЫ

НАУЧНЫЙ ЖУРНАЛ

1(92)/2018

INFORMATSIONNO- UPRAVLIAYUSHCHIE SISTEMY (INFORMATION AND CONTROL SYSTEMS)

Founder
«Information and Control Systems», Ltd.

Publisher
Saint-Petersburg State University
of Aerospace Instrumentation

Editor-in-Chief
M. Sergeev
Dr. Sc., Tech., Professor, St. Petersburg, Russia

Deputy Editor-in-Chief
E. Krouk
Dr. Sc., Tech., Professor, Moscow, Russia

Executive secretary
O. Muravtsova

Editorial Council
C. Christodoulou
PhD, Professor, Albuquerque, New Mexico, USA
L. Chubraeva
RAS Corr. Member, Dr. Sc., Tech., Professor, St. Petersburg, Russia
L. Fortuna
PhD, Professor, Catania, Italy
A. Fradkov
Dr. Sc., Tech., Professor, St. Petersburg, Russia
V. Kozlov
Dr. Sc., Tech., Professor, St. Petersburg, Russia
B. Meyer
Dr. Sc., Professor, Zurich, Switzerland
A. Ovodenko
Dr. Sc., Tech., Professor, St. Petersburg, Russia
Y. Podoplyokin
Dr. Sc., Tech., Professor, St. Petersburg, Russia
Yu. Shokin
RAS Academician, Dr. Sc., Phys.-Math., Novosibirsk, Russia
V. Simakov
Dr. Sc., Tech., Professor, Moscow, Russia
V. Vasilev
RAS Corr. Member, Dr. Sc., Tech., Professor, St. Petersburg, Russia
R. Yusupov
RAS Corr. Member, Dr. Sc., Tech., Professor, St. Petersburg, Russia

Editorial Board
V. Anisimov
Dr. Sc., Tech., Professor, St. Petersburg, Russia
B. Bezruchko
Dr. Sc., Phys.-Math., Saratov, Russia
N. Blaunstein
Dr. Sc., Phys.-Math., Professor, Beer-Sheva, Israel
A. Dudin
Dr. Sc., Tech., Professor, Minsk, Belarus
I. Dumér
PhD., Professor, Riverside, USA
V. Khimenko
Dr. Sc., Tech., Professor, St. Petersburg, Russia
G. Mal'tsev
Dr. Sc., Tech., Professor, St. Petersburg, Russia
G. Matvienko
Dr. Sc., Phys.-Math., Professor, Tomsk, Russia
V. Melekhin
Dr. Sc., Tech., Professor, St. Petersburg, Russia
A. Shalyto
Dr. Sc., Tech., Professor, St. Petersburg, Russia
A. Shelupanov
Dr. Sc., Tech., Professor, Tomsk, Russia
A. Shepeta
Dr. Sc., Tech., Professor, St. Petersburg, Russia
A. Smirnov
Dr. Sc., Tech., Professor, St. Petersburg, Russia
Z. Yuldashev
Dr. Sc., Tech., Professor, St. Petersburg, Russia
A. Zeifman
Dr. Sc., Phys.-Math., 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

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

INFORMATION PROCESSING AND CONTROL

- Karasev E. Yu., Vatamaniuk I. V., Saveliev A. I., Ronzhin A. L.**
*Architectural Solutions for Integrating a Video Conferencing Module into
Cyberphysical Intelligent Space* 2
- Krasilnikov N. N., Krasilnikova O. I.** *Providing Compatibility for 2D and 3D
Television Broadcasting Systems* 11

INFORMATION AND CONTROL SYSTEMS

- Gorodetskiy A. E., Kurbanov V. G., Tarasova I. L.** *Decision-making in
Central Nervous System of a Robot* 21
- Martynova L. A.** *Underwater Observation under Intellectual
Interference* 31

SYSTEM AND PROCESS MODELING

- Khorev V. S., Ishbulatov J. M., Lapsheva E. E., Kiselev A. R.,
Gridnev V. I., Bezruchko B. P., Butenko A. A., Ponomarenko V. I.,
Karavaev A. S.** *Diagnostics of Directional Coupling between Blood
Circulation Regulation Loops using Analysis of Time Series of
Mathematical Model of Human Cardiovascular System* 42
- Maltsev G. N., Yakimov V. L.** *Reliability of Multi-Stage Control over
Technical Condition of Tested Objects* 49

INFORMATION SECURITY

- Fedorchenko A. V., Kotenko I. V.** *Correlation of Information in
SIEM Systems based on Event Type Relation Graph* 58

INFORMATION CODING AND TRANSMISSION

- Engelberg S., Keren O.** *Error-correcting Codes for Ternary Content
Addressable Memories: a New Perspective* 68

INFORMATION CHANNELS AND MEDIUM

- Blaunstein N. S., Sergeev M. B., Nenashev V. A.** *Impact of Built-up Terrain
on Operational Parameters of Signals in Land-Satellite Communication
Links* 74
- Blaunstein N. S., Ben-Shimol Y.** *Prediction of Operational Parameters of
Radio Signals Passing a Land-Satellite Link through StormTime
Ionosphere* 85

INFORMATION AND MEASURING SYSTEMS

- Inochkin F. M., Kruglov S. K., Bronshtein I. G.** *Breaking the Diffraction
Limit in Optical Edge Estimation Problem* 96

CONTROL IN MEDICAL AND BIOLOGICAL SYSTEMS

- Runnova A. E., Maksimenko V. A., Pchelintseva S. V., Kulanin R. A.,
Hramov A. E.** *Wavelet Method for Motor Activity Pattern Analysis on Experimen-
tal Data from Human Multichannel Electroencephalography
for External Device Control* 106

CONTROL IN SOCIAL AND ECONOMIC SYSTEMS

- Suvorova A. V., Tulup'yev A. L.** *Bayesian Belief Network Structure
Synthesis for Risky Behavior Rate Estimation* 116

BRIEF SCIENTIFIC REPORTS

- Lyandres V.** *Model of Deep Fading* 123

INFORMATION ABOUT THE AUTHORS

128

Submitted for publication 09.01.18. Passed for printing 20.02.18. 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.

1(92)/2018

РЕЦЕНЗИРУЕМОЕ ИЗДАНИЕ

ИНФОРМАЦИОННО-
УПРАВЛЯЮЩИЕ
СИСТЕМЫУчредитель
ООО «Информационно-управляющие системы»Издатель
Санкт-Петербургский государственный университет
аэрокосмического приборостроенияГлавный редактор
М. Б. Сергеев,
д-р техн. наук, проф., С.-Петербург, РФЗам. главного редактора
Е. А. Крук,
д-р техн. наук, проф., Москва, РФОтветственный секретарь
О. В. МуравцоваРедакционный совет:
Председатель А. А. Оводенко,
д-р техн. наук, проф., С.-Петербург, РФ
В. Н. Васильев,
чл.-корр. РАН, д-р техн. наук, проф., С.-Петербург, РФ
В. Н. Козлов,
д-р техн. наук, проф., С.-Петербург, РФ
К. Кристофолу,
д-р наук, проф., Альбукерке, Нью-Мексико, США
Б. Мейер,
д-р наук, проф., Цюрих, Швейцария
Ю. Ф. Подоллэкин,
д-р техн. наук, проф., С.-Петербург, РФ
В. В. Симаков,
д-р техн. наук, проф., Москва, РФ
Л. Фортуна,
д-р наук, проф., Катания, Италия
А. Л. Фрадков,
д-р техн. наук, проф., С.-Петербург, РФ
Л. И. Чубраева,
чл.-корр. РАН, д-р техн. наук, С.-Петербург, РФ
Ю. И. Шокин,
акад. РАН, д-р физ.-мат. наук, проф., Новосибирск, РФ
Р. М. Юсупов,
чл.-корр. РАН, д-р техн. наук, проф., С.-Петербург, РФРедакционная коллегия:
В. Г. Анисимов,
д-р техн. наук, проф., С.-Петербург, РФ
Б. П. Безручко,
д-р физ.-мат. наук, проф., Саратов, РФ
Н. Блаунштейн,
д-р физ.-мат. наук, проф., Беэр-Шева, Израиль
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
Г. Н. Мальцев,
д-р техн. наук, проф., С.-Петербург, РФ
Г. Г. Матвиенко,
д-р физ.-мат. наук, проф., Томск, РФ
В. Ф. Мелехин,
д-р техн. наук, проф., С.-Петербург, РФ
А. В. Смирнов,
д-р техн. наук, проф., С.-Петербург, РФ
В. И. Хименко,
д-р техн. наук, проф., С.-Петербург, РФ
А. А. Шалыто,
д-р техн. наук, проф., С.-Петербург, РФ
А. А. Шелупанов,
д-р техн. наук, проф., Томск, РФ
А. П. Шепета,
д-р техн. наук, проф., С.-Петербург, РФ
З. М. Юлдашев,
д-р техн. наук, проф., С.-Петербург, РФРедактор: А. Г. Ларионова
Корректор: Т. В. Звертановская
Дизайн: М. Л. Черненко, Ю. В. Умницына
Компьютерная верстка: Ю. В. УмницынаАдрес 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.ruЖурнал 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.Журнал входит в «Перечень ведущих рецензируемых научных журналов и изданий,
в которых должны быть опубликованы основные научные результаты диссертации
на соискание ученой степени доктора и кандидата наук».
© Коллектив авторов, 2018

ОБРАБОТКА ИНФОРМАЦИИ И УПРАВЛЕНИЕ

Карасёв Е. Ю., Ватаманюк И. В., Савельев А. И., Ронжин А. Л.
Архитектурные решения интеграции модуля видео-конференц-связи
в киберфизическое интеллектуальное пространство 2**Красильников Н. Н., Красильникова О. И.** Проблема совместимости
систем телевизионного вещания 2D и 3D и метод ее решения 11

ИНФОРМАЦИОННО-УПРАВЛЯЮЩИЕ СИСТЕМЫ

Городецкий А. Е., Курбанов В. Г., Тарасова И. Л. Принятие
решений в центральной нервной системе робота 21**Мартынова Л. А.** Решение задачи подводного наблюдения
в условиях применения интеллектуальных помех 31

МОДЕЛИРОВАНИЕ СИСТЕМ И ПРОЦЕССОВ

**Хорев В. С., Ишбулатов Ю. М., Лапшева Е. Е., Киселев А. Р.,
Гриднев В. И., Безручко Б. П., Бутенко А. А., Пономаренко В. И.,
Караваев А. С.** Диагностика направленной связи контуров регуляции
кровообращения по временным рядам математической модели
сердечно-сосудистой системы человека 42**Мальцев Г. Н., Якимов В. Л.** Достоверность многоэтапного контроля
технического состояния объектов испытаний 49

ЗАЩИТА ИНФОРМАЦИИ

Федорченко А. В., Котенко И. В. Корреляция информации
в SIEM-системах на основе графа связей типов событий 58

КОДИРОВАНИЕ И ПЕРЕДАЧА ИНФОРМАЦИИ

Engelberg S., Keren O. Error-correcting Codes for Ternary Content
Addressable Memories: a New Perspective 68

ИНФОРМАЦИОННЫЕ КАНАЛЫ И СРЕДЫ

Blaunstein N. S., Sergeev M. B., Nenashv V. A. Impact of Built-up
Terrain on Operational Parameters of Signals in Land-Satellite
Communication Links 74**Blaunstein N. S., Ben-Shimol Y.** Prediction of Operational Parameters
of Radio Signals Passing a Land-Satellite Link through StormTime
Ionosphere 85

ИНФОРМАЦИОННО-ИЗМЕРИТЕЛЬНЫЕ СИСТЕМЫ

Иночкин Ф. М., Круглов С. К., Бронштейн И. Г. Преодоление дифрак-
ционного предела при оптических измерениях границ объектов 96

УПРАВЛЕНИЕ В МЕДИЦИНЕ И БИОЛОГИИ

**Руннова А. Е., Максименко В. А., Пчелинцева С. В., Куланин Р. А.,
Храмов А. Е.** Метод вейвлет-анализа паттернов двигательной
активности на экспериментальных данных многоканальной электро-
энцефалографии человека для управления внешними устройствами 106

УПРАВЛЕНИЕ В СОЦИАЛЬНО-ЭКОНОМИЧЕСКИХ СИСТЕМАХ

Суворова А. В., Тулупьев А. Л. Синтез структур байесовской сети
доверия для оценки характеристик рискованного поведения 116

КРАТКИЕ СООБЩЕНИЯ

Lyandres V. Model of Deep Fading 123

СВЕДЕНИЯ ОБ АВТОРАХ 128

Сдано в набор 09.01.18. Подписано в печать 20.02.18. Формат 60×84/8.
Гарнитура SchoolBook. Печать цифровая.
Усл. печ. л. 15,7. Уч.-изд. л. 21,5. Тираж 1000 экз (1-й завод 140 экз). Заказ № 51.
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.

АРХИТЕКТУРНЫЕ РЕШЕНИЯ ИНТЕГРАЦИИ МОДУЛЯ ВИДЕО-КОНФЕРЕНЦ-СВЯЗИ В КИБЕРФИЗИЧЕСКОЕ ИНТЕЛЛЕКТУАЛЬНОЕ ПРОСТРАНСТВО

Е. Ю. Карасёв^а, программист, gnomskg@gmail.com

И. В. Ватаманюк^а, аспирант, vatamaniuk@iias.spb.su

А. И. Савельев^а, канд. техн. наук, старший научный сотрудник, saveliev@iias.spb.su

А. Л. Ронжин^а, доктор техн. наук, профессор, ronzhin@iias.spb.su

^аСанкт-Петербургский институт информатики и автоматизации РАН, 14-я линия В.О., 39, Санкт-Петербург, 199178, РФ

Введение: киберфизические системы — широкий класс систем, основной характеристикой которых является высокая степень интеграции вычислительных ресурсов в каждый физический компонент системы. Одна из задач киберфизического интеллектуального пространства — обеспечение удобного взаимодействия пользователей друг с другом посредством различных узлов данной системы. Для этого современные приложения видео-конференц-связи должны не только использовать различные модальности: текст, аудио- и видеоданные, графические объекты, — но и соответствовать критериям модульности и кроссплатформенности. **Цель:** разработка архитектуры приложения видео-конференц-связи, предоставляющего сервис видео-конференц-связи пользователям киберфизического интеллектуального пространства. **Результаты:** предложены архитектурные решения модуля видео-конференц-связи, обеспечивающие возможность его интеграции в киберфизическую многомодальную информационно-навигационную облачную систему, а также обобщенная структура модуля видео-конференц-связи с реализацией частных и контролируемых публичных аккаунтов. Сетевое взаимодействие представлено в виде трех отдельных слоев, отвечающих за управление сетевым соединением с сервером, соединением пользователей, а также соединением для передачи мультимедийных данных участников видео-конференц-связи. Произведены замеры времени отклика соединения при организации видео-конференц-связи. В среднем отклик для двух пользователей составил 5,81 мс.

Ключевые слова — пиринговое взаимодействие, клиент-серверное взаимодействие, теле-конференц-связь, видео-конференц-связь, модульная архитектура программного обеспечения, киберфизические системы, интеллектуальное пространство.

Цитирование: Карасёв Е. Ю., Ватаманюк И. В., Савельев А. И., Ронжин А. Л. Архитектурные решения интеграции модуля видео-конференц-связи в киберфизическое интеллектуальное пространство // Информационно-управляющие системы. 2018. № 1. С. 2–10. doi:10.15217/issn1684-8853.2018.1.2

Citation: Karasev E. Yu., Vatamaniuk I. V., Saveliev A. I., Ronzhin A. L. Architectural Solutions for Integrating a Video Conferencing Module into Cyberphysical Intelligent Space. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 1, pp. 2–10 (In Russian). doi:10.15217/issn1684-8853.2018.1.2

Введение

Видео-конференц-связь (ВКС) сегодня востребована как государственными министерствами и ведомствами, к примеру, арбитражным и гражданским судом РФ [1], медицинскими организациями [2], образовательными учреждениями [3], так и крупными корпорациями, общественно-политическими организациями — теми институтами, которым необходим простой и удобный инструмент, позволяющий оперативно проводить совещания с удаленными коллегами [4]. Внедрение модуля ВКС в распределенную киберфизическую систему (КФС) предприятия либо организации существенно расширяет возможности пользователей, а также предоставляет дополнительные инструменты для контроля безопасности и управления рабочим процессом.

Кроссплатформенные веб-технологии позволяют адаптировать работу приложения ВКС под

различные программно-аппаратные платформы и обеспечивать контроль над обычными и управляемыми аккаунтами в единообразной адаптируемой среде в режиме реального времени. Специализированные системы совместной работы, помимо передачи аудиовизуальных потоков данных, захватываемых микрофонами и видеокамерами, обеспечивают обмен дополнительной мультимедийной информацией, включая [5]:

- совместный доступ к экрану или отдельным приложениям (screen sharing);
- интерактивную доску (whiteboard);
- демонстрацию презентаций;
- синхронный просмотр веб-страниц (co-browsing);
- аннотацию экрана;
- мониторинг присутствия участников;
- текстовый чат;
- интегрированную VoIP-связь;
- видео-конференц-связь;

- возможность менять ведущего;
- возможность отдавать контроль над мышью и клавиатурой;
- модерацию онлайн-встреч;
- обратную связь (например, опросы или оценки);
- планирование встреч и приглашение участников;
- запись хода веб-конференции.

Эффективная передача мультимедийных данных в реальном времени без потери пакетов — одна из немаловажных задач при организации ВКС. В статье [6] исследуется проблема пакетных передач от одного источника нескольким при организации ВКС в реальном времени. Схема передачи TRAding DELay for Reliability (TRADER), предложенная авторами, распределяет отправленные пакеты с учетом ограничений по задержке, что существенно улучшает качество передаваемого видеопотока с точки зрения отношения сигнал/шум.

В последние годы мультимедийные приложения на мобильных устройствах пользуются большой популярностью. Благодаря быстрому развитию технологий беспроводной связи пропускная способность мобильных соединений становится достаточной для поддержания потокового видео в реальном времени, что позволяет интегрировать пользовательские мобильные устройства в КФС. Однако из-за их ограниченных возможностей потоковая передача видео высокой четкости (HD) по-прежнему представляет собой непростую задачу. Зачастую мобильные клиенты имеют ограниченный размер экрана, относительно медленное сетевое соединение, а также ограниченное время работы от аккумулятора, поэтому при разработке приложений ВКС для мобильных устройств необходимо учитывать следующие аспекты [7].

1. Изменчивость сетевой среды: в отличие от проводной сети, мобильное соединение зачастую неустойчиво, а пропускная способность может значительно изменяться. Чтобы гарантировать бесперебойную и высококачественную доставку видео для мобильного клиента, система должна динамически регулировать скорость передачи видеопотока, отправленного клиенту.

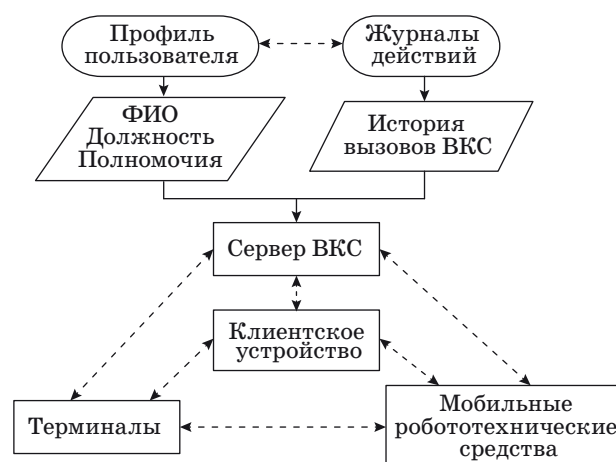
2. Разнообразие возможностей мобильных устройств: мобильные устройства очень разнообразны с точки зрения аппаратных и программных решений для обработки видео. У высокопроизводительных смартфонов часто есть аппаратный ускоритель для декодирования видео, тогда как большинство мобильных телефонов не могут поддерживать аппаратное декодирование. Для повышения производительности видеопроцессорной системы должна отправлять видео в подходящем разрешении в соответствии с возможностями клиентов.

При разработке модуля ВКС КФС следует учитывать эти особенности функционирования мобильных клиентских устройств, чтобы оптимизировать обработку мультимедийных данных соответствующим образом. Как правило, средства ВКС ориентированы на захват клиентским устройством аудиовидеоданных только одного пользователя, и чаще всего сеанс связи организован между двумя устройствами. Однако при проведении сеансов связи между несколькими пользователями такой подход зачастую недостаточен, и требуется многоканальная обработка аудиовизуальных сигналов на стороне каждого клиента [8]. В следующем разделе рассмотрены особенности разрабатываемого модуля ВКС, которые позволяют решить проблему обработки и хранения разнородной информации, поступающей в различные моменты времени, обеспечить модульность и простое взаимодействие логических частей КФС [9, 10].

Архитектурные решения модуля ВКС, разрабатываемого для КФС МИНОС

Рассмотрим структуру модуля ВКС, разрабатываемого в рамках многомодальной информационно-навигационной облачной системы (МИНОС) (рис. 1) [9, 10]. С точки зрения КФС обобщенный профиль пользователя ВКС состоит из следующих компонентов:

- профиля пользователя в системе (если пользователь зарегистрирован);
- журналов действий пользователя, включая историю его запросов, историю перемещений, историю проложенных маршрутов, историю вызовов ВКС;
- данных о его клиентском устройстве.



■ Рис. 1. Обобщенная структура модуля ВКС
 ■ Fig. 1. The general structure of the videoconferencing module

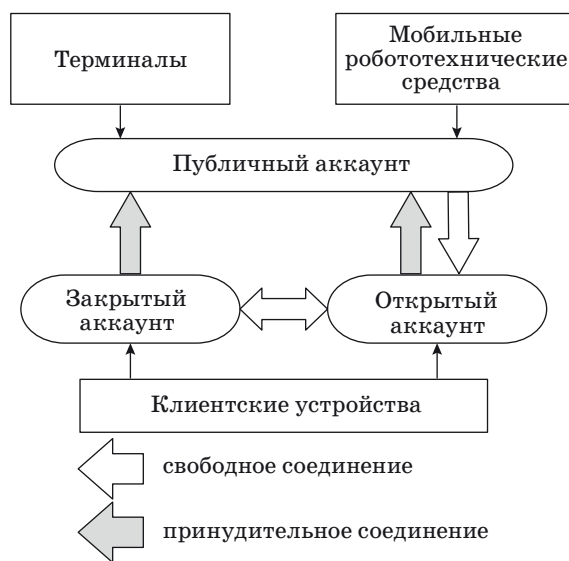
Киберфизическая система располагает о пользователе следующей информацией. Если пользователь зарегистрирован в системе, то в его профиле хранятся данные о его ФИО, его статусе (сотрудник, посетитель, прочее), его полномочиях в системе, а также, если это сотрудник, то о подразделении организации, в котором он работает, и занимаемой должности. Кроме того, в профиль может быть добавлена и другая информация, например, о дне рождения сотрудника — для показа поздравительных слайдов в модуле корпоративного телевидения, уведомлений сотрудников по ВКС и т. д.

Отталкиваясь от предложенной схемы, рассмотрим подробнее возможную реализацию модуля ВКС в рамках КФС. Модуль ВКС должен быть доступен на устройствах, интегрированных в КФС, включая терминалы, мобильные робототехнические платформы, рабочие компьютеры сотрудников, а также клиентские устройства пользователей. Объединение перечисленных устройств в единую локальную сеть позволит модулю ВКС функционировать в рамках организации независимо от сети Интернет, что может быть полезно при потере связи во внештатных ситуациях или для обеспечения безопасности предприятия.

Архитектура модуля ВКС пиринговая, что позволяет пользователям подключаться друг к другу напрямую, не используя сервер в качестве промежуточного звена связи. Однако в разрабатываемом решении для КФС серверная часть используется для хранения промежуточных данных, данных пользователей и соединения с другими облачными сервисами, в частности, для авторизации пользователей и обмена «сигнальными» данными. Подобная реализация передачи и хранения «сигнальных» данных на клиенте и сервере обеспечивает буферизацию и последующую обработку «сигнальных» данных, исключая их потерю и поддерживая взаимодействие между группами клиентов [11].

Разрабатываемый модуль ВКС предполагает наличие так называемых публичных аккаунтов (рис. 2), обладающих двумя важными особенностями. Во-первых, публичный аккаунт предоставляет доступ к ВКС незарегистрированным пользователям (с некоторыми ограничениями); во-вторых, публичные аккаунты являются контролируемыми, т. е. зарегистрированные пользователи могут принудительно к ним подключиться. Контролируемые аккаунты в архитектурном плане отличаются от обычных только взаимодействием с сервером. В случае с пиринговым соединением и передачей мультимедийных данных они подобны обычным аккаунтам [12].

Аккаунты пользователей подразделяются на публичные и персональные, последние в свою



■ Рис. 2. Типы и способы взаимодействия аккаунтов ВКС

■ Fig. 2. Types and ways of interactions of videoconferencing accounts

очередь могут быть открытыми или закрытыми. Так, публичные аккаунты по умолчанию установлены на терминалах КФС и мобильных платформах, предусматривающих возможность использования ВКС. Они не требуют регистрации в системе ВКС и предназначены для оперативной связи с сотрудниками, имеющими открытый персональный аккаунт. Персональные аккаунты зарегистрированных в системе пользователей могут соединяться друг с другом без ограничений (при наличии вызываемого пользователя онлайн и принятии им вызова), кроме того, они обладают возможностью установить принудительное соединение с любым публичным аккаунтом системы. В отличие от открытых персональных аккаунтов, закрытые не доступны для связи со стороны публичных аккаунтов. При желании зарегистрированный пользователь может войти в свою учетную запись через терминал или мобильную платформу, чтобы воспользоваться своим персональным аккаунтом.

Архитектура модуля ВКС

Реализация предложенной гибридной схемы организации сетевого соединения пользователей приложения ВКС представлена рядом программных компонентов, объединенных в одну архитектуру. Чтобы разрабатываемое программное обеспечение (ПО) соответствовало критерию пластичности, сетевое взаимодействие представляется в виде трех отдельных слоев:

1) слой, отвечающий за управление сетевым соединением с сервером приложения;

2) слой, отвечающий за управление соединением пользователей, вступающих в ВКС;

3) слой, отвечающий за управление соединением для передачи мультимедийных данных участников ВКС.

Каждый слой выражен совокупностью объектов, функций, формирующих выраженную абстракцию над логикой обработки запросов к графическому интерфейсу и обращений к соответствующему программному интерфейсу (WebRTC, WebSocket, GraphQL).

Каждый из программных компонентов, представляющих описываемые слои, в своей основе имеет схожую логику управления соединением — операции по открытию, закрытию транспортных каналов, передаче и обработке поступающих сообщений между соединившимися пользователями. В разработанном ПО данная логика отражена в классах *Transport*, *Channel* и его подтипов.

Transport инкапсулирует логику работы с создаваемым соединением, предоставляя клиентам обобщенный программный интерфейс управления (рис. 3).

Статический метод *create(pattern)* класса *Transport* используется для инстанцирования новых объектов класса. Метод принимает стратегию создания канала для передачи данных между пользователями [13]. Стратегия представляет собой функцию, возвращающую объект типа *Channel* и принимающую в качестве аргументов соединяемых пользователей. При создании объект находится в состоянии ожидания подключения, что выражено значением “inactive” свойства *status*. Метод *connect(connecting, connected)* позволяет клиенту выполнить соединение пользователей, информация о которых была передана в аргументах, с использованием переданной ранее стратегии. Информация о пользователях должна быть передана посредством объектов типа *PeerData*. При удачном соединении объект класса переходит в состояние “active”, записы-

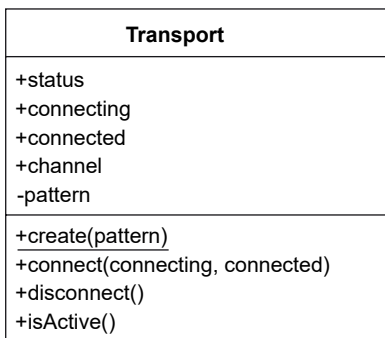
вает информацию о соединенных пользователях в свойствах *connecting* и *connected* и сохраняет созданный канал в свойстве *channel*. В случае возникновения ошибок метод *connect* генерирует исключение с соответствующим проблеме сообщением. Для отключения соединения пользователей класс предоставляет метод *disconnect()*, при использовании которого объект класса удаляет сохраненные данные о созданном соединении и переводит *status* в состояние “inactive”. Таким образом, объект класса *Transport* может быть использован в различных сеансах связи.

Класс *Channel* является базовым для объектов, скрывающих в себе логику создания канала связи для обмена данными между двумя пользователями. *Channel* сводит различия в управлении каналах разных типов к общему для всех интерфейсу, обеспечивая единообразную работу с каналами на всех сетевых уровнях приложения (рис. 4).

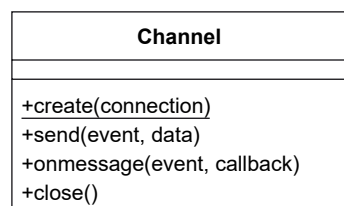
Инстанцирование объекта класса *Channel* выполняется посредством статического метода класса *create(connection)*. В качестве аргумента метод ожидает простой объект *connection* с методами *send(event, data)*, *onmessage(event, callback)* и *close()*, оперирующими активным соединением пользователей. *Channel* выступает в роли класса-адаптера, предоставляя потомкам возможность адаптации интерфейсов создаваемых каналов к общему для всех виду. Подтипы *Channel* инкапсулируют логику создания каналов, расширяя конструктор базового класса. В конструкторе подтипов происходит выполнение соединения пользователей, переданных в аргументе, и последующий вызов конструктора базового класса с передачей созданного *connection* в аргументе.

В разработанном ПО для реализации слоев сети использованы следующие подтипы, образованные от *Channel*: *IOChannel*, *P2PChannel*, *MP2PChannel* (рис. 5).

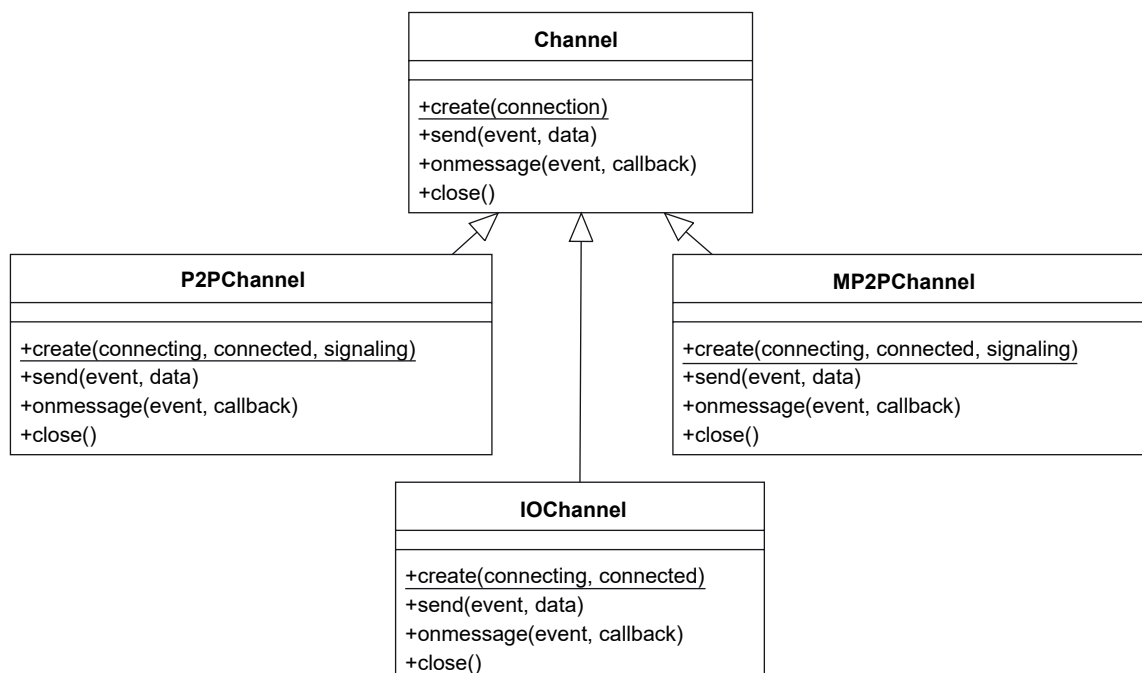
IOChannel принимает в аргументах конструктора данные о соединяемых пользователях, представленных объектами с полями *name* и *address*. При создании объекта выполняется соединение клиента, представленного в *connecting*, с *connected* по WebSocket, с использованием записанного в *connected* адреса.



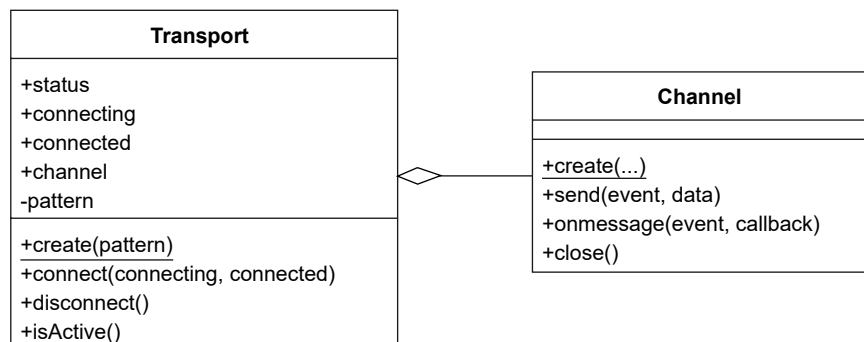
■ Рис. 3. Диаграмма класса *Transport*
■ Fig. 3. Diagram of class *Transport*



■ Рис. 4. Диаграмма класса *Channel*
■ Fig. 4. Diagram of class *Channel*



■ **Рис. 5.** Диаграмма иерархии классов, образованных от *Channel*
 ■ **Fig. 5.** Hierarchy diagram of classes inherited from *Channel*



■ **Рис. 6.** Диаграмма связи *Transport* и *Channel*
 ■ **Fig. 6.** Diagram of relationship between *Transport* and *Channel*

P2PChannel использует в качестве механизма соединения API *RTCPeerConnection* и *RTCDataChannel* из WebRTC. *P2PChannel* при создании объекта принимает в аргументах конструктора данные о соединяемых пользователях и канал передачи сигнализирующих сообщений. Канал передачи представляет собой объект, образованный от *Channel*. Канал передачи используется на этапе открытия соединения для обмена между клиентами информацией, необходимой для установления сеанса связи. После завершения соединения канал может быть закрыт.

Класс *MP2PChannel* имеет схожую с *P2PChannel* логику работы. Для соединения пользователей он использует API *RTCPeerConnection* из

WebRTC и *MediaStream* API, соединяя пользователей для обмена мультимедийной информацией.

При вызове *connect(connecting, connected)* объект *Transport* создает канал связи между пользователями с использованием переданной в аргументе конструктора стратегии. Таким образом, *Transport* агрегирует *Channel* или его подтипы (рис. 6).

Объекты класса *PeerData* в разработанном ПО используются как контейнеры для хранения информации о пользователях, участвующих в сетевом взаимодействии. *PeerData* позволяет компонентам использовать в качестве исходных данных единообразные персистентные структу-

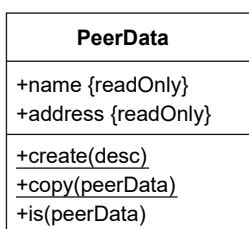
ры данных, обеспечивая упрощение взаимодействия внутри ПО (рис. 7).

PeerData обладает свойствами *name* и *address*: *name* является уникальным идентификатором пользователя, участвующего в межсетевом взаимодействии; *address* предоставляет адрес, по которому пользователь может быть локализован. При создании объекта класса конструктор *PeerData* принимает объект с информацией о пользователе, ожидая при этом, что *desc* имеет свойства *name* и *address*. В случае получения некорректного объекта-дескриптора конструктор генерирует исключение с описанием обнаруженной проблемы.

PeerData предоставляет ряд методов, упрощающих работу с созданным контейнером. Статический метод *copy(peerData)* позволяет сделать чистую копию существующего объекта *PeerData*. Метод *is(peerData)* позволяет сравнивать объекты *PeerData*. Данные методы позволяют обеспечивать неизменяемость данных, сохраненных в *PeerData*, в процессе их использования компонентами ПО.

При успешном соединении пользователей объект *Transport* сохраняет копии переданных *connecting* и *connected* в своих свойствах. Таким образом, поля *connecting* и *connected* объекта класса *Transport* имеют тип *PeerData*. Описанные отношения между классами *Transport* и *PeerData* выражены в виде композиции, представленной на рис. 8.

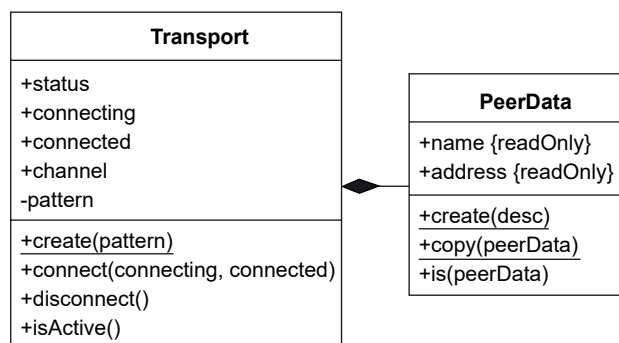
Объекты *Transport* отражают связь между парами пользователей, участвующих в сетевом взаимодействии. Каждый слой ПО сетевого взаимодействия выражен в виде отдельного модуля и сформирован определенным набором объектов *Transport*. Первый слой ПО сети использует объект *Transport* с переданной ему при создании фабричной функцией *createIOChannel(connecting, connected)*. Функция является оберткой для вызова метода *create(connecting, connected)* класса *IOChannel*. Клиентское приложение вызывает метод *connect(connecting, connected)* объекта *Transport*, используя в качестве аргументов объекты типа *PeerData* с информацией о себе и о сервере приложения, тем самым выполняя их соеди-



■ **Рис. 7.** Диаграмма класса *PeerData*
 ■ **Fig. 7.** Diagram of class *PeerData*

нение. Модуль слоя сети устанавливает ряд обработчиков сообщений на объект *Channel*, записанный в объекте *Transport*, необходимых для синхронизации состояния клиентского приложения с состоянием других пользователей. Кроме того, модуль устанавливает ряд обработчиков событий на действия пользователя, выполняемые с интерфейсом приложения, для отправления соответствующих сообщений на сервер. Второй слой при инициации нового сеанса ВКС попарно соединяет каждого из участников связи, создавая для этого объекты типа *Transport* с передачей им фабричной функции *createP2PChannel(connecting, connected, signaling)*. В качестве сигнального сервера будет использован созданный на прошлом слое WebSocket-канал. Третий слой выполняет соединение участников ВКС согласно сформированным в процессе инициации ВКС отношениям между участниками, используя для соединения объект *Transport* с переданной ему функцией *createMP2PChannel(connecting, connected, signaling)*. В качестве сигнального сервера будет использован объект *P2PChannel*, созданный на втором слое сети.

Представленная сетевая архитектура ПО приложения ВКС обладает иерархией отношения между отдельными компонентами. Решения, принимаемые на первом слое сетевой архитектуры, например, об отключении соединения с сервером приложения, могут повлиять на работу второго и третьего слоев. В свою очередь второй слой может повлиять на работу третьего. Иерархические отношения между слоями сетевой архитектуры в ПО выражены посредством композиции. На стороне клиента слои представлены объектами класса *NetworkNode*. При создании объект *NetworkNode* принимает ссылку на слой выше по иерархии. Класс предоставляет метод *createChild()*, создающий дочерний объект *NetworkNode*. Объект класса содержит информацию о всех соединениях на своем уровне (вклю-



■ **Рис. 8.** Диаграмма связи *Transport* и *PeerData*
 ■ **Fig. 8.** Diagram of relationship between *Transport* and *PeerData*

чая информацию о самих пользователях) и может оперировать дочерним *NetworkNode* на свое усмотрение.

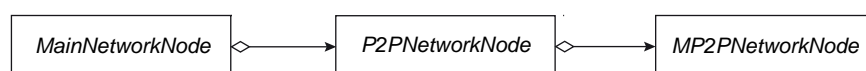
Подтипы класса *NetworkNode* инкапсулируют логику работы для каждого слоя сетевой архитектуры. Отношения между уровнями сети, сформированные в приложении ВКФ, отражены на рис. 9.

Разделение сети на несколько уровней, действующих в своих рамках практически независимо от остальных, позволяет приложению продолжать осуществлять свои функции при отказе ключевых узлов. В случае отказа сервера приложения работа пользователей не будет нарушена, поскольку между ними по-прежнему будет существовать прямое соединение. Таким образом, активные видеоконференции не будут прерваны. Помимо этого, дополнительная функциональность может быть включена в реализацию каждого из уровней сети как часть композиции моду-

ля, обеспечивая простой алгоритм масштабирования сетевой архитектуры приложения.

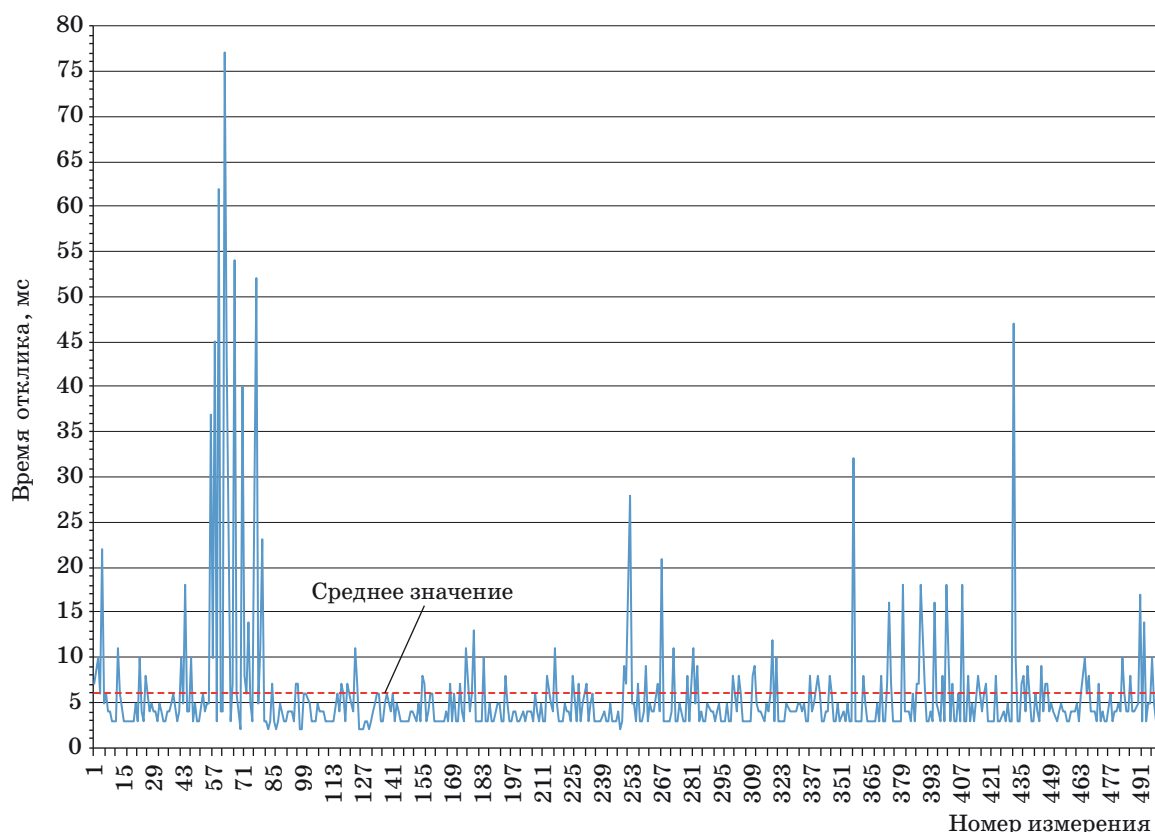
В то же время использование описанной схемы соединения обладает недостатками. Вторым слоем сетевого соединения, обеспечивающим прямое соединение пользователей приложения, подразумевает создание сетевого соединения по схеме «каждый-с-каждым», формируя при этом избыточное количество активных соединений участников. Тем не менее данный подход обеспечивает надежное соединение пользователей в сети, поскольку отключение отдельных узлов сети не сможет повлиять на соединение оставшихся.

Предложенная в работе схема соединения пользователей приложения ВКС, используемая в разработанном ПО, позволила добиться значительного времени отклика соединения, представленных на рис. 10. Измерения отклика выполнялись с интервалом 500 мс. В среднем отклик для двух пользователей, находящихся в одном здании, со-



■ **Рис. 9.** Организация слоев архитектуры сетевого соединения на стороне клиента

■ **Fig. 9.** Organization of architectural layers of the network connection on the client side



■ **Рис. 10.** Время отклика пирингового соединения, установленного посредством разработанного ПО

■ **Fig. 10.** The response time of the peer-to-peer connection established by means of the developed software

ставил 5,81 мс. Передача осуществлялась по Wi-Fi внутри локальной сети.

Возникновение пиков (выбросов) на графике, наблюдаемое для измерений под номерами 40–76, 252, 357, может быть связано с особенностями порядка исполнения программы интерпретатором кода. Кроме того, поскольку передача осуществлялась по Wi-Fi, увеличение задержки сигнала могло произойти вследствие возникновения помех в окружающей среде.

Заключение

Разрабатываемая архитектура соединения пользователей в приложении ВКС обладает рядом преимуществ, способствующих конкуренто-

способности разработанного приложения на фоне аналогов. Предложенная организация модулей ПО сетевого соединения приложения ВКС предоставляет пространство для внедрения дополнительных функций в приложение. В рамках КФС управляемые аккаунты могут использоваться в модулях обеспечения безопасности и управления доступом, например, для обеспечения контроля за рабочими местами, производством и другими объектами. Организация хранения и обработки данных позволяет внедрить модуль ВКС в КФС, а также сделать удобным его использование как на терминалах и мобильных робототехнических платформах, так и на клиентских мобильных устройствах.

Исследование выполнено при поддержке Российского научного фонда (грант № 16-19-00044).

Литература

1. Решетняк В. И. Применение видеоконференц-связи в арбитражном судопроизводстве // Российский юридический журнал. 2013. № 1 (88). С. 154–159.
2. Борисов Д. Н., Иванов В. В. Организационная телемедицина // Врач и информационные технологии. 2017. № 3. С. 112–120.
3. Тулемисова Д. З. Использование видеоконференций в современном образовании // Современные проблемы и направления развития системы подготовки кадров: сб. науч. тр. по материалам I Междунар. науч.-практ. конф./ НОО «Профессиональная наука», 2016. С. 129–135.
4. Марченков С. А., Вдовенко А. С., Корзун Д. Ж. Расширение возможностей совместной деятельности в интеллектуальном зале на основе сервисов электронного туризма // Тр. СПИИРАН. 2017. № 1 (50). С. 165–189.
5. Савельев А. И. Архитектуры, алгоритмы и программные средства обработки потоков многомодальных данных в пиринговых веб-приложениях видеоконференцсвязи: дис. ... канд. техн. наук / СПИИРАН. — СПб., 2016. — 135 с.
6. Wu J., Shang Y., Yuen C., Cheng B., Chen J. TRADER: A Reliable Transmission Scheme to Video Conferencing Applications over the Internet // Journal of Network and Computer Applications. 2014. Vol. 44. P. 161–171.
7. Cheng R., Wu W., Lou Y., Chen Y. A Cloud-based Transcoding Framework for Real-Time Mobile Video Conferencing System // Proc. of 2nd IEEE Intern. Conf. “Mobile Cloud Computing, Services, and Engineering” (MobileCloud). 2014. P. 236–245.
8. Saveliev A. I., Ronzhin A. L., Vatamaniuk I. V. Architecture of Data Exchange with Minimal Client-Server Interaction at Multipoint Video Conferencing // Lecture Notes in Computer Science. 2014. Vol. 8638 LNCS. P. 164–174.
9. Левоневский Д. К., Ватаманюк И. В., Савельев А. И. Многомодальная информационно-навигационная облачная система МИНОС для корпоративного киберфизического интеллектуального пространства // Программная инженерия. 2017. № 3. С. 120–128.
10. Vatamaniuk I., Levonevskiy D., Saveliev A., Denisov A. Scenarios of Multimodal Information Navigation Services for users in Cyberphysical Environment // Proc. of SPECOM-2016, Budapest, Hungary. — Springer, 2016. Vol. 9811 LNAI. P. 588–595.
11. Савельев А. И., Соменков Н. В. Архитектура клиентской части пирингового приложения видеоконференцсвязи // Математические методы в технике и технологиях. 2016. № 9(91). С. 176–179.
12. Савельев А. И. Алгоритмы обработки данных в контролируемых аккаунтах системы видеоконференцсвязи // Информационно-управляющие системы. 2016. № 3. С. 15–23. doi:10.15217/issn1684-8853.2016.3.15
13. Гамма Э., Хелм Р., Джонсон Р., Влиссидес Дж. Приемы объектно-ориентированного проектирования. Паттерны проектирования. — СПб.: Питер, 2015. — 368 с.

UDC 004.773.5

doi:10.15217/issn1684-8853.2018.1.2

Architectural Solutions for Integrating a Video Conferencing Module into Cyberphysical Intelligent Space

Karasev E. Yu.^a, Programmer, gnomskg@gmail.com

Vatamaniuk I. V.^a, Post-Graduate Student, vatamaniuk@iias.spb.su

Saveliev A. I.^a, PhD, Tech., Senior Researcher, saveliev@iias.spb.su

Ronzhin A. L.^a, Dr. Sc., Tech., Professor, ronzhin@iias.spb.su

^aSaint-Petersburg Institute for Informatics and Automation of the RAS, 39, 14 Line, V. O., 199178, Saint-Petersburg, Russian Federation

Introduction: Cyberphysical systems are a wide class of systems whose main characteristic is high integration of computing resources into each physical component of a system. One of the tasks of a cyberphysical intelligent space is to provide convenient interaction between users via various system nodes. Modern videoconferencing applications are expected not only to use various modalities like text, audio/video data or graphic objects, but also to meet the criteria of modularity and cross-platformness. **Purpose:** To develop an application architecture which would provide a video conferencing service for cyberphysical intelligent space users. **Results:** The paper discusses architectural solutions for a videoconferencing module, providing the possibility of its integration into a MINOS cyberphysical system, as well as a generalised structure of a videoconferencing module with the implementation of private and controlled public accounts. The network interaction is presented in the form of three separate layers responsible for managing the network connection to the server, the connection between the users, and the connection for transferring multimedia data to the video conference participants. The connection response time during the organization of the video conferencing has been measured. On average, the response for two users is 5.81 ms.

Keywords — Peering Interaction, Client-Server Interaction, Teleconferencing, Videoconferencing, Modular Software Architecture, Cyberphysical Systems, Intellectual Space.

Citation: Karasev E. Yu., Vatamaniuk I. V., Saveliev A. I., Ronzhin A. L. Architectural Solutions for Integrating a Video Conferencing Module into Cyberphysical Intelligent Space. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 1, pp. 2–10 (In Russian). doi:10.15217/issn1684-8853.2018.1.2

References

1. Reshetnyak V. I. The use of Videoconferencing in Arbitration Proceedings. *Rossiiskii iuridicheskii zhurnal* [Russian Juridical Journal], 2013, no. 1(88), pp. 154–159 (In Russian).
2. Borisov D. N., Ivanov V. V. Organizational Telemedicine. *Vrach i informatsionnye tekhnologii* [Physicians and IT], 2017, no. 3, pp. 112–120 (In Russian).
3. Tulemisova D. Z. The use of Videoconferencing in Modern Education. *Sbornik nauchnykh trudov po materialam I Mezhdunarodnoi nauchno-prakticheskoi konferentsii "Sovremennye problemy i napravleniia razvitiia sistemy podgotovki kadrov"* [Proc. on the Materials of the First Intern. Scientific and Practical Conf. "Modern Problems and Directions of Development of the System of Personnel Training"]. NOO "Professional'naiia nauka" Publ., 2016, pp. 129–135 (In Russian).
4. Marchenkov S. A., Vdovenko A. S., Korzun D. G. Enhancing the Opportunities of Collaborative Work in an Intelligent Room using e-Tourism Services. *Trudy SPIIRAN* [SPIIRAS Proceedings], 2017, vol. 50, pp. 165–189 (In Russian).
5. Saveliev A. I. *Arkhitektury, algoritmy i programmnye sredstva obrabotki potokov mnogomodal'nykh dannykh v piringovykh veb-prilozheniiakh videokonferentssviasi* Dis. kand. tehn. nauk [Architectures, Algorithms and Software for Processing Multimodal Data Streams in Peer-To-Peer Web Applications of Video Conferencing. PhD tech. sci. diss.]. SPIIRAN Publ., 2016. 135 p. (In Russian).
6. Wu J., Shang Y., Yuen C., Cheng B., Chen J. TRADER: A Reliable Transmission Scheme to Video Conferencing Applications over the Internet. *Journal of Network and Computer Applications*, 2014, vol. 44, pp. 161–171.
7. Cheng R., Wu W., Lou Y., Chen Y. A Cloud-based Transcoding Framework for Real-Time Mobile Video Conferencing System. *Proc. of 2nd IEEE Intern. Conf. "Mobile Cloud Computing, Services, and Engineering" (MobileCloud)*, 2014, pp. 236–245.
8. Saveliev A. I., Ronzhin A. L., Vatamaniuk I. V. Architecture of Data Exchange with Minimal Client-Server Interaction at Multipoint Video Conferencing. *Lecture Notes in Computer Science*, 2014, vol. 8638 LNCS, pp. 164–174.
9. Levonevskiy D. K., Vatamaniuk I. V., Saveliev A. I. MINOS Multimodal Information and Navigation Cloud System for the Corporate Cyber-Physical Smart Space. *Programmnaia inzheneriia* [Software Engineering], 2017, vol. 8, no. 3, pp. 120–128 (In Russian).
10. Vatamaniuk I., Levonevskiy D., Saveliev A., Denisov A. Scenarios of Multimodal Information Navigation Services for users in Cyberphysical Environment. *Proc. of SPECOM-2016*, Budapest, Hungary, Springer, 2016, vol. 9811 LNAI, pp. 588–595.
11. Saveliev A. I., Somenkov N. V. Architecture of Client Part of Peer-To-Peer Videoconferencing Application. *Matematicheskie metody v tekhnike i tekhnologiiakh* [Mathematical Methods in Technique and Technologies], 2016, no. 9, pp. 176–179 (In Russian).
12. Saveliev A. I. Algorithms of Data Processing in Supervised Accounts of a Videoconferencing System. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2016, no. 3, pp. 15–23 (In Russian). doi:10.15217/issn1684-8853.2016.3.15
13. Gamma E., Helm R., Johnson R., Vlissides J. *Design Patterns: Elements of Reusable Object-Oriented Software*. Addison-Wesley, 2000. 403 p.

УДК 612.843.721

doi:10.15217/issn1684-8853.2018.1.11

ПРОБЛЕМА СОВМЕСТИМОСТИ СИСТЕМ ТЕЛЕВИЗИОННОГО ВЕЩАНИЯ 2D И 3D И МЕТОД ЕЕ РЕШЕНИЯ

Н. Н. Красильников^а, доктор техн. наук, профессор, NNKrasilnikov@yandex.ru

О. И. Красильникова^а, канд. техн. наук, доцент, OIKrasilnikova@yandex.ru

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

Постановка проблемы: внедрение систем 3D телевизионного вещания в настоящее время затормозилось отсутствием совместимого формата телевизионного вещания, который позволил бы передачу 3D-контента смотреть на 2D-телевизорах в формате 2D. С экономической точки зрения важно, чтобы любой создаваемый телевизионный контент, в том числе фильмы, всевозможные передачи, реклама и т. д., был доступен для просмотра на экранах любых телевизоров, как 2D, так и 3D. **Цель исследования:** поиск принципа построения алгоритма обработки стереоскопического изображения для совместимого формата. **Методы:** решение проблемы базируется на учете свойств зрения человека, на методах конвертации 2D-изображений в 3D, компрессии изображений, а также большом объеме экспериментальных измерений, выполненных с группой зрителей, которым предъявлялись для наблюдения стереоскопические изображения, синтезированные при различных форматах передачи, для определения заметности возникающих при этом искажений. **Результаты:** разработан принцип построения совместимого формата для 3D телевизионного вещания. В рамках этого формата в качестве основного передается левый компонент изображения стереопары, но при этом малая часть времени передачи кадра, что соответствует малой площади передаваемого изображения, отводится для передачи сигнала, содержащего информацию, которая, используя левый компонент, позволяет сформировать правый компонент стереопары. В качестве такого сигнала используются пиксели карты экранных параллаксов, передача которых растягивается на 3–4 кадра. В 3D телевизионном приемнике из этих пикселей методом интерполяции нулевого порядка восстанавливаются карты экранных параллаксов, используя которые совместно с левыми компонентами стереопары, можно синтезировать недостающие правые компоненты стереопар, обеспечивая воспроизведение стереоизображения. Телевизионные приемники 2D просто ограничиваются воспроизведением левого компонента стереопары. **Практическая значимость:** разработка принципа построения совместимого формата 3D телевизионного вещания позволяет многократно сократить трудоемкость разработки соответствующих стандартов.

Ключевые слова — стереоскопическое изображение, совместимость, карта экранных параллаксов.

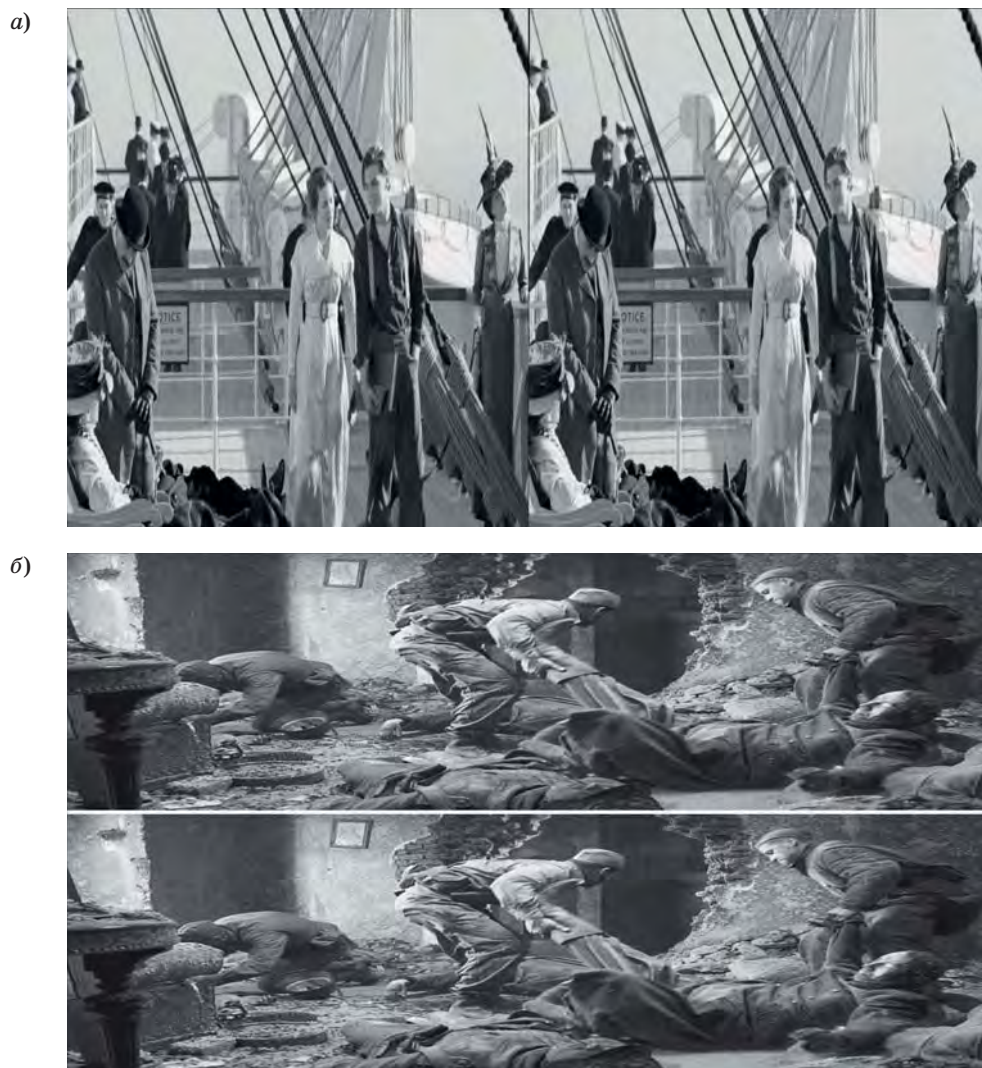
Цитирование: Красильников Н. Н., Красильникова О. И. Проблема совместимости систем телевизионного вещания 2D и 3D и метод ее решения // Информационно-управляющие системы. 2018. № 1. С. 11–20. doi:10.15217/issn1684-8853.2018.1.11

Citation: Krasilnikov N. N., Krasilnikova O. I. Providing Compatibility for 2D and 3D Television Broadcasting Systems. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 1, pp. 11–20 (In Russian). doi:10.15217/issn1684-8853.2018.1.11

Введение

Дальнейшее развитие и внедрение систем 3D телевизионного вещания тормозится нерешенностью проблемы их совместимости с существующими системами 2D телевизионного вещания. В настоящее время форматы Anamorph SideBySide (анаморфная горизонтальная стереопара) и Anamorph OverUnder (анаморфная вертикальная стереопара) являются стандартами 3D телевизионного вещания. При использовании этих форматов оба изображения стереопары записываются в одном файле. При этом площадь, отводимая для передачи одного кадра в системе 2D телевизионного вещания, в системе 3D телевизионного вещания делится поровну между левым и правым изображениями стереопары, которые располагаются рядом вплотную друг к другу в одном кадре [1].

При передаче 3D телевизионного изображения с использованием формата Anamorph SideBySide левый компонент стереопары располагается слева, а правый — справа, при этом обе картинка сжимаются в горизонтальном направлении в два раза так, что получившаяся в результате картинка принимает размер стандартного кадра, например 1920 × 1080. Вследствие этих преобразований каждое изображение стереопары выглядит вытянутым в вертикальном направлении, а разрешение каждого ракурса в горизонтальном направлении уменьшается в два раза. На рис. 1, а в качестве примера приведено такое изображение. При просмотре 3D телевизионного изображения с использованием этого формата выполняются обратные преобразования, а именно, суммарная картинка растягивается по горизонтали в два раза и разделяется на два изображения стереопары — левое и правое, и затем они выводятся на экран од-



■ *Рис. 1.* Вид изображения на экране 2D-телевизора при просмотре передачи с использованием формата Anamorph SideBySide (а) и Anamorph OverUnder (б)

■ *Fig. 1.* The image on the 2D TV screen when viewing the transmission using the format Anamorph SideBySide (a) and Anamorph OverUnder (б)

ним из известных способов. В случае телевизоров с чересстрочной поляризацией (LG Cinema 3D) использование этого формата приводит к потере разрешения в два раза не только в вертикальном, но и в горизонтальном направлении. В телевизорах с затворными очками использование этого формата приводит к потере разрешения в два раза только в горизонтальном направлении.

При передаче 3D телевизионного изображения с использованием формата Anamorph OverUnder левый ракурс кадра располагается вверху, а правый внизу. Обе картинка сжимаются в вертикальном направлении в два раза, при этом получившаяся в результате картинка принимает размер стандартного кадра, например 1920×1080 . Вследствие этих преобразований каждое из изображений стереопары оказывается вытянутым

в горизонтальном направлении, а их разрешение в вертикальном направлении уменьшается в два раза (рис. 1, б). При просмотре 3D телевизионного изображения с использованием этого формата выполняются обратные преобразования: результирующая картинка растягивается по вертикали в два раза и разделяется на левое и правое изображения стереопары, которые выводятся на экран.

Таким образом, современные стандарты 3D телевизионного вещания имеют два очень серьезных недостатка. Во-первых, при реализации любого из них имеет место уменьшение разрешения каждого из изображений стереопары по крайней мере в два раза. Во-вторых, и это наиболее существенно, при передаче 3D телевизионной программы по каналу ее нельзя смотреть на

2D-телевизоре, так как на его экране будет отображаться вдвоенное изображение. В результате возникает замкнутый круг. С одной стороны, экономически нецелесообразно наполнять рынок 3D-телевизорами, так как в настоящее время практически полностью отсутствуют телевизионные передачи в 3D-формате. С другой стороны, даже несмотря на наличие 3D-контента, экономически невыгодно осуществлять телевизионное вещание в формате 3D, поскольку у населения очень мало 3D-телевизоров. В связи с этим компании, выпускающие 3D-телевизоры, например Phillips, Sony, LG, Samsung и др., либо сильно сокращают их выпуск, либо прекращают совсем [2]. Выходом из сложившейся ситуации являются разработка и внедрение совместимого формата телевизионного вещания, при использовании которого можно будет не только по-прежнему смотреть на экранах 3D телевизоров 2D-контент в формате 2D, но и смотреть на экранах 2D-телевизоров 3D-контент в 2D-формате. Сложившаяся ситуация стимулировала исследования и разработки совместимых форматов [3–5].

Совместимый формат 3D телевизионного вещания

Принципиальная возможность разработки такого формата следует из того, что изображения стереопары мало различаются между собой, поскольку их съемка производится из близко расположенных точек пространства. Поэтому, разрабатывая совместимый формат, за основу можно взять одно из изображений стереопары, например левое, а правое изображение стереопары фор-

мировать из левого, корректируя его сигналом разницы между левым и правым изображениями, поскольку резонно ожидать, что малое различие между изображениями стереопары может быть представлено малым цифровым потоком. Дальнейшие экспериментальные исследования подтвердили это предположение. Исходя из изложенного в простейшем случае совместимый формат 3D телевизионного вещания может быть следующим. За основу берется один из форматов, принятых в цифровых системах телевизионного вещания, и в рамках этого формата почти полностью передается левый компонент изображения стереопары, и лишь незначительная часть времени, отведенная на его передачу, расходуется на передачу сигнала, содержащего информацию о разнице между левым и правым изображениями. Эксперименты показывают, что эта часть составляет всего несколько процентов от длительности передачи кадра, а следовательно, от площади левого компонента изображения. При этом изображение на экране 2D-телевизора будет иметь вид, показанный на рис. 2.

При использовании описываемого формата в случае наблюдения на 2D-телевизорах 3D-контента зрители будут видеть левое изображение стереопары, у которого незначительная часть кадра, например расположенная в его нижней части, будет зашторена, так как эта площадь кадра отведена для передачи информации о разнице между левым и правым изображениями. Предварительные эксперименты с аудиторией зрителей показали, что потери полезной площади кадра порядка 5 % ими практически не замечаются. Возможны и другие варианты включения в кадр информации о разнице между левым



■ **Рис. 2.** Вид изображения на экране 2D-телевизора при просмотре передачи с использованием 3D совместимого формата

■ **Fig. 2.** The image on the 2D TV screen when viewing the transmission using a 3D compatible format

и правым изображениями. При таком подходе, по сравнению с форматами 3D Tile Format и 3DZ Tile [3–6], обеспечивается более высокое разрешение как 2D-, так и 3D-изображения. При этом пикселей в кадре оказывается в 2,14 раза больше.

Рассмотрим возможные варианты представления информации, необходимой для синтеза правого компонента стереопары при использовании в качестве опорного левого компонента стереопары.

1. В виде разности компонентов стереопары:

$$\Delta L(k, n) = L_{\text{п}}(k, n) - L_{\text{л}}(k, n),$$

где $L_{\text{п}}(k, n)$ — правый компонент стереопары; $L_{\text{л}}(k, n)$ — левый компонент стереопары; k — номер строки; n — номер столбца.

2. В виде карты глубины (карты экранных параллаксов): как известно, координата глубины d_{ob} и величина экранного параллакса p связаны соотношением

$$p = \frac{b(d_{ob} - d_{sc})}{d_{ob}},$$

где b — расстояние между зрачками глаз зрителя (база); d_{sc} — расстояние от зрителя до экрана, поэтому карта глубины элементарно преобразуется в карту экранных параллаксов. Тем не менее предпочтение следует отдать карте экранных параллаксов, поскольку при равномерной шкале квантования, которая имеет место при цифровом представлении информации о глубине, она оказывается лучше согласована с разрешением зрения по координате глубины, чем карта глубины.

При выполнении работы были исследованы оба варианта представления и, следовательно, передачи информации о различии между компонентами стереопары.

Первый вариант заключался в вычислении «разностного изображения» путем вычитания из правого изображения стереопары левого изображения стереопары и дальнейшей его компрессии путем преобразования в формат JPEG при различных значениях качества q (Quality), что, соответственно, приводило к получению различных степеней сжатия. При этом в компрессируемое разностное изображение вносились искажения, поскольку используемый метод компрессии является методом с потерями информации. После этого «разностное изображение» восстанавливалось в формате BMP. Затем с использованием левого изображения стереопары в качестве опорного и «разностного изображения» восстанавливалось правое изображение стереопары. Таким образом, изготавливались группы искаженных стереопар.

Второй вариант заключался в вычислении карты экранных параллаксов (векторов смеще-

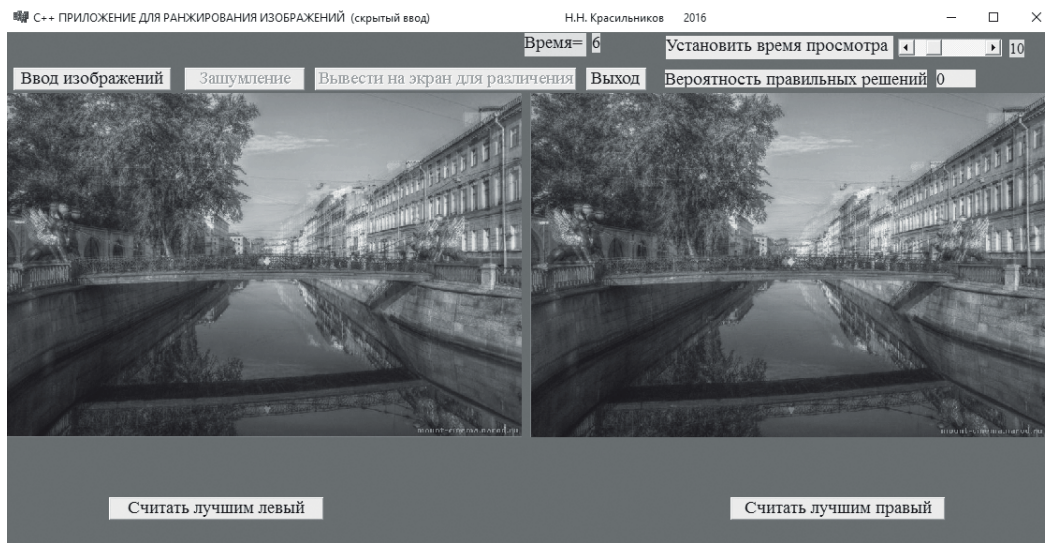
ний) с последующим ее сжатием при использовании для этой цели, помимо записи в формат JPEG, еще двух методов компрессии с потерями информации. В первом случае сжатие цифрового потока, представляющего карту экранных параллаксов, достигалось уменьшением количества двоичных разрядов кода на пиксель, которое в конечном итоге приводило к уменьшению количества уровней квантования, а следовательно, и к менее точной передаче информации о глубине. Во втором случае сжатие цифрового потока, представляющего карту экранных параллаксов, достигалось уменьшением плотности отсчетов на карте экранных параллаксов, которое влекло за собой размытие границ представленных на ней объектов.

Естественно, что при использовании метода сжатия с потерями информации возникают погрешности при восстановлении стереопары, однако до определенного уровня они не приводят к заметным для зрителей искажениям стереоскопических изображений. Целью описываемых ниже исследований было определение экспериментальным путем такой величины компрессии, при которой эти искажения можно еще считать для зрителя незаметными. Для проведения исследований были отобраны 3D-изображения различного сюжетного содержания и различной структуры.

Методика выполнения эксперимента со зрителями

Для определения допустимой величины компрессии цифрового потока, которым передается информация о различии между компонентами стереопары, независимо от того, делается это посредством «разностного изображения» или карты экранных параллаксов, было разработано специальное приложение, благодаря которому на экран для просмотра одновременно выводились два стереоскопических изображения в анаглифном формате: эталонное, при формировании которого компрессия не применялась, и изображение, содержащее искажения. При формировании последнего цифровой поток, которым передается информация о различии между компонентами стереопары, подвергался компрессии. На рис. 3 приведено изображение интерфейса этого приложения. При выполнении эксперимента зритель рассматривал и сравнивал между собой оба стереоскопических изображения через анаглифные очки с красно-голубыми фильтрами.

Зритель ранжировал сравниваемые изображения, т. е. определял, какое изображение, справа или слева, имеет более высокое качество, т. е. является эталонным, и в зависимости от принятого



■ **Рис. 3.** Интерфейс приложения для ранжирования стереоскопических изображений зрителями
 ■ **Fig. 3.** The interface of the application used to rank stereoscopic images by viewers

решения нажимал специальные кнопки, расположенные в нижней части окна приложения, соответственно справа и слева. Перед выводом стереоскопических изображений на экран зрителю предъявлялись стирающие шумовые изображения, а при выводе расположение эталонного и искаженного изображений (справа, слева) делалось случайным, неизвестным зрителю. Время предъявления изображений для их сравнения при выполнении эксперимента могло быть установлено на один из следующих интервалов: 1, 2, 3, 5, 10, 15 и 20 с, — после чего оба изображения с экрана исчезали. С каждой парой стереоскопических изображений (эталонное — искаженное) зритель выполнял заданное экспериментатором количество сравнений, при этом приложение автоматически вычисляло относительное количество правильных решений и выводило эту величину на экран. Полученные данные записывались экспериментатором в таблицу. Данный метод оценки качества 3D-изображений, по сравнению с использованными для аналогичных целей и описанными в работах [7–11], обладает тем достоинством, что в нем оценка дается самим зрителем.

Результаты экспериментальных исследований

Метод передачи разности компонентов стереопары

Для оценки ожидаемой относительной величины цифрового потока $\Delta L(k, n) / L_{\Pi}(k, n)$, которым может быть представлена разность компонентов стереопары, был поставлен эксперимент,

при выполнении которого для семи стереопар были изготовлены разности левого и правого компонентов. Далее динамический диапазон каждой разности был ограничен интервалом от -127 до 127 квантовых уровней, а получившийся сигнал центрирован относительно 127-го уровня квантования, т. е. вычислялся сигнал

$$\Delta L_N(k, n) = \Delta L(k, n) / 2 + 127.$$

Далее этот сигнал, как уже было сказано, подвергался компрессии путем преобразования в формат JPEG при различных значениях качества q , и, соответственно, при этом получались различные степени сжатия. Величина сжатия в значительной степени зависела от структуры сжимаемого изображения. На этапе компрессии в сигнал вносились искажения, поскольку использовался алгоритм компрессии с потерями информации.

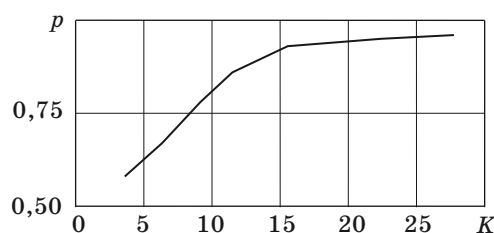
На следующем этапе обработки выполнялась декомпрессия сигнала $\Delta L_N(k, n)$, а затем восстановление сигнала $\Delta L(k, n)$:

$$\Delta L(k, n) = 2[\Delta L_N(k, n) - 127].$$

Затем восстанавливался правый компонент стереопары путем прибавления к левому компоненту сигнала разности:

$$L_{\Pi}(k, n) = L_{\text{л}}(k, n) + \Delta L(k, n).$$

Полученная таким образом стереопара предъявлялась зрителю одновременно с исходной стереопарой на предмет обнаружения различий



■ **Рис. 4.** Зависимость вероятности p , с которой зритель обнаруживал искажения на изображении, от величины компрессии K при длительности наблюдения 10 с

■ **Fig. 4.** Dependence of the probability p with which the viewer detected distortions in the image on the value of compression K at the observation time 10 s

в соответствии с описанной выше методикой. На рис. 4 приведена усредненная зависимость вероятности, с которой зритель обнаруживал изображение, полученное рассматриваемым способом, в зависимости от величины компрессии K , имеющей место при используемом параметре q .

При вычислении этой зависимости усреднение производилось по результатам экспериментальных исследований, выполненных 12 зрителями с семью изображениями. Зрители отмечали, что искажения на стереоскопических изображениях проявлялись в виде постороннего шума, а также в размытии границ на мелких объектах. Из экспериментов следует, что уже при малых величинах компрессии, например $K = 3,9$, в 58 % случаев зрители оценивали преобразованное рассматриваемым способом стереоизображение как изображение худшее, чем исходное.

Метод передачи карты экранных параллаксов с использованием формата JPEG

Для оценки ожидаемой величины цифрового потока, которым может быть представлена карта экранных параллаксов при использовании для этой цели формата JPEG, был поставлен эксперимент, при выполнении которого были отобраны шесть 2D-изображений различного сюжетного содержания и различной структуры, а также соответствующие им карты экранных параллаксов. Далее каждая из карт экранных параллаксов была подвергнута сжатию путем ее записи в файл в формате JPEG при опциях $q = 0$ и $q = 1$, а затем «разжатию» с записью в файл в формате BMP. В результате такого преобразования в карты экранных параллаксов были внесены искажения (шум преобразования), величина которых была тем выше, чем выше степень сжатия. Величина сжатия файла карты экранных параллаксов при этом определялась как отношение размера файла карты экранных параллаксов в формате BMP к размеру файла карты экранных параллаксов в формате JPEG. Обработанные та-

ким образом карты экранных параллаксов были использованы для конвертации каждого из шести 2D-изображений в соответствующие им стереоскопические изображения в анаглифном формате. Кроме стереоскопических изображений, полученных описанным способом, были также изготовлены эталонные стереоскопические изображения с использованием карт экранных параллаксов, не подвергавшихся операции сжатия.

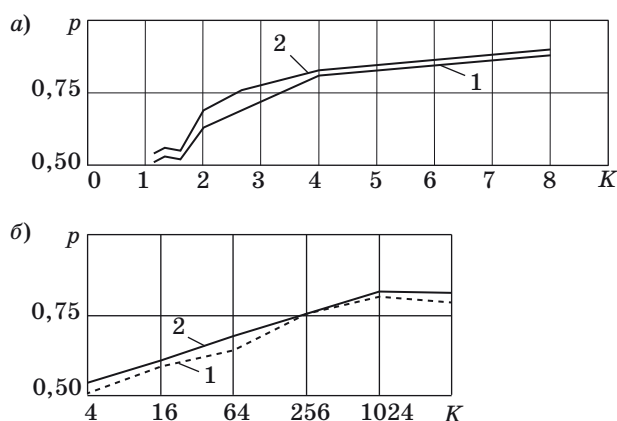
В результате обработки данных эксперимента было найдено, что при компрессиях, усредненные значения которых по шести изображениям составляли $K = 73$ и $K = 71$, что соответствовало значениям параметра $q = 0$ и $q = 1$, вероятности, с которой зритель обнаруживал изображение, полученное рассматриваемым способом, усредненная по результатам экспериментов, выполненных 12 зрителями с шестью изображениями, составила $p = 0,613$ ($q = 0$) и $p = 0,622$ ($q = 1$) при времени наблюдения 10 с.

Метод сжатия карты экранных параллаксов путем сокращения количества разрядов кода

Для оценки ожидаемой относительной величины цифрового потока, которым может быть представлена карта экранных параллаксов, был поставлен эксперимент, при выполнении которого были отобраны шесть 2D-изображений и соответствующих им карт экранных параллаксов.

Далее карты экранных параллаксов были проквантованы на 2, 4, 8, 16, 32 и 64 уровня яркости с использованием равномерной шкалы квантования, после чего для каждой пары «2D-изображение — проквантованная карта экранных параллаксов» были получены соответствующие им стереоскопические изображения в анаглифном формате. Естественно, что чем меньшее количество уровней была проквантована карта экранного параллакса, тем больший уровень искажений возникал в стереоскопическом изображении, которое было получено с ее использованием, поскольку при этом с большей погрешностью представлялось расположение объектов по дальности. Полученная таким образом стереопара предъявлялась зрителю одновременно с исходной стереопарой в целях обнаружения различий в соответствии с описанной выше методикой. На рис. 5, а приведена усредненная зависимость вероятности, с которой зритель обнаруживал изображение, полученное рассматриваемым способом, в зависимости от величины компрессии K , имеющей место при используемом количестве уровней квантования. При вычислении этой зависимости усреднение производилось по результатам экспериментов, выполненных 11 зрителями с шестью изображениями.

Обращает на себя внимание то, что уже при пяти разрядах кода, что соответствует $K = 1,6$,



■ **Рис. 5.** Зависимость вероятности p , с которой зритель обнаруживал искажения на изображении, от величины компрессии K за счет сокращения количества разрядов кода (а) и плотности отсчетов (б): кривая 1 и кривая 2 соответствуют длительности наблюдения 10 и 20 с соответственно

■ **Fig. 5.** Dependence of the probability p with which the viewer detected distortions in the image on the value of compression K by reducing the number of bits of code (a) and the sample density (b): curve 1 and curve 2 correspond to observation durations of 10 s and 20 s respectively

более 60 % зрителей замечают разного рода артефакты (см. также [12, 13]).

Метод сжатия карты экранных параллакс путем сокращения плотности отсчетов

Для оценки ожидаемой относительной величины цифрового потока, которым может быть представлена карта экранных параллакс при методе сокращенной плотности отсчетов, был поставлен эксперимент, при выполнении которого были отобраны шесть 2D-изображений различного сюжетного содержания и различной структуры, а также соответствующие им карты глубины.

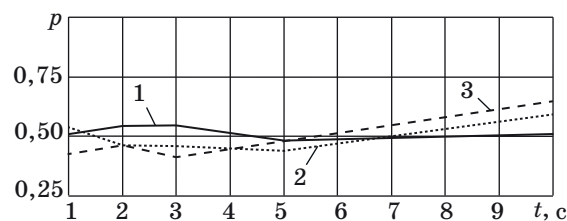
С картами экранных параллакс в этом эксперименте были выполнены следующие преобразования. Вначале из каждой карты посредством графического редактора Adobe PhotoShop были изготовлены версии, размеры которых были уменьшены в 2, 4, 8, 16 и 32 раза. При уменьшении линейного размера карты экранных параллакс количество отсчетов, которым она представляется, уменьшается пропорционально квадрату величины этого уменьшения, т. е. происходит сжатие цифрового потока, которым она представляется. Затем каждая из уменьшенных версий карт экранных параллакс посредством этого же редактора была восстановлена до своего первоначального размера, при этом недостающие отсчеты восстанавливались путем интерполяции.

После этого для каждой пары «2D-изображение — преобразованная карта экранных параллакс» были получены соответствующие им стереоскопические изображения в анаглифном формате.

Естественно, что чем меньшим количеством отсчетов представляется карта экранных параллакс, тем больший уровень искажений возникает в стереоскопическом изображении, которое получается с использованием этой карты.

Полученная таким образом стереопара предъявлялась зрителю одновременно с исходной стереопарой в целях обнаружения различий в соответствии с описанной выше методикой. Время предъявления изображений для их сравнения при выполнении эксперимента ограничивалось одним из следующих интервалов, которые устанавливались перед началом эксперимента: 1, 2, 3, 5, 10 и 20 с. На рис. 5, б приведены зависимости вероятности, с которой зритель обнаруживал изображение, полученное рассматриваемым способом, от величины компрессии K . В эксперименте приняли участие 10 зрителей.

На рис. 6, в отличие от рис. 5, б, представлена зависимость вероятности, с которой зритель обнаруживал изображение, полученное рассматриваемым способом, но от времени наблюдения для различных величин компрессии. При выполнении этого эксперимента приняли участие 10 зрителей. При рассмотрении графиков рис. 6 обращает на себя внимание тот факт, что при малых значениях времени наблюдения вероятности p оказываются близкими к 0,5, т. е. зритель не замечает искажений в стереоскопическом изображении, которые обусловлены погрешностями карты экранных параллакс. Этот феномен объясняется тем, что для определения глубины любой точки на стереоскопическом изображении необходимо, чтобы зрительная система произвела хотя бы одну саккаду, на которую затрачивается время около 0,2–0,3 с [14–17]. Для обнаружения искажений при передаче глубины в стереоскопическом изображении необходимо несколько саккад, чтобы можно было сравнить между собой пе-



■ **Рис. 6.** Зависимость вероятности p , с которой зритель обнаруживал искажения на изображении, от длительности наблюдения при величинах компрессии K , равных 4 (кривая 1), 16 (кривая 2) и 64 (кривая 3)

■ **Fig. 6.** The dependence of the probability p , with which the viewer detected distortions in the image on the duration of the observation at the compression values K : 4 (curve 1), 16 (curve 2) and 64 (curve 3)

редачу глубины различных частей изображения, на что и затрачивается 2–4 с.

Результаты исследований влияния уменьшенного пространственного и временного разрешения карты глубины на итоговый коэффициент сжатия приведены в работе [18].

Анализ полученных результатов

Проанализируем полученные экспериментальные результаты исходя из того, что требуется разработать формат представления контента, обеспечивающий совместимость 2D- и 3D-систем телевизионного вещания, и сформулируем предъявляемые к нему требования. Создаваемый формат должен обеспечивать возможность просмотра 3D-контента на 2D-телевизорах в 2D-формате, а на 3D-телевизорах — 2D-контента в формате 2D. Кроме того, поскольку в настоящее время у населения уже имеется небольшое количество 3D-телевизоров, желательно также обеспечить возможность смотреть 3D-контент в новом формате на этих телевизорах путем применения несложных перекодирующих приставок. Будем иметь в виду, что поскольку компрессия для передачи парного компонента стереопары в том или ином виде необходима, то для ее реализации желательно не применять сложные алгоритмы компрессии. Отметим также, что для того, чтобы затраты полезной площади кадра, используемой для передачи информации о различии между левым и правым изображениями стереопары, не превышали 5 %, величина компрессии должна быть не менее 20 при незаметности ее последствий для зрителя.

Рассматривая результаты экспериментов по определению вероятности, с которой зритель обнаруживает изображение, полученное рассматриваемым способом, мы сразу же должны отказаться от метода передачи разности компонентов стереопары по причине малой степени компрессии даже при использовании для ее реализации формата JPEG, которая в среднем при $q = 0$ составляет $K = 3,9$, что недостаточно. Вторая причина, делающая нежелательным применение этого метода, состоит в использовании сравнительно сложного алгоритма компрессии.

Далее мы вынуждены также отказаться от применения метода передачи карты экранных параллаксов путем сокращения количества рядов кода. Из экспериментов следует, что для того, чтобы вероятность, с которой зритель обнаруживает изображение, полученное рассматриваемым способом, была меньше 0,6, необходимо карту глубины передавать, используя для этой цели не менее пяти двоичных единиц кода, что соответствует компрессии $K = 1,6$.

Обращаясь к методу передачи карты экранных параллаксов с использованием формата JPEG, укажем его один, но существенный недостаток, а именно сложность алгоритма компрессии, хотя требованию обеспечения необходимой величины компрессии он удовлетворяет. Следует сказать, что применение этого алгоритма компрессии непосредственно к передаче парного компонента стереопары также обеспечивает требования совместимости, но тоже нежелательно вследствие сложности алгоритма. Отметим, что такие и аналогичные методы сжатия с потерями описаны, например, в работах [19, 20].

Таким образом, для рассмотрения остается метод передачи карты экранных параллаксов путем сокращения плотности отсчетов. Достоинством этого метода является простота декодирования сигнала, представляющего карту экранных параллаксов, которое по существу может быть сведено к интерполяции нулевого порядка, при его способности обеспечить необходимую степень компрессии, как это следует из графиков, представленных на рис. 5, б и 6. Более того, данный метод позволяет растянуть формирование карты экранных параллаксов в 3D телевизионном приемнике на несколько кадров (см. рис. 6), передавая отсчеты группами по k отсчетов, где $k = m/n$, m — количество отсчетов в карте экранных параллаксов, n — количество кадров, необходимое для передачи всей карты экранных параллаксов. В данном случае важно, чтобы время передачи этих n кадров не превышало 3–4 с, в течение которых зритель, как это следует из графика на рис. 6, не успевает заметить искажения, обусловленные погрешностями представления карты экранных параллаксов.

Заключение

Завершая изложение полученных результатов, отметим, что выполненные исследования, несмотря на их предварительный, оценочный характер, разрешают существенно сократить объем дальнейшей работы по созданию совместимого формата телевизионного вещания, позволяющего на 2D-телевизорах смотреть 3D-контент в 2D-формате, а на 3D-телевизорах — 2D-контент в формате 2D, исключив из дальнейшего рассмотрения целый ряд неперспективных, но весьма трудоемких для исследования методов передачи информации о недостающем компоненте стереопары. Действительно, для получения одной точки на любом из приведенных в статье графиков было выполнено более 1000 замеров.

Следует также отметить, что предлагаемый метод передачи карты экранных параллаксов, заключающийся в передаче ее в течение ряда кадров, по-

зволяет, помимо решения поставленной проблемы создания совместимого формата телевизионного вещания, решить еще одну проблему совместимости, а именно, обеспечить возможность просмотра передач в новом стандарте на 3D-телевизорах, которые в настоящее время хотя и в небольшом количестве,

но имеются у населения. Для этого можно создать сравнительно несложные индивидуальные приставки, посредством которых сигнал нового совместимого формата будет перекодироваться для этих телевизоров в один из имеющихся форматов — Anamorph SideBySide или Anamorph OverUnder.

Литература

1. Форматы 3D. <http://3d-roliki.ru/formati-3d/> (дата обращения: 25.11.2017).
2. 3D технологиям в телевизорах пришел конец. <http://nig.mirtesen.ru/blog/43919072251/3D-tehnologiyam-v-televizorah-prishel-konets> (дата обращения: 25.11.2017).
3. 3D Tile Format and 3DZ Tile Format Implementation Notes. <http://www.sisveltechnology.com/files/2013/3dtv/3D-3DZ-Tile-Format-implementation-notes.pdf> (дата обращения: 25.11.2017).
4. Sisvel Technology и Triaxes Vision представляют универсальный формат 3DZ Tile. <https://triaxes.com/downloads/info-3dz-tile-ru.pdf> (дата обращения: 25.11.2017).
5. Томской компанией Triaxes Vision разработан новый универсальный 3D-формат. <https://sdelanounas.ru/blogs/21419/> (дата обращения: 25.11.2017).
6. Visintinh F. How does Sisvel Technology's 3D Tile Format Work? <http://www.sisveltechnology.com/files/2013/3dtv/SPECIALReportAugust2012.pdf> (дата обращения: 25.11.2017).
7. Боков А. А. Методы объективной оценки качества видео, конвертированного в стереоформат // Материалы XXII Междунар. науч. конф. студентов, аспирантов и молодых ученых «Ломоносов-2015», Москва, 13–17 апреля 2015 г. М., 2015. С. 11–14.
8. Ватолин Д. С., Боков А. А., Фёдоров А. А. Тенденции изменения технического качества стереокино — 5 лет после «Аватара» // Мир техники кино. 2015. № 3(9). С. 17–28.
9. Ватолин Д. С., Боков А. А. Переход индустрии 3D-фильмов к конвертации и методы оценки ее качества // Мир техники кино. 2016. № 3 (10). С. 14–23.
10. Kakaletsis E., and Nikolaidis N. A Technique for Fake 3D (2D-to-3D Converted) Video Recognition // Proc. 23rd European Signal Processing Conf. (EUSIPCO). Nice, 2015. P. 106–109.
11. Winkler S. Efficient Measurement of Stereoscopic 3D Video Content Issues. <http://vintage.winklerbros.net/Publications/ei2014iqsp.pdf> (дата обращения: 25.11.2017).
12. Раев О. Н. Дискретизация глубины киноизображения при цифровой стереосъемке // Мир техники кино. 2012. № 3 (6). С. 10–16.
13. Белозеров А. Е., Рожкова Г. И. Вариабельность предельной величины положительных экранных параллаксов // Мир техники кино. 2011. № 3 (5). С. 22–26.
14. Физиология сенсорных систем. Ч. I. Физиология зрения / под ред. Г. В. Гершуни. — Л.: Наука, 1971. — 416 с.
15. Ярбус А. Л. Роль движения глаз в процессе зрения. — М.: Наука, 1965. — 166 с.
16. Красильников Н. Н., Красильникова О. И. Исследование критического значения экранного параллакса // Информационно-управляющие системы. 2016. № 2. С. 19–25. doi:10.15217/issn1684-8853.2016.2.19
17. Красильников Н. Н. Влияние расстояния наблюдения на глубину пространства, воспроизводимую стереоскопическим изображением // Оптический журнал. 2016. № 8(83). С. 68–76.
18. Матюнин С. Б., Ватолин Д. С. Сжатие карт глубины при передаче стереовидео для цифрового телевизионного вещания // Мир техники кино. 2012. № 3 (6). С. 2–9.
19. Fecker U., Guenegues A., Scholz I., Kaup A. Depth Map Compression for Unstructured Lumigraph Rendering. <https://pdfs.semanticscholar.org/e6e3/d4c29b44c01093a691cec43ba13ec6b28e71.pdf> (дата обращения: 25.11.2017).
20. Ancush Rai, Jagadeesh Kannan R. Compressed Transmission of Depth Maps in 3D Stream Services for Robotics and Surveillance. <https://innovareacademics.in/journals/index.php/ajpcr/article/view-File/19644/11819> (дата обращения: 25.11.2017).

UDC 612.843.721

doi:10.15217/issn1684-8853.2018.1.11

Providing Compatibility for 2D and 3D Television Broadcasting Systems

Krasilnikov N. N.^a, Dr. Sc., Tech., Professor, NNKrasilnikov@yandex.ru

Krasilnikova O. I.^a, PhD, Tech., Associate Professor, OIKrasilnikova@yandex.ru

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

Problem: The introduction of 3D television broadcasting systems is now stalled by the lack of a compatible television broadcasting format which would allow 3D content to be viewed on 2D TVs in 2D format. From the economic point of view, it is important that any television content produced, including films, all kinds of broadcasts, advertisements, etc., can be watched on any TV screen, either 2D or 3D. **Purpose:** Search for a way to construct a stereoscopic image processing algorithm for a compatible format. **Methods:** The solution is based on the consideration of human vision features, on the methods of converting 2D images into 3D, on image compression, and also on a large amount of experimental measurements made in a group of spectators who were asked to observe stereoscopic images synthesized in various transmission formats, examining the visibility of the resulting distortions. **Results:** A principle of constructing a compatible format for 3D television broadcasting has been developed. In this format, the left component of a stereopair image is transmitted as the main one, but a small part of the frame transmission time, corresponding to a small area of the transmitted image, is allocated to transmit a signal containing the information which, using the left component, forms the right component of the stereopair. As such a signal, the screen parallax map pixels are used, whose transmission is stretched to 3–4 frames. In a 3D television receiver, using a zero-order interpolation method, screen parallax maps are reconstructed from these pixels. Using these maps along with the left components of the stereo pair, you can synthesize the missing right components of the stereo pairs, reproducing the stereoscopic image. 2D TV receivers simply play back the left component of the stereo pair. **Practical relevance:** Constructing a compatible format for 3D television broadcasting makes it possible to significantly reduce the complexity of developing broadcasting standards.

Keywords — Stereoscopic Image, Compatibility, Screen Parallax Map.

Citation: Krasilnikov N. N., Krasilnikova O. I. Providing Compatibility for 2D and 3D Television Broadcasting Systems. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 1, pp. 11–20 (In Russian). doi:10.15217/issn1684-8853.2018.1.11

References

1. *Format 3D* [3D Format]. Available at: <http://3d-roliki.ru/formati-3d/> (accessed 25 November 2017).
2. *3D tekhnologiyam v televizorakh prishel konets* [3D Technologies in TVs Came to an End]. Available at: <http://nig.mir-tesen.ru/blog/43919072251/3D-tehnologiyam-v-televizorah-prishel-konets> (accessed 25 November 2017).
3. *3D Tile Format and 3DZ Tile Format Implementation Notes*. Available at: <http://www.sisveltechnology.com/files/2013/3dtv/3D-3DZ-Tile-Format-implementation-notes.pdf> (accessed 25 November 2017).
4. *Sisvel Technology i Triaxes Vision predstavliaiut universal'nyi format 3DZ* [Sisvel Technology and Triaxes Vision Represent the Universal 3DZ Tile Format]. Available at: <https://triauxes.com/downloads/info-3dz-tile-ru.pdf> (accessed 25 November 2017).
5. *Tomskoi kompaniei Triaxes Vision razrabotan novyi universal'nyi 3D-format* [Tomsk Company Triaxes Vision Developed a New Universal 3D Format]. Available at: <https://sdelanounas.ru/blogs/21419/> (accessed 25 November 2017).
6. Visintinh F. *How does Sisvel Technology's 3D Tile Format Work?* Available at: <http://www.sisveltechnology.com/files/2013/3dtv/SPECIALReportAugust2012.pdf> (accessed 25 November 2017).
7. Bokov A. A. Methods for an Objective Evaluation of the Quality of Video Converted to a Stereo Format. *Trudy XXII Mezhdunarodnoi nauchnoi konferentsii studentov, aspirantov i molodykh uchenykh "Lomonosov-2015"* [Proc. 22th Intern. Conf. of Students, Graduate Students and Young Scientists "Lomonosov-2015"], Moscow, April 13–17, 2015, pp. 11–14 (In Russian).
8. Vatolin D. S., Bokov A. A., Fedorov A. A. Trends in the Change of Technical Quality Stereokino — 5 Years after the "Avatar". *Mir tekhniki kino*, 2015, no. 3 (9), pp. 17–28 (In Russian).
9. Vatolin D. S., Bokov A. A. Transition of the 3D Film Industry to the Conversion and Methods for Assessing its Quality. *Mir tekhniki kino*, 2016, no. 3 (10), pp. 14–23 (In Russian).
10. Kakaletsis E. and Nikolaidis N. A Technique for Fake 3D (2D-to-3D Converted) Video Recognition. *Proc. 23rd European Signal Processing Conference (EUSIPCO)*, Nice, 2015, pp. 106–109.
11. Winkler S. *Efficient Measurement of Stereoscopic 3D Video Content Issues*. Available at: <http://vintage.winklerbros.net/Publications/ei2014iqsp.pdf> (accessed 25 November 2017).
12. Raev O. N. Discretization of Motion Picture Depth on Digital Stereo Shooting. *Mir tekhniki kino*, 2012, no. 3 (6), pp. 10–16 (In Russian).
13. Belozarov A. E., Rozhkova G. I. Variability of the Limiting Value of Positive Screen Parallax. *Mir tekhniki kino*, 2011, no. 3 (5), pp. 22–26 (In Russian).
14. *Fiziologiya sensorykh sistem. Ch. I. Fiziologiya zreniia* [Physiology of Sensory Systems. Part 1. Vision Physiology]. Ed. by G. B. Gershuni. Leningrad, Nauka Publ., 1971. 416 p. (In Russian).
15. Yarbus A. L. *Rol' dvizheniia glaz v protsesse zreniia* [The Role of Eye Movements in the Visual Process]. Moscow, Nauka Publ., 1965. 166 p. (In Russian).
16. Krasilnikov N. N., Krasilnikova O. I. Study of Critical Values of On-Screen Parallax. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2016, no. 2, pp. 19–25 (In Russian). doi:10.15217/issn1684-8853.2016.2.19
17. Krasilnikov N. N. Effect of Viewing Distance on Spatial Depth Reproduced by a Stereoscopic Image. *Opticheskii zhurnal* [Journal of Optical Technology], 2016, no. 8(83), pp. 502–507 (In Russian).
18. Matiunin S. B., Vatolin D. S. Compression of Depth Maps for the Transmission of Stereo Video for Digital Television Broadcasting. *Mir tekhniki kino*, 2012, no. 3 (6), pp. 2–9 (In Russian).
19. Ulrich Fecker, Audrey Guenegues, Ingo Scholz, Andre Kaup. *Depth Map Compression for Unstructured Lumigraph Rendering*. Available at: <https://pdfs.semanticscholar.org/g/e6e3/d4c29b44c01093a691cec43ba13ec6b28e71.pdf> (accessed 25 November 2017).
20. Ancush Rai, Jagadeesh Kannan R. *Compressed Transmission of Depth Maps in 3D Stream Services for Robotics and Surveillance*. Available at: <https://innovareacademics.in/journals/index.php/ajpcr/article/viewFile/19644/11819> (accessed 25 November 2017).

ПРИНЯТИЕ РЕШЕНИЙ В ЦЕНТРАЛЬНОЙ НЕРВНОЙ СИСТЕМЕ РОБОТА

А. Е. Городецкий^а, доктор техн. наук, профессор, g27764@yandex.ru

В. Г. Курбанов^{а, б}, канд. физ.-мат. наук, доцент, vugar_borchali.yahoo.com

И. Л. Тарасова^{а, в}, канд. техн. наук, доцент, g17265@yandex.ru

^аИнститут проблем машиноведения РАН, Большой пр. В. О., 61, Санкт-Петербург, 199178, РФ

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

^вСанкт-Петербургский политехнический университет Петра Великого, Политехническая ул., 29, Санкт-Петербург, 195251, РФ

Введение: умные электромеханические системы имеют возможность осуществлять параллельные вычисления, групповое управление, коммуникацию, хранение информации, мониторинг, измерение и контроль собственных параметров и параметров окружающей среды при параллельной кинематике исполнительных механизмов. Это позволяет подобным системам с точки зрения точности и жесткости манипулировать большими нагрузками. Поведение такой системы основывается на информации, полученной от входящей в ее состав центральной нервной системы о состоянии окружающей среды и собственном состоянии. **Цель исследования:** разработка алгоритмов принятия поведенческих решений роботов, подобных человеку и построенных на базе модулей интеллектуальных электромеханических систем на основе информации, получаемой от центральной нервной системы. **Результаты:** рассмотрены дедуктивный, индуктивный и абдукционный типы принятия поведенческих решений в центральной нервной системе робота, строящегося на базе модулей интеллектуальных электромеханических систем. Показано, что абдукционный метод является наиболее быстрым по аналогии с интуицией, но его надежность зависит от полноты базы данных хороших решений из прошлого опыта, т. е. от времени эксплуатации подобных роботов в похожих условиях окружения. Дедуктивный метод при большом числе ограничений более быстрый по сравнению с индуктивным, так как не требует проверки ограничений для всех решений. При сложных критериях качества и малом числе ограничений индуктивный метод может быстрее дать результат, так как отбрасывает поиск решения по сложным критериям качества для неприемлемых по ограничениям решений. **Практическая значимость:** исходя из рассмотренных типов принятия поведенческих решений в центральной нервной системе робота, предложенные алгоритмы можно использовать для формирования стратегии и тактики управления интеллектуальными роботами.

Ключевые слова — умные электромеханические системы, органы чувств человека, центральная нервная система человека, центральная нервная система робота, измерения, вычисления, управление, дедукция, индукция, абдукция.

Цитирование: Городецкий А. Е., Курбанов В. Г., Тарасова И. Л. Принятие решений в центральной нервной системе робота // Информационно-управляющие системы. 2018. № 1. С. 21–30. doi:10.15217/issn1684-8853.2018.1.21

Citation: Gorodetskiy A. E., Kurbanov V. G., Tarasova I. L. Decision-Making in Central Nervous System of a Robot. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 1, pp. 21–30 (In Russian). doi:10.15217/issn1684-8853.2018.1.21

Введение

Интеллектуальные электромеханические системы (SEMS) используются в киберфизических системах, в частности, в интеллектуальных роботах. Формализация задачи структурно-функционально-параметрического синтеза киберфизической системы и формирования соответствующего киберфизического пространства рассмотрена в работе [1]. Для правильного функционирования система автоматического управления (САУ) робота должна получать сведения об окружающей обстановке и «поведении» самого робота [2]. Полученные сведения позволяют САУ прогнозировать поведение киберфизической системы [3] и принимать решения о целях функционирования и управляющих воздействиях, необходимых для достижения этих целей, т. е. обе-

спечивать целесообразное или целеустремленное поведение робота [4]. В этом случае САУ роботов должны не только быть снабжены более совершенными датчиками ощущений (сенсорами), но и иметь способность понимать язык ощущений, т. е. обладать чувствами типа «свой — чужой», «опасно — безопасно», «любимый — нелюбимый», «приятно — неприятно» и др., формируемыми в результате решения систем логических уравнений. Такие САУ становятся похожими на центральную нервную систему человека [5]. Поэтому их можно по аналогии называть центральной нервной системой робота (ЦНСР), в которой в результате решения систем логических уравнений, формируемых на базе языка ощущений робота, появляется возможность самостоятельно принимать решения относительно целесообразного поведения [6, 7].

Принципы построения ЦНСР

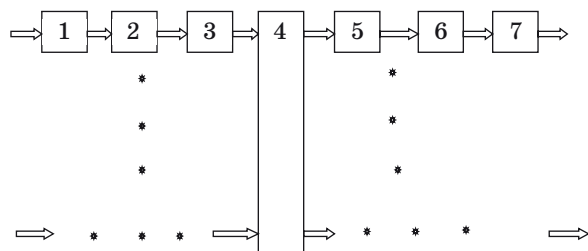
Построение ЦНСР сводится прежде всего к созданию следующей цепочки: (измерительные системы, включающие различные сенсоры и датчики) → (измерительный интерфейс, включающий информационные каналы приема сигналов и их первичной обработки) → (формирователь поведенческих решений, включающий фаззификацию сигналов, распознавание и классификацию образов, формирование бинарных оценок образов и принятие решений) → (формирователь управляющих воздействий, включающий выбор целей функционирования робота на основе выбора рефлексивных рассуждений, выработку управляющих воздействий, обеспечивающих целесообразное поведение робота, и передачу их на рабочие органы) → (рабочие органы робота, обеспечивающие различные перемещения, растяжения, сжатие, захваты и другие действия рабочих органов робота) (рисунок).

При этом решаются две основные взаимосвязанные задачи:

1) создание более совершенных датчиков и технических систем ощущений;

2) создание программных средств, обеспечивающих роботов способностью понимать язык ощущений и формировать поведенческие процессы на основе анализа ощущений, т. е. обеспечивать роботов возможностью рефлексивно рассуждать, что существенно приближает их знаковые системы к тем, которыми пользуется в своей повседневной практике человек.

Разработано достаточно много измерительных систем и сенсоров, имитирующих зрение (блок 1): системы технического зрения, ПЗС-матрицы,



■ Схема ЦНСР: 1 — измерительная система; 2 — канал передачи измерительных сигналов; 3 — блок предобработки измерительных сигналов; 4 — блок фаззификации, распознавания и принятия решения; 5 — канал передачи управляющих сигналов; 6 — блок формирования управляющих воздействий; 7 — рабочие органы робота

■ CNSR scheme: 1 — measuring system; 2 — measuring signal transmission channel; 3 — preprocessing block for measuring signals; 4 — block for fuzzification, recognition and decision making; 5 — control signal transmission channel; 6 — control action forming unit; 7 — robot working parts

оптико-электронные и телевизионные системы измерения перемещений, размеров, яркостей и других параметров [8]. То же можно сказать и о системах, имитирующих слух [9]. Имеются определенные успехи в области создания измерительных систем и сенсоров, имитирующих осязание [10]. Мало изученными и практически слабо используемыми являются измерительные системы, имитирующие обоняние и вкус [11]. Отсутствуют практические разработки, касающиеся создания технических систем, имитирующих чувство равновесия, интуицию и телепатию. Следовательно, решение первой задачи во многом еще не завершено.

Методы и средства передачи и преобразования сигналов в целях выделения информационных параметров из смеси сигнал-шум изучены достаточно хорошо, и имеется значительное количество различных технических средств приема-передачи и фильтрации (блоки 2, 3, 5) сигналов. Так же хорошо развиты технические средства формирования управляющих воздействий (блок 6).

Известны методы фаззификации данных (сигналов), логического вывода и логико-интервальные, логико-вероятностные и логико-лингвистические методы распознавания и принятия решения (блок 6) [12]. Однако решение задачи создания математических и программных средств, позволяющих обеспечивать роботов возможностью рефлексивно рассуждать, находится еще на самой начальной стадии, ограничивающейся моделированием поведенческих процессов на основе анализа ощущений.

Таким образом, имеющиеся технические решения на данном этапе позволяют приступить к созданию упрощенных прототипов ЦНСР. Одним из наиболее перспективных вариантов математической реализации блока 4 может быть логико-вероятностная реализация формирования поведенческих процессов [13] на основе анализа ощущений в виде сигналов от сенсорной системы робота. При этом первой и очень важной для дальнейших логических построений операцией является фаззификация данных, поступающих по сенсорным информационным каналам от различных датчиков.

Фаззификация сенсорной информации

Операция физзификации численных данных широко используется в интеллектуальных системах управления и, соответственно, в системах управления интеллектуальными роботами. После выполнения операции фаззификации для каждого сенсорного измерительного канала образуются множества X_i , содержащие наборы логических переменных x_{ij} . Например, для канала измерения яркости изображения можно

получить следующие логические переменные: x_{11} — «очень темное», x_{12} — «темное», x_{13} — «полутемное», x_{14} — «полусветлое», x_{15} — «светлое», x_{16} — «полуюркое», x_{17} — «яркое», x_{18} — «очень яркое». Полученные логические переменные для различных точек окружающего робот пространства могут быть истинными ($x_{ij} = 1$) или ложными ($x_{ij} = 0$). При этом часто могут возникать ситуации, когда при фаззификации численных данных об истинности или ложности получаемых тех или иных логических переменных можно говорить лишь с некоторой долей уверенности. В этом случае каждая полученная логическая переменная x_{ij} снабжается соответствующим атрибутом в виде вероятности $P\{x_{ij} = 1\}$ или функции $\mu(x_{ij})$, который хранится в базе данных вместе с логической переменной. Там же хранятся координаты точек окружающего робот пространства, соответствующие каждой логической переменной. Естественно, что при изменении окружения робота содержание базы данных обновляется.

При фаззификации логические переменные, например, характеризующие температуру, v_{3j} , образуются путем квантования всего диапазона датчика температуры и присвоения полученным квантам Δ_{3i} имен логических переменных v_{3j} , принимающих значения истина (1) или ложно (0). Тогда, если входная переменная — температура t — может изменяться в пределах от -20 до $+20$ °C, то, введя квант в 10 °C, можно весь диапазон изменения температуры разбить на четыре кванта: $\Delta_{31} = [-20, -10]$, $\Delta_{32} = [-10, 0]$, $\Delta_{33} = [0, +10]$, $\Delta_{34} = [+10, +20]$. Затем кванту Δ_{31} можно присвоить имя v_{31} {очень холодно}, кванту Δ_{32} — имя v_{32} {холодно}, кванту Δ_{33} — v_{33} {прохладно} и кванту Δ_{34} — имя v_{34} {тепло}. При этом:

— логической переменной v_{31} будет соответствовать интервал

$$\begin{aligned} & [(-20 + 20)/(20 + 20), \\ & (-10 + 20)/(20 + 20)] = [0; 0,25]; \end{aligned}$$

— логической переменной v_{32} будет соответствовать интервал

$$\begin{aligned} & [(-10 + 20)/(20 + 20), \\ & (0 + 20)/(20 + 20)] = [0,25; 0,5]; \end{aligned}$$

— логической переменной v_{33} будет соответствовать интервал

$$\begin{aligned} & [(0 + 20)/(20 + 20), \\ & (10 + 20)/(20 + 20)] = [0,5; 0,75]; \end{aligned}$$

— логической переменной v_{34} будет соответствовать интервал

$$\begin{aligned} & [(10 + 20)/(20 + 20), \\ & (20 + 20)/(20 + 20)] = [0,75; 1]. \end{aligned}$$

В частности, если, например, датчик показывает температуру $t = +50$ °C, то после фаззифика-

ции в базу данных ЦНСР будут занесены следующие значения логических переменных: $v_{31} = 0$, $v_{32} = 0$, $v_{33} = 1$, $v_{34} = 0$ — и соответствующие им описанные выше интервалы как атрибуты этих логических переменных.

В более сложных случаях в качестве атрибутов могут быть вероятности $P\{x_{ij} = 1\}$ или функции принадлежности $\mu(x_{ij})$. Тогда данные от сенсоров органов чувств роботов хранятся в памяти в виде логико-вероятностных либо логико-лингвистических переменных, когда в качестве атрибутов получаемых после фаззификации логических переменных в базе данных ЦНСР будут храниться вероятности либо функции принадлежности. В этом случае машина логического вывода ЦНСР практически всегда будет получать не одно решение, а несколько с разной степенью уверенности. Например:

если $x_{ij} = 1$ с вероятностью P_x , и $y_{ij} = 1$ с вероятностью P_y , и $z_i = 1$ с вероятностью P_z , и $u_{ij} = 1$ с вероятностью P_u , и $v_{ij} = 1$ с вероятностью P_v , и $w_{ij} = 1$ с вероятностью P_w , то $q_{ij} = 1$ с вероятностью P_q .

Сенсоры целесообразно объединять в группы, образующие следующие органы чувств робота подобно органам чувств человека: зрение в виде множества X , слух в виде множества Y , обоняние в виде множества Z , вкус в виде множества U , осязание в виде множества V , равновесие в виде множества W и телепатию в виде множества Q .

В каждом из введенных множеств можно выделить образующие их подмножества, характеризующие свойства наблюдаемого или изучаемого объекта: $X_i \subset X$, $Y_i \subset Y$, $Z_i \subset Z$, $U_i \subset U$, $V_i \subset V$, $W_i \subset W$, $Q_i \subset Q$.

Набор таких подмножеств зависит от набора сенсоров, образующих органы чувств конкретного робота. Например, могут быть введены следующие подмножества:

— для зрения: X_1 — контур изображения, X_2 — размер изображения, X_3 — яркость изображения, X_4 — цвет изображения, X_5 — расстояние до объекта, X_6 — скорость приближения, X_7 — скорость удаления;

— для слуха: Y_1 — громкость, Y_2 — тональность, Y_3 — интервал, Y_4 — скорость приближения, Y_5 — скорость удаления, Y_6 — направление;

— для обоняния: Z_1 — тип запаха, Z_2 — интенсивность запаха, Z_3 — направление запаха, Z_4 — скорость приближения, Z_5 — скорость удаления;

— для вкуса: U_1 — тип вкуса, U_2 — сила вкуса, U_3 — направление;

— для осязания: V_1 — ровность поверхности, V_2 — сухость поверхности, V_3 — температура поверхности.

Чувство равновесия у роботов обеспечивается обычно гироскопами. В этом случае могут быть введены следующие подмножества: W_1 — откло-

нение «вверх — вниз»; W_2 — отклонение «вперед — назад»; W_3 — отклонение «влево — вправо»; W_4 — скорость отклонения «вверх — вниз»; W_5 — скорость отклонения «вперед — назад»; W_6 — скорость отклонения «влево — вправо».

Телепатия у роботов, в отличие от человека, вполне объяснимое явление. У них прием и передача сообщений производятся по каналам беспроводной связи. В настоящее время наиболее широко для этого используется Wi-Fi. Наиболее просто подмножества Q_i , образующие множество телепатии Q , сформировать заранее, на стадии проектирования робота, предназначенного для выполнения тех или иных технологических операций. В этом случае такими подмножествами будут инструкции или типы реакций ($q_{ij} \in Q_i$), которые хранятся в памяти робота и извлекаются оттуда машиной логического вывода.

Принятие решений в ЦНСР

При принятии решения о целесообразном поведении робота необходимо в определенной последовательности выполнить следующие процедуры: выделение и распознавание образов в окружении робота; получение бинарных оценок этих образов; формирование рефлексивных или поведенческих рассуждений робота на основе логического анализа бинарных оценок образов; определение целей функционирования робота на основе выбора рефлексивных рассуждений и выбор оптимального решения, соответствующего целям функционирования и текущим ограничениям.

Выделение и распознавание образов

Операция выделения образов M_i в окружающем робот пространстве широко используется в системах технического зрения интеллектуальных роботов. В простейшем случае эта операция сводится к объединению в одно множество M_i тех точек пространства, которые обладают одинаковым набором логических переменных с одинаковыми атрибутами, при условии, что расстояние до ближайшей соседней точки с теми же параметрами не превышает некоторой заранее заданной величины. При этом определяются координаты центра тяжести получаемых изображений (образов). После объединения точек в множества последние могут получить дополнительные качественные параметры в виде логических переменных y_{ij} после, например, анализа геометрических параметров (площадей, объемов, контуров и др.) изображений. Эти дополнительные параметры: y_{11} — «большой объем», y_{21} — «гладкий контур» и др. — заносятся в базу данных в раздел «множество образов» вместе с другими логическими

параметрами множеств и координатами их центров тяжести. Содержание этого раздела базы данных обновляется при изменении окружения робота. При возникновении ситуаций не полной определенности в процессе объединения точек пространства в некоторое множество (образ) из-за, например, вероятностных атрибутов логических переменных необходимо, помимо геометрических мер близости точек, вводить дополнительные меры близости, например, допустимый разброс значений вероятности логических переменных у соседних точек.

При формировании образов в окружающем робот пространстве обычно используются следующие правила работы с данными:

$$\text{If } (x_{ij} = 1 \wedge y_{ij} = 1 \wedge z_{ij} = 1 \wedge u_{ij} = 1 \wedge v_{ij} = 1 \wedge w_{ij} = 1), \\ \text{then } q_{ij} = 1. \quad (1)$$

При большом количестве логических переменных таких правил может быть очень много. Тогда последовательный перебор правил в целях выявления их выполнимости будет занимать значительное время. В этом случае желательно использовать параллельные вычисления. Для этого, как показано в работе [14], можно использовать процедуру алгебраизации логических выражений, заключающуюся в следующем.

Правила вида (1) представляют собой импликации на языке алгебры логики или булевой алгебры:

$$x_{ij} \wedge y_{ij} \wedge z_{ij} \wedge u_{ij} \wedge v_{ij} \wedge w_{ij} \rightarrow q_{ij}. \quad (2)$$

Выражения вида (2) можно преобразовать в форму алгебры Жегалкина или в эквивалентные алгебраические уравнения по mod 2

$$s_{ij} \oplus s_{ij} * q_{ij} \oplus 1 = b_{ij}, \quad (3)$$

где \oplus — знак сложения по mod 2; $*$ — знак умножения по mod 2; $b_{ij} = 0$ либо 1 и $s_{ij} = x_{ij} \wedge y_{ij} \wedge z_{ij} \wedge u_{ij} \wedge v_{ij} \wedge w_{ij}$.

Тогда полученную систему логических уравнений можно записать в матричной форме по mod 2 [14]

$$\mathbf{A} * \mathbf{F} = \mathbf{B}, \quad (4)$$

где \mathbf{A} — прямоугольная двоичная матрица размерности $[n, m]$, элементы которой состоят из нулей и единиц [14]; \mathbf{F} — фундаментальный вектор логической системы размерностью n , строящийся из комбинаций логических переменных $x_{ij}, y_{ij}, z_{ij}, u_{ij}, v_{ij}, w_{ij}, q_{ij}$, полученных при фазсификации сенсорных данных, и дополненный 1 на месте последнего элемента; \mathbf{B} — двоичный вектор размерности n .

Процедура получения системы уравнений (4) легко формализуется [14].

Естественно, что полученная таким способом матричная система уравнений по mod 2 вида (4) будет иметь большую размерность. Однако в реальной ЦНСР не все компоненты этого уравнения (не все сочетания логических переменных) физически реализуемы и могут быть отброшены. В результате такой редукции мы получим матричную систему уравнений по mod 2 меньшей размерности

$$\mathbf{C} * \mathbf{R} = \mathbf{G}, \quad (5)$$

где размеры матриц \mathbf{C} , \mathbf{R} , \mathbf{G} меньше размеров матриц \mathbf{A} , \mathbf{F} , \mathbf{B} соответственно.

Распознавание сформированных ЦНСР образов M_i — это отнесение их к тем или иным классам образов C_j^M , содержащим так называемые идеальные образы M_i , для которых известны принятые ранее оптимальные решения ($M_i^* \in C_j^M$). При этом можно существенно ускорить процесс принятия решения при формировании целей функционирования и целесообразного поведения робота.

Процедура распознавания образов в случае представления их в алгебре по mod 2 требует задания правил или алгоритмов обработки атрибутивной части, характеризующей логические переменные, при проведении над ними операций сложения и умножения по mod 2. Лингвистические атрибуты, характеризующие образы, образуют в общем случае неметризуемые множества V_i . В этом случае при распознавании выбор наилучшего класса из множества альтернативных может опираться на процедуру поиска бинарных отношений $M_i g M_j^*$, g — двуместный предикат на анализируемых множествах, который, например, может быть задан указанием формул логико-математического языка или формализованного лингвистического выражения [15]. При этом проблема выявления наилучшего приближения сводится к двум задачам: получению множеств M_i , M_j^* и конструированию оптимальной процедуры g , позволяющей получить количественную оценку близости M_i к M_j^* .

Создание исходной базы для конструирования g целесообразно начинать с выделения в каждом из сравниваемых множеств метризуемых подмножеств (например, подмножества вероятностей решений), для элементов которых могут быть указаны отношения и числовые меры близости. Следующим, наиболее сложным шагом является упорядочивание элементов неметризуемых подмножеств. Весьма вероятно, что для решения этой задачи понадобится построение новой системы логических уравнений, решение которой приведет либо к метризуемым множествам, либо к упорядоченным. В первом случае мы сразу получаем числовые меры близости, во втором эти меры придется строить заново. В качестве возмож-

ных числовых оценок могут быть использованы мощности множеств, количество совпадающих элементов, число групп совпавших элементов и др. Каких-либо рекомендаций по выбору тех или иных оценок в настоящее время нет в связи со слабой изученностью подобных моделей. В случае невозможности упорядочивания неметризуемых множеств решение о наибольшей близости какого-либо множества к эталону должен принимать сам разработчик или оператор, исходя из своих предпочтений, опыта и интуиции.

К наиболее часто используемым и легко конструируемым бинарным функциональным отношениям можно отнести:

- оценку по максимальному отклонению мощностей множеств;
- оценку по среднеквадратическому отклонению мощностей множеств;
- вероятностную оценку по максимальному отклонению мощностей множеств;
- вероятностную оценку по среднеквадратическому отклонению мощностей множеств.

Использование указанных бинарных функциональных отношений позволяет легко ранжировать образы M_i модели по их близости к эталонам M_j^* и при этом ввести числовую оценку близости.

Формирование бинарных оценок

Бинарные оценки образов можно получить путем логического анализа параметров, характеризующих образы. Для этого вначале необходимо составить правила присвоения данному образу той или иной бинарной оценки, например, если образ очень яркий, большой и быстро перемещается в сторону робота, то этот образ (объект) очень опасный. Система таких правил вносится в базу знаний ЦНСР на стадии создания робота. В ряде случаев она может корректироваться в процессе эксплуатации робота путем обучения или самообучения. При большом количестве таких правил целесообразно их привести к системе алгебраических уравнений по mod 2 или к алгебре логики Жегалкина (4). В результате мы получаем матричные уравнения, решение которых легко распараллеливается матричными процессорами. Это позволяет резко ускорить логический анализ параметров образов. Полученные бинарные оценки образов также заносятся в базу данных образов в окружающем робот пространстве. При изменении окружения робота бинарные оценки и сами образы обновляются.

Выбор бинарной оценки осуществляется путем проверки выполнимости тех или иных правил работы с данными из подмножеств ($x_{ij} \in X_i$, $y_{ij} \in Y_i$, $z_{ij} \in Z_i$, $u_{ij} \in U_i$, $v_{ij} \in V_i$, $w_{ij} \in W_i$). Подобные правила обычно имеют вид:

если $x_{ij} = 1$, и $y_{ij} = 1$, и $z_{ij} = 1$, и $u_{ij} = 1$, и $v_{ij} = 1$, и $w_{ij} = 1$, то $g_{ij} \in G_i$.

При большом количестве логических переменных таких правил может быть также очень много. Тогда последовательный перебор правил в целях выявления их выполнимости будет занимать значительное время. В этом случае желательно использовать параллельные вычисления. Для этого можно использовать процедуру алгебраизации логических выражений [14].

Данные логического типа $(x_{ij}, y_{ij}, z_{ij}, u_{ij}, v_{ij}, w_{ij})$ извлекаются из данных или сигналов от сенсоров органов чувств роботов путем их фаззификации. При этом с каждой логической переменной будет связан присущий ей атрибут. В простейшем случае таким атрибутом является интервал. В более сложных случаях в качестве атрибутов могут быть вероятности $P\{x_{ij} = 1\}$ или функции принадлежности $\mu(x_{ij})$. Поэтому данные от сенсоров органов чувств роботов хранятся в памяти в виде логико-интервальных, логико-вероятностных либо логико-лингвистических переменных. Машина логического вывода относительно бинарных оценок образов практически всегда будет получать не одно решение, а несколько с разной степенью уверенности. Например, если $x_{ij} = 1$ с вероятностью P_x , и $y_{ij} = 1$ с вероятностью P_y , и $z_{ij} = 1$ с вероятностью P_z , и $u_{ij} = 1$ с вероятностью P_u , и $v_{ij} = 1$ с вероятностью P_v , и $w_{ij} = 1$ с вероятностью P_w , то $g_{ij} = 1$ с вероятностью P_{ij} или $g_{kj} = 1$ с вероятностью P_{kj} .

Это, естественно, будет приводить к неоднозначности поведения робота. Человек в данной ситуации ведет себя целесообразно или целеустремленно интуитивно, опираясь на собственный опыт либо генетически заложенный поведенческий стереотип [15]. Задача надления роботов навыками целесообразного поведения пока находится на самом начальном этапе. Обычно в таких случаях в системе управления роботом должна запускаться процедура поиска оптимального решения.

В настоящее время наиболее полно изучены проблемы выбора оптимальных решений в условиях неполной определенности интервального, вероятностного либо лингвистического типа [16]. При этом требуется использовать известные [17, 18] алгоритмы вычисления таких атрибутов логических функций, как интервалы, вероятности и функции принадлежности.

Формирование рефлексивных рассуждений

Рефлексивные или поведенческие рассуждения робота можно получить на основе логического анализа бинарных оценок образов, находящихся в его окружении. Для этого необходимо составить правила типа «если — то» реакции на ту или иную бинарную оценку образа с учетом местоположения и состояния самого робота. Например:

если образ очень опасный и находится рядом, то робот должен удалиться от него;

если образ очень опасный, находится рядом и рядом находится большой хороший образ, то робот должен спрятаться за него.

Такие правила составляются и заносятся в базу знаний на стадии создания робота. Их может быть очень много, и они могут корректироваться в процессе эксплуатации. При этом целесообразно привести их к системе алгебраических уравнений по mod 2 или к алгебре логики Жегалкина для распараллеливания вычислений. Программа перевода системы правил в алгебраические уравнения по mod 2 должна входить в математическое обеспечение ЦНСР.

Определение целей функционирования

Формирование целей функционирования робота на основе выбора рефлексивных рассуждений, полученных после анализа бинарных оценок образов, окружающих робота, является сложной проблемой, связанной с решением плохо формализуемых многокритериальных оптимизационных задач. При этом часто требуется выбирать не одну конкретную цель, а последовательность следующих друг за другом целей при успешном выполнении предыдущих целей. На стадии проектирования робота невозможно предусмотреть все ситуации, в которых может находиться робот при принятии решения о выборе цели функционирования. Поэтому в память робота вносятся возможные по предполагаемым условиям эксплуатации ситуации и соответствующие возможные цели с индексом их эффективности. Тогда в ЦНСР должно быть такое программное обеспечение, которое могло бы путем оценки допустимых рефлексивных рассуждений и имеющихся пригодных наиболее эффективных целей функционирования в данной ситуации выбора составить последовательность целей, которая обеспечивала бы экстремум критерия качества (функционала). Формирование такого критерия качества является сложной и трудоемкой задачей, решение которой связано прежде всего с формированием и решением ряда логических задач, приводящих к формуле вычисления критерия качества [17]. Наиболее хорошие результаты при решении этой проблемы можно получить при использовании многошагового обобщенного программирования и программных сред типа A-life.

Выбор оптимального решения

Человек в процессе мышления и принятия решения на основе обработки имеющейся информации обычно придерживается одного из двух стилей — дедуктивного или индуктивного. Имеется и третий, плохо изученный и редко

встречающийся, тип мышления — абдуктивный. Рассмотрим варианты использования в ЦНСР каждого из перечисленных типов принятия решений.

При *дедуктивном* принятии решения процесс мышления в ЦНСР начинается с глобального уровня, и далее идет движение вниз к локальному. Техническим аналогом такого вида мышления может быть процесс оптимизации, когда вначале на основе имеющейся информации ищется наилучшее решение из всех возможных, а затем путем проверки на основе имеющейся информации всех ограничений производится коррекция решения.

Выбор оптимального решения из всех решений q_{ij} , полученных из системы (5), может осуществляться различными способами. Наиболее просто в этом случае использовать методы математического программирования. Тогда при логико-вероятностном описании неопределенностей [16], т. е. когда атрибутами логических переменных в уравнениях (4), (5) являются вероятности $P\{q_{ij} = 1\}$, критерий качества можно выразить следующим образом:

$$f_0(Q_i) = P\{q_{ij} = 1\} \rightarrow \max. \quad (6)$$

При этом значения вероятностей $P\{q_{ij} = 1\}$ можно вычислять приближенно по описанному в работах [15, 18] алгоритму.

Если в результате анализа какой-либо ЦНСР будет выявлено, что влияние тех или иных компонент q_{ij} на ее поведение различно, то критерий качества (6) целесообразно привести к виду

$$f_0(Q_i) = \beta_i P\{q_{ij} = 1\} \rightarrow \max, \quad (7)$$

где β_i — назначаемые весовые коэффициенты.

При логико-лингвистическом описании неопределенностей, т. е. когда атрибутами логических переменных в уравнениях (4), (5) являются функции принадлежности $\mu(q_{ij})$, критерий качества можно выразить следующим образом:

$$f_0(Q_i) = \mu(q_{ij}) \rightarrow \max. \quad (8)$$

При этом значения функций принадлежности $\mu(q_{ij})$ можно вычислять по описанным в работе [15] алгоритмам.

Если в результате анализа какой-либо ЦНСР будет выявлено, что влияние тех или иных компонент q_{ij} на ее поведение различно, то критерий качества (8) целесообразно привести к виду

$$f_0(Q_i) = \beta_i \mu(q_{ij}) \rightarrow \max. \quad (9)$$

При логико-интервальном описании неопределенностей, т. е. когда атрибутами логических

переменных в уравнениях (4), (5) являются интервалы $\Delta_{ij} = [a_{ji}, b_{ji}]$, критерий качества можно выразить в виде следующих выражений:

$$f_0(Y) = k_{ji}(b_{ji} - a_{ji}) \rightarrow \min; \quad (10)$$

$$f_0(Y) = k_{ji}[(b_{ji} - a_{ji}) - c_{ji}]^2 \rightarrow \min; \quad (11)$$

$$f_0(Y) = k_{ji}^b[(b_{ji} - b_{ji}^0)^2 + (a_{ji} - a_{ji}^0)^2] \rightarrow \min; \quad (12)$$

$$f_0(Y) = [k_{ji}^b(b_{ji} - b_{ji}^0)^2 + k_{ji}^a(a_{ji} - a_{ji}^0)^2] \rightarrow \min, \quad (13)$$

где k_{ji} , k_{ji}^b , k_{ji}^a — коэффициенты предпочтения лица, принимающего решения, об оптимальности; c_{ji} — желаемая для лица, принимающего решения, ширина интервала; b_{ji}^0 , a_{ji}^0 — желаемые лицом, принимающим решения, границы интервалов.

После вычисления критериев качества всех возможных решений в соответствии с уравнениями (6), (7) при логико-вероятностном описании неопределенностей, либо (8), (9) при логико-лингвистическом описании неопределенностей, либо (10)–(13) при логико-интервальном описании неопределенностей все найденные решения ранжируются. Затем решения проверяются на выполнимость ограничений для уравнений (4), (5), начиная с первого, имеющего наивысший критерий качества. При этом первое из проверяемых решений, удовлетворяющее ограничениям, считается оптимальным.

При *индуктивном* принятии решения процесс мышления в ЦНСР начинается с анализа отдельных решений, и далее идет поиск общего, глобального вывода. Техническим аналогом такого вида мышления может быть процесс оптимизации, когда вначале на основе имеющейся информации проверяются все решения на выполнимость ограничений, а затем ищется наилучшее решение из всех возможных по условиям ограничений решений по критериям типа (6)–(13).

При *абдуктивном* принятии решения, согласно Пирсу, познавательная деятельность в ЦНСР есть взаимодействие абдукции, индукции и дедукции [19]. При этом абдукция осуществляет принятие правдоподобных гипотез посредством объяснения фактов, с помощью индукции реализуется тестирование выдвинутых гипотез, а путем дедукции из принятых гипотез выводятся следствия. Техническим аналогом такого вида мышления может быть процесс поиска оптимального решения по аналогии, когда из всех возможных решений, получаемых из уравнений (5), вначале отбираются методами распознавания образов те решения, которые наиболее близки к уже имеющимся решениям, хранящимся в базе данных ЦНСР и дававшим в прошлом хорошие результаты. Затем можно методами дедуктивного и (или) индуктивного принятия решений по критериям качества (6)–(13) выбрать лучшее.

В более сложных случаях, характерных для интеллектуальных систем, когда не удается сформировать скалярный критерий качества, выбор оптимального решения из всех решений q_{ij} , полученных из системы (5), может осуществляться методами математического программирования в порядковых шкалах, обобщенного математического программирования или многошагового обобщенного математического программирования. Сравнивая описанные методы принятия решения, можно сделать вывод, что абдукционный метод является наиболее быстрым по аналогии с интуицией, но его надежность зависит от полноты базы данных хороших решений из прошлого опыта, т. е. от времени эксплуатации подобных роботов в похожих условиях окружения. Дедуктивный метод более быстрый по сравнению с индуктивным при большом числе ограничений, так как не требует проверки ограничений для всех решений. При сложных критериях качества и малом числе ограничений индуктивный метод может быстрее дать результат, так как отбросит поиск решения по сложным критериям качества для неприемлемых по ограничениям решений.

После выбора оптимального решения в блоке 4 (см. рисунок) запускается операция дефаззификации, под которой понимается процедура преобразования нечетких величин, получаемых в результате нечеткого вывода, в четкие. Эти величины по каналу передачи управляющих сигналов 5 передаются в блок 6 формирования управляющих воздействий. Обычно в блоке 6 формируются опти-

мальные законы управления рабочими органами 7, которые поступают туда в виде управляющих сигналов (напряжений). Методы поиска оптимальных управляющих воздействий изучены достаточно полно, например, в работе [20].

Заключение

Предложенные принципы дедуктивного, индуктивного и абдукционного принятия решений и соответствующие процедуры обработки информации в ЦНСР с использованием алгебраизации и матричного решения систем логических уравнений эффективно применять при формировании стратегии и тактики управления интеллектуальными роботами в условиях неполной определенности. При этом наиболее быстро будет приниматься решение при использовании абдукционного принципа, включающего элементы дедуктивного и индуктивного мышления. Достоверность и надежность принятия решения при таком подходе могут быть повышены в процессе эксплуатации робота, если включить в систему управления элементы самообучения, пополняющие базу отбираемых хороших решений, дававших в прошлом правильные решения.

Работа выполнена при финансовой поддержке РФФИ (проект № 16-29-04424, 18-01-00076, 18-51-06003 и 15-07-04760), РНФ (проект № 18-19-00005) и Программы Президиума РАН 1.31П «Актуальные проблемы робототехники».

Литература

1. Ронжин А. Л., Басов О. О., Соколов Б. В., Юсупов Р. М. Концептуальная и формальная модели синтеза киберфизических систем и интеллектуальных пространств // Изв. вузов. Приборостроение. 2016. № 11(59). С. 897–905. doi: 10.17586/0021-3454-2016-59-11-897-905
2. Distefano S., Merlino G., Puliafito A. Sensing and Actuation as a Service: A New Development for Clouds // Proc. of the 11th Intern. Symp. on Network Computing and Applications, NCA'12. IEEE Computer Society, Washington, DC, USA, 2012. P. 272–275.
3. Li B. S. X., Wan B., Wang C., Zhou X., Chen X. Definitions of Predictability for Cyber Physical Systems // J. of Systems Architecture. 2016. N 63. P. 48–60. doi: 10.1016/j.sysarc.2016.01.007
4. Добрынин Д. А. Интеллектуальные роботы вчера, сегодня, завтра // X национальная конф. по искусственному интеллекту с международным участием КИИ-2006, Обнинск, 25–28 сентября 2006 г.: тр. конф. в 3 т. Т. 2. М.: Физматлит, 2006. С. 20–32.
5. Rogatkin D. A., Kulikov D. A., Ivlieva A. L. Three Views on Current Data of Neuroscience for the Purposes of Intelligent Robotics // Modeling of Artificial Intelligence. 2015. Vol. 6. Iss. 2. P. 98–136.
6. Хахалин Г. К. Прикладная онтология на языке гиперграфов // Тр. второй Всерос. конф. с международным участием «Знания-Онтологии-Теории» (ЗОНТ-09), Новосибирск, 20–22 октября 2009 г. 2009. С. 223–231.
7. Аккоф Р., Эмери Ф. О целеустремленных системах. — М.: Сов. радио, 1974. — 269 с.
8. Поливцев С. А., Хашан Т. С. Исследование геометрических и акустических свойств сенсоров для системы технического слуха роботов // Проблемы бионики. 2003. № 6. С. 63–69.
9. Ming Ying, Andrew P. Bonifas, Nanshu Lu, Yewang Su, Rui Li, Huanyu Cheng, Abid Ameen, Yonggang Huang, and John A. Rogers. Silicon Nanomembranes for Fingertip Electronics // Nanotechnology. 2012. Vol. 23. N34. https://doi.org/10.1088/0957-4484/23/34/344004 (дата обращения: 15.01.2017).
10. Юревич Е. И. Сенсорные системы в робототехнике. — СПб.: Изд-во Политехнического ун-та, 2013. — 100 с.
11. Carolyn Mathas. The Five Senses of Sensors. Part 1. Smell, Taste and Hearing // Article Library. Digi-Key

- Corporation, 2015. www.digikey.com (дата обращения: 15.01.2017).
12. Городецкий А. Е., Курбанов В. Г., Тарасова И. Л. Эргатические методы анализа процессов эксплуатации и принятия решений при повреждениях и авариях энергообъектов // Информационно-управляющие системы. 2013. № 6. С. 29–36.
 13. Gorodetskiy A. E., Tarasova I. L., Kurbanov V. G. Behavioral Decisions of a Robot Based on Solving of Systems // Smart Electromechanical Systems: The Central Nervous System. 2017. P. 61–71. doi:10.1007/978-3-319-53327-8_5
 14. Дубаренко В. В., Курбанов В. Г. Метод приведения систем логических уравнений к форме линейных последовательных машин // Информационно-измерительные и управляющие системы. Т. 7. № 4. С. 37–40.
 15. Городецкий А. Е., Тарасова И. Л. Нечеткое математическое моделирование плохо формализуемых процессов и систем. — СПб.: Изд-во Политехнического ун-та, 2010. — 336 с.
 16. Gorodetskiy A. E., Kurbanov V. G., Tarasova I. L. Methods of Synthesis of Optimal Intelligent Control Systems SEMS // Smart Electromechanical Systems. 2016. P. 25–45. http://dx.doi.org/ 10.1007/978-3-319-27547-5_4 (дата обращения: 15.01.2017).
 17. Заде Л. А. Понятие лингвистической переменной и его применение к принятию приближенных решений. — М.: Мир, 1976. — 168 с.
 18. Дубаренко В. В., Курбанов В. Г., Кучмин А. Ю. Об одном методе вычисления вероятностей логических функций // Информационно-управляющие системы. 2010. № 5. С. 2–7.
 19. Пирс Ч. С. Рассуждение и логика вещей: лекции для Кембриджских конференций 1898 года. — М.: Изд-во РГГУ, 2005. — 371 с.
 20. Городецкий А. Е., Дубаренко В. В., Курбанов В. Г. Метод поиска оптимальных управляющих воздействий на динамические объекты с адаптацией к изменениям внешней среды // Тр. 6-го Санкт-Петербургского симп. по теории адаптивных систем (SPAS'99). СПб., 1999. С. 228–232.

UDC 62-52

doi:10.15217/issn1684-8853.2018.1.21

Decision-making in Central Nervous System of a RobotGorodetskiy A. E.^a, Dr. Sc., Tech., Professor, g27764@yandex.ruKurbanov V. G.^{a,b}, PhD, Phys.-Math., Associate Professor, vugar_borchali.yahoo.comTarasova I. L.^{a,c}, PhD, Tech., Associate Professor, g17265@yandex.ru^aInstitute of Problems of Mechanical Engineering of RAS, 61, Bolshoi Pr. V.O., 199178, Saint-Petersburg, Russian Federation^bSaint-Petersburg State University of Aerospace Instrumentation, 67, B. Morskaia St., 190000, Saint-Petersburg, Russian Federation^cPeter the Great St. Petersburg Polytechnic University, 29, Politechnicheskaja St., 195251, Saint-Petersburg, Russian Federation

Introduction: Smart electromechanical systems have the ability to perform parallel calculations, group control, communication, information storage, monitoring, measurement and control of their own or environmental parameters with parallel kinematics of the actuators. This allows such systems to manipulate large loads in terms of accuracy and rigidity. The behavior of such a system is based on the information it obtains from its own central nervous system about the state of the environment and the state of the system itself. **Purpose:** Developing algorithms of individual behavioral decisions for humanoid robots built from modules of intelligent electromechanical systems. The decisions should be based on the information received from the central nervous system. **Results:** The paper discusses the deductive, inductive and abductive types of behavioral decision-making in the central nervous system of a robot built from modules of intelligent electromechanical systems. It is shown that the abductive method is the fastest one by analogy with intuition, but its reliability depends on the completeness of the database of good solutions from the past experience, i.e. it strongly depends on the time of operating similar robots in similar environments. The deductive method, with a large number of constraints, is faster than the inductive one, as it does not require that the constraints for all the solutions are verified. Under complex quality criteria and a small number of constraints, the inductive method can produce faster results, as it discards the search for solutions by complex quality criteria for the decisions unacceptable by constraints. **Practical relevance:** Based on the considered types of behavioral decision-making in the central nervous system of a robot, the proposed algorithms can be used to formulate the strategy and tactics of the control over intelligent robots.

Keywords — Intelligent Electromechanical Systems, Human Sense Organs, Human Central Nervous System, Central Nervous System of a Robot, Measurements, Calculations, Control, Deduction, Induction, Abduction.

Citation: Gorodetskiy A. E., Kurbanov V. G., Tarasova I. L. Decision-Making in Central Nervous System of a Robot. *Informatsionno- upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 1, pp. 21–30 (In Russian). doi:10.15217/issn1684-8853.2018.1.21

References

- Ronzhin A. L., Basov O. O., Sokolov B. V., Yusupov R. M. Conceptual and Formal Models of Synthesis of Cyber-Physical Systems and Cyber-Physical Intellectual Spaces. *Izvestia vysshikh uchebnykh zavedenii. Priborostroneniye* [Journal of Instrument Engineering], 2016, vol. 59, no. 11, pp. 897–905 (In Russian). doi:10.17586/0021-3454-2016-59-11-897-905
- Distefano S., Merlino G., Puliafito A. Sensing and Actuation as a Service: A New Development for Clouds. *Proc. of the 11th Intern. Symp. on Network Computing and Applications, NCA'12*, IEEE Computer Society, Washington, DC, USA, 2012, pp. 272–275.
- Li B. S. X., Wan B., Wang C., Zhou X., Chen X. Definitions of Predictability for Cyber Physical Systems. *J. of Systems Architecture*, 2016, no. 63, pp. 48–60. doi:10.1016/j.sysarc.2016.01.007
- Dobrynin D. A. Intellectual Robots Yesterday, Today, Tomorrow. *Trudy X natsional'noi konferentsii po iskusstvennomu intellektu s mezhdunarodnym uchastiem KII-2006* [Proc. of the X National Conf. on Artificial Intelligence with International Participation KII-2006, Obninsk, September 25–28, 2006]. Moscow, Fizmatlit Publ., 2006, vol. 2, pp. 20–32 (In Russian).
- Rogatkin D. A., Kulikov D. A., Ivlieva A. L. Three Views on Current Data of Neuroscience for the Purposes of Intelligent Robotics. *Modeling of Artificial Intelligence*, 2015, iss. 2, vol. 6, pp. 98–136.
- Khakhalin G. K. Applied Ontology in the Language of Hypergraphs. *Trudy vtoroi Vserossiiskoi Konferentsii s mezhdunarodnym uchastiem "Znaniia-Ontologii-Teorii" (ZONT-09)* [Proc. of the second All-Russian Conf. with Intern. Participation "Knowledge-Ontology-Theory" (ZONT-09), Novosibirsk, October 20–22, 2009], 2009, pp. 223–231 (In Russian).
- Ackoff R., Emery F. *O tselestremlynykh sistemakh* [On Purposeful Systems]. Moscow, Sovetskoe radio Publ., 1974. 269 p. (In Russian).
- Polivtsev S. A., Khashan T. S. The Study of Geometric and Acoustic Properties of the Sensors for the Technical Hearing System. *Problemy bioniki* [Problems Robots Bionics], 2003, no. 6, pp. 63–69 (In Russian).
- Ming Ying, Andrew P. Bonifas, Nanshu Lu, Yewang Su, Rui Li, Huanyu Cheng, Abid Ameen, Yonggang Huang, and John A. Rogers. Silicon Nanomembranes for Fngertip Electronics. *Nanotechnology*, 2012, vol. 23, no. 34. Available at: <https://doi.org/10.1088/0957-4484/23/34/344004> (accessed 15 January 2017).
- Yurevich E. I. *Sensornye sistemy v robototekhnike* [Sensor Systems in Robotics]. Saint-Petersburg, Politekhnikeskii universitet Publ., 2013. 100 p. (In Russian).
- Carolyn Mathas. The Five Senses of Sensors. Part 1. Smell, Taste and Hearing. *Article Lirary*, 2015. Available at: www.digikay.com (accessed 15 January 2017).
- Gorodetskiy A. E., Kurbanov V. G., Tarasova I. L. Ergatic Operating Instructions Manual Methods of Analysis and Decision-Making Processes in Injuries and Accidents of Power. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2013, no. 6, pp. 29–36 (In Russian).
- Gorodetskiy A. E., Tarasova I. T., Kurbanov V. G. Behavioral Decisions of a Robot Based on Solving of Systems. *Smart Electromechanical Systems: The Central Nervous System*, 2017, pp. 61–71. doi:10.1007/978-3-319-53327-8_5
- Dubarenko V. V., Kurbanov V. G. The Method of Bringing the Systems of Logical Equations in the Form of Linear Sequential Machines. *Informatsionno-izmeritel'nye i upravliaiushchie sistemy* [Information-Measuring and Control Systems], 2009, vol. 7, no. 4, pp. 37–40 (In Russian).
- Gorodetskiy A. E., Tarasova I. L. *Nechetkoe matematicheskoe modelirovanie plokhho formalizuemykh protsessov i sistem* [Fuzzy Mathematical Modeling of Poorly Formalized Processes and Systems]. Saint-Petersburg, Politekhnikeskii universitet Publ., 2010. 336 p. (In Russian).
- Gorodetskiy A. E., Kurbanov V. G., Tarasova I. L. Methods of Synthesis of Optimal Intelligent Control Systems SEMS. *Smart Electromechanical Systems*, 2016, pp. 25–45. Available at: http://dx.doi.org/10.1007/978-3-319-27547-5_4 (accessed 15 January 2017).
- Zadeh L. A. The Concept of a Linguistic Variable and its Application to Approximate Reasonins. *Information Sciences*, 1975, no. 8(3), pp. 199–249.
- Dubarenko V. V., Kurbanov V. G., Kuchmin A. Y. A Method of Calculating the Probabilities of Logic Functions. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2010, no. 5, pp. 2–7 (In Russian).
- Peirce Ch. S. *Rassuzhdenie i logika veshchei: Lektsii dlia Kembridzhskikh konferentsii 1898 goda* [Reasoning and Logic of Things: The Cambridge Conference Lectures of 1898]. Moscow, RGGU Publ., 2005. 371 p. (In Russian).
- Gorodetskiy A. E., Dubarenko V. V., Kurbanov V. G. Method of Searching for Optimal Control Actions on Objects with Dynamic Adaptation to Changes in the Environment]. *Trudy 6-go Sankt-Peterburgskogo simpoziuma po teorii adaptivnykh sistem (SPAS'99)*. [Proc. 6th Saint-Petersburg Symp. on Adaptive Systems Theory (SPAS'99)]. Saint-Petersburg, 1999, pp. 228–232 (In Russian).

УДК 519.87

doi:10.15217/issn1684-8853.2018.1.31

РЕШЕНИЕ ЗАДАЧИ ПОДВОДНОГО НАБЛЮДЕНИЯ В УСЛОВИЯХ ПРИМЕНЕНИЯ ИНТЕЛЛЕКТУАЛЬНЫХ ПОМЕХ

Л. А. Мартынова^а, доктор техн. наук, старший научный сотрудник, martynowa999@bk.ru
^аАО «Концерн «ЦНИИ «Электроприбор», Малая Посадская ул., 30, Санкт-Петербург, 197046, РФ

Введение: автономные необитаемые подводные аппараты выполняют разнообразные задачи, в том числе наблюдение за другими автономными необитаемыми подводными аппаратами, представляющими определенный интерес, например, с коммерческой точки зрения. Задача аппарата-наблюдателя заключается в поддержании постоянного контакта с наблюдаемым объектом и получении информации о его координатах и параметрах движения. В то же время наблюдаемый аппарат стремится избавиться от наблюдения, для чего использует интеллектуальную помеху в виде высокоскоростного аппарата. Интеллектуальность помехи заключается в ее способности распознавать средства противодействия, которые использует против нее наблюдатель, и снижать их влияние выбранной стратегией своего поведения. До сих пор в задачах поиска и преследования объекта, в том числе и в условиях помех, поведение помехи не носило интеллектуальный характер. **Цель исследования:** оценка влияния поведения интеллектуальной помехи на возможность сохранения контакта с наблюдаемым аппаратом в условиях, когда ряд параметров помехи и стратегия ее поведения заранее не определены. **Результаты:** сформирован тактический эпизод наблюдения за сторонним аппаратом, использующим интеллектуальную помеху; выработаны стратегия использования наблюдателем средств противодействия и стратегия поведения высокоскоростного аппарата для снижения влияния средств противодействия. Для оценки влияния определен показатель эффективности, для его расчета разработана математическая модель. В основу расчета показателя эффективности наблюдения положен метод статистических испытаний (метод Монте-Карло). С помощью разработанной программной реализации математической модели проведены численные эксперименты, что позволило оценить влияние интеллектуальной помехи на возможность сохранения контакта наблюдателя с наблюдаемым аппаратом. **Практическая значимость:** результаты могут быть использованы при планировании стратегии поведения наблюдателя в условиях помех, при выборе технических решений по обеспечению наблюдения за сторонним аппаратом в условиях помех, при выборе средств противодействия интеллектуальной помехе, применяемой наблюдаемым аппаратом.

Ключевые слова — автономный необитаемый подводный аппарат, математическое моделирование, метод статистических испытаний, интеллектуальная помеха.

Цитирование: Мартынова Л. А. Решение задачи подводного наблюдения в условиях применения интеллектуальных помех// Информационно-управляющие системы. 2018. № 1. С. 31–41. doi:10.15217/issn1684-8853.2018.1.31

Citation: Martynova L. A. Underwater Observation under Intellectual Interference. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 1, pp. 31–41 (In Russian). doi:10.15217/issn1684-8853.2018.1.31

Введение

В сетевидной системе защиты заданного морского района одним из элементов являются подводные необитаемые аппараты, проводящие освещение обстановки, в том числе обнаружение аппаратов, оказавшихся в заданном районе, которые требуют наблюдения и сопровождения: один из аппаратов — наблюдатель — ведет наблюдение за другим аппаратом — наблюдаемым. Задача аппарата-наблюдателя заключается в поддержании постоянного контакта с наблюдаемым объектом и получении информации о его координатах и параметрах движения. Однако наблюдение со стороны наблюдателя может противоречить интересам наблюдаемого аппарата, поэтому для уклонения от наблюдения он может использовать помеху в виде высокоскоростного самоходного необитаемого подводного аппарата, существенно превосходящего в скорости перемещения наблюдателя. Высокоскоростной аппарат создает невыносимые условия наблюдения, по

крайней мере на некоторое время, и тем самым дает возможность наблюдаемому аппарату оторваться от наблюдателя. В таких условиях задача наблюдателя — продолжать вести наблюдение в условиях помех. При этом наблюдателю необходимо решить две задачи:

- максимально снизить эффект создания высокоскоростным аппаратом помехи;
- не потерять контакт с наблюдаемым аппаратом.

Последнее означает, что в случае уклонения от помехи наблюдатель может отойти на дистанцию, прерывающую его контакт с наблюдаемым аппаратом и не позволяющую вести наблюдение. В этом случае после преодоления помехи наблюдателю вновь придется вести поиск аппарата наблюдения и только после обнаружения перейти в режим его наблюдения.

Для того чтобы максимально снизить эффект создания помехи, наблюдателю необходимо, прежде всего, использовать маневр уклонения, однако этого недостаточно из-за большого превосход-

ства по скорости высокоскоростного аппарата. Поэтому дополнительным способом снижения влияния помехи является использование некоторого арсенала средств противодействия высокоскоростному аппарату. Высокоскоростной аппарат способен в силу своего интеллекта, в свою очередь, преодолеть эти средства, однако на это будет затрачено некоторое время, а время его функционирования ограничено энергоресурсом.

В связи с этим возникает задача оценки влияния интеллектуальной помехи на возможность сохранения контакта наблюдателя с наблюдаемым аппаратом.

Подобным задачам, классифицируемым как задачи поиска и преследования, в том числе в условиях помех [1–10], в литературе уделено достаточно внимания. Например, из теории игр известна задача о водителе-убийце [11], в которой гипотетический убегающий может двигаться медленно, но маневренно, пытаясь уйти от водителя, ведущего машину куда быстрее, но существенно ограниченного в маневре. При этом автор Р. Айзекс вместо «водителя» и «пешехода» подразумевал торпеду и увертывающийся от нее небольшой катер [12]. Дискретная версия задачи описана в работе [13].

В указанных работах формирование оптимального поведения помехи рассматривалось без присущего ей искусственного интеллекта, позволяющего менять тактику и параметры создания помехи в зависимости от сложившейся ситуации, связанной с использованием против них способов снижения формируемой помехи. До настоящего времени в такой постановке задача формирования помехи со сложным поведением и в условиях противодействия помехе не рассматривалась.

Целью настоящей работы является оценка влияния стратегии поведения наблюдателя в условиях интеллектуальной помехи, создаваемой высокоскоростным аппаратом, и стратегии использования против высокоскоростного аппарата мер противодействия, снижающих эффект влияния помехи на возможность наблюдателя удерживать контакт с наблюдаемым аппаратом.

Для достижения цели в работе были решены следующие задачи:

- определена тактическая ситуация, учитывающая различные стратегии поведения высокоскоростного аппарата и средств противодействия помехам;

- определены показатели эффективности сохранения контакта с наблюдаемым аппаратом путем создания противодействия помехам;

- разработана математическая модель поведения наблюдателя в условиях помех;

- разработан план численного эксперимента и проведены расчеты с использованием программной реализации разработанной математической модели;

- на основе полученных результатов выданы рекомендации по стратегии поведения наблюдателя в условиях помех.

Описание тактической ситуации наблюдения одним автоматическим аппаратом другого

Рассматривается задача ведения наблюдения одним автономным необитаемым подводным аппаратом другого. Аппарат-наблюдатель ведет наблюдение на дистанции, не превышающей $D_{\text{набл}}$, исходя из своих технических возможностей.

Наблюдаемый аппарат стремится уйти в отрыв от наблюдателя, для чего использует высокоскоростной аппарат, создающий помехи наблюдателю и вынуждающий его маневрировать. В результате наблюдаемому аппарату удается увеличить дистанцию, с которой наблюдение за ним становится проблематичным.

Предполагается, что в ответ на действия высокоскоростного аппарата наблюдатель для снижения создаваемой высокоскоростным аппаратом помехи использует:

- маневр уклонения по курсу, скорости и глубине, в ходе которого в максимально короткие сроки пытается выйти из трехмерного коридора, просматриваемого высокоскоростным аппаратом в процессе создания помех наблюдателю;

- специальные средства противодействия высокоскоростному аппарату для того, чтобы высокоскоростной аппарат потерял ориентацию в пространстве и затратил определенное время на ее восстановление.

Привести высокоскоростной аппарат к потере ориентации способны устройства, которые:

- создают помехи обнаружения своего сигнала дрейфующим прибором помех;

- создают искаженное поле, в результате чего скоростной аппарат сходит с истинного маршрута и движется в ложном направлении (имитаторы);

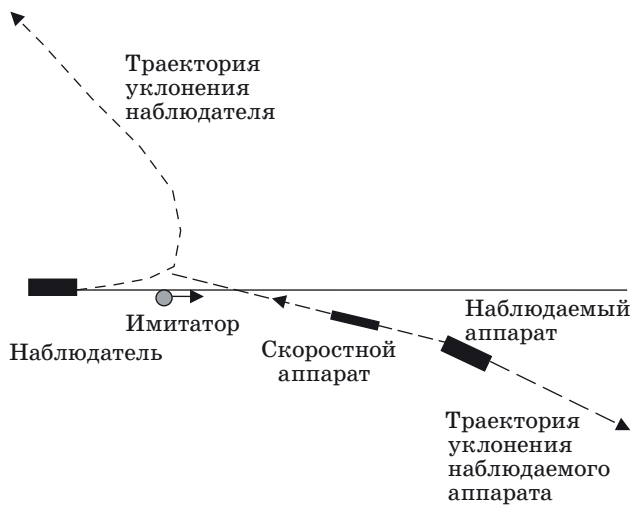
- создают физические препятствия, в которых скоростной аппарат «вязнет» и не имеет возможности выбраться (анти-аппарат).

Относительно высокоскоростного аппарата предполагается, что он достаточно интеллектуален, чтобы избегать создаваемые ему трудности, и постоянно пытается помешать наблюдателю путем сближения с ним на дистанцию контакта. Сближение на дистанцию контакта означает, что после этого наблюдатель на некоторое время не способен вести наблюдение за интересующим его аппаратом.

Относительно наиболее целесообразной стратегии использования средств, создающих трудности высокоскоростному аппарату, предполагается, что до того, как начать маневрирование,

наблюдатель выпускает самоходное средство, создающее ложное поле (на циркуляции выпуск самоходного средства (имитатора) невозможен по техническим причинам). На рис. 1 показана траектория движения имитатора, отводящего высокоскоростной аппарат на ложное направление.

Через некоторое время происходит выпуск дрейфующего прибора помех, в результате чего в зоне вокруг него создается высокоуровневый шум, затрудняющий обнаружение высокоскоростным аппаратом наблюдателя (рис. 2).

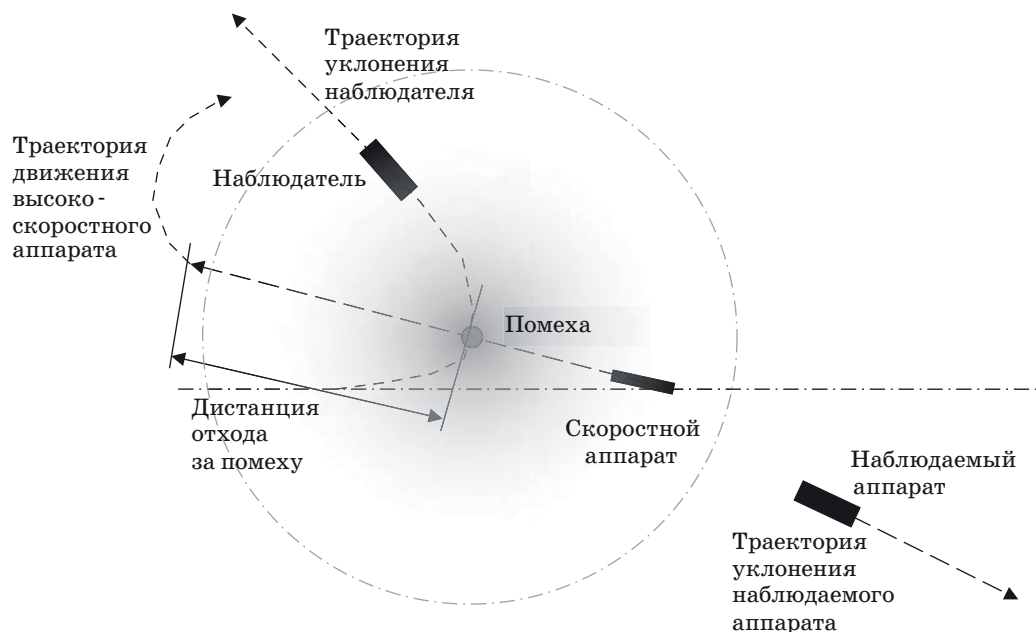


■ **Рис. 1.** Схема взаимного положения наблюдателя и наблюдаемого аппарата
 ■ **Fig. 1.** Diagram of the relative position of the observer and the observed apparatus

Наблюдатель использует дрейфующий прибор помех так, чтобы шум помехи не заглушал сигнал имитатора. Принимаем, что помеха может находиться и в стороне от направления движения скоростного аппарата. Если в результате прибор помех отвлек на некоторое время высокоскоростной аппарат, то последний возвращается и вновь осуществляет поиск наблюдателя. В случае его обнаружения высокоскоростной аппарат переходит в режим сопровождения наблюдателя. Тогда наблюдателем используется самоходный аппарат (анти-аппарат), приводящий в негодность высокоскоростной аппарат. Применение анти-аппарата наиболее эффективно на коротких дистанциях, так как на больших дистанциях может не хватить ресурса, и система наведения анти-аппарата не способна обнаружить и попасть в высокоскоростной аппарат. Дистанция использования анти-аппарата около 400 м.

Предполагается, что скоростной аппарат достаточно «умный», чтобы с течением времени разобраться в том, что перед ним ложный аппарат, или преодолеть создаваемое помехой поле. При возникновении имитатора, который вызывает его отведение на ложное направление, высокоскоростной аппарат с некоторой дистанции $D_{кл}$ классифицирует имитатор как ложную цель и понимает, что сошел с истинной траектории. После этого высокоскоростной аппарат возвращается в точку последнего контакта с наблюдателем, чтобы попытаться снова создать ему помеху.

При возникновении шумовой помехи резко сокращается дальность обнаружения, пока пре-



■ **Рис. 2.** Схема применения помехи, препятствующей обнаружению высокоскоростным аппаратом наблюдателя
 ■ **Fig. 2.** Schematic of the use of interference preventing detection of an observer by a high-speed device

следователь находится в зоне действия помех. За счет собственного интеллекта и проводимых измерений уровня шума самоходный аппарат находит оптимальный выход из района создаваемых ему помех. По уровню шума он выбирает в качестве наиболее целесообразной тактики движение в эпицентр источника помех с тем, чтобы пройти его насквозь, поскольку такой подход представляет собой кратчайший путь преодоления зоны действия помехи в виде круга. После преодоления помехи самоходный аппарат переходит на траекторию поиска наблюдателя для продолжения создания ему помех и отрыва от наблюдения с его стороны.

Принимается, что при близкой дистанции до высокоскоростного аппарата наблюдатель использует анти-аппарат, поскольку на таких дистанциях другие средства (имитатор и дрейфующий прибор помех) неэффективны. При этом высокоскоростной аппарат не имеет возможности без сторонней помощи выбраться и продолжить создание помехи наблюдателю.

Задача состояла в оценке возможности сохранения контакта с наблюдаемым аппаратом после уклонения от высокоскоростного аппарата.

Влияние на сохранение контакта с наблюдаемым аппаратом оказывает то, что стратегия применения мешающих средств со стороны наблюдателя выбирается таким образом, чтобы свести к минимуму попытки высокоскоростного аппарата помешать наблюдению.

После потери контакта с наблюдателем высокоскоростному аппарату придется возобновить поиск наблюдателя, однако ограниченность его энергоресурса не позволит ему вести этот поиск бесконечно.

Для проведения оценки возможности сохранения контакта с наблюдаемым аппаратом после его уклонения от высокоскоростного аппарата определен показатель эффективности его наблюдения.

Поскольку задача состоит в оценке влияния стратегии наблюдателя на сохранение контакта с наблюдаемым аппаратом после уклонения от высокоскоростного аппарата, то в качестве показателя эффективности выбрана вероятность сохранения контакта наблюдателя с наблюдаемым аппаратом.

Расчет вероятности сохранения контакта происходил с учетом того, что задача рассматривается в условиях неопределенности, т. е. при неизвестной заранее стратегии поведения высокоскоростного аппарата, которая зависит в свою очередь от используемых против него заранее известных средств противодействия.

Поскольку ряд параметров заранее неизвестен, включая параметры поведения высокоскоростного аппарата, то для расчета вероятности сохранения контакта наблюдателя с наблюдаемым аппаратом используем метод статистических ис-

пытаний (метод Монте-Карло) [14, 15], применяемый авторами ранее при проведении подобных исследований [16–20].

Проводилась серия из $N_{\text{исп}}$ испытаний. В каждом испытании определялось прерывание контакта наблюдателя с наблюдаемым аппаратом или сохранение контакта. В случае сохранения контакта испытание считалось результативным, в случае прерывания контакта — нерезультативным. Вероятность сохранения контакта определялась как отношение результативных испытаний к общему количеству испытаний.

Для расчета вероятности сохранения контакта была разработана математическая модель.

Описание математической модели

В математической модели рассматриваемые объекты: наблюдатель, наблюдаемый объект, высокоскоростной аппарат, самоходный имитатор — формализуем как материальные точки, движение которых определяется курсом, скоростью, дифферентом, а их положение характеризуется глубиной и координатами. При использовании помехи оценивалось время окончания маневра по курсу, скорости и глубине. При проведении расчетов производился переход из географической системы координат (широта, долгота) в местную (локальную) систему координат, определяемую следующим образом: центр земной системы координат совпадет с математическим ожиданием положения наблюдателя, первая ось направлена в сторону математического ожидания положения высокоскоростного аппарата на тот же момент времени, вторая ось направлена вертикально вверх, третья ось образует правую тройку с первыми двумя осями.

Время хода звука используемого скоростного аппарата до наблюдателя определяется выражением

$$t_{sndm} = \frac{r}{V_{ssnd}},$$

где V_{ssnd} — скорость звука в воде; r — расстояние между скоростным аппаратом и наблюдателем, которое определяется по формуле

$$r = \sqrt{x_1^2 + x_3^2}.$$

Здесь x_1 и x_3 — полярные координаты наблюдателя относительно высокоскоростного аппарата. Через стартовые декартовы координаты корабля $k_{k0[1]}$, $k_{k0[3]}$ и стартовые декартовы координаты k_{t01} , k_{t03} скоростного аппарата они вычисляются по формулам

$$x_1 = k_{k01} - k_{t01} + t_{beg} \cdot V_{ss} \cdot \cos(q_k);$$

$$x_3 = k_{k03} - k_{t03} + t_{beg} \cdot V_{ss} \cdot \sin(q_k),$$

где t_{beg} — время начала поиска; V_{ss} — текущая скорость аппарата; q_k — текущий курс наблюдателя.

Определение времени начала уклонения наблюдателя от высокоскоростного аппарата определяется выражением

$$t_y = t_{beg} + t_{sol} + t_{ndm},$$

где t_{sol} — время принятия решения.

Определение угла отворота наблюдателя Q_{otv} (значение угла — случайное, поскольку зависит от ряда факторов, которые заранее неизвестны) определяется выражением

$$Q_{otv} = \frac{\pi}{2} + p - q_k + Q_{mn} + x \cdot (Q_{mx} - Q_{mn}),$$

где p — угол между первой осью земной системы координат и направлением от высокоскоростного аппарата на наблюдателя; x — случайное число, разыгрываемое равномерно; Q_{mn} , Q_{mx} — минимальное и максимальное значение из заданного диапазона углов отворота.

Расчет времени t_V увеличения скорости до максимально допустимой V_{ssyf} происходил по формуле

$$t_V = t_y + b_{ass_1} \cdot \frac{V_{ssyf} - V_{ss}}{b_{ass_2} \cdot a_{ss}},$$

где b_{ass} , a_{ss} — коэффициенты движения скоростного аппарата по окружности.

Путь s_V , проходимый высокоскоростным аппаратом в ходе набора скорости до предельно допустимого значения, определяется выражением

$$s_V = b_{ass_1} \cdot V_{ss} \cdot (t_V - t_y) + b_{ass_2} a_{ss} \frac{(t_V - t_y)^2}{2}.$$

Пусть скорость s_{V1} высокоскоростного аппарата на циркуляции при повороте от текущего угла на угол отворота определяется выражением

$$s_{V1} = |Q_{otv}| \cdot r_z.$$

Сравниваем, какой путь при маневре больше: по скорости t_V или по курсу t_ψ ; если по скорости, т. е. $s_V < s_{V1}$, то путь s_{k1} , проходимый наблюдателем при ускорении в режиме циркуляции от начальной скорости до максимальной, определяется выражениями

$$s_{k1} = s_V;$$

$$t_\psi = t_V + t_{sol} + \frac{s_{V1} - s_v}{b_{ass_1} V_{ssyf}};$$

если по курсу:

$$s_{k1} = s_{V1};$$

$$t_\psi = t_y + t_{sol} + \frac{-b_{ass_1} \cdot V_{ss} + \sqrt{(b_{ass_1} \cdot V_{ss})^2 + 2b_{ass_2} \cdot a_{ss} \cdot s_{V1}}}{b_{ass_2} \cdot a_{ss}}.$$

Пересчет времени t_V , затрачиваемого на разгон до максимальной скорости с учетом потери скорости на циркуляции, определяется выражением

$$t_V = t_\psi + \frac{ssyf - V_{ss}}{a_{ss}} - b_{ass_2} \cdot \frac{t_\psi - t_y}{b_{ass_1}} - t_{sol}.$$

Тогда максимальное время, затрачиваемое на маневр, определяется соотношениями

$$t_{y_{mx}} = t_V \text{ при } t_\psi < t_V;$$

$$t_{y_{mx}} = t_\psi \text{ при } t_\psi \geq t_V.$$

Маневр по глубине считается следующим образом:

$$t_h = t_{y_{mx}} + \frac{h_{mn} - hk_2}{V_{ssyf} \cdot \sin \theta} \text{ при } \theta > 0;$$

$$t_h = t_{y_{mx}} + \frac{h_{mx} - hk_2}{V_{ssyf} \cdot \sin \theta} \text{ при } \theta \leq 0,$$

где θ — дифферент наблюдателя.

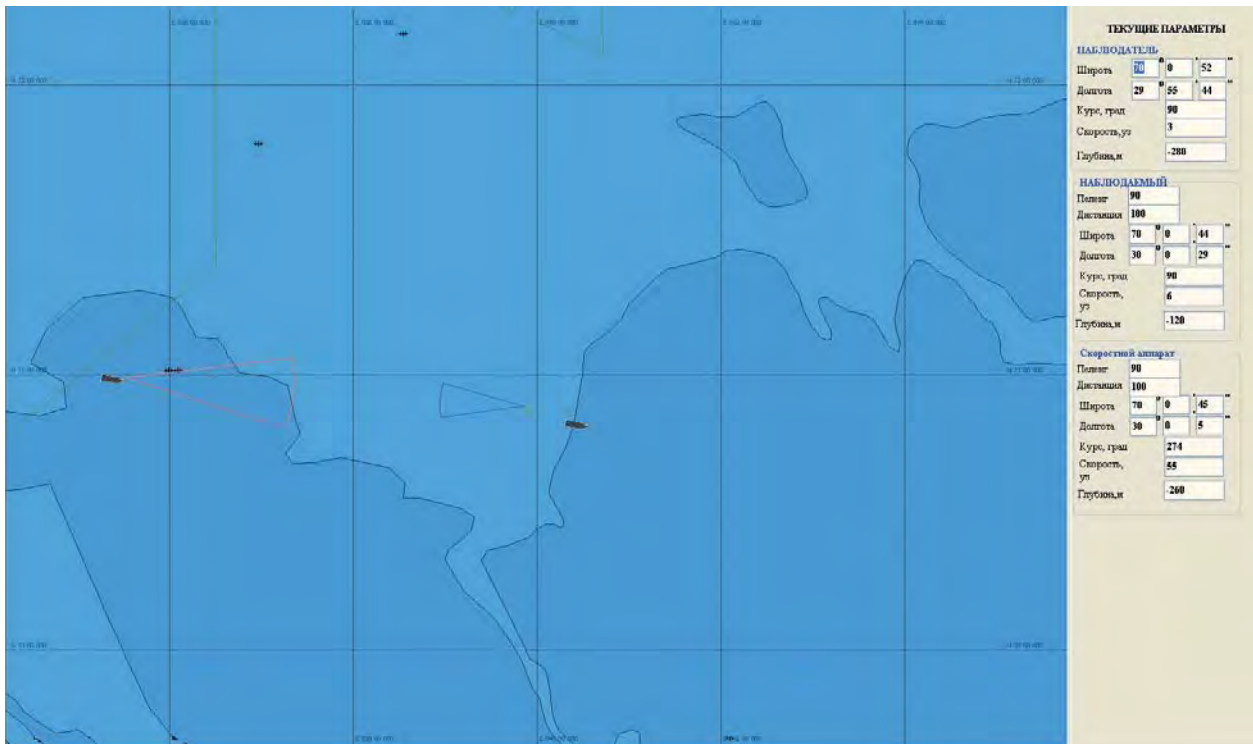
На этом расчет времени окончания маневра уклонения наблюдателя по курсу, скорости и глубине от высокоскоростного аппарата закончен. Полученные значения использованы затем в математической модели при определении текущих координат наблюдателя в период его маневрирования.

В каждом цикле имитации воспроизводились процессы перемещения наблюдателя, наблюдаемого объекта, высокоскоростного аппарата, самоходного имитатора в соответствии с принятой стратегией их поведения. Кроме того, воспроизводились результаты функционирования средств противодействия высокоскоростному аппарату и его поведение для снижения эффекта их воздействия.

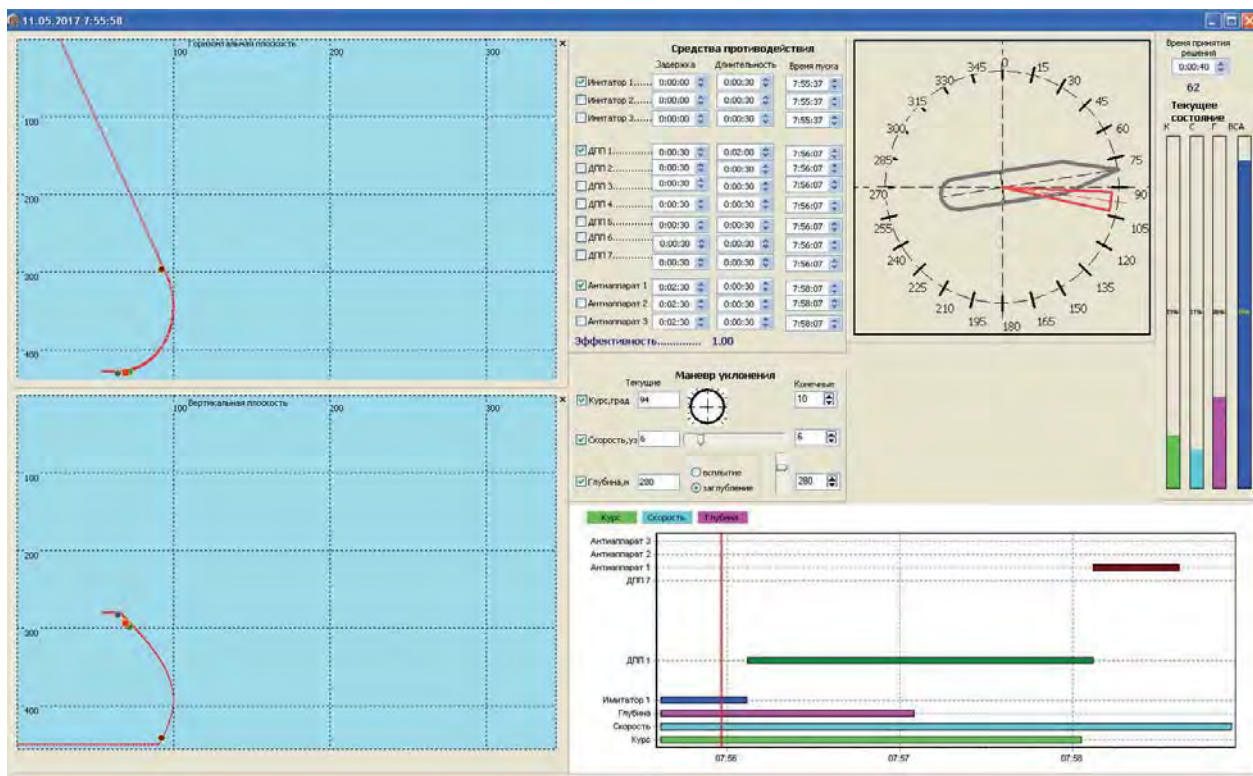
Описание программной реализации

Для проведения исследований по математической модели была разработана программа, внешний вид главного окна которой представлен на рис. 3.

На вспомогательном экране (рис. 4) показан инструментарий для проведения исследований



■ **Рис. 3.** Интерфейс главного окна программы для проведения исследований
 ■ **Fig. 3.** Appearance of the main window of the program for research



■ **Рис. 4.** Интерфейс окна программы с инструментарием для оценки эффективности ведения наблюдения в условиях помех
 ■ **Fig. 4.** Appearance of the window of the program with the tools provided to the researcher for assessing the effectiveness of conducting observations under interference conditions

по оценке эффективности ведения наблюдения за наблюдаемым аппаратом в условиях помех. В инструментарий входят:

— отображение траектории перемещения наблюдателя в ходе выполнения маневра уклонения;

— соотношение времени, затрачиваемого на маневр уклонения наблюдателя по курсу, скорости и глубине, на специальной диаграмме с указанием текущего момента времени;

— количество и перечень используемых наблюдателем средств противодействия высокоскоростному аппарату с указанием рекомендованного момента времени их использования и продолжительности использования;

— временная диаграмма соотношения маневра и функционирования средств противодействия высокоскоростному аппарату;

— эффективность противодействия высокоскоростному аппарату, рассчитанная по определенной методике, описание которой приведено ниже.

Арсенал и количество используемых средств при оценке эффективности определяются с учетом времени, на которое необходимо отвлечь высокоскоростной аппарат от наблюдателя с учетом того, что энергоресурс высокоскоростного аппарата ограничен. Каждое средство «заберет» у высокоскоростного аппарата определенное время на разбор со средством, тем самым расходуя дополнительно его энергоресурс.

Пусть высокоскоростному аппарату, движущемуся со скоростью V , необходимо время T_D на преодоление дистанции D : $T_D = D/V$. Пусть запас времени по энергоресурсу высокоскоростного аппарата составляет T_m . Если $T_m < T_D$, то высокоскоростной аппарат не сможет на такой дистанции создавать помеху наблюдателю. Если же $T_m > T_D$, то у высокоскоростного аппарата возникает «запас» времени, который необходимо «закрыть» средствами противодействия. Тогда, отвлекаясь на средства противодействия, высокоскоростной аппарат израсходует ресурс, и даже если все средства противодействия преодолеет в соответствии с заложенным в нем интеллектом, ресурса все равно может оказаться недостаточно для того, чтобы сблизиться с наблюдателем на необходимую дистанцию. Промежуток времени Δt , который необходимо заполнить средствами противодействия, определяется выражением

$$\Delta t = T_m - T_D.$$

Если предположить, что в среднем высокоскоростной аппарат тратит время $t_{c.п}$ на то, чтобы разобраться со средством противодействия и вернуться на позицию создания помехи наблюдателю, то количество средств противодействия определяется отношением

$$N = [\Delta t/t_{c.п}] + 1,$$

где символ $[\]$ означает «целая часть» числа.

Таким образом, инструментарий, представленный на рис. 4, позволяет исследователю изменить порядок применения имеющихся средств противодействия и временные характеристики — задержку начала и продолжительность функционирования.

С использованием разработанной математической модели и ее программной реализации были проведены исследования, в ходе которых выполнялся численный эксперимент. Исследования были направлены на оценку влияния параметров наблюдателя и высокоскоростного аппарата на эффективность сохранения процесса наблюдения.

Результаты численного эксперимента

В численном эксперименте задействованы два аппарата: наблюдатель и наблюдаемый, — расположенные в точках с координатами 70.029146701° с. ш., 29.98280885° в. д. и 70.012483446° с. ш., 30.06972222° в. д. соответственно (см. рис. 3).

При проведении численного эксперимента разыгрывается:

— стартовый курс движения наблюдателя; диапазон разыгрывания 0° – 90° относительно наблюдаемого аппарата;

— глубина и угол отворота наблюдателя при выполнении маневра уклонения из диапазонов:

глубина 20° – 450 м;

угол отворота $\pm 60^\circ$.

Стартовую дистанцию «наблюдатель — наблюдаемый аппарат» в каждом испытании считаем величиной постоянной и равной 3050 м.

Если высокоскоростной аппарат потерял свою работоспособность в результате действия прибора противодействия, используемого со стороны наблюдателем: израсходовал ресурс или ушел в ложном направлении — испытание считается завершенным.

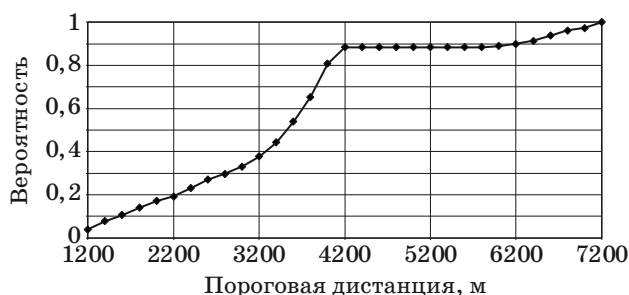
В ходе численного эксперимента проводилась серия испытаний и определялись значения дистанции от наблюдателя до наблюдаемого аппарата в момент, когда помехи от высокоскоростного аппарата преодолевались. Затем полученные в испытаниях дистанции сортировались по возрастанию, формировался диапазон дистанций, ограниченный пороговым значением дистанции, за пределами которой наблюдатель терял контакт с наблюдаемым аппаратом. После сортировки полученных дистанций определялось количество дистанций, попадающих в рассматриваемый диапазон.

На рис. 5 приведена вероятность сохранения контакта наблюдателя с наблюдаемым аппара-

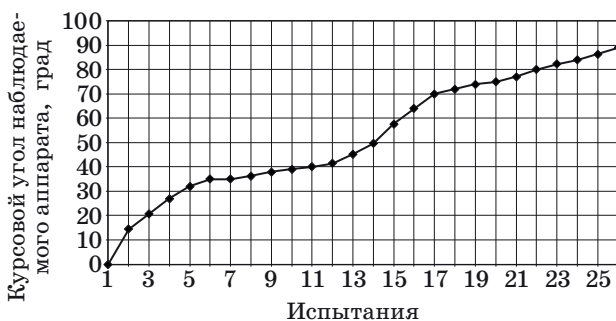
том в зависимости от порогового значения дистанции, при котором контакт еще возможен: по горизонтальной оси отложено пороговое значение дистанции, на которой контакт еще возможен, по вертикальной — вероятность. Среднее время отвлечения составило 235 с, минимальное — 212 с, максимальное — 249 с, среднеквадратическое отклонение составило 13,8 с.

Из рисунка видно, что гарантированное сохранение контакта наблюдается в основном на дистанциях свыше 3800 м. Поэтому на следующем этапе исследований анализировались причины полученных значений дистанции. При проведении исследований исходили из того, что дистанция до наблюдаемого аппарата после преодоления помехи определяется временем, затраченным наблюдателем на устранение помехи. Кроме того, на дистанцию влияние оказывает курс движения наблюдаемого аппарата.

Значения пеленгов, полученные в каждом испытании серии, упорядочены по возрастанию. Из рис. 6 видно, что характер кривой близок к линейной зависимости, что позволяет говорить о равномерном распределении пеленгов в серии испытаний.



■ **Рис. 5.** Вероятность сохранения контакта в зависимости от порогового значения дистанции
 ■ **Fig. 5.** The probability of maintaining a contact as a function of the distance threshold



■ **Рис. 6.** Изменение пеленга от наблюдателя на наблюдаемый аппарат в зависимости от курса движения наблюдаемого аппарата
 ■ **Fig. 6.** Bearing change from the observer to the observed device depending on the course of movement of the monitored device

На дистанцию влияние оказывают курс наблюдаемого аппарата и время, затрачиваемое наблюдателем на подавление помехи, создаваемой высокоскоростным аппаратом.

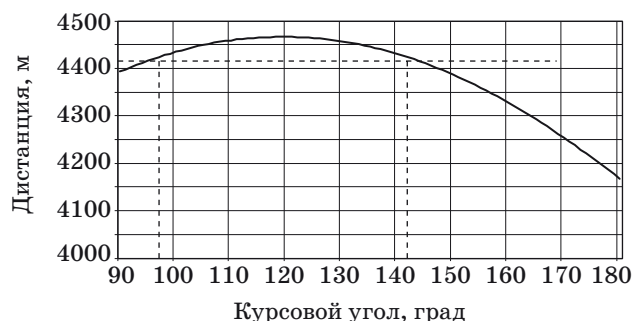
На следующем этапе исследований было проанализировано, существуют ли курсовые углы перемещения наблюдаемого аппарата, которые приводят к прерыванию наблюдения.

Наблюдаемый аппарат мог двигаться произвольным курсом, в том числе «диаметрально» расходясь с наблюдателем. Поэтому для дальнейшего анализа было пересчитано положение наблюдателя к моменту окончания его уклонения от высокоскоростного аппарата.

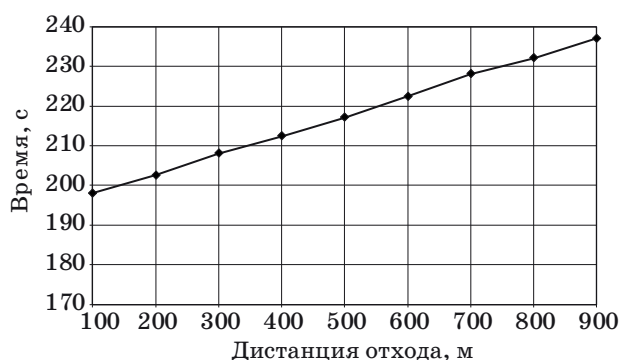
Для этого определялись положение наблюдателя и время, которое он потратил на борьбу с помехой. Далее перебирались курсовые углы и определялось положение наблюдаемого аппарата, в котором он мог оказаться через рассчитанный промежуток времени. Результаты моделирования представлены на рис. 7.

Из рисунка видно, что существует диапазон курсовых углов, при движении с которыми наблюдаемый аппарат выходит из-под наблюдения за то время, пока наблюдатель «разбирается» со скоростным аппаратом. Такой вариант является самым неблагоприятным для наблюдателя, поскольку придется заново вести поиск, а в случае обнаружения — «догонять» наблюдаемого исключительно за счет разности скоростей.

При анализе стратегии поведения скоростного аппарата по преодолению противодействия, созданного наблюдателем, выявлено, что время может зависеть, в частности, от дистанции отхода высокоскоростного аппарата за дрейфующий прибор помех для того, чтобы выйти из шумового поля, в котором невозможно обнаружить наблюдателя. Для этого при фиксированных значениях положения наблюдателя и наблюдаемого аппарата



■ **Рис. 7.** Изменение дистанции от наблюдателя до наблюдаемого в зависимости от курса движения наблюдаемого
 ■ **Fig. 7.** Changing the distance from the observer to the observed, depending on the course of motion of the observed



■ Рис. 8. Зависимость времени отвращения наблюдателя на скоростной аппарат от дистанции отхода скоростного аппарата за дрейфующий прибор помех

■ Fig. 8. Dependence of the observer's distraction time on the velocity apparatus from the distance of the departure of the velocity apparatus for the drifting interference device

та перебирались дистанции отхода высокоскоростного аппарата от 100 до 900 м.

Результаты влияния дистанции отхода скоростного аппарата за дрейфующий прибор помех на время отвращения наблюдателя на скоростной аппарат приведены на рис. 8.

Видно, что с увеличением дистанции отхода высокоскоростного аппарата за прибор помех происходит увеличение времени, затрачиваемого наблюдателем на преодоление интеллектуальной помехи, вызванной поведением высокоскоростного аппарата.

Литература

1. Ким Д. П. Методы поиска и преследования подвижных объектов. — М.: Наука, 1989. — 336 с.
2. Мартынова Л. А., Шебалов А. Н. Условия взаимодействия объектов в дифференциальной игре преследования // Проблемы математического моделирования и разработка САПР для судостроения: науч.-техн. сб. (труды). — Л.: ЛКИ, 1990. С. 55–60.
3. Понтрягин Л. С. Линейные дифференциальные игры. I // Докл. АН СССР. 1967. Т. 174. № 6. С. 1278–1281.
4. Понтрягин Л. С. Линейные дифференциальные игры. II // Докл. АН СССР. 1967. Т. 175. № 4. С. 764–767.
5. Абрамов Н. С., Хачумов В. М. Моделирование проводки по маршруту беспилотного летательного аппарата как задачи преследования цели // Авиакосмическое приборостроение. 2013. № 9. С. 9–22.
6. Мельц И. О., Сурженко А. С. Игровая задача преследования–уклонения для трех участников // Уч. зап. ЦАГИ. 2005. Т. XXXVI. № 3–4. С. 92–102.
7. Дзюбенко Г. Ц., Пшеничный Б. Н. Дискретные дифференциальные игры с запаздыванием информации // Кибернетика. 1972. Т. 8. № 6. С. 65–71.
8. Сатимов Н., Азамов А. Нелинейные дискретные игры убегающего // Кибернетика. 1976. Т. 12. № 4. 1976. С. 70–74.
9. Петров Н. Н. Задача группового преследования в классе импульсных стратегий преследователей // Изв. РАН. Теория и системы управления. 2009. Т. 48. № 2. С. 38–44
10. Ухоботов В. И., Зайцева О. В. Одна задача импульсного преследования при ограниченной скорости убегающего // Вестник Юж.-Урал. гос. ун-та. Сер. Компьютерные технологии, управление, радиоэлектроника. 2010. № 2 (178). Вып. 11. С. 29–32.
11. Isaacs R. Differential Games: A Mathematical Theory with Applications to Warfare and Pursuit, Control and Optimization. — N. Y.: John Wiley & Sons, 1965. — P. 349–350.
12. Пацко В. С., Турова В. Л. Игра «Шофер-убийца» и ее модификации // Вестник Удмуртского универ-

Таким образом, по результатам проведенных исследований показано, что наблюдаемый аппарат может выбрать стратегию (курс или параметры — дистанцию отхода), которая позволяет ему оторваться от наблюдателя и тем самым прервать наблюдение за ним со стороны наблюдателя.

Заключение

Для решения задачи подводного наблюдения в условиях применения против наблюдателя интеллектуальных помех в работе рассмотрены варианты поведения наблюдателя для сохранения контакта с наблюдаемым аппаратом. Для проведения сопоставления полученных результатов предложен показатель эффективности наблюдения — вероятность сохранения контакта.

Для расчета вероятности сохранения контакта разработана математическая модель, учитывающая особенности поведения наблюдателя, наблюдаемого аппарата, а также аппаратов, создающих помехи наблюдателю и скоростному объекту, используемому наблюдаемым аппаратом в качестве помехи наблюдателю.

Результаты моделирования позволили прогнозировать возможность сохранения контакта и выбирать тактику поведения, обеспечивающую эффективное наблюдение.

Работа выполнена при поддержке Российского фонда фундаментальных исследований (проект № 17-08-00666).

- ситета. Математика. Механика. Компьютерные науки. 2008. Вып. 2. С. 105–110.
13. Гарднер М. Оптимальные стратегии для игр с двумя участниками // Математические новеллы. — М.: Мир, 1974. — С. 214–230.
 14. Гмурман В. Е. Теория вероятности и математическая статистика. — М.: Высш. шк., 2000. — 305 с.
 15. Бусленко Н. П. Моделирование сложных систем. — М.: Наука, 1978. — 400 с.
 16. Мартынова Л. А. Имитационная модель оценки эффективности функционирования АНПА // Тр. Междунар. конф. по судостроению и океанотехнике, Санкт-Петербург, 6–8 июня 2016 г.; сб. ст. — СПб.: СПбГМТУ, 2016. С. 455–469.
 17. Мартынова Л. А., Розенгауз М. Б. К вопросу о надежности автономного необитаемого подводного аппарата с мультиагентной архитектурой системы управления // Информационно-управляющие системы. 2016. № 5. С. 25–34. doi:10.15217/issn1684-8853.2016.5.25
 18. Мартынова Л. А., Машошин А. И., Пашкевич И. В. Подходы к оценке эффективности автономного необитаемого подводного аппарата // Управление в морских и аэрокосмических системах (УМАС-2016): материалы конф., Санкт-Петербург, 4–6 октября 2016 г. — СПб.: Концерн «Электроприбор», 2016. С. 205–209.
 19. Мартынова Л. А., Машошин А. И. Особенности оценки эффективности функционирования автономных необитаемых подводных аппаратов в нештатных ситуациях // Экстремальная робототехника: тр. Междунар. науч.-техн. конф. — СПб.: АПАПринт, 2016. С. 86–91.
 20. Безрук Г. Г., Мартынова Л. А. Оценка эффективности образцов робототехнических комплексов в условиях противодействия // Вопросы оборонной техники. Сер. 16. 2016. № 11–12 (101–102). С. 45–49.

UDC 519.87

doi:10.15217/issn1684-8853.2018.1.31

Underwater Observation under Intellectual InterferenceMartynova L. A.^a, Dr. Sc., Tech., Associate Professor, martynowa999@bk.ru^aConcern CSRI Elektropribor, JSC State Research Center of Russia, 30, Malaya Posadskaya St., 197046, Saint-Petersburg, Russian Federation

Introduction: Autonomous uninhabited submersibles perform a variety of tasks, including monitoring other autonomous uninhabited submersibles which are of interest, for example commercially. The task of an observing device is to maintain a constant contact with the observed object and to obtain the information about its coordinates and motion parameters. At the same time, the observed device tries to get rid of the surveillance, for which it uses an intellectual interference in the form of a high-speed apparatus. The interference is intelligent because it can recognize the countermeasures used against it by the observer and to reduce their influence by a chosen strategy of its behavior. So far, in the problems of search or pursuit of an object, potentially under an interference, the behavior of the interference has never been intellectual. **Purpose:** Evaluating the influence of an intellectual interference behavior upon the possibility to maintain the contact with a monitored device under the conditions when a number of the interference parameters and behavioral strategy are not determined in advance. **Methods:** The efficiency indicator calculation is based on the method of statistical tests (the Monte Carlo method). **Results:** A tactical episode was formed of observing an object which uses an intellectual disturbance. A strategy has been developed for the observer to use countermeasures, along with a strategy for the behavior of a high-speed device to reduce the influence of the countermeasures. To assess the impact, performance indicators were defined, and a mathematical model developed for their calculation. The developed software implementation of the mathematical model made it possible to conduct numerical experiments which allowed us to evaluate the influence of an intellectual interference on the possibility of keeping the contact between the observer and the monitored device. **Practical relevance:** The results can be used for planning an observer's behavior strategy under interference, for choosing technical solutions in order to ensure monitoring an object under interference, or for choosing a strategy of anti-interference measures against the monitored device.

Keywords — Autonomous Uninhabited Submersible, Mathematical Modeling, Method of Statistical Tests, Intellectual Interference.

Citation: Martynova L. A. Underwater Observation under Intellectual Interference. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 1, pp. 31–41 (In Russian). doi:10.15217/issn1684-8853.2018.1.31

References

1. Kim D. P. *Metody poiska i presledovaniia podvizhnykh ob'ektov* [Methods of Searching and Prosecuting Mobile Objects]. Moscow, Nauka Publ., 1989. 336 p. (In Russian).
2. Martynova L. A., Shebalov A. N. Conditions of Interaction of Objects in a Differential Game of Pursuit. In: *Problemy matematicheskogo modelirovaniia i razrabotka SAPR dlia sudostroeniia* [Problems of Mathematical Modeling and Development of CAD for Shipbuilding]. Leningrad, Leningradskii korablestroitel'nyi institute Publ., 1990. Pp. 55–60 (In Russian).
3. Pontryagin L. S. Linear Differential Games. Part I. *Doklady AN SSSR* [Reports of the USSR Academy of Sciences], 1967, vol. 174, no. 6, pp. 1278–1281 (In Russian).
4. Pontryagin L. S. Linear Differential Games. Part II. *Doklady AN SSSR* [Reports of the USSR Academy of Sciences], 1967, vol. 175, no. 4, pp. 764–767 (In Russian).
5. Abramov N. S., Khachumov M. V. Air Route Simulation of Unmanned Aerial Vehicles as the Pursuit-Evasion Problem. *Aviakosmicheskoe priborostroenie* [Aerospace Instrument-Making], 2013, no. 9, pp. 9–22 (In Russian).
6. Melts I. O., Surzhenko A. S. Game Problem of Prosecution of Deviation for Three Participants. *Uchenye zapiski TsAGI*, 2005, vol. XXXVI, no. 3–4, pp. 92–102 (In Russian).
7. Dzyubenko G. T., Pshenichny B. N. Discrete Differential Games with Time Lag of Information. *Kibernetika* [Cybernetics], Kiev, 1972, vol. 8, no. 6, pp. 65–71 (In Russian).

8. Satimov N., Azamov A. Nonlinear Discrete Games of Evasion. *Kibernetika* [Cybernetics], Kiev, 1976, vol. 12, no. 4, pp. 70–74 (In Russian).
9. Petrov N. N. A Problem of Group Pursuit in the Class of Impulse Strategies of Pursuers. *Journal of Computer and Systems Sciences International*, 2009, vol. 48, no. 2, pp. 199–205. doi:10.1134/S106423070902004X
10. Ukhobotov V. I., Zayceva O. V. About One Problem of Impulse Pursuit at the Limited Velocity of the Escaping. *Vestnik Iuzhno-Ural'skogo gosudarstvennogo universiteta. Seriya "Komp'yuternye tekhnologii, upravlenie, radioelektronika"* [Bulletin South Ural State University. Series "Computer Technologies, Automatic Control & Radioelectronics"], 2010, no. 2 (178), iss. 11, pp. 29–32 (In Russian).
11. Isaacs R. *Differential Games: A Mathematical Theory with Applications to Warfare and Pursuit, Control and Optimization*. N. Y., John Wiley & Sons, 1965. Pp. 349–350.
12. Patsko V. S., Turova V. L. Homicidal Chauffeur Game and its Modifications. *Vestnik Udmurtskogo universiteta. Matematika. Mekhanika. Komp'yuternye nauki*, 2008, no. 2, pp. 105–110 (In Russian).
13. Gardner M. *Optimal'nye strategii dlia igr s dvumia uchastnikami*. In: *Matematicheskie novelly* [Optimal Strategies for Games with Two Participants. In: Mathematics Games]. Moscow, Mir Publ., 1974. Pp. 214–230 (In Russian).
14. Gnurman V. E. *Fundamentals of Probability Theory & Mathematical Statistics Hardcover*. Piiffe Books Ltd, 1968.
15. Buslenko N. P. *Modelirovanie slozhnykh sistem* [Modeling of Complex Systems]. Moscow, Nauka Publ., 1978. 400 p. (In Russian).
16. Martynova L. A. Imitation Model for Evaluation of the Operational Efficiency of AUV. *Trudy Mezhdunarodnoi konferentsii po sudostroeniiu i okeanotekhnike* [Proc. of the Intern. Conf. on Naval Architecture and Ocean Engineering], June 6–8, 2016, Saint-Petersburg, Russia, 2016, pp. 455–469 (In Russian).
17. Martynova L. A., Rozengauz M. B. Reliability of an Autonomous Underwater Vehicle with a Multiagent Control System. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2016, no. 5, pp. 25–34 (In Russian). doi:10.15217/issn1684-8853.2016.5.25
18. Martynova L. A., Mashoshin A. I., Pashkevich I. V. Approaches to Assessing the Effectiveness of an Autonomous Uninhabited Underwater Vehicle. *Materialy konferentsii "Upravlenie v morskikh i aerokosmicheskikh sistemakh (UMAS-2016)"* [Proc. of the Conf. "Management in Marine and Aerospace Systems" (UMAS-2016)], October 4–6, 2016, Saint-Petersburg, pp. 205–209 (In Russian).
19. Martynova L. A., Mashoshin A. I. Features of Evaluating the Performance of Autonomous Underwater Vehicle in Emergency Situations. *Trudy Mezhdunarodnoi nauchno-tekhnicheskoi konferentsii "Ekstremal'naia robototekhnika"* [Proc. of Intern. Scientific and Technological Conf. "Extreme Robotics"], November 24–25, 2016, Saint-Petersburg, Russia, 2016, pp. 86–91 (In Russian).
20. Bezruk G. G., Martynova L. A. Evaluation of Samples Robotic Systems in the Face of Opposition. *Voprosy oboronnoi tekhniki. Seriya 16* [Military Enginery. Series 16], 2016, no. 11–12 (101–102), pp. 45–49 (In Russian).

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/>

Полнотекстовые версии журнала за 2002–2016 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-upravlyayushchiesistemy).

УДК 577.31

doi:10.15217/issn1684-8853.2018.1.42

ДИАГНОСТИКА НАПРАВЛЕННОЙ СВЯЗИ КОНТУРОВ РЕГУЛЯЦИИ КРОВООБРАЩЕНИЯ ПО ВРЕМЕННЫМ РЯДАМ МАТЕМАТИЧЕСКОЙ МОДЕЛИ СЕРДЕЧНО-СОСУДИСТОЙ СИСТЕМЫ ЧЕЛОВЕКА

В. С. Хорев^а, канд. физ.-мат. наук, доцент, khorevvs@gmail.com

Ю. М. Ишбулатов^а, аспирант, ishbulatov95@mail.ru

Е. Е. Лапшева^а, старший преподаватель, lapsheva@yandex.ru

А. Р. Киселев^{б, в}, доктор мед. наук, профессор, antonkis@list.ru

В. И. Гриднев^{а, в}, доктор мед. наук, профессор, gridnev@cardio-it.ru

Б. П. Безручко^а, доктор физ.-мат. наук, профессор, bezruchkobp@gmail.com

А. А. Бутенко^а, ассистент, butenkoaan@yandex.ru

В. И. Пономаренко^{а, г}, доктор физ.-мат. наук, профессор, ponomarenkovi@gmail.com

А. С. Караваев^{а, г}, канд. физ.-мат. наук, доцент, karavaevas@gmail.com

^аСаратовский национальный исследовательский государственный университет им. Н. Г. Чернышевского, Астраханская ул., 83, Саратов, 410012, РФ

^бНациональный медицинский исследовательский центр сердечно-сосудистой хирургии им. А. Н. Бакулева Минздрава России, Рублевское ш., 135, Москва, 121552, РФ

^вСаратовский государственный медицинский университет им. В. И. Разумовского, Б. Казачья ул., 112, Саратов, 410012, РФ

^гСаратовский филиал Института радиотехники и электроники им. В. А. Котельникова РАН, Зеленая ул., 38, Саратов, 410019, РФ

Постановка проблемы: развитие методов оценки взаимодействия между сигналами биологической природы зачастую осложнено необходимостью настройки и апробации параметров методик ввиду отсутствия детальных представлений о функционировании объекта исследования, объективных данных о значениях параметров объекта. Эффективному решению данной проблемы способствует использование математических моделей. **Цель исследования:** анализ направленной связанности по сигналам математической модели контуров симпатической регуляции частоты сердечных сокращений и барорефлекторной регуляции тонуса артериальных сосудов, а также по экспериментальным данным. **Методы:** использован метод, основанный на построении индивидуальных и совместных моделей динамики фаз сигналов колебаний в виде гармонических полиномов третьей степени. Сила воздействия одного элемента на другой оценивалась по изменению точности прогноза при включении в полином членов, учитывающих динамику колебаний второго элемента. **Результаты:** на основе сравнительного анализа оценок силы воздействия и запаздывания в связи между модельными сигналами кардиоинтервалограмм и тонуса дистальных сосудов, вычисленных в результате обработки записей в скользящих окнах, были получены зависимости динамики средних значений и разброса оценок свойств связи от длины анализируемого участка ряда. Показано, что использование рядов длительностью более 1200 с существенно увеличивает точность оценки свойств связи и уменьшает дисперсию оценок. Продемонстрировано качественное соответствие модельных и экспериментальных результатов. **Практическая значимость:** результаты исследований могут быть использованы для развития методов диагностики патологий сердечно-сосудистой системы человека. Они имеют важное значение для изучения особенностей использования методов диагностики связанности, основанных на динамике мгновенных фаз колебаний, по экспериментальным данным сложных систем.

Ключевые слова — анализ сигналов, математическая модель, мгновенная фаза, направленная связь, временное запаздывание, сердечно-сосудистая система, вегетативная регуляция.

Цитирование: Хорев В. С., Ишбулатов Ю. М., Лапшева Е. Е., Киселев А. Р., Гриднев В. И., Безручко Б. П., Бутенко А. А., Пономаренко В. И., Караваев А. С. Диагностика направленной связи контуров регуляции кровообращения по временным рядам математической модели сердечно-сосудистой системы человека // Информационно-управляющие системы. 2018. № 1. С. 42–48. doi:10.15217/issn1684-8853.2018.1.42

Citation: Khorev V. S., Ishbulatov J. M., Lapsheva E. E., Kiselev A. R., Gridnev V. I., Bezruchko B. P., Butenko A. A., Ponomarenko V. I., Karavaev A. S. Diagnostics of Directional Coupling between Blood Circulation Regulation Loops Using Analysis of Time Series of Mathematical Model of Human Cardiovascular System. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 1, pp. 42–48 (In Russian). doi:10.15217/issn1684-8853.2018.1.42

Введение

Нарушение вегетативной регуляции тонуса артериальных сосудов и частоты сердечных сокращений может приводить к ряду опасных

патологий сердечно-сосудистой системы [1]. Перспективным способом диагностики и терапии некоторых патологий сердца и сосудов являются развиваемые нами подходы, основанные на оценке степени взаимодействия контуров вегетатив-

ной регуляции кровообращения с помощью диагностики подстройки фаз колебаний [2, 3] и количественной оценке степени связанности этих контуров методами фазовой динамики [4–6]. Такая оценка затруднена наличием свободных параметров используемых нелинейных методов, выбор которых при анализе сложных нестационарных сигналов систем биологической природы является самостоятельной нетривиальной задачей [7–10].

Апробация и настройка параметров методов анализа данных биологической природы существенно упрощается при наличии адекватной математической модели исследуемого объекта, параметры и свойства динамики которой задаются исследователем априори [11].

В рамках данной работы производится оценка направленной связанности по сигналам предложенной нами ранее математической модели сердечно-сосудистой системы [12], отражающей динамику контуров симпатической регуляции частоты сердечных сокращений и барорефлекторной регуляции тонуса артериальных сосудов, и сопоставление полученных оценок модельных и экспериментальных временных реализаций.

Данные и методы проводимых исследований

В работе были использованы данные по 10 испытуемым, не имеющим патологий сердечно-сосудистой системы. Всем обследованным проводилась одновременная регистрация электрокардиограммы (ЭКГ), фотоплетизмограммы (ФПГ) и механической записи дыхания в горизонтальном положении тела. Продолжительность каждой записи составляла 7200 с. Во время регистрации сигналов дыхание у всех обследованных было произвольным.

Регистрация ЭКГ, ФПГ и дыхания осуществлялась при помощи многоканального электроэнцефалографа-анализатора ЭЭГА-21/26 «Энцефалан-131-03» модель 10 (НПКФ «Медиком МТД», Россия) с комплектом стандартных датчиков. Регистрация сигналов производилась с частотой 250 Гц при 12-разрядном разрешении. Запись ФПГ проводилась пульсоксиметрическими датчиками (в проходящем свете), помещенными на дистальной фаланге указательного пальца руки, пальца ноги, а также на мочке уха. По записи дыхания контролировались спонтанность дыхания, отсутствие форсированных вдохов и задержек дыхания во время регистрации записей, тем самым из анализа исключались случайные влияния на сердечно-сосудистую систему со стороны дыхания, т. е. все исследования проводились в относительно стандартных условиях спонтан-

ного дыхания. Для дальнейшего анализа отбирались записи сигналов ЭКГ и ФПГ, не содержащие помех, экстрасистол, заметного линейного тренда и переходных процессов.

Также в работе были использованы сигналы математической модели, предложенной в работе [12], описывающей основной сердечный ритм, работу контура симпатической регуляции частоты сердечных сокращений, контура барорефлекторной регуляции тонуса артериальных сосудов, влияние барорефлекторной регуляции на частоту и силу сердечных сокращений. Процесс формирования артериального давления (АД, blood pressure — BP) учитывает влияние дыхания на динамику перечисленных систем. Длина реализаций и параметры модели были выбраны таким образом, чтобы динамика систем соответствовала наблюдаемой у здорового человека в состоянии покоя. Модель представляет собой систему из пяти дифференциальных уравнений 1-го порядка с запаздываниями [12].

Информация о вариабельности сердечного ритма была получена путем выделения из ЭКГ RR-последовательности (кардиоинтервалограммы — КИГ, R-R intervalogram — RRI), аппроксимации полученной реализации кубическими сплайнами и ресемплирования до частоты дискретизации 10 Гц. Медленные 0,1 Гц-колебания подсистем регуляции сосудистого тонуса и частоты сердечных сокращений выделялись с помощью полосовой фильтрации 0,05–0,15 Гц. Фильтрованный сигнал ФПГ прореживался до частоты дискретизации 10 Гц. Временные реализации математической модели подвергались аналогичной обработке. Такая предварительная обработка экспериментальных данных соответствует методическим рекомендациям, изложенным Р. М. Баевским в работе [13].

Исследования направлены на сопоставление при анализе модельных и экспериментальных данных нескольких подходов. Применялись метод, основанный на моделировании фазовой динамики исследуемых систем [14], а также широко известный метод расчета индекса фазовой когерентности.

Метод, основанный на моделировании фазовой динамики, состоит в построении индивидуальных и совместных моделей динамики фаз сигналов колебаний в виде гармонических полиномов третьей степени. Сила воздействия одного элемента на другой оценивалась по изменению точности прогноза при включении в полином членов, учитывающих динамику колебаний второго элемента. Уменьшение ошибки прогноза совместной модели по отношению к индивидуальной является признаком влияния второй системы на первую. Время запаздывания в связи между элементами оценивалось по положению,

соответствующему максимуму оценки силы воздействия со стороны одного элемента системы на другой по диапазону пробных задержек

$$G_1^2 \approx \frac{1}{2\pi^2} \int_0^{2\pi} \int_0^{2\pi} \partial F_{1,2}(\varphi_1, \varphi_2, \mathbf{a}_{1,2}) / \partial \varphi_{1,2}^2 d\varphi_1 d\varphi_2, \quad (1)$$

где G_1 — интенсивность воздействия второй системы на первую; $F_{1,2}$ — тригонометрические многочлены третьего порядка; $\varphi_{1,2}$ — развернутые фазы исследуемых временных рядов; $\mathbf{a}_{1,2}$ — вектор коэффициентов.

Также был использован распространенный [15–19] метод расчета индекса фазовой когерентности [20]:

$$\rho(d) = \left| \frac{1}{N} \sum_{j=1}^N e^{i(\varphi_1(j) - \varphi_2(j-d))} \right|, \quad (2)$$

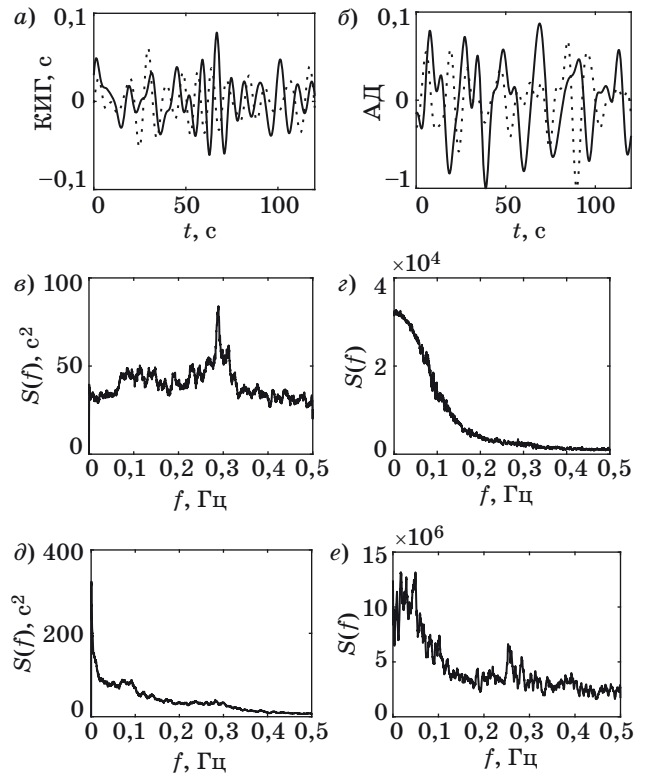
где ρ — индекс фазовой когерентности, рассчитываемый для пробного времени задержки d ; N — длина рядов.

Результаты исследований и выводы

На рис. 1 представлены фрагменты временных рядов кардиоинтервалограммы (рис. 1, а) и фотоплетизмограммы (рис. 1, б) здорового испытуемого (пунктиром) и модели (сплошной линией), сигналы АД нормированы к единичной дисперсии и приведены в безразмерных единицах. На рис. 1, в–е приведены фурье-спектры модельных (рис. 1, в, з) и экспериментальных (рис. 1, д, е) сигналов, на которых можно заметить подъемы вблизи частотных диапазонов, соответствующих полосам активности подсистем симпатической (0,05–0,15 Гц) и парасимпатической (0,25–0,35 Гц) регуляции.

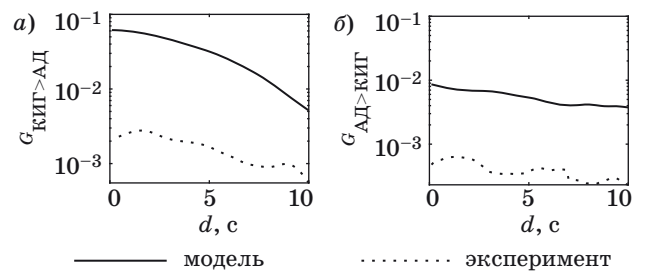
Оценки силы воздействия и запаздывания в связи между модельными сигналами КИГ и тонуса дистальных сосудов получены в результате обработки двухчасовой записи. Максимумы зависимости оценки силы связи в направлении «КИГ → АД» (рис. 2, а), выведенные с помощью метода моделирования фазовой динамики, приходятся на нулевую задержку при силе воздействия $G_1 = 0,06$, для обратного направления связи «АД → КИГ» (рис. 2, б) задержка составляет 0,7 с при силе воздействия $G_2 = 0,14$. Задержка, оцениваемая с помощью индекса фазовой когерентности, принимает значение 2,8 с при максимальном значении индекса $\rho = 0,76$.

Динамика изменения максимального значения уровня связи, полученная с помощью метода моделирования фазовой динамики в скользящих неперекрывающихся окнах, представлена на рис. 3. Длина окна для получения оценок со-



■ **Рис. 1.** Временные ряды: а — КИГ, фильтрованной в полосе 0,05–0,15 Гц; б — АД, фильтрованного в полосе 0,05–0,15 Гц; в, д — фурье-спектры сигналов КИГ; з, е — фурье-спектры сигналов АД

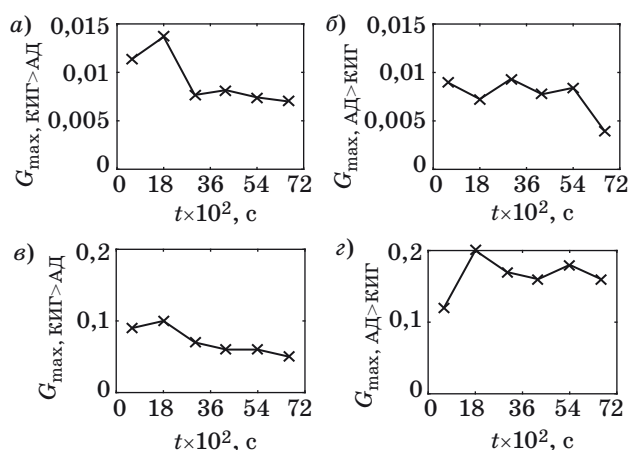
■ **Fig. 1.** Time series: а — RRI, filtered in the band 0.05–0.15 Hz; б — BP, filtered in the band 0.05–0.15 Hz; в, д — Fourier spectra of the RRI-signals; з, е — Fourier spectra of the BP-signals



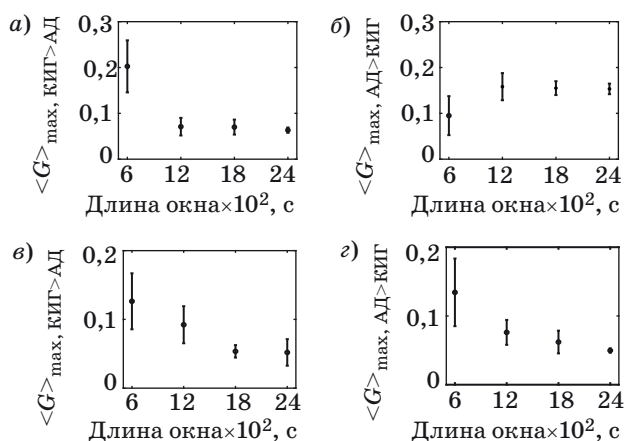
■ **Рис. 2.** Оценки силы воздействия и запаздывания: а — «КИГ → АД»; б — «АД → КИГ»

■ **Fig. 2.** Estimates of coupling strength and delay time: а — RRI → BP; б — BP → RRI

ставляла 1200 с. Оценки, полученные по экспериментальным данным, показаны на рис. 3, а, б, по модельным — на рис. 3, в, г. Видно, что уровень связи в направлении «КИГ → АД» плавно уменьшается с течением времени в случае использования окон длиной 1200 с. Также следует отметить уменьшение дисперсии оценок, получаемых для более длинных рядов, что свидетельствует



■ **Рис. 3.** Динамика максимального уровня связи: а, в — «КИГ → АД»; б, г — «АД → КИГ»
 ■ **Fig. 3.** Dynamics of maximum coupling strength: а, в — RRI → BP; б, г — BP → RRI



■ **Рис. 4.** Динамика среднего уровня связи: а, в — «КИГ → АД»; б, г — «АД → КИГ»
 ■ **Fig. 4.** Dynamics of average coupling strength: а, в — RRI → BP; б, г — BP → RRI

в пользу того, что для получения корректных результатов желательно использовать временные ряды длиной не менее 1200 с.

Таким образом, в результате обработки двухчасовых записей здоровых испытуемых с помощью описанных выше методов были получены оценки силы воздействия и запаздывания в связи между сигналами КИГ и тонуса дистальных сосудов. По ансамблю 20 двухчасовых записей было установлено, что средние значения задержек между сигналами, полученные с помощью метода моделирования фазовой динамики, при анализе воздействия контура регуляции сердечных сокращений на контур регуляции тонуса дистальных сосудов составляют $(2,13 \pm 0,14)$ с при средней силе воздействия $G_1 = 0,09 \pm 0,08$, для воздействия в противоположном направлении — $(2,12 \pm 0,17)$ с при средней силе воздействия $G_2 = 0,11 \pm 0,12$. По методу расчета индекса фазовой когерентности средние значения задержки составили $(3,88 \pm 0,76)$ с, а средние значения максимума индекса фазовой когерентности — $0,22 \pm 0,09$.

Полученные результаты согласуются с наблюдаемыми ранее в работах [12]. На графиках оценок по методу моделирования фазовой динамики наблюдаются значимые максимумы, соответствующие задержкам порядка секунды, в обоих направлениях. На рис. 4 приведены зависимости среднего максимального значения оценки уровня связи, полученного с помощью метода моделирования фазовой динамики, от длины окна. Значения приведены в форме среднее \pm среднеквадратическое отклонение. Как по модельным (рис. 4, а, б), так и по экспериментальным (рис. 4, в, г) данным абсолютные значения оценок сил связи указывают на то, что преимущественное направление связи соответствует направлению «АД → КИГ», что осо-

бенно хорошо видно на графиках динамики уровня связи. То есть имеет место качественное соответствие результатов, полученных для модельных и реальных данных. Однако абсолютные значения оценок для модели оказываются существенно завышены, что указывает на необходимость совершенствования модельных представлений в направлении уточнения структуры и значений параметров модели. Использование полученных результатов позволит в дальнейшем провести совершенствование модели для количественного соответствия показателей, получаемых при анализе записей реальных людей.

Заключение

Проведен анализ взаимодействия между сигналами, отражающими динамику отдельных контуров регуляции на модели сердечно-сосудистой системы, а также сопоставление с результатами, полученными на здоровых субъектах. Полученные результаты согласуются с современными представлениями об устройстве и организации взаимодействия подсистем регуляции в сердечно-сосудистой системе. Продемонстрировано качественное соответствие результатов оценки преимущественного направления, а также оценки силы воздействия и задержки в связи между системами. Полученные результаты свидетельствуют о наличии симметричного двунаправленного взаимодействия изучаемых контуров регуляции при анализе как экспериментальных данных, так и записей модели.

Работа выполнена при поддержке грантов РФФИ № 16-32-00326 и 15-02-03061, гранта Президента РФ МД-3318.2017.7.

Литература

1. Kiselev A. R., Gridnev V. I., Prokhorov M. D., Karavaev A. S., Posnenkova O. M., Ponomarenko V. I., Bezruchko B. P. Selection of Optimal Dose of Beta-blocker Treatment in Myocardial Infarction Patients Basing on Changes in Synchronization Between 0.1 Hz Oscillations in Heart Rate and Peripheral Microcirculation // *Journal of Cardiovascular Medicine*. 2012. Vol. 13. Iss. 8. P. 491–498. doi:10.2459/JCM.0b013e3283512199
2. Kiselev A. R., Karavaev A. S., Gridnev V. I., Prokhorov M. D., Ponomarenko V. I., Borovkova E. I., Shvartz V. A., Ishbulatov Y. M., Posnenkova O. M., Bezruchko B. P. Method of Estimation of Synchronization Strength Between Low-Frequency Oscillations in Heart Rate Variability and Photoplethysmographic Waveform Variability // *Russian Open Medical Journal*. 2016. Vol. 5. Iss. 1. P. e0101. doi:10.15275/rusomj.2016.0103
3. Караваев А. С., Киселев А. Р., Гриднев В. И., Боровкова Е. И., Прохоров М. Д., Посненкова О. М., Пономаренко В. И., Безручко Б. П., Шварц В. А. Фазовый и частотный захват 0.1 Гц-колебаний в ритме сердца и барорефлекторной регуляции артериального давления дыханием с линейно меняющейся частотой у здоровых лиц // *Физиология человека*. 2013. Т. 39. № 4. С. 93–104. doi:10.7868/S0131164613010049
4. Боровкова Е. И., Караваев А. С., Пономаренко В. И., Прохоров М. Д. Сопоставление методов диагностики фазовой синхронизованности по тестовым данным, моделирующим нестационарные сигналы биологической природы // *Изв. Саратовского университета. Новая серия. Серия Физика*. 2015. Вып. 3. № 15. С. 36–42. doi:10.18500/1817-3020-2015-15-3-36-42
5. Shvartz V. A., Karavaev A. S., Borovkova E. I., Mironov S. A., Ponomarenko V. I., Prokhorov M. D., Ishbulatov Y. M., Lapsheva E. E., Gridnev V. I., Kiselev A. R. Investigation of Statistical Characteristics of Interaction between the Low-frequency Oscillations in Heart Rate Variability and Photoplethysmographic Waveform Variability in Healthy Subjects and Myocardial Infarction Patients // *Russian Open Medical Journal*. 2016. Vol. 5. Iss. 1. P. e0203. doi:10.15275/rusomj.2016.0203
6. Smirnov D. A., Marwan N., Breitenbach S. F. M., Lechleitner F., Kurths J. Coping with Dating Errors in Causality Estimation // *Europhysics Letters*. 2017. Vol. 117. P. 10004. doi:10.1209/0295-5075/117/10004
7. Syssoeva M. V., Luttjohann A., G. van Luijtelaa, Syssoev I. V. Dynamics of Directional Coupling Underlying Spike-Wave Discharges // *Neuroscience*. 2016. Vol. 314. P. 75–89. doi:10.1016/j.neuroscience.2015.11.044
8. Moskalenko O., Koronovskii A., Hramov A. Inapplicability of an Auxiliary-System Approach to Chaotic Oscillators with Mutual-type Coupling and Complex Networks // *Phys Rev E*. 2013. Vol. 87. P. 064901. doi:10.1103/PhysRevE.87.064901
9. Luckett P., McDonald J. T., Hively L. M. Dissimilarity of Graph Invariant Features from EEG Phase-space Analysis // *J Comput Eng Inf Technol*. 2017. Vol. 6. Iss. 3. P. 1–10. doi:10.4172/2324-9307.1000172
10. Руннова А. Е., Журавлев М. О., Ситникова Е. Ю., Короновский А. А., Храмов А. Е. Метод удаления глазодвигательных артефактов на ЭЭГ человека при распознавании неоднозначного зрительного образа // *Информационно-управляющие системы*. 2017. № 5. С. 105–112. doi:10.15217/issn1684-8853.2017.5.105
11. Безручко Б. П., Смирнов Д. А. Математическое моделирование и хаотические временные ряды. — Саратов: Колледж, 2005. — 320 с.
12. Karavaev A. S., Ishbulatov J. M., Ponomarenko V. I., Prokhorov M. D., Gridnev V. I., Bezruchko B. P., Kiselev A. R. Model of Human Cardiovascular System with a Loop of Autonomic Regulation of the Mean Arterial Pressure // *Journal of the American Society of Hypertension*. 2016. Vol. 10. Iss. 3. P. 235–243. doi:10.1016/j.jash.2015.12.014
13. Баевский Р. М., Никулина Г. А., Тазетдинов И. Г. Математический анализ сердечного ритма в оценке особенностей адаптации организма к условиям космического полета // *Вестник АМН СССР*. 1984. № 4. С. 62–69.
14. Smirnov D. A., Bezruchko B. P. Revealing Mutual Influence of Oscillatory Systems from the Observation Data // *Radiophysics and Quantum Electronics*. 2013. Vol. 55. P. 662–675.
15. Hong H., O'Keefe K. P., Strogatz S. H. Phase Coherence Induced by Correlate Disorder // *Phys Rev E*. 2016. Vol. 93. Iss. 2. P. 022219. doi:10.1103/PhysRevE.93.022219
16. Kuklik P., Bidar E., Gharaviri A., Maessen J., Schotten U. Application of Phase Coherence in Assessment of Spatial Alignment of Electrodes during Simultaneous Endocardial-Epicardial Direct Contact Mapping of Atrial Fibrillation // *Europace*. 2014. Vol. 16. Suppl. 4. P. 135–140. doi:10.1093/europace/euu247
17. Kvandal P., Sheppard L., Landsverk S. A., Stefanovska A., Kirkeboen K. A. Impaired Cerebrovascular Reactivity after Acute Traumatic Brain Injury can be Detected by Wavelet Phase Coherence Analysis of the Intracranial and Arterial Blood Pressure Signals // *J Clin Monit Comput*. 2013. Vol. 27. Iss. 4. P. 375–383. doi:10.1007/s10877-013-9484-z
18. Tian F., Tarumi T., Liu H., Zhang R., Chalak L. Wavelet Coherence Analysis of Dynamic Cerebral Autoregulation in Neonatal Hypoxic-Ischemic Encephalopathy // *Neuroimage Clin*. 2016. Vol. 11. P. 124–132. doi:10.1016/j.nicl.2016.01.020.
19. Gordon S. M., Franaszczuk P. J., Hairston W. D., Vindiola M., McDowell K. Comparing Parametric and Nonparametric Methods for Detecting Phase

Synchronization in EEG // Journal of Neuroscience Methods. 2013. Vol. 212. P. 247–258. doi:10.1016/j.jneumeth.2012.10.002

20. Mormann F., Lehnertz K., David P., Elger C. Mean Phase Coherence as a Measure for Phase Synchroni-

zation and its Application to the EEG of Epilepsy Patients // Physica D. 2000. N 144. P. 358–369. doi:10.1016/S0167-2789(00)00087-7

UDC 577.31

doi:10.15217/issn1684-8853.2018.1.42

Diagnostics of Directional Coupling between Blood Circulation Regulation Loops using Analysis of Time Series of Mathematical Model of Human Cardiovascular System

Khorev V. S.^a, PhD, Phys.-Math., Associate Professor, khorevvs@gmail.com

Ishbulatov J. M.^a, Post-Graduate Student, ishbulatov95@mail.ru

Lapsheva E. E.^a, Senior Lecturer, lapsheva@yandex.ru

Kiselev A. R.^{b,c}, Dr. Sc., Med., Professor, antonkis@list.ru

Gridnev V. I.^{a,c}, Dr. Sc., Med., Professor, gridnev@cardio-it.ru

Bezruchko B. P.^a, Dr. Sc., Phys.-Math., Professor, bezruchkobp@gmail.com

Butenko A. A.^a, Assistant Professor, butenkoaan@yandex.ru

Ponomarenko V. I.^{a,d}, Dr. Sc., Phys.-Math., Professor, ponomarenkovi@gmail.com

Karavaev A. S.^{a,d}, PhD, Phys.-Math., Associate Professor, karavaevas@gmail.com

^aSaratov State University named after N. G. Chernyshevsky, 83, Astrakhanskaya St., 410012, Saratov, Russian Federation

^bBakoulev SCCVS, Department of Surgical Treatment for Interactive Pathology, 135, Roublyevskoe Sh., 121552, Moscow, Russian Federation

^cSaratov State Medical University named after V. I. Razumovsky, 112, Bolshaya Kazachia St., 410012, Saratov, Russian Federation

^dSaratov Branch of the Kotelnikov Institute of Radioengineering and Electronics of RAS, 38, Zelyonaya St., 410019, Saratov, Russian Federation

Introduction: The development of methods for assessing the interaction between signals of biological nature is often complicated by the need to adjust and approximate certain technique parameters, as detailed representations of the investigated object functioning are lacking. Using mathematical models is an effective solution to this problem. **Purpose:** Analyzing the directional coupling between the signals reflecting the dynamics of the contours of sympathetic regulation in human heart rate and baroreflex regulation of blood pressure, taking into account experimental data. **Methods:** Our method is based on constructing individual and joint models of oscillation signal phase dynamics in the form of harmonic polynomials of the 3rd degree. The main idea is to evaluate how much the future evolution of a phase of one system depends on the current phase value of the other system. If taking into account the dynamics of the first system reduces the prediction error of the model for the individual dynamics of the second system, the first system has an influence on the second one. **Results:** Comparative analysis of estimates of coupling strength between the model signals of the cardiointervalograms and the blood pressure allowed us to obtain the dependencies of mean values and deviation of estimates on the length of the analyzed part of the series. We analyzed the data in non-overlapping windows from 10 to 40 minutes long. The use of time series less than 20 minutes substantially increases the deviation of the estimates and thus negatively affects the possibility of obtaining an accurate estimate of the coupling strength. The results obtained for the model are qualitatively consistent with the results obtained from an experiment on healthy subjects, however the quantitative discrepancy requires further improvement of the model. **Practical relevance:** The results can be used to further improve the model of baroreflex regulation of the cardiovascular system.

Keywords — Signal Analysis, Mathematical Model, Instantaneous Phase, Directional Coupling, Delay Time, Cardiovascular System, Autonomic Regulation of Heart.

Citation: Khorev V. S., Ishbulatov J. M., Lapsheva E. E., Kiselev A. R., Gridnev V. I., Bezruchko B. P., Butenko A. A., Ponomarenko V. I., Karavaev A. S. Diagnostics of Directional Coupling between Blood Circulation Regulation Loops Using Analysis of Time Series of Mathematical Model of Human Cardiovascular System. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 1, pp. 42–48 (In Russian). doi:10.15217/issn1684-8853.2018.1.42

References

1. Kiselev A. R., Gridnev V. I., Prokhorov M. D., Karavaev A. S., Posnenkova O. M., Ponomarenko V. I., Bezruchko B. P. Selection of Optimal Dose of Beta-blocker Treatment in Myocardial Infarction Patients Basing on Changes in Synchronization Between 0.1 Hz Oscillations in Heart Rate and Peripheral Microcirculation. *Journal of Cardiovascular Medicine*, 2012, vol. 13, iss. 8, pp. 491–498. doi:10.2459/JCM.0b013e3283512199
2. Kiselev A. R., Karavaev A. S., Gridnev V. I., Prokhorov M. D., Ponomarenko V. I., Borovkova E. I., Shvartz V. A., Ishbulatov Y. M., Posnenkova O. M., Bezruchko B. P. Method of Estimation of Synchronization Strength Between Low-Frequency Oscillations in Heart Rate Variability and Photoplethysmographic Waveform Variability. *Russian Open Medical Journal*, 2016, vol. 5, iss. 1, pp. e0101. doi:10.15275/rusomj.2016.0103
3. Karavaev A. S., Kiselev A. R., Gridnev V. I., Borovkova E. I., Prokhorov M. D., Posnenkova O. M., Ponomarenko V. I., Bezruchko B. P., Shvartz V. A. Phase and Frequency Locking of 0.1 Hz Oscillations in Heart Rhythm and Baroreflex Control of Arterial Pressure by Respiration with Linearly Varying Frequency in Healthy Subjects. *Fiziologiya cheloveka*, 2013, vol. 39, no. 4, pp. 93–104 (In Russian). doi:10.7868/S0131164613010049
4. Borovkova E. I., Karavaev A. S., Ponomarenko V. I., Prokhorov M. D. Comparison of Methods for Phase Syn-

- chronization Diagnostics from Test Data Modelling Nonstationary Signals of Biological Nature. *Izvestiia Saratovskogo universiteta. Novaia seriia. Serii Fizika*, 2015, iss. 3, no. 15, pp. 36–42 (In Russian). doi:10.18500/1817-3020-2015-15-3-36-42
5. Shvartz V. A., Karavaev A. S., Borovkova E. I., Mironov S. A., Ponomarenko V. I., Prokhorov M. D., Ishbulatov Y. M., Lapsheva E. E., Gridnev V. I., Kiselev A. R. Investigation of Statistical Characteristics of Interaction between the Low-Frequency Oscillations in Heart Rate Variability and Photoplethysmographic Waveform Variability in Healthy Subjects and Myocardial Infarction Patients. *Russian Open Medical Journal*, 2016, vol. 5, iss. 1, pp. e0203. doi:10.15275/rusomj.2016.0203
 6. Smirnov D. A., Marwan N., Breitenbach S. F. M., Lechleitner F., Kurths J. Coping with Dating Errors in Causality Estimation. *Europhysics Letters*, 2017, vol. 117, pp. 10004. doi:10.1209/0295-5075/117/10004
 7. Sysoeva M. V., Luttjohann A., G. van Luijelaar, Sysoev I. V. Dynamics of Directional Coupling Underlying Spike-Wave Discharges. *Neuroscience*, 2016, vol. 314, pp. 75–89. doi:10.1016/j.neuroscience.2015.11.044
 8. Moskalenko O., Koronovskii A., Hramov A. Inapplicability of an Auxiliary-System Approach to Chaotic Oscillators with Mutual-type Coupling and Complex Networks. *Phys Rev E*, 2013, vol. 87, pp. 064901. doi:10.1103/PhysRevE.87.064901
 9. Luckett P., McDonald J. T., Hively L. M. Dissimilarity of Graph Invariant Features from EEG Phase-space Analysis. *J Comput Eng Inf Technol*, 2017, vol. 6, iss. 3, pp. 1–10. doi:10.4172/2324-9307.1000172
 10. Runnova A. E., Zhuravlev M. O., Sitnikova E. Y., Koronovskii A. A., Hramov A. E. Removing Oculomotor Artifacts in Human EEG during Ambiguous Image Recognition. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2017, no. 5, pp. 105–112 (In Russian). doi:10.15217/issn1684-8853.2017.5.105
 11. Bezruchko B. P., Smirnov D. A. *Matematicheskoe modelirovanie i khaoticheskie vremennye riady* [Mathematical Modeling and Chaotic Time Series]. Saratov, College Publ., 2005. 320 p. (In Russian).
 12. Karavaev A. S., Ishbulatov J. M., Ponomarenko V. I., Prokhorov M. D., Gridnev V. I., Bezruchko B. P., Kiselev A. R. Model of Human Cardiovascular System with a Loop of Autonomic Regulation of the Mean Arterial Pressure. *Journal of the American Society of Hypertension*, 2016, vol. 10, iss. 3, pp. 235–243. doi:10.1016/j.jash.2015.12.014
 13. Baevsky R. M., Nikulina G. A. and Tazetdinov I. G. Mathematical Analysis of Cardiac Rhythm for Estimation of Specific Features of Organism Adaptation to the Space Flight Conditions. *Vestnik AMN SSSR*, 1984, no. 4(62), pp. 1–18 (In Russian).
 14. Smirnov D. A., Bezruchko B. P. Revealing Mutual Influence of Oscillatory Systems from the Observation Data. *Radio-physics and Quantum Electronics*, 2013, vol. 55, pp. 662–675.
 15. Hong H., O’Keeffe K. P., Strogatz S. H. Phase Coherence Induced by Correlate Disorder. *Phys Rev E*, 2016, vol. 93, iss. 2, pp. 022219. doi:10.1103/PhysRevE.93.022219
 16. Kuklik P., Bidar E., Gharaviri A., Maessen J., Schotten U. Application of Phase Coherence in Assessment of Spatial Alignment of Electrodes During Simultaneous Endocardial-Epicardial Direct Contact Mapping of Atrial Fibrillation. *Europace*, 2014, vol. 16, suppl. 4, pp. 135–140. doi:10.1093/europace/euu247
 17. Kvandal P., Sheppard L., Landsverk S. A., Stefanovska A., Kirkeboen K. A. Impaired Cerebrovascular Reactivity after Acute Traumatic Brain Injury can be Detected by Wavelet Phase Coherence Analysis of the Intracranial and Arterial Blood Pressure Signals. *J Clin Monit Comput*, 2013, vol. 27, iss. 4, pp. 375–383. doi:10.1007/s10877-013-9484-z
 18. Tian F., Tarumi T., Liu H., Zhang R., Chalak L. Wavelet Coherence Analysis of Dynamic Cerebral Autoregulation in Neonatal Hypoxic-Ischemic Encephalopathy. *Neuroimage Clin*, 2016, vol. 11, pp. 124–132. doi: 10.1016/j.nicl.2016.01.020
 19. Gordon S. M., Franaszczuk P. J., Hairston W. D., Vindiola M., McDowell K. Comparing Parametric and Nonparametric Methods for Detecting Phase Synchronization in EEG. *Journal of Neuroscience Methods*, 2013, vol. 212, pp. 247–258. doi:10.1016/j.jneumeth.2012.10.002
 20. Mormann F., Lehnertz K., David P., Elger C. Mean Phase Coherence as a Measure for Phase Synchronization and its Application to the EEG of Epilepsy Patients. *Physica D*, 2000, no. 144, pp. 358–369. doi:10.1016/S0167-2789(00)00087-7

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: входите на страницу <http://www.researcherid.com>, слева под надписью «New to ResearcherID?» нажимаете на синюю кнопку «Join Now It’s Free» и заполняете короткую анкету. По указанному электронному адресу получаете сообщение с предложением по ссылке заполнить полную регистрационную форму на ORCID. Получаете ID.

УДК 62.001.4

doi:10.15217/issn1684-8853.2018.1.49

ДОСТОВЕРНОСТЬ МНОГОЭТАПНОГО КОНТРОЛЯ ТЕХНИЧЕСКОГО СОСТОЯНИЯ ОБЪЕКТОВ ИСПЫТАНИЙ

Г. Н. Мальцев^а, доктор техн. наук, профессор, georgy_maltsev@mail.ru

В. Л. Якимов^а, канд. техн. наук, доцент, yakim78@yandex.ru

^аВоенно-космическая академия им. А. Ф. Можайского, Ждановская наб., 13, Санкт-Петербург, 197198, РФ

Введение: обеспечение достоверности контроля технического состояния объектов испытаний является обязательным условием эффективного проведения любых испытаний технических систем. Оценка достоверности контроля технического состояния объекта испытаний при планировании испытаний и анализе их результатов требует выбора показателя достоверности и разработки методик его расчета на основе моделей процесса испытаний. При этом используемый показатель должен быть ориентирован прежде всего на достоверное заключение об исправности объекта испытаний. **Цель исследования:** обоснование интегрального показателя достоверности контроля технического состояния объектов испытаний и методик его расчета при проведении отдельного этапа испытаний и многоэтапных испытаний. **Результаты:** обосновано использование в качестве интегрального показателя достоверности контроля технического состояния объекта испытаний апостериорной вероятности правильного заключения о его исправности, представлены аналитические выражения для его расчета при проведении отдельного этапа испытаний и при проведении многоэтапных однородных испытаний, исследовано влияние на выбранный показатель достоверности вероятностей ошибок первого и второго рода при контроле технического состояния объекта испытаний. **Практическая значимость:** представленные результаты определяют принципы анализа с помощью интегральных показателей достоверности контроля технического состояния объектов испытаний при планировании испытаний и анализе их результатов. Приведенные выражения позволяют рассчитать рассмотренный в работе интегральный показатель достоверности при проведении отдельного этапа испытаний и при проведении многоэтапных испытаний.

Ключевые слова — испытания, контроль технического состояния, достоверность, ошибки первого и второго рода.

Цитирование: Мальцев Г. Н., Якимов В. Л. Достоверность многоэтапного контроля технического состояния объектов испытаний // Информационно-управляющие системы. 2018. № 1. С. 49–57. doi:10.15217/issn1684-8853.2018.1.49

Citation: Maltsev G. N., Yakimov V. L. Reliability of Multi-Stage Control over Technical Condition of Tested Objects. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 1, pp. 49–57 (In Russian). doi:10.15217/issn1684-8853.2018.1.49

Введение

Контроль технического состояния (КТС) является одной из основных операций при проведении испытаний технических систем [1–3]. Под техническим состоянием объекта испытаний (ОИ) в общем случае понимается совокупность его свойств, характеризуемая в определенный момент времени совокупностью значений показателей и (или) качественных признаков, установленных техническими требованиями. Число видов технического состояния ОИ, о принадлежности к одному из которых принимается решение в результате КТС, определяется целью испытаний. Наиболее общими видами технического состояния являются исправное и неисправное, при этом КТС и определение исправности ОИ может составлять основную цель испытаний или являться первым этапом технического диагностирования ОИ.

Контроль технического состояния сложных технических систем является многоуровневым и многоэтапным [3–5]. На каждом этапе испытаний в качестве ОИ выступают сложная техническая система в целом или ее составные части и имеет место определенная детализация техниче-

ского состояния ОИ в соответствии с учитываемыми признаками. При этом по мере усложнения технических систем роль испытаний в процессе их создания становится все более значительной. По имеющимся оценкам, для сложных технических систем, создаваемых в аэрокосмической отрасли, из-за невозможности получить адекватное теоретическое описание до 40 % всех возникающих проблем решаются при помощи испытаний [3]. Кроме того, процесс испытаний сложных технических систем может рассматриваться в широком смысле — относительно не только этапа их создания и постановки на производство с экспериментальной отработкой [4, 5], но и этапа эксплуатации, на котором подготовка к применению сложных технических систем, как правило, включает проведение многочисленных автономных и комплексных проверок [6, 7].

Основным результатом КТС при его проведении в целях определения исправности ОИ является принятие решения о его годности (готовности): на этапе экспериментальной отработки — к последующим испытаниям или к постановке на производство, а на этапе эксплуатации — к применению по назначению. Однако в силу действия случайных факторов принимаемые решения о

техническом состоянии ОИ характеризуются конечной достоверностью, зависящей от программы и объема испытаний. Определение объема испытаний, исходя из требуемой достоверности их результатов, является одной из важных задач при планировании испытаний сложных технических систем. В настоящей статье исследуются показатели достоверности КТС ОИ при проведении многоэтапных испытаний.

Показатели достоверности отдельного этапа испытаний

Под достоверностью КТС ОИ в общем случае понимается степень объективного соответствия результата контроля его истинному техническому состоянию. Достоверность результатов контроля является интегральной характеристикой качества КТС любой технической системы. При проведении испытаний она отражает степень обоснованности решений о техническом состоянии ОИ и определяет уровень доверия к полученным результатам. При проведении КТС ОИ с целью определить его исправность по результатам испытаний принимаются решения «годен» или «негоден», означающие оценку ОИ как исправного и неисправного соответственно. Таким образом, достоверность КТС ОИ определяет достоверность оценки свойств ОИ по результатам испытаний, являющуюся одним из основных критериев эффективности испытаний [1, 3].

Случайный характер возникновения неисправностей и ошибок при оценке технического состояния ОИ приводит к тому, что основными характеристиками достоверности КТС любых технических систем являются вероятностные характеристики. Они могут быть априорными и апостериорными. При планировании испытаний и прогнозировании ожидаемой достоверности результатов испытаний используются априорные вероятностные характеристики. В то же время задачам КТС на различных этапах жизненного цикла технических систем в наибольшей степени соответствуют апостериорные вероятностные характеристики, характеризующие вероятность того, что ОИ, признанный по результатам КТС исправным и соответствующим техническим требованиям, действительно является таковым.

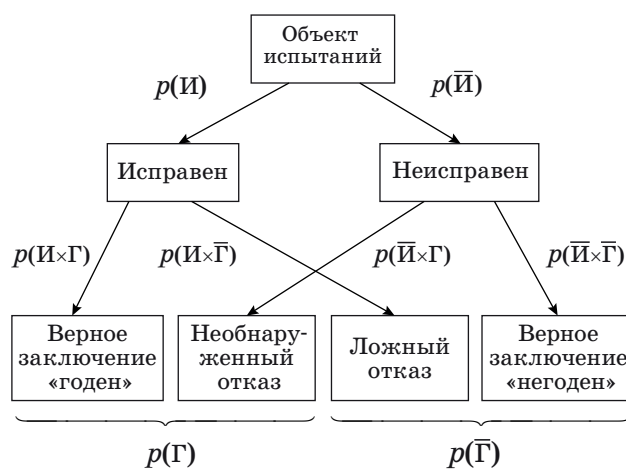
Задача КТС ОИ в целях определения его исправности сводится к задаче распознавания двух объективно возможных технических состояний ОИ — исправного (И) и неисправного (Ӣ) и принятия решения о результатах контроля ОИ — годен (Г) или негоден (Г̄). Исходное техническое состояние ОИ характеризуется вероятностями исправного $P = p(И)$ и неисправного $Q = p(Ӣ)$ состояний в момент контроля, которые могут быть

различными на различных этапах контроля. Результаты контроля характеризуются вероятностями заключения «годен» $p(Г)$ и заключения «негоден» $p(Г̄)$. Очевидно, что: $P + Q = 1$, $p(Г) + p(Г̄) = 1$.

Схема, представленная на рис. 1, соответствует двухальтернативному контролю технического состояния ОИ и при соответствующем задании вероятностей, характеризующих достоверность результатов контроля, может быть отнесена к оценке как отдельных или обобщенных параметров ОИ, так и технического состояния ОИ в целом.

Возможны четыре ситуации результатов контроля. В двух случаях по результатам КТС ОИ даются верные заключения: 1) признанный годным ОИ действительно с вероятностью верного заключения о годности $P_{в.з.г} = p(И \times Г)$ является исправным; 2) признанный негодным ОИ действительно с вероятностью верного заключения о негодности $P_{в.з.г̄} = p(Ӣ \times Г̄)$ является неисправным. В двух случаях из-за наличия различного рода ошибок заключения по результатам КТС ОИ являются неверными: 1) признанный негодным ОИ с вероятностью $P_{л.о} = p(И \times Г̄)$ является исправным — ложный отказ; 2) признанный годным ОИ с вероятностью $P_{н.о} = p(Ӣ \times Г)$ является неисправным — необнаруженный отказ. Суммарная вероятность верного заключения: $P_{в.з} = P_{в.з.г} + P_{в.з.г̄}$. Суммарная вероятность неверного заключения: $P_{н.з} = P_{л.о} + P_{н.о}$. Рассмотренные четыре ситуации составляют полную группу событий: $P_{в.з.г} + P_{в.з.г̄} + P_{л.о} + P_{н.о} = 1$.

Вероятности верных заключений по результатам КТС ОИ $P_{в.з.г}$ и $P_{в.з.г̄}$ определяются безусловными вероятностями $p(И \times Г)$ и $p(Ӣ \times Г̄)$ правильных заключений системы контроля об



■ Рис. 1. Обобщенная схема формирования результатов контроля технического состояния

■ Fig. 1. Generalized scheme for the formation of control results technical condition

исправности или неисправности ОИ. Они объективно характеризуют качество системы контроля и могут быть использованы при аналитическом описании и моделировании процессов КТС и диагностирования сложных технических систем [8, 9]. В то же время достоверность результатов конкретных испытаний, в каждом из которых исходные вероятности исправного и неисправного состояния ОИ и вероятности признания ОИ годным или негодным по результатам испытаний могут быть различными, вероятности $P_{в.з.г}$ и $P_{в.з.г}$ отражают не в полной мере. Систему контроля принято характеризовать априорными вероятностями ошибок первого и второго рода α и β , непосредственно связанными с условиями проведения испытаний и не зависящими от вероятностей исправного и неисправного состояний ОИ в момент КТС.

Вероятность ошибки первого рода $\alpha = p(\bar{\Gamma}|\text{И}) = \frac{p(\text{И} \times \bar{\Gamma})}{p(\text{И})}$ есть вероятность наступле-

ния события ложного отказа, при котором в результате контроля исправный ОИ будет признан негодным. Появление ложного отказа приводит к неоправданным затратам изготовителя на восстановление ОИ или его перепроверку (поиск несуществующего дефекта). Поэтому ошибку первого рода также называют риском изготовителя. Вероятность ошибки второго рода

$\beta = p(\Gamma|\bar{\text{И}}) = \frac{p(\bar{\text{И}} \times \Gamma)}{p(\bar{\text{И}})}$ есть вероятность наступле-

ния события необнаруженного отказа, при котором неисправный ОИ в результате контроля будет признан годным. Наличие необнаруженного отказа ОИ приводит к выходу его из строя в процессе эксплуатации потребителем и срыву выполнения поставленных задач. Поэтому ошибку второго рода также называют риском потребителя. Очевидно, что необнаруженные отказы наиболее опасны.

Вероятности ошибок первого и второго рода α и β широко используются как показатели достоверности результатов КТС ОИ [1–3]. При этом, хотя понятия рисков поставщика и потребителя в буквальном смысле слова относятся к приемосдаточным испытаниям, по результатам которых изделие принимается потребителем у поставщика, соответствующие показатели достоверности и их интерпретация используются по отношению к испытаниям любого уровня, вплоть до стандартных операций допускового контроля отдельных параметров. В последнем случае определение вероятностей ошибок первого и второго рода строго формализуется для заданного статистического описания разброса значений контролируемого параметра и погрешностей его измерения.

Пусть осуществляется допусковый контроль параметра x . ОИ признается годным, если оценка параметра не выходит за допустимые пределы: $c < x < d$, где c и d — соответственно нижняя и верхняя границы контрольного поля допуска. При этом ОИ действительно является исправным на момент окончания контроля, если истинное значение параметра находится в области работоспособности: $a < x < b$, где a и b — соответственно нижняя и верхняя границы области работоспособности. Вероятности ошибок первого и второго рода при допусковом контроле параметра y определяются следующими выражениями [10]:

$$\alpha = \int_a^b w(x) \left[1 - \int_{c-x}^{d-x} q(\xi) d\xi \right] dx; \quad (1)$$

$$\beta = \int_{-\infty}^a w(x) \left[\int_{c-x}^{d-x} q(\xi) d\xi \right] dx + \int_b^{\infty} w(x) \left[\int_{c-x}^{d-x} q(\xi) d\xi \right] dx, \quad (2)$$

где $w(x)$ — априорная плотность вероятности значений измеряемого параметра x на момент начала контроля; $q(\xi)$ — плотность вероятности погрешности измерения ξ . Выражения (1) и (2) являются примерами аналитических выражений для расчета вероятностей ошибок первого и второго рода при проведении отдельной операции КТС ОИ и будут использованы для определения этих вероятностей в случае многоэтапных однородных испытаний.

Достоверность результатов контроля как степень объективности отражения в результатах контроля истинного технического состояния ОИ характеризуют безусловные вероятности правильных заключений системы контроля, нормированные к соответствующим вероятностям признания ОИ по результатам испытаний годным или негодным: $D_{\Gamma} = \frac{p(\text{И} \times \Gamma)}{p(\Gamma)}$ и $D_{\bar{\Gamma}} = \frac{p(\bar{\text{И}} \times \bar{\Gamma})}{p(\bar{\Gamma})}$.

Введенные величины D_{Γ} и $D_{\bar{\Gamma}}$ есть не что иное, как апостериорные вероятности правильных заключений системы контроля об исправности или неисправности ОИ: $D_{\Gamma} = p(\text{И}|\Gamma)$ и $D_{\bar{\Gamma}} = p(\bar{\text{И}}|\bar{\Gamma})$.

Представление апостериорных вероятностей правильных заключений системы контроля $D_{\Gamma} = p(\text{И}|\Gamma)$ и $D_{\bar{\Gamma}} = p(\bar{\text{И}}|\bar{\Gamma})$ по формуле Байеса дает

$$D_{\Gamma} = p(\text{И}|\Gamma) = \frac{p(\text{И})p(\Gamma|\text{И})}{p(\text{И})p(\Gamma|\text{И}) + p(\bar{\text{И}})p(\Gamma|\bar{\text{И}})}; \quad (3)$$

$$D_{\bar{\Gamma}} = p(\bar{\text{И}}|\bar{\Gamma}) = \frac{p(\bar{\text{И}})p(\bar{\Gamma}|\bar{\text{И}})}{p(\bar{\text{И}})p(\bar{\Gamma}|\bar{\text{И}}) + p(\text{И})p(\bar{\Gamma}|\text{И})}. \quad (4)$$

Подставляя в выражения (3) и (4) значения вероятностей $p(\text{И}) = P$, $p(\bar{\text{И}}) = 1 - P$, $p(\Gamma|\text{И}) = \alpha$, $p(\Gamma|\bar{\text{И}}) = \beta$, после преобразований получаем

$$D_{\Gamma} = \frac{P(1-\alpha)}{P(1-\alpha) + (1-P)\beta}; \quad (5)$$

$$D_{\bar{\Gamma}} = \frac{(1-P)(1-\beta)}{(1-P)(1-\beta) + P\alpha}. \quad (6)$$

Выражения (5) и (6) определяют обобщенные показатели достоверности КТС ОИ, учитывающие исходное техническое состояние ОИ, характеризующее вероятностью его исправного состояния к началу контроля P , и характеристики системы контроля — вероятности ошибок первого и второго рода α и β . Наибольший интерес и для поставщика, и для потребителя представляет достоверное заключение о годности исправного объекта испытаний. Достоверность такого результата контроля характеризует показатель D_{Γ} (5). К такому же представлению показателя достоверности можно прийти на основании анализа графа состояний процесса КТС ОИ, приведенного на рис. 1. При этом введением соответствующих вероятностей в показателе D_{Γ} дополнительно могут быть учтены отказ объекта испытаний в процессе и после окончания испытаний:

$$D_{\Gamma} = \frac{P(1-q)(1-\alpha)(1-\gamma)}{P[(1-q)(1-\alpha) + q\beta] + (1-P)[(1-q)\beta + q\beta]}, \quad (7)$$

где q — вероятность отказа ОИ в процессе контроля; γ — вероятность отказа ОИ после окончания испытаний (до начала следующего этапа испытаний). В выражении (7) параметры P , q и γ характеризуют ОИ, а параметры α и β — систему испытаний. При отсутствии отказов ОИ в процессе и после окончания испытаний $q = 0$ и $\gamma = 0$, и выражение (7) сводится к выражению (5).

Значения показателя достоверности D_{Γ} изменяются в пределах от 0 до 1. Полная достоверность контроля ($D_{\Gamma} = 1$) достигается в трех случаях: 1) при $q = 0$, $\gamma = 0$, $P = 1$ и ненулевых значениях α и β ; 2) при $\alpha = 0$, $\beta = 0$, $\gamma = 0$, ненулевом значении q и отличном от единичного значении P ; 3) при $q = 0$, $\beta = 0$, $\gamma = 0$, ненулевом значении α и отличном от единичного значении P . При $\gamma > 0$ достижение полной достоверности невозможно и максимально возможное значение D_{Γ} составляет $(1 - \gamma)$. Полная недостоверность контроля ($D_{\Gamma} = 0$) имеет место, если $q = 1$, или $\alpha = 1$, или $\gamma = 1$. Показатель $D_{\bar{\Gamma}}$, наоборот, уменьшается с увеличением вероятности P и в области ее значений, близких к единице, стремится к нулю и может быть интересен лишь в задачах выявления неисправностей аппаратуры с высокой вероятностью отказов.

Показатель D_{Γ} , определяемый выражениями (5) или (7), может быть рекомендован к использованию в качестве интегрального вероятност-

ного показателя достоверности КТС высоконадежных технических систем при проведении как одноэтапных, так и многоэтапных испытаний. В последнем случае результаты испытаний характеризует показатель достоверности, относящийся к заключительному этапу испытаний. Условия проведения испытаний технических систем должны выбираться таким образом, чтобы при расчетном обосновании достоверности КТС ОИ для показателя достоверности $D = D_{\Gamma}$ выполнялось условие $D > D_{\text{треб}}$, где $D_{\text{треб}}$ — требуемое значение показателя достоверности контроля.

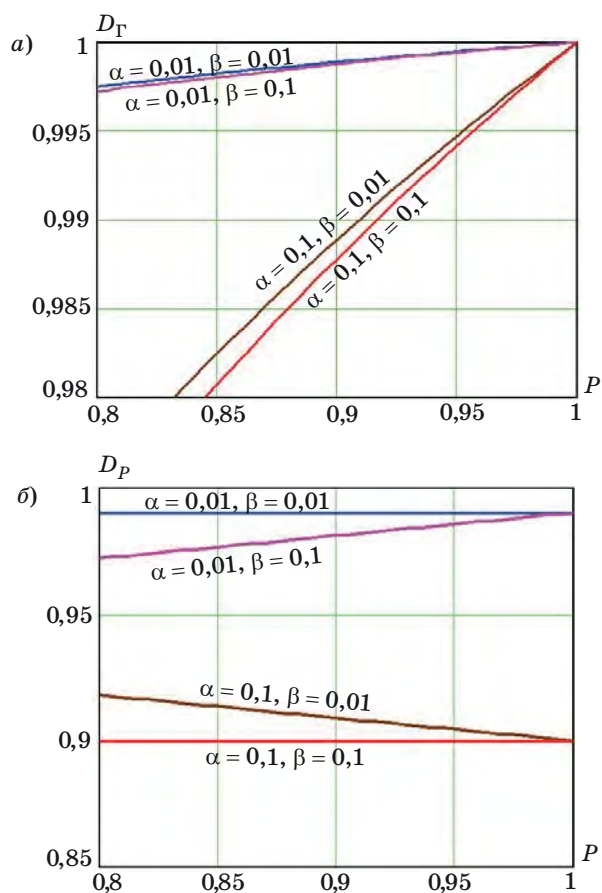
Особенностью показателя достоверности контроля D_{Γ} является критичность к вероятности ошибок второго рода β и, соответственно, к риску потребителя. Это вполне оправдано тем, что ошибки контроля технического состояния типа необнаруженных отказов наиболее опасны. Кроме того, показатель достоверности D_{Γ} существенным образом зависит от априорной вероятности исправного состояния ОИ до начала испытаний P . Чем больше P , тем выше показатель достоверности контроля D_{Γ} . Влияние на достоверность контроля вероятности ошибок первого рода α менее значительно. Эти ошибки приводят к менее опасным ошибкам КТС типа ложных отказов. Увеличение вероятностей q и γ во всех случаях приводит к уменьшению показателя достоверности контроля D_{Γ} . В дальнейшем будем полагать, что $q = 0$, $\gamma = 0$ и величина D_{Γ} зависит от вероятностей P , α и β .

На рис. 2, а приведены зависимости показателя достоверности D_{Γ} от вероятности исправного технического состояния ОИ в начале контроля P , рассчитанного в соответствии с выражением (5), при различных сочетаниях ошибок первого и второго рода α и β .

Определяющее влияние на показатель достоверности D_{Γ} оказывают начальная вероятность исправного технического состояния ОИ P и величина вероятности ошибки второго рода β . В области высоких значений вероятности P показатель достоверности контроля D_{Γ} стремится к единице, что отражает очевидный физический принцип: чем выше надежность ОИ, тем выше достоверность контроля его технического состояния. Это позволяет достигать близких к единице значений показателя достоверности D_{Γ} при ненулевых значениях вероятностей α и β .

Для сравнения на рис. 2, б приведены зависимости от вероятности P при различных сочетаниях ошибок первого и второго рода α и β так называемой достоверности разбраковки $D_p = 1 - P\alpha - (1 - P)\beta$, относящейся к числу безусловных показателей достоверности и также используемой при оценке достоверности КТС ОИ [4, 6].

Показатель достоверности D_p не столь критичен к вероятности исправного технического со-



■ **Рис. 2.** Показатель достоверности КТС (а) и забраковки (б) в зависимости от вероятности исправного состояния ОИ в начале контроля при различных вероятностях ошибок первого и второго рода

■ **Fig. 2.** Indicator of technical condition control reliability (a) and culling (b) depending on the probability of operational state of tested object at the beginning of the control for different error probabilities of type I and II

стояния ОИ P и достигает единичного значения только при нулевых значениях вероятностей ошибок первого и второго рода α и β , а при различных сочетаниях этих вероятностей может как увеличиваться, так и уменьшаться с увеличением вероятности P .

В области значений вероятности P , близких к единице, показатель D_P определяется значением вероятности ошибки первого рода α .

Показатели достоверности многоэтапных испытаний

При анализе достоверности многоэтапных испытаний необходимо определять показатели достоверности КТС ОИ после заключительного этапа испытаний. Многократные испытания могут быть однородными и неоднородными. При однородных испытаниях на различных этапах ис-

пытаний проводятся одинаковые проверки ОИ, что соответствует использованию одной и той же системы КТС ОИ с не изменяющимися от этапа к этапу параметрами α и β . При неоднородных испытаниях порядок проведения проверок ОИ на различных этапах испытаний отличается, что соответствует использованию на различных этапах различных систем КТС ОИ, каждой со своими параметрами α и β .

Определение показателя достоверности контроля D_{Γ} формализуется в виде достаточно простых расчетных выражений в случае многократного КТС ОИ при проведении однородных испытаний. Рассмотрим однородные испытания, включающие N независимых проверок. В общем случае по результатам N проверок ОИ может быть признан годным, если результат контроля «годен» был получен $n \leq N$ раз. В большинстве случаев в ходе испытаний сложных технических систем при наблюдении отрицательного исхода проверки осуществляется отбраковка (снятие с испытаний) ОИ, и после ремонта, настройки или регулировки его испытания начинаются сначала. Следовательно, годным признается только тот ОИ, который по результатам N -кратной проверки признан годным $n = N$ раз. Показатель достоверности контроля D_{Γ} после N проверок определяется в соответствии с выражением (5) для исходной вероятности исправного технического состояния ОИ в начале контроля P и вероятности ошибок первого и второго рода $\alpha_{(N)}$ и $\beta_{(N)}$ после N проверок.

Пусть известны $\alpha = \alpha_{(1)}$ и $\beta = \beta_{(1)}$ — вероятности ошибок первого и второго рода при однократной проверке ($N = 1$). Рассмотрим порядок определения вероятностей ошибок первого и второго рода $\alpha_{(N)}$ и $\beta_{(N)}$ после N -кратной проверки с отбраковкой, которыми будет определяться достижимое значение показателя достоверности контроля технического состояния ОИ D_{Γ} . Для определения значений $\alpha_{(N)}$ и $\beta_{(N)}$ используем подход, основанный на анализе условной вероятности ошибки типа ложного отказа $p(\bar{\Gamma}|\text{И})$ и апостериорной вероятности ошибки типа необнаруженного отказа $p(\bar{\text{И}}|\Gamma)$ при N -кратной проверке с использованием допускового контроля [10, 11].

Априорная вероятность ошибки первого рода α определяется условной вероятностью ошибки типа ложного отказа: $\alpha = p(\bar{\Gamma}|\text{И})$. Определив для N -кратной проверки вероятность $p_{(N)}(\bar{\Gamma}|\text{И})$, мы непосредственно получаем значение $\alpha_{(N)}$. Априорная вероятность ошибки второго рода β определяется условной вероятностью ошибки типа необнаруженного отказа: $\beta = p(\Gamma|\bar{\text{И}})$. Определив для N -кратной проверки вероятность $p_{(N)}(\bar{\text{И}}|\Gamma)$, для получения значения $\beta_{(N)}$ необходимо установить взаимосвязь между вероятностями $p_{(N)}(\bar{\text{И}}|\Gamma)$ и $p_{(N)}(\Gamma|\bar{\text{И}})$.

Если входящую в выражения (1) и (2) плотность вероятности погрешности измерения $q(\xi)$ можно полагать симметричной гауссовой плотностью вероятности с нулевым математическим ожиданием, то при условии, что поле рассеяния погрешности измерений меньше поля допуска на измеряемый параметр, а области принятия решения и работоспособности совпадают, условные вероятности $p(\bar{\Gamma}|\text{И})$ и $p(\bar{\text{Й}}|\Gamma)$ при многократных измерениях для очередного i -го измерения приближенно определяются по формуле [10]

$$p_i(\bar{\Gamma}|\text{И}) = p_i(\bar{\text{Й}}|\Gamma) \approx \frac{w_{i-1}(c) + w_{i-1}(d)}{d} \int_c^d q(\xi) d\xi, \quad (8)$$

где $w_{i-1}(x)$ — плотность вероятности значений параметра на момент начала i -й проверки, $i = 1, 2, \dots, N$; c и d — соответственно нижняя и верхняя границы контрольного поля допуска, совпадающие с нижней и верхней границей области работоспособности ($c = a, d = b$). Формула (8) тем точнее, чем меньше погрешности отдельных измерений и чем больше их число N .

Ложный отказ ОИ может возникнуть при каждой из N проверок. При контроле с отбраковкой плотность распределения значений контролируемого параметра признанных годными ОИ после $i - 1$ проверок определяется выражением

$$w_{i-1}(x) = \frac{w(x)}{\prod_{k=1}^{i-1} p_{(k)}(\Gamma)} \left[\int_{c-x}^{d-x} q(\xi) d\xi \right]^{i-1}, \quad (9)$$

где $p_{(i)}(\Gamma)$ — вероятность признания ОИ годным после i -й проверки. При высокой точности измерений контролируемых параметров можно полагать, что вероятность признания ОИ годным после первой i -й проверки составляет

$p_{(1)}(\Gamma) \approx \int_c^d w(x) dx$, а после последующих проверок выполняется $\int_c^d w_{i-1}(x) dx \approx 1$ и $p_{(i)}(\Gamma) \approx 1, i = 2, 3, \dots, N$. Тогда, подставив в выражение (8) соответствующую указанным допущениям плотность распределения $w_{i-1}(x)$ вида (9), после преобразований получаем

$$p_i(\bar{\Gamma}|\text{И}) = \frac{1}{2^{i-1}} p_{(1)}(\bar{\Gamma}|\text{И}), \quad (10)$$

где условная вероятность ошибки типа ложного отказа $p_{(1)}(\bar{\Gamma}|\text{И})$ есть вероятность ошибки первого рода при одном измерении: $\alpha = p_{(1)}(\bar{\Gamma}|\text{И})$.

Переходя от определяемых выражением (10) условных вероятностей ошибок типа ложного отказа для каждой проверки $p_i(\bar{\Gamma}|\text{И}), i = 1, 2, \dots, N$, к результирующей условной вероятности ошибки типа ложного отказа, после N -кратной проверки с отбраковкой $\alpha_{(N)} = p_{(N)}(\bar{\Gamma}|\text{И})$ получаем

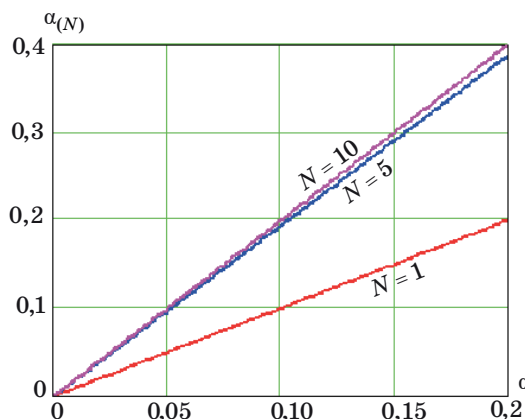
$$\alpha_{(N)} = 2\alpha \left(1 - \frac{1}{2^N} \right). \quad (11)$$

Выражение (11) устанавливает зависимость $\alpha_{(N)}$ от $\alpha = \alpha_{(1)}$ и от N : при $N \gg 1$ можно полагать $\alpha_{(N)} \approx 2\alpha$.

Анализ выражения (11) показывает, что многократный контроль с отбраковкой приводит к увеличению вероятности ошибки первого рода $\alpha_{(N)}$ по сравнению с вероятностью ошибки первого рода α при однократном контроле. Однако порядок величины вероятности ошибки первого рода при этом сохраняется, и максимальное ее увеличение не превышает двух раз. Рассчитанные по формуле (12) зависимости $\alpha_{(N)}$ от α при различных N показаны на рис. 3.

Видно, что зависимости, соответствующие $N = 5$ и $N = 10$, отличаются незначительно и практически совпадают с предельной зависимостью $\alpha_{(N)} = 2\alpha$.

Необнаруженный отказ ОИ при многократных проверках возникает после последней N -й проверки. Отказы, не обнаруженные при каждой из предыдущих $N - 1$ проверок, могут быть обнаружены при последующих проверках и не учитываются при определении результирующей условной вероятности необнаруженного отказа при N -кратной проверке $p_{(N)}(\bar{\text{Й}}|\Gamma)$. Следовательно, $p_{(N)}(\bar{\text{Й}}|\Gamma) = p_N(\bar{\text{Й}}|\Gamma)$, где $p_N(\bar{\text{Й}}|\Gamma)$ — апостериор-



■ Рис. 3. Вероятность ошибки первого рода при многоэтапных испытаниях в зависимости от вероятности ошибки первого рода на отдельном этапе

■ Fig. 3. Probability of type I error in multi-stage tests, depending on the probability of type I error at a particular stage

ная вероятность ошибки типа необнаруженного отказа при N -й проверке.

Плотность распределения значений контролируемого параметра признанных годными ОИ после $(N - 1)$ -й проверки определяется выражением (9) при $i = N$. Так же, как и при определении $p_{(N)}(\bar{\Gamma}|\bar{\text{И}})$, будем полагать, что имеет место высокая точность измерений контролируемых параметров и при первой проверке $p_{(1)}(\Gamma) \approx \int_c^d w(x)dx$,

а при последующих проверках выполняется $\int_c^d w_{i-1}(x)dx \approx 1$ и $p_{(i)}(\Gamma) \approx 1$, $i = 1, 2, \dots, N$. Тогда,

подставив в выражение (8) соответствующую указаным допущениям плотность распределения $w_{N-1}(x)$ вида (9), после преобразований получаем

$$p_N(\bar{\text{И}}|\Gamma) = \frac{1}{2^{N-1}} p_{(1)}(\bar{\text{И}}|\Gamma), \quad (12)$$

где $p_{(1)}(\bar{\text{И}}|\Gamma)$ — апостериорная вероятность ошибки типа необнаруженного отказа при одном измерении.

Переход от апостериорной вероятности ошибки типа необнаруженного отказа при N -кратной проверке $p_{(N)}(\bar{\text{И}}|\Gamma) = p_N(\bar{\text{И}}|\Gamma)$, определяемой выражением (12), к условной вероятности ошибки типа необнаруженного отказа $p_{(N)}(\Gamma|\bar{\text{И}})$, определяющей вероятность ошибки второго рода $\beta_{(N)}$, осуществляется по выражению

$$p(\bar{\text{И}}|\Gamma) = \frac{p(\Gamma|\bar{\text{И}})[1 - p(\text{И})]}{p(\text{И})[1 - p(\bar{\Gamma}|\text{И})] + p(\Gamma|\bar{\text{И}})[1 - p(\text{И})]}. \quad (13)$$

Используя выражения (11)–(13) с учетом того, что $P = p(\text{И})$, $\alpha = p(\bar{\Gamma}|\text{И})$, $\beta = p(\Gamma|\bar{\text{И}})$, для результирующей условной вероятности ошибки типа необнаруженного отказа после N -кратной проверки с отбраковкой $\beta_{(N)} = p_{(N)}(\Gamma|\bar{\text{И}})$ после преобразований получаем

$$\beta_{(N)} = \frac{\beta P}{2^{N-1} [P(1 - \alpha) + \beta(1 - P)] - \beta(1 - P)} \times \left[1 - 2\alpha \left(1 - \frac{1}{2^N} \right) \right]. \quad (14)$$

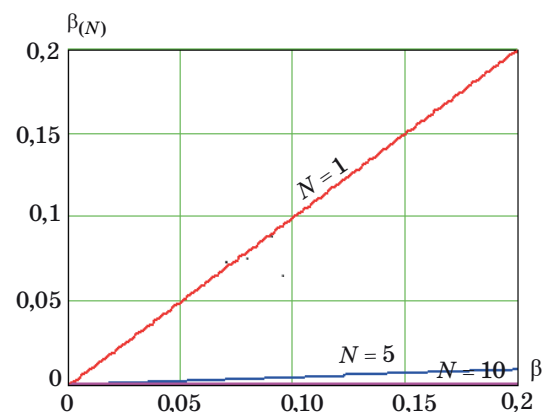
Выражение (14) устанавливает зависимость $\beta_{(N)}$ от β и от N , кроме того, $\beta_{(N)}$ зависит от α и от P .

Анализ выражения (14) показывает, что многократный контроль с отбраковкой приводит к уменьшению вероятности ошибки второго рода $\beta_{(N)}$ по сравнению с вероятностью ошибки второго рода β при однократном контроле, причем при $N \gg 1$ обеспечивается $\beta_{(N)} \rightarrow 0$. При этом величина

$\beta_{(N)}$ определяется в основном значениями β и N , а зависимостью $\beta_{(N)}$ от α и от P при $\alpha < 0,2$ и $P > 0,8$ можно пренебречь и полагать $\beta_{(N)} = \beta/2^{N-1}$, что соответствует $\alpha = 0$ и $P = 1$. Из этого приближенного выражения непосредственно следует, что с увеличением числа проверок величина вероятности ошибки второго рода быстро уменьшается. Так, при $N = 5$ уменьшение $\beta_{(N)}$ по сравнению с β составляет 16 раз, а при $N = 10$ — 512 раз. На рис. 4 приведены рассчитанные по формуле (14) зависимости $\beta_{(N)}$ от β при различных N . Полагалось $\alpha = 0,2$ и $P = 0,8$. Из графиков видно, что с увеличением N величина $\beta_{(N)}$ резко уменьшается и стремится к нулю, при этом влияние параметров α и P на величину $\beta_{(N)}$ незначительно, и приведенные зависимости $\beta_{(N)}$ от β практически совпадают с предельной зависимостью $\beta_{(N)} = \beta/2^{N-1}$.

Уменьшение вероятности ошибки второго рода с увеличением числа проверок оказывает более значительное влияние на повышение достоверности контроля, чем увеличение вероятности ошибки первого рода, по двум причинам. Во-первых, показатель достоверности контроля D_Γ более критичен к величине вероятности ошибки второго рода β , чем к величине вероятности ошибки первого рода α . Во-вторых, при $\beta = 0$ независимо от величины α достигается максимальное значение показателя достоверности контроля $D_\Gamma = 1$, а при многократном контроле с отбраковкой величина $\beta_{(N)}$ уменьшается до значений, близких к нулю.

Полученные выражения (11) и (14) для вероятностей ошибок первого и второго рода $\alpha_{(N)}$ и $\beta_{(N)}$ позволяют рассчитать в соответствии с выражением (5) показатель достоверности КТС ОИ D_Γ после N проверок при проведении однородных испытаний, когда на всех этапах испытаний параметры α и β одинаковы. В случае неоднородных испытаний,



■ Рис. 4. Вероятность ошибки второго рода при многоэтапных испытаниях в зависимости от вероятности ошибки второго рода на отдельном этапе

■ Fig. 4. Probability of type II error in multi-stage tests, depending on the probability of type II error at a particular stage

когда на различных этапах испытаний параметры α и β различаются, показатель достоверности КТС ОИ D_T после N проверок может быть рассчитан только в результате его последовательного определения в соответствии с выражением (5) для каждого этапа. При этом значение показателя достоверности КТС ОИ D_T на предыдущем $(i - 1)$ -м этапе испытаний рассматривается в качестве априорной вероятности пребывания ОИ в исправном состоянии P перед началом анализируемого i -го этапа испытаний, что отражает повышение качества ОИ от одного этапа испытаний к другому. Каждый этап характеризуется вероятностями ошибок первого и второго рода α_i и β_i , $i = 1, 2, \dots, N$. Последовательный расчет значений показателя достоверности D_T от первого этапа к последнему соответствует прямой задаче анализа достоверности, а от последнего этапа к первому — обратной задаче анализа достоверности [11]. При этом совокупность используемых для расчетов зависимостей может быть представлена в виде номограммы [12].

Заключение

Испытания и экспериментальная обработка сложных технических систем представляют собой сложный технологический процесс, реали-

зуемый на основе научных подходов, формализованных в виде программ и методик испытаний, обеспечивающих достижение цели испытаний. Оценивание достоверности КТС ОИ является одной из важных задач, решаемых при планировании испытаний и анализе их результатов, поскольку достоверность КТС ОИ определяет достоверность оценки свойств ОИ по результатам испытаний, характеризующую эффективность испытаний, и, исходя из требуемой достоверности результатов испытаний, осуществляется определение объема испытаний.

В общем случае достоверность КТС сложных технических систем определяется целым рядом факторов, связанных как с ОИ, так и с системой испытаний в целом. Рассмотренный показатель достоверности КТС ОИ — апостериорная вероятность правильного заключения об его исправности — является интегральным показателем, ориентированным на достоверное заключение об исправности объекта испытаний, что является основной целью любых испытаний. Представленные результаты характеризуют свойства данного показателя и определяют принципы его использования при анализе достоверности КТС ОИ при проведении как отдельного этапа испытаний, так и многоэтапных испытаний в целом.

Литература

1. Городецкий В. И., Дмитриев А. К., Марков В. М. Элементы теории испытаний и контроля технических систем. — Л.: Энергия, 1978. — 192 с.
2. Дмитриев А. К., Мальцев П. А. Основы теории построения и контроля сложных систем. — Л.: Энергоатомиздат, 1988. — 192 с.
3. Александровская Л. Н., Круглов В. И., Кузнецов А. Г. Теоретические основы испытаний и экспериментальная обработка сложных технических систем. — М.: Логос, 2003. — 736 с.
4. Афанасьев В. А., Барсуков В. С., Гофин М. Я. Экспериментальная обработка космических летательных аппаратов. — М.: Изд-во МАИ, 1994. — 412 с.
5. Боев С. Ф. Управление рисками проектирования и создания радиолокационных станций дальнего обнаружения. — М.: Изд-во МГТУ им. Н. Э. Баумана, 2017. — 430 с.
6. Северцев Н. А. Надежность сложных систем в эксплуатации и обработке. — М.: Высш. шк., 1989. — 432 с.
7. Дорохов А. Н., Керножицкий В. А., Миронов А. Н., Шестопалова О. Л. Обеспечение надежности сложных технических систем. — СПб.: Лань, 2017. — 352 с.
8. Назаров А. В., Якимов В. Л., Авдеев В. А. Алгоритм максимизации энтропии обучающей выборки и его использование при разработке прогнозной модели дискретных состояний нелинейной динамической системы // Информационно-управляющие системы. 2015. № 2. С. 57–66. doi:10.15217/issn1684-8853.2015.2.57
9. Мальцев Г. Н., Назаров А. В., Якимов В. Л. Имитационное моделирование процесса диагностирования сложной технической системы с высоким уровнем автономности функционирования // Информационно-управляющие системы. 2016. № 4. С. 34–43. doi:10.15217/issn1684-8853.2016.4.34
10. Савин С. К., Никитин А. А., Кравченко В. И. Достоверность контроля сложных радиоэлектронных систем летательных аппаратов. — М.: Машиностроение, 1984. — 168 с.
11. Мальцев Г. Н., Карасев А. С., Гришин В. В. Определение требуемого объема испытаний на отдельном этапе наземной экспериментальной обработки сложного технического изделия по заданным вероятностям ошибок первого и второго рода // Надежность и качество: тр. Второго Междунар. симп. Пенза: ПГУ, 2006. Т. 2. С. 36–39.
12. Калинин А. А., Меньшиков В. А., Мирош Ю. М., Медушевский Л. С. Квадранты качества. Снижение уровня дефектности систем // Электроника: наука, технология, бизнес. 2001. № 4. С. 42–44.

UDC 62.001.4

doi:10.15217/issn1684-8853.2018.1.49

Reliability of Multi-Stage Control over Technical Condition of Tested ObjectsMaltsev G. N.^a, Dr. Sc., Tech., Professor, georgy_maltsev@mail.ruYakimov V. L.^a, PhD, Tech., Associate Professor, yakim78@yandex.ru^aA. F. Mozhaiskii Military Space Academy, 13, Zhdanovskaia St., 197198, Saint-Petersburg, Russian Federation

Introduction: When testing technical systems, the reliability of control over the technical condition of the tested objects must be ensured. Evaluation of this reliability when you plan the test or analyze the test results requires choosing a reliability indicator and methods for its calculation based on test process models. The chosen indicator should ensure a reliable conclusion about the serviceability of the tested object. **Purpose:** Substantiating an integral indicator of the reliability of control over the tested object condition, and methods for calculating it during a separate stage of a test or multi-stage tests. **Results:** Aposterior probability of a correct conclusion about the tested object serviceability can be used as an integral indicator of reliability. In the paper, we present analytical expressions for its calculation at a separate stage of testing or in multi-stage homogeneous tests. The chosen reliability indicator is influenced by the probability of errors of types I and II. **Practical relevance:** The presented results determine the principles of using integral indicators to analyze the reliability of control over the tested object condition when planning the tests or analyzing their results. The given expressions allow you to calculate the integral reliability indicator during a separate stage of a test or in multi-stage tests.

Keywords — Tests, Control of Technical Condition, Validity, Errors of Types I and II.

Citation: Maltsev G. N., Yakimov V. L. Reliability of Multi-Stage Control over Technical Condition of Tested Objects. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 1, pp. 49–57 (In Russian). doi:10.15217/issn1684-8853.2018.1.49

References

1. Gorodetskii V. I., Dmitriev A. K., Markov V. M. *Elementy teorii ispytaniy i kontrolya tekhnicheskikh sistem* [Elements of Technical Systems Testing and Control Theory]. Leningrad, Energia Publ., 1978. 192 p. (In Russian).
2. Dmitriev A. K., Mal'tsev P. A. *Osnovy teorii postroeniia i kontrolya slozhnykh sistem* [Fundamentals of Complex Systems Construction and Control Theory]. Leningrad, Energoatomizdat Publ., 1988. 192 p. (In Russian).
3. Aleksandrovskaya L. N., Kruglov V. I., Kuznetsov A. G. *Teoreticheskie osnovy ispytaniy i eksperimental'naya otrabotka slozhnykh tekhnicheskikh sistem* [Theoretical Bases of Tests and Experimental Testing of Complex Technical Systems]. Moscow, Logos Publ., 2003. 736 p. (In Russian).
4. Afanas'ev V. A., Barsukov V. S., Gofin M. Ia. *Eksperimental'naya otrabotka kosmicheskikh letatel'nykh apparatov* [Experimental Spacecraft Development]. Moscow, MAI Publ., 1994. 412 p. (In Russian).
5. Boev S. F. *Upravlenie riskami proektirovaniia i sozdaniia radioloka-tсионnykh stantsii dal'nego obnaruzheniia* [Risk Management Design and Creation of Early Warning Radar Stations]. Moscow, MGTU im. N. E. Baumana Publ., 2017. 430 p. (In Russian).
6. Severtsev N. A. *Nadezhnost' slozhnykh sistem v ekspluatatsii i otrabotke* [Reliability of Complex Systems in Operation and Development]. Moscow, Vysshiaia shkola Publ., 1989. 432 p. (In Russian).
7. Dorokhov A. N., Kernozhitskii V. A., Mironov A. N., Shestopalova O. L. *Obespechenie nadezhnosti slozhnykh tekhnicheskikh sistem* [Ensuring the Reliability of Complex Technical Systems]. Saint-Petersburg, Lan' Publ., 2017. 352 p. (In Russian).
8. Nazarov A. V., Yakimov V. L., Avdeev V. A. Training Sample Entropy Maximization Algorithm and its Use in Synthesis of Predictive Models of Nonlinear Dynamical System Discrete States. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2015, no. 2, pp. 57–66 (In Russian). doi:10.15217/issn1684-8853.2015.2.57
9. Mal'tsev G. N., Nazarov A. V., Yakimov V. L. Simulation Modeling of Diagnostics of a Highly Autonomous Complex Technical System. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2016, no. 4, pp. 34–43 (In Russian). doi:10.15217/issn1684-8853.2016.4.34
10. Savin S. K., Nikitin A. A., Kravchenko V. I. *Dostovernost' kontrolya slozhnykh radioelektronnykh sistem letatel'nykh apparatov* [Reliability of Aircrafts Complex Radio-Electronic Systems Control]. Moscow, Mashinostroenie Publ., 1984. 168 p. (In Russian).
11. Mal'tsev G. N., Karasev A. S., Grishin V. V. Determination of Tests Required Volume at Ground-Based Experimental Testing Particular Stage of Complex Technical Product Based on Specified Probabilities of First and Second Kind Errors. *Trudy Vtorogo Mezhdunarodnogo simpoziuma "Nadezhnost' i kachestvo"* [Proc. 2th Int. Symp. "Reliability and Quality"]. Penza, PGU Publ., 2006, vol. 2, pp. 36–39 (In Russian).
12. Men'shikov V. A., Mirosh Yu. M., Medushevsky L. S., Kalinin A. A. Quality Quadrants. Reduction in the Level of Systems Defectiveness. *Elektronika: nauka, tekhnologiia, biznes*, 2001, no. 4, pp. 42–44 (In Russian).

КОРРЕЛЯЦИЯ ИНФОРМАЦИИ В SIEM-СИСТЕМАХ НА ОСНОВЕ ГРАФА СВЯЗЕЙ ТИПОВ СОБЫТИЙ

А. В. Федорченко^{а, б}, аспирант, младший научный сотрудник, fedorchenko@comsec.spb.ru

И. В. Котенко^{а, б}, доктор техн. наук, главный научный сотрудник, ivkote@comsec.spb.ru

^аСанкт-Петербургский институт информатики и автоматизации РАН, 14-я линия В. О., 39, Санкт-Петербург, 199178, РФ

^бУниверситет ИТМО, Кронверкский пр., 49, Санкт-Петербург, 197101, РФ

Постановка проблемы: в настоящее время системы управления информацией и событиями безопасности являются активно развивающимся и широко распространенным классом средств мониторинга безопасности различных инфраструктур. Неотъемлемым процессом, реализуемым системами данного класса, является корреляция информации для выявления предупреждений и инцидентов безопасности. С учетом роста различных видов источников исходных данных, а также их количества и сложности взаимосвязей между ними существующие подходы не в состоянии обеспечивать эффективное выполнение процесса корреляции. **Цель исследования:** разработка методики корреляции событий безопасности с автоматизированной адаптацией к анализируемой инфраструктуре, а также создание модели анализа событий безопасности на основе их типов. **Результаты:** разработана модель корреляции для выполнения структурного анализа входных данных, на основе которой производится построение графа типов событий с прямыми и косвенными связями между ними. Сформулированы требования к нормализации исходных данных по наличию равнозначных свойств в форматах типов событий, а также к полноте и временной синхронизации журналов. Приведен пример анализа журнала событий безопасности, а также полученный в результате граф связей типов событий. **Практическая значимость:** предлагаемый подход основан на учете различных свойств типов отношений и связей между ними и позволяет использовать ранее не применяемый метод ранговой корреляции наряду с другими методами интеллектуальной обработки информации, что обеспечивает выполнение процесса корреляции событий и информации безопасности с возможностью адаптации к инфраструктуре.

Ключевые слова — защита информации, системы мониторинга и управления безопасностью, корреляция событий безопасности.

Цитирование: Федорченко А. В., Котенко И. В. Корреляция информации в SIEM-системах на основе графа связей типов событий // Информационно-управляющие системы. 2018. № 1. С. 58–67. doi:10.15217/issn1684-8853.2018.1.58

Citation: Fedorchenko A. V., Kotenko I. V. Correlation of Information in SIEM Systems based on Event Type Relation Graph. *Informatsionno-upravliayushchie sistemy* [Information and Control Systems], 2018, no. 1, pp. 58–67 (In Russian). doi:10.15217/issn1684-8853.2018.1.58

Введение

В настоящее время на мировом рынке присутствует достаточно большое количество разнообразных классов средств по обеспечению безопасности различных инфраструктур. Данные средства направлены на предупреждение, обнаружение и предотвращение компьютерных атак и вредоносной активности, а также мониторинг и управление текущим уровнем безопасности. Одним из классов таких средств являются SIEM-системы (Security Information and Event Management), развивающиеся уже на протяжении более 10 лет [1, 2].

Основная задача SIEM-систем заключается в сборе определенной разнородной информации и выявлении в ней высокоуровневых инцидентов и предупреждений о нарушении безопасности [1–4]. Для достижения указанной цели, как правило, применяются методы нормализации, агрегации, фильтрации и корреляции событий. Однако с ростом сложности проводимых атак (например, целевых атак), защищаемых объектов (напри-

мер, киберфизических систем) и используемых технологий (например, технологии Интернета вещей) применяемые методы и подходы зачастую не в состоянии обеспечивать должный уровень защищенности. Данная тенденция усугубляется возрастающим объемом данных, обработка которых становится все более затруднительной.

Процесс корреляции данных в средствах защиты класса SIEM играет основополагающую роль. Данный процесс в основном направлен на определение причинно-следственных связей между поступающими на обработку событиями. Он позволяет выполнять обнаружение вредоносной и аномальной активности, определение источника и цели атаки, выявление многошаговых атак и зависит от конкретной реализации [5]. Несмотря на разнообразие методов и подходов, применимых в процессе корреляции, на данный момент наиболее распространенным остается правило-ориентированный метод [6–8].

В данной работе представлен подход к корреляции на основе анализа типов событий безопасности в целях определения связей между ними.

Особенностями данного подхода является использование формируемого графа типов событий с прямыми и косвенными связями между собой для выполнения функционального и поведенческого анализа за счет вычисления частотно-временных характеристик событий, определения причинно-следственных связей, ранжирования событий и построения шаблонов поведения.

В работе рассматривается место, роль и общие принципы процесса корреляции, ставится глобальная задача разработки методик корреляции и частная задача построения модели анализа событий безопасности. Рассматриваются отдельные этапы применения предлагаемого подхода. В том числе раскрывается процесс формирования неориентированного графа отношений типов событий, выполняемого на основе анализа типов событий в пределах одного журнала событий, с учетом его нормализованного представления. Рассматриваются особенности использования полученных данных для проведения функционального и поведенческого анализа. Для корректной работы предлагаемого подхода формулируются необходимые и достаточные условия использования исходных (входных) данных. В заключении описывается эксперимент по анализу типов событий журнала безопасности операционной системы (ОС) Windows. Приводятся результаты эксперимента и оценка возможностей применения предлагаемого подхода для корреляции событий безопасности.

Релевантные работы

Корреляция данных первоначально была применена в средствах обнаружения вторжений (Intrusion Detection Systems) для выявления связей между сетевыми событиями с целью их агрегации и последующего детектирования атак (в том числе распределенных и многошаговых) [5]. Именно из систем данного класса методы корреляции были адаптированы для корреляции информации в SIEM-системах.

При общем рассмотрении процесс корреляции может быть разделен на следующие этапы [5]: 1) нормализация; 2) агрегация; 3) фильтрация; 4) анонимизация; 5) приоритезация и 6) собственно корреляция. Наличие и дополнительная декомпозиция данных этапов зависит от конкретной реализации. С нашей точки зрения, каждый из указанных этапов необходим для полноценного выполнения процесса корреляции.

На данный момент существует множество методов корреляции событий безопасности со своими достоинствами и недостатками. На конкретном этапе процесса корреляции целесообразно применять методы, в наибольшей сте-

пени удовлетворяющие его задачам в текущей инфраструктуре. Методики выполнения общего процесса корреляции в существующих решениях, как правило, комбинируют несколько методов. Все методы можно условно разделить на сигнатурные и эвристические. Данные методы могут применять различные алгоритмы, основанные на анализе схожести, статистическом анализе, интеллектуальном анализе данных и др.

Сложность оценки качества используемых методик корреляции данных заключается в том, что производители SIEM-систем в целях защиты интеллектуальной собственности не разглашают особенности технологических решений, применяемых в своих системах. К тому же даже при покупке SIEM-системы исследование модуля корреляции затруднено тем, что его настройка заключается преимущественно в формировании новых (дополнительных) правил и исключений.

Однако наряду с платными решениями SIEM-систем существуют также проекты с открытым исходным кодом, а также множество научных публикаций по методам и подходам к корреляции событий безопасности.

Наиболее популярным и простым при реализации является правило-ориентированный метод [6–8], основанный на фиксированном соотношении событий друг с другом при выполнении определенных условий. Данные условия могут содержать логические операции над данными, их свойствами и вычисляемыми показателями. Главным недостатком данного метода является сложность и длительность составления правил администратором безопасности. Причем эффективность выполнения корреляции правило-ориентированным методом напрямую зависит от квалификации специалиста по внедрению.

Многие методы, такие как шаблонно-ориентированный (сценарно-ориентированный) [6], графо-ориентированный [9, 10], на основе машины конечных состояний [9, 11], на основе схожести [12, 13] и другие, по своей сути имеют различные модели представления событий и их связей, однако в конечном счете они также могут быть выражены в виде правил.

Современным направлением развития методов корреляции событий является применение методов интеллектуального анализа данных, таких как байесовские сети [6, 9, 14], иммунные сети [9, 14], искусственные нейронные сети [9, 14–16] и др. Достоинство данных подходов заключается в возможности самостоятельной (безусловной) корреляции событий с минимизацией ручной настройки. Однако для построения предсказательных моделей на основе интеллектуальных методов требуется предварительный анализ самих данных, который далеко не всегда можно автоматизировать. Кроме того, применение интеллектуальных

подходов определяет требование по оценке адекватности и качества моделей, а исходные данные обучения должны быть достаточно полными.

Подход к корреляции событий

Процесс корреляции

Для постановки задачи исследований следует определить место и роль процесса корреляции в SIEM-системах. Считается, что процесс корреляции направлен на решение следующих задач:

- 1) группировку низкоуровневых событий в более высокоуровневые события;
- 2) определение взаимосвязей между разноуровневыми событиями и информацией безопасности;
- 3) определение важности событий и их групп в рамках задачи обеспечения безопасности;
- 4) обнаружение предупреждений и инцидентов безопасности.

Таким образом, выполнение процесса корреляции начинается со сбора данных из разнородных источников и заканчивается на этапе формирования отчета о текущем состоянии защищенности анализируемых инфраструктур. Стоит также отметить, что процесс корреляции является непрерывным и должен быть рассчитан на выполнение в реальном масштабе времени.

Основная (глобальная) задача проводимых исследований заключается в разработке методики автоматизированной корреляции разнородной информации безопасности. Для достижения данной задачи предлагается использовать результаты структурного, функционального, поведенческого и эволюционного анализа защищаемых объектов [17]. Представленное разделение задачи обусловлено соответствующими аспектами сложности анализируемых инфраструктур как сложных динамических систем [17]. В рамках текущего исследования частной задачей является разработка модели на основе анализа типов событий для выполнения процесса корреляции данных. Новизна предлагаемого подхода заключается в автоматизации поиска прямых и косвенных связей между разнородными событиями для выполнения процесса корреляции. Данный подход должен обеспечить автоматическое включение в модель корреляции типов событий, которые ранее не были известны, но только после приведения данных о событиях к нормализованному виду.

В качестве входных данных процесса корреляции могут выступать различные источники информации (внутренние и внешние): сенсоры (датчики) измерений, агенты сбора данных, журналы событий, конфигурации объектов инфраструктуры и многие другие. В общей схеме



■ **Рис. 1.** Общая схема использования входных данных для процесса корреляции в SIEM-системах
 ■ **Fig. 1.** General scheme of using input data for correlation process in SIEM-systems

использования различных источников данных (рис. 1) приведены базы программно-аппаратного обеспечения ПАО — это хранилища информации об идентификационных характеристиках установленных программно-аппаратных средств в анализируемой инфраструктуре. На данной схеме входные («сырые») данные представлены внутренней информацией с динамичным содержанием и внутренней и внешней информацией с условно-статичным содержанием.

Данное разделение необходимо ввиду сложности корреляции в одном процессе информации из разных категорий, главным различием которых является привязка ко времени (для динамичного содержания). Также следует отметить, что на данном этапе разрабатываемый подход ориентирован преимущественно на входные данные с динамичным содержанием, поскольку любое изменение условно-статичной информации может быть также представлено как событие. Однако данный факт не исключает учет данных с условно-статичным содержанием при анализе состояния безопасности. Также представленная схема включает средства защиты, осуществляющие промежуточную обработку входной информации и генерирующие более высокоуровневые события. Однако связь источников данных с приведенными средствами не фиксирована, т. е. использование того или иного источника конкретным средством зависит от его реализации. Таким образом, исходными данными для выполнения процесса корреляции являются разнородные и разноуровневые события безопасности, что должно быть обязательно учтено при решении глобальной задачи.

Анализ журналов событий

Под событием понимается результат действия (выполненного, отклоненного, завершенного с ошибкой) или попытки совершения действия, генерируемый источником действия или системой его обработки, имеющий predetermined формат описания, понятный системе обработки, а также обладающий специфическими свойствами, описывающими само действие.

События различных типов в пределах одного журнала являются исходными данными для проведения исследований и могут быть выражены как

$$\{e_1, e_2, \dots, e_k\} = E_L, \{t_1, t_2, \dots, t_n\} = T_L,$$

где E — множество событий журнала L , а T — множество типов событий журнала L .

Анализ типов событий предлагается производить на основе реальных исходных данных (журналов событий). В данном случае исключается вероятность возникновения ошибок, связанных

с изменением формата типов, а при наличии такого изменения подобные события будут зафиксированы.

На основе анализа журнала событий производится формирование множества свойств множества типов событий:

$$\{p_1, p_2, \dots, p_m\} = P_T,$$

где P — множество свойств множества типов T .

Свойства событий можно условно разделить на следующие группы:

- *идентификационные* свойства, значения которых для каждого события уникальны в пределах множества событий одного журнала (группы журналов) или системы (например, идентификаторы записей событий);

- свойства *принадлежности*, значения которых указывают на содержание событий в определенных множествах, таких как тип, провайдер, хост;

- *временные* свойства, отражающие значение времени создания, записи, старта, окончания и других временных характеристик действия;

- свойства *аудита*, определяющие результат выполнения действия, который описывает действие как успешное, запрещенное, завершившееся с ошибкой и др.;

- *информационные* свойства, отражающие специфические характеристики действия, описываемого в событии (является наиболее обширной группой атрибутов).

Таким образом, анализ журнала событий (с целью выявления структур типов и их свойств) можно представить как отображение множества событий в множества типов и свойств типов событий:

$$E \rightarrow T \times P.$$

Модель корреляции

на основе анализа типов событий

Поскольку выявленные типы событий определяются свойствами, характеризующими описываемое в событии действие, связи между типами событий с помощью анализа их структур формируются за счет связей между их свойствами. Для определения места структурного анализа в задаче определения связей необходимо ввести классификацию отношений между свойствами типов событий. Выделим *отношения* по равнозначным и неравнозначным свойствам.

Равнозначное свойство p_r — это одинаковое свойство двух различных типов событий t_1 и t_2 :

$$\forall p_r \in P_T : p_r \in T t_1, p_r \in T t_2, \text{ где } t_1, t_2 \in T.$$

В свою очередь отношения по *неравнозначным свойствам* делятся на однотипные и разнотипные.

Однотипные неравнозначные свойства p_1 и p_2 — это свойства, эквивалентные по типу содержимого:

$$p_1, p_2 \in P_T : p_1 \sim p_2.$$

Разнотипные неравнозначные свойства — это свойства, эквивалентные по значениям содержимого при явной разнице между типами содержимого.

Кроме того, один тип событий t может содержать несколько однотипных и разнотипных неравнозначных свойств p в своей структуре:

$$\forall \{p_1, p_2, \dots, p_s\} \subset P_t, p_1 \sim p_2 \sim \dots \sim p_s, t \in T,$$

где s — количество однотипных или разнотипных неравнозначных свойств типа t .

При последующем анализе наличие у различных типов событий равнозначных свойств будет рассматриваться как *прямая связь* между свойствами типов событий, а наличие однотипных и разнотипных неравнозначных свойств — как однотипные и разнотипные *косвенные связи* соответственно. В рамках структурного анализа рассматриваются только прямые связи между типами событий, тогда как функциональный и поведенческий анализы подразумевают определение косвенных однотипных и разнотипных связей соответственно.

Например, при сравнении структур двух типов событий безопасности для ОС Windows «Завершение процесса» (4689) и «Вызвана привилегированная служба» (4673) одним из равнозначных свойств обоих типов является «ProcessId» (процесс-инициатор), что является прямой связью между указанными типами событий.

Событие типа «Запуск процесса» (4688), помимо свойства «NewProcessID», содержит свойство «CreatorProcessID». Оба свойства инициализируют идентификатор процесса, только в первом случае — дочернего процесса (наследника), а во втором — процесса-инициатора (предка). Указанная связь является косвенной однотипной по типу содержимого (тип — «Идентификатор процесса»), поэтому позволяет проследить функциональные связи между событиями разных типов и, в данном случае, выявлять события рабочих сессий процессов и их иерархии наследования.

Событие типа «Запуск процесса» (4688) содержит также свойство «ProcessName». При рассмотрении свойств «ProcessId» и «ProcessName» типы их содержимого явно отличаются: в первом случае — это «Идентификатор процесса», а во втором — «Имя исполняемого модуля в файловой системе». Однако оба свойства описывают идентификационные характеристики процесса. В первом случае данная характеристика имеет привязку ко времени — идентификатор присва-

ивается системой каждому создаваемому процессу и имеет уникальное случайное значение в текущий момент времени в рамках отдельной сессии процесса (от создания до завершения). Во втором случае идентификационная характеристика является статичной и не имеет привязки ко времени. При дальнейшей корреляции указанных неравнозначных свойств разного типа может быть определена косвенная разнотипная связь.

В результате анализа структур типов событий, в соответствии с предлагаемой моделью корреляции, формируется неориентированный граф связей G , используемый при дальнейшем анализе входных данных и выполнении процесса корреляции:

$$G = (T, P, \varphi), \varphi : P \rightarrow T \times T.$$

Следует учитывать, что полученная модель не является окончательной, а наоборот, является первичным представлением отношений между событиями. Далее, после функционального и поведенческого анализа, данная модель корректируется за счет учета силы связей (весов) между типами событий. Вычисление весов ребер полученного графа производится на основе относительных и абсолютных частотных характеристик по принципу схожести значений связующих свойств. В свою очередь направленность связей, выражающая причинно-следственные отношения между типами событий, определяется за счет частотно-временных характеристик с определенной долей вероятности.

Также стоит отметить, что при последующем выделении групп связанных событий, а также определении классов объектов со временем жизни и других сущностей будут получены модели последующих (более высоких) уровней. Под классами объектов со временем жизни понимаются высокоуровневые структуры, определяемые с помощью происходящих событий. Например, ОС, как правило, имеют в качестве неделимого оперативного объекта исполняемый поток. Группа потоков формирует процесс. Процесс в свою очередь может определяться одним или несколькими сервисами. Таким образом, более высокоуровневая модель представления событий должна содержать подобные классы объектов и позволять в дальнейшем связывать между собой данные с динамичным и условно-статичным содержанием.

Статистические методы корреляции

Методы корреляции могут подразделяться на параметрические и непараметрические. В связи с тем, что любое событие характеризуется набором качественных и количественных специфич-

ных свойств, применение параметрических методов (ковариации, линейной корреляции) для определения коэффициента корреляции между событиями затруднено. Однако параметрические методы могут быть использованы для определения линейных зависимостей между свойствами и типами событий. В свою очередь при ранжировании входных данных определенным образом возможно использование непараметрических методов ранговой корреляции. Данные методы предлагается использовать для решения следующих задач: 1) определения связей между свойствами внутри типов событий; 2) вычисления силы и направления связей между типами событий; 3) определения связей между группами событий и 4) поиска взаимосвязей между классами объектов. Первые три задачи реализуют функциональный анализ защищаемой инфраструктуры, а последняя задача необходима для мониторинга поведения отдельных объектов и, следовательно, системы в целом.

Одним из способов ранжирования событий для проведения поведенческого анализа является определение силы связей между типами событий, а также между самими экземплярами событий.

Предлагается выделить следующие виды весов связей:

1) удельные веса прямой, косвенной однотипной и косвенной разнотипной связей между типами событий, задающиеся количеством равнозначных, неравнозначных однотипных и неравнозначных разнотипных свойств соответственно;

2) относительные веса связей между экземплярами событий, определяющиеся отношением количества совпадающих значений свойств к соответствующим удельным весам.

В результате анализа выбранного временного окна в пределах анализируемого журнала формируется набор пар значений относительного веса и интервала времени. Частотный анализ полученных наборов между типами событий, а также применение методов ранговой корреляции позволяют определить причинно-следственные отношения между типами событий и между конкретными экземплярами событий.

Выбор конкретного метода корреляции (например, метода Кендалла, метода Спирмена, метода множественной ранговой корреляции) зависит от этапа, на котором он применяется, и возможности применения исходя из входных данных. Он выбирается экспериментальным путем с учетом эффективности использования.

Эволюционный анализ защищаемой инфраструктуры подразумевает выделение классов типовых элементов (хостов, серверов, ОС, сервисов и др.) и обучение модуля корреляции по вычисленным показателям. В качестве «учителя»

в данном случае предполагается использовать компоненты защиты информации, обнаруживающие текущие угрозы и подающие их в виде высокоуровневых событий. Ввиду вероятного наличия циклов в направленном графе связей типов, групп типов и классов объектов применение байесовских сетей затруднено. Данный факт обусловлен тем, что цикличность связей элементов графа теоретически не имеет возможности устранения:

— удаление маловероятных циклических связей может привести к искажению результата корреляции и пропуску аномальных групп событий;

— упрощение невозможно в связи с использованием низкоуровневых и неделимых (элементарных) событий.

Требования к исходным данным

Предлагаемый подход имеет ряд ограничений на входные данные. Так, предполагается, что перед началом выявления структур типов событий в рамках одной модели формат событий является нормализованным. Нормализация структур преимущественно отражается в следующем условии: структура одного типа события t не должна иметь равнозначных свойств p_1 и p_2 :

$$\forall p_1, p_2 \in P_t : p_1 \neq p_2, t \in T.$$

Данное ограничение необходимо для исключения зацикливания на одном событии в ходе использования предлагаемого подхода. Вместе с тем следует соблюдать нормализованный (однозначный) формат записи свойств событий разных типов.

Стоит отметить, что исходные данные также должны удовлетворять необходимому условию по полноте всевозможных типов событий в рамках рассматриваемой модели и достаточному условию по полноте количества разнотипных событий для выполнения структурного, функционального, поведенческого и эволюционного анализа. Также следует учитывать необходимость соблюдения показателей полноты исходных данных по предоставляемой информации. Данный факт обусловлен условием достаточности анализируемой выборки с помощью метода ранговой корреляции, а также обучаемой выборки. Кроме того, в связи с чувствительностью и привязкой предлагаемого подхода к реальному масштабу времени значения временных свойств событий в рамках одной модели должны быть синхронизированы. Таким образом, для корректного применения предлагаемого подхода к журналу, системе, сегменту или инфраструктуре временные показатели событий должны быть синхронизированы в пределах данного журнала, системы, сегмента и инфраструктуры соответственно.

Определение временного окна журнала событий для выполнения поведенческого анализа следует производить также с учетом требования репрезентативности выборки. Поэтому размер выборки предлагается определять на основе:

1) частотно-временного анализа входных типов событий, использование которого обусловлено периодичностью выполнения ряда процессов и задач в различных системах;

2) учета динамики (частоты) изменения значений равнозначных, неравнозначных однотипных и неравнозначных разнотипных свойств, что обуславливается как случайностью (работой пользователя), так и периодичностью (работой системы) описываемых в журналах действий.

Эксперименты и оценка результатов

В настоящее время анализ структур разных типов событий используется, например, средством аналитического анализа данных безопасности Splunk [18]. В этой системе свойства типов событий применяются для нормализации информации после ее загрузки и последующего индексирования для выполнения задач обработки. В данном случае прямые связи, полученные при анализе исходной информации, могут использоваться в запросах обработки, которые в свою очередь составляются на основе экспертных знаний и их ручной корректировки. Для получения косвенных связей необходимо также проводить дополнительный частотный анализ однотипного содержимого. В указанном решении прямые связи только подразумевают их применение, тогда как в предлагаемом авторами статьи подходе прямые связи берутся за основу модели взаимосвязей между событиями.

В рамках проведенных исследований в качестве исходных данных был проанализирован журнал событий безопасности ОС Windows 8 офисного компьютера, не включенного в локальный домен.

Экспериментальный набор данных обладал следующими характеристиками:

- размер журнала ~7 ГБ в формате XML;
- время записи журнала ~1 мес.;
- время обработки журнала ~50 мин;
- количество событий журнала ~6 700 000;
- количество выявленных типов событий —

80 из 418 заявленных в документации [19] для версии данной ОС (типы событий предыдущей версии не учитывались; число экземпляров событий данной версии составило не больше 20);

— количество выявленных свойств — 158, из них 14 — общие (встречаются во всех типах событий), 53 — уникальные (встречаются только в од-

ном типе событий), 89 — смежные (встречаются более чем в одном типе событий).

В результате анализа представленного журнала был сформирован граф прямых связей типов событий (рис. 2). На рисунке видно, что большинство выявленных типов событий имеют достаточно большое количество прямых связей. Всего граф содержит 1309 узлов. Также присутствуют типы событий, не имеющие прямых связей ни с одним другим типом.

Фрагмент графа типов событий ОС MS Windows 8 с учетом вычисленных удельных весов связей между ними представлен на рис. 3. В уточненном графе типов событий сила связи отображается за счет толщины дуг.

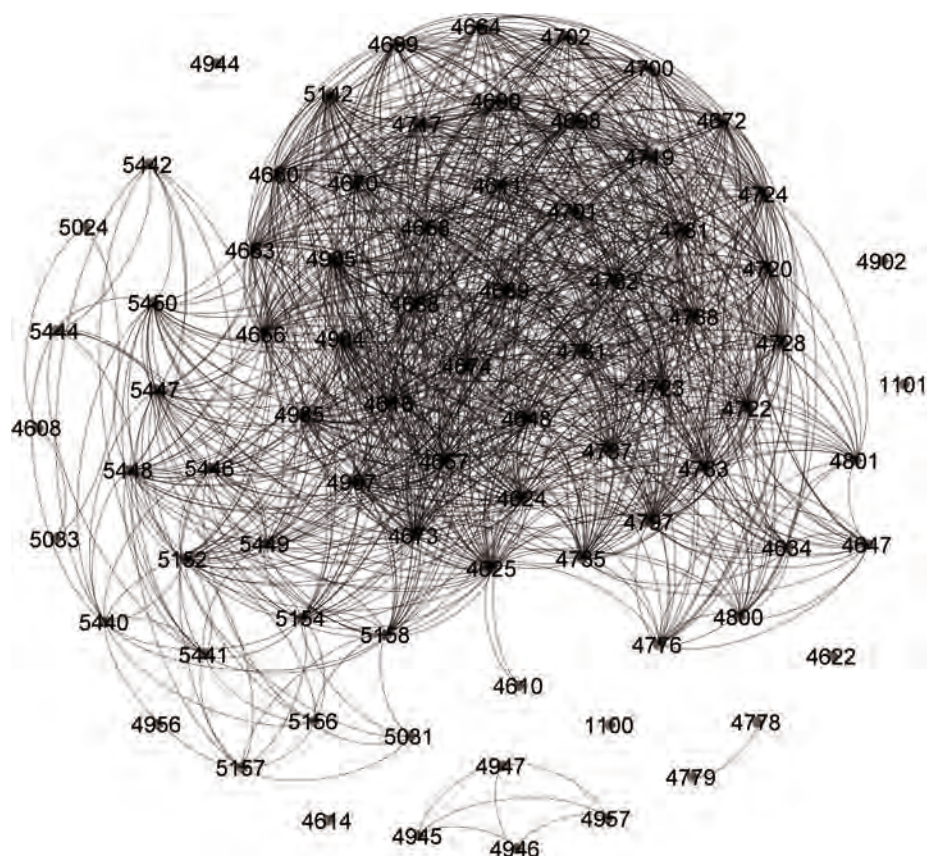
В таблице приведены наиболее распространенные смежные свойства типов событий; количество типов отображает число типов событий, обладающих конкретным свойством.

Предлагаемый подход на основе анализа структур типов событий является составной частью общей методики корреляции событий безопасности. Он может быть применен для формирования модели взаимосвязей (корреляции) событий, которая будет уточняться на каждом последующем шаге методики.

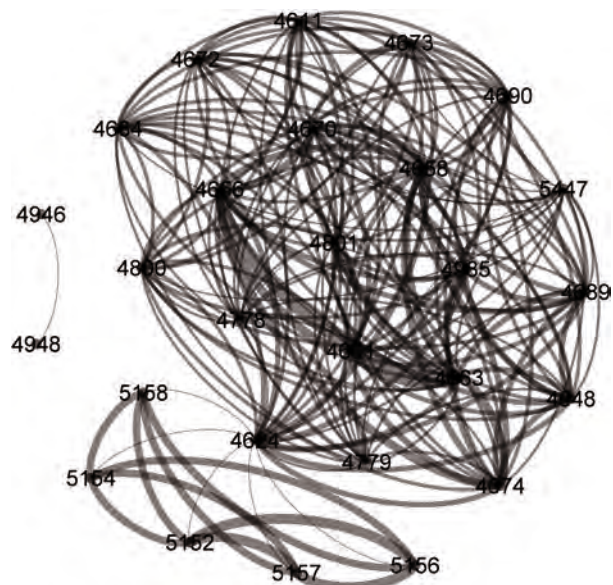
Процесс анализа структур типов событий также может рассматриваться как завершающая стадия процесса нормализации. В свою очередь события, коррелируемые в результате работы методики по определенным свойствам их типов, позволят генерировать события более высоко-

- Результаты анализа распространенности свойств типов событий
- Results of the prevalence analysis of event types properties

Свойство	Количество типов	Свойство	Количество типов
SubjectDomainName	28	ProviderKey	9
SubjectLogonId	28	ProviderName	9
SubjectUserName	28	ObjectServer	8
SubjectUserSid	28	TargetUserName	8
ProcessId	25	HandleId	7
ProcessName	16	TargetDomainName	7
LayerName	9	Application	6



■ *Рис. 2.* Пример графа связей типов событий
 ■ *Fig. 2.* Example of a events types relations graph



■ *Рис. 3.* Пример фрагмента графа типов событий с вычисленными удельными весами связей между ними
 ■ *Fig. 3.* Example of a event types graph fragment with computed the specific weights of the relations between them

го уровня, которые будут содержать составные свойства.

Для обеспечения работы в реальном масштабе времени с учетом требования по полноте исходных данных и теоретически неограниченного размера анализируемой инфраструктуры реализация методики будет нуждаться в применении технологий параллельной и многопоточной обработки больших данных.

Заключение

Проведенные исследования в области корреляции событий безопасности для SIEM-систем показали необходимость разработки новой методики корреляции событий безопасности. На основе анализа общего процесса корреляции и формулировки глобальной задачи в статье предложена модель корреляции на основе анализа типов событий, а также подход к корреляции событий на основе методов ранговой корреляции.

Разрабатываемая методика главным образом направлена на решение таких актуальных задач в области обеспечения безопасности, как монито-

ринг состояния безопасности киберфизических инфраструктур; обнаружение целевых атак (на основе выявления и автоматизированной псевдоклассификации аномалий), а также автоматизированная оценка состояния безопасности инфраструктур неограниченного размера с применением устройств Интернета вещей.

В дальнейшем планируется продолжение разработки общей методики корреляции на основе определения функциональных связей между событиями и построения шаблонов поведения анализируемых инфраструктур.

Работа выполнена при поддержке гранта РФ № 15-11-30029 в СПИИРАН.

Литература

1. **Kotenko I. V., Chechulin A. A.** A Cyber Attack Modeling and Impact Assessment Framework // Proc. of 5th Intern. Conf. on Cyber Conflict 2013 (CyCon 2013). 2013. P. 119–142.
2. **Kotenko I. V., Polubelova O. V., Saenko I. B.** The Ontological Approach for SIEM Data Repository Implementation // IEEE Intern. Conf. on Green Computing and Communications, Conference on Internet of Things, and Conference on Cyber, Physical and Social Computing. 2012. P. 761–766.
3. **Дойникова Е. В., Котенко И. В.** Методики и программный компонент оценки рисков на основе графов атак для систем управления информацией и событиями безопасности // Информационно-управляющие системы. 2016. № 5. С. 54–65. doi:10.15217/issn1684-8853.2016.5.54
4. **Котенко И. В., Дойникова Е. В.** Методика выбора контрмер на основе комплексной системы показателей защищенности в системах управления информацией и событиями безопасности // Информационно-управляющие системы. 2015. № 3. С. 60–69. doi:10.15217/issn1684-8853.2015.3.60
5. **Kruegel C., Valeur F., Vigna G.** Intrusion Detection and Correlation. Challenges and Solutions. — Springer, 2004. — 118 p.
6. **Sadoddin R., Ghorbani A.** Alert Correlation Survey: Framework and Techniques // Proc. of 2006 Intern. Conf. on Privacy, Security and Trust: Bridge the Gap Between PST Technologies and Business Services (PST'06). 2006. Article no. 37.
7. **Hanemann A., Marcu P.** Algorithm Design and Application of Service-Oriented Event Correlation // 3rd IEEE/IFIP Intern. Workshop on Business-Driven IT Management: Proc. of Conf. BDIM 2008. 2008. P. 61–70.
8. **Limmer T., Dressler F.** Survey of Event Correlation Techniques for Attack Detection in Early Warning Systems: Tech report. — University of Erlangen, Dept. of Computer Science 7, 2008. — 37 p.
9. **Muller A.** Event Correlation Engine: Master's Thesis. — Swiss Federal Institute of Technology Zurich, 2009. — 165 p.
10. **Ning P., Xu D.** Correlation Analysis of Intrusion Alerts // Intrusion Detection Systems. Ser.: Advances in Information Security. Springer, 2008. Vol. 38. P. 65–92.
11. **Ghorbani A. A., Lu W., Tavallaee M.** Network Intrusion Detection and Prevention. — Springer, 2010. — 224 p.
12. **Hasan M.** A Conceptual Framework for Network Management Event Correlation and Filtering Systems // Proc. of the Sixth IFIP/IEEE Intern. Symp. on Integrated Network Management. 1999. P. 233–246.
13. **Zurutuza U., Uribeetxeberria R.** Intrusion Detection Alarm Correlation: A Survey // Proc. of IADAT Intern. Conf. on Telecommunications and Computer Networks. 2004. P. 1–3.
14. **Guerer D. W., Khan I., Oglar R., Keffer R.** An Artificial Intelligence Approach to Network Fault Management. — SRI International, 1996. — 10 p.
15. **Tiffany M.** A Survey of Event Correlation Techniques and Related Topics. <http://www.tiffman.com/netman/netman.html> (дата обращения: 14.02.2017).
16. **Elshoush H. T., Osman I. M.** Alert Correlation in Collaborative Intelligent Intrusion Detection Systems — a Survey // Applied Soft Computing. 2011. P. 4349–4365.
17. **Охтилев М. Ю., Соколов Б. В., Юсупов Р. М.** Интеллектуальные технологии мониторинга и управления структурной динамикой сложных технических объектов. — М.: Наука, 2006. — 410 с.
18. **Splunk Security.** https://www.splunk.com/en_us/solutions/solution-areas/security-and-fraud.html (дата обращения: 21.06.2017).
19. **Windows Security Log Events.** <https://www.ultimatewindowssecurity.com/securitylog/encyclopedia/Default.aspx> (дата обращения: 21.06.2017).

UDC 004.056

doi:10.15217/issn1684-8853.2018.1.58

Correlation of Information in SIEM Systems based on Event Type Relation GraphFedorchenko A. V.^{a,b}, Post-Graduate Student, Junior Researcher, fedorchenko@comsec.spb.ruKotenko I. V.^{a,b}, Dr. Sc., Tech., Chief Researcher, ivkote@comsec.spb.ru^aSaint-Petersburg Institute for Informatics and Automation of the RAS, 39, 14 Line, V. O., 199178, Saint-Petersburg, Russian Federation^bITMO University, 49, Kronverkskii Pr., 197101, Saint-Petersburg, Russian Federation

Introduction: Security information and event management systems are now an actively developing and widely distributed class of tools for security monitoring of various computing infrastructures. An essential process implemented by systems of this class is information correlation in order to detect security alarms and incidents. Considering the growth of initial data source types, as well as their number and the complexity of their relationships, the existing approaches cannot provide efficient correlation. **Purpose:** The development of an event correlation technique which automatically adapts to the analyzed infrastructure; in particular, the development of a security event analysis model on the base of event types. **Results:** A correlation model has been developed for structural analysis of the input data. On the basis of this model, a graph of event types is constructed with direct and indirect links between the events. The paper specifies requirements to the initial data normalization concerning the existence of equivalent properties within the formats of the event types, as well as the completeness and temporal synchronization of logs. An example of security events log analysis is provided, along with the resulting graph of the event type links. **Practical relevance:** The proposed approach is based on taking into account various properties of relation types and links between them. It allows you to apply the previously unused method of rank correlation along with other intelligent methods. This provides event correlation and the adaptation to the infrastructure.

Keywords — Information Security, SIEM Systems, Security Event Correlation.

Citation: Fedorchenko A. V., Kotenko I. V. Correlation of Information in SIEM Systems based on Event Type Relation Graph. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 1, pp. 58–67 (In Russian). doi:10.15217/issn1684-8853.2018.1.58

References

1. Kotenko I. V., Chechulin A. A. A Cyber Attack Modeling and Impact Assessment Framework. *Proc. of 5th Intern. Conf. on Cyber Conflict 2013 (CyCon 2013)*, 2013, pp. 119–142.
2. Kotenko I. V., Polubelova O. V., Saenko I. B. The Ontological Approach for SIEM Data Repository Implementation. *IEEE Intern. Conf. on Green Computing and Communications, Conference on Internet of Things, and Conf. on Cyber, Physical and Social Computing*, 2012, pp. 761–766.
3. Doynikova E. V., Kotenko I. V. Techniques and Software Tool for Risk Assessment on the Base of Attack Graphs in Information and Security Event Management Systems. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2016, no. 5, pp. 54–65 (In Russian). doi:10.15217/issn1684-8853.2016.5.54
4. Kotenko I. V., Doynikova E. V. Countermeasure Selection in Security Management Systems. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2015, no. 3, pp. 60–69 (In Russian). doi:10.15217/issn1684-8853.2015.3.60
5. Kruegel C., Valeur F., Vigna G. *Intrusion Detection and Correlation. Challenges and Solutions*. Springer, 2004. 118 p.
6. Sadoddin R., Ghorbani A. Alert Correlation Survey: Framework and Techniques. *Proc. of 2006 Intern. Conf. on Privacy, Security and Trust: Bridge the Gap Between PST Technologies and Business Services (PST'06)*, 2006, article no. 37.
7. Hanemann A., Marcu P. Algorithm Design and Application of Service-Oriented Event Correlation. *Proc. of Conf. BDIM 2008, 3rd IEEE/IFIP Intern. Workshop on Business-Driven IT Management*, 2008, pp. 61–70.
8. Limmer T., Dressler F. *Survey of Event Correlation Techniques for Attack Detection in Early Warning Systems*. Tech report. University of Erlangen, Dept. of Computer Science 7, 2008. 37 p.
9. Muller A. *Event Correlation Engine*. Master's Thesis. Swiss Federal Institute of Technology Zurich, 2009. 165 p.
10. Ning P., Xu D. Correlation Analysis of Intrusion Alerts. *Intrusion Detection Systems. Series Advances in Information Security*, Springer, 2008, vol. 38, pp. 65–92.
11. Ghorbani A. A., Lu W., Tavallaee M. *Network Intrusion Detection and Prevention*. Springer, 2010. 224 p.
12. Hasan M. A Conceptual Framework for Network Management Event Correlation and Filtering Systems. *Proc. of the Sixth IFIP/IEEE Intern. Symp. on Integrated Network Management*, 1999, pp. 233–246.
13. Zurutuza U., Uribeetxeberria R. Intrusion Detection Alarm Correlation: A Survey. *Proc. of IADAT Intern. Conf. on Telecommunications and Computer Networks*, 2004, pp. 1–3.
14. Guerer D. W., Khan I., Ogler R., Keffer R. *An Artificial Intelligence Approach to Network Fault Management*. SRI International, 1996. 10 p.
15. Tiffany M. *A Survey of Event Correlation Techniques and Related Topics*. Available at: <http://www.tiffman.com/netman/netman.html> (accessed 14 February 2017).
16. Elshoush H. T., Osman I. M. Alert Correlation in Collaborative Intelligent Intrusion Detection Systems — a Survey. *Applied Soft Computing*, 2011, pp. 4349–4365.
17. Okhtilev M. Y., Sokolov B. V., Yusupov R. M. Intelligent Technologies for Monitoring and Management of Structural Dynamics of Complex Technical Objects. Moscow, Nauka Publ., 2006. 410 p. (In Russian).
18. *Splunk Security*. Available at: https://www.splunk.com/en_us/solutions/solution-areas/security-and-fraud.html (accessed 21 June 2017).
19. *Windows Security Log Events*. Available at: <https://www.ultimatewindowssecurity.com/securitylog/encyclopedia/Default.aspx> (accessed 21 June 2017).

UDC 621

doi:10.15217/issn1684-8853.2018.1.68

ERROR-CORRECTING CODES FOR TERNARY CONTENT ADDRESSABLE MEMORIES: A NEW PERSPECTIVE

S. Engelberg^a, PhD, Math., Associate Professor and Dean, shlomoe@jct.ac.

O. Keren^b, PhD, Senior Lecturer, Osnat.Keren@biu.ac.il

^aJerusalem College of Technology – Lev Academic Center, 21 Havaad Haleumi, P.O.B. 16031, Jerusalem, 91160, Israel

^bBar-Ilan University, Ramat Gan, 5290002, Israel

Introduction: In the work “Error-Correcting Codes for Ternary Content Addressable Memories”, Krishnan et al. show that under certain assumptions, using $2t + 1$ copies of a word is an optimal strategy for guaranteeing the reliable operation of a ternary content addressable memory in the presence of up to t errors. **Purpose:** To present a new proof of the results of Krishnan et al. about coding for ternary content addressable memories and to extend these results somewhat. **Results:** A new logic-oriented extension of the Hamming distance is presented. Making use of this new distance, an alternate proof that repetition-based coding is optimal over the set of non-context-oriented codes is provided. The new proof allows the results of Krishnan et al. to be extended to cases where some information about the memory organization is available to the code designer. It is shown, for example, that the number of necessary redundancy bits in a non-context-oriented code cannot be reduced by assuming that the memory organizer stores codes in a particularly effective order. **Practical relevance:** The results described in this paper make clear that a repetition code is the optimal code for protecting the data stored in a ternary content addressable memories from errors. The new proof presented in this paper allows the results of Krishnan et al. to be extended to certain cases where some information about the memory organization is available to the code designer.

Keywords — Content Addressable Memories, Ternary Content Addressable Memories, Context-Oriented Codes, Non-Context-Oriented Codes.

Citation: Engelberg S., Keren O. Error-Correcting Codes for Ternary Content Addressable Memories: a New Perspective. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 1, pp. 68–73. doi:10.15217/issn1684-8853.2018.1.68

Overview

Ordinary content addressable memories (CAMs) store completely specified words (i. e., sequences of zeros and ones) [1]. Ternary content addressable memories (TCAMs) store completely specified words or incompletely specified words, words which are sequences of ones, zeros and wildcards (which are denoted by asterisks, “*”)s). Wildcards, *s, match ones and zeros. The input to a TCAM is always a completely specified word. In its standard operating mode, a T/CAM notifies its user if a given word matches one or more completely or incompletely specified words stored in it and, if there is a match, informs the user of the address of the first location whose contents the word matches [2].

As the size of individual memory cells shrinks and the number of cells on a single chip grows, the probability of one or more errors in the contents of the memory increases [3, 4]. When a symbol in the memory is corrupted, it can go from being a zero, one, or wildcard to being any element of that set. Error correcting codes for T/CAMs are usually designed under the assumption that while a T/CAM is operating in its standard operating mode, *there is no simple way to read the contents of the T/CAM*.

There are several ways to make T/CAMs more tolerant of soft errors [5–10]. One of which is to add a sense amplifier (and some additional logic) at the

end of each match line [7]. When the sense amplifier is present, T/CAMs declare a match between the input to the unit and a value stored in the unit as long as the two items differ in t or fewer places. Throughout this paper, it is assumed that T/CAMs with sense amplifiers are being used.

In [11], Krishnan et al. prove that under certain assumptions, a simple repetition code, one that consists of $2t + 1$ copies of a word, is an optimal strategy for guaranteeing the reliable operation of a TCAM in the presence of up to t errors. In this note, a clear distinction is made between context- and non-context-oriented coding and a weaker extension of the Hamming distance is described. Using this new logic-oriented distance, a simpler proof that the central theorem of [11] is optimal over the set of non-context-oriented codes is provided. Using the new proof the results of [11] are then extended to cases where some information about the memory organization is available to the code designer.

Context-Oriented and Non-Context-Oriented Codes

T/CAMs are often used to associate rules with headers. If each header of interest, say h , is associated with a separate set of rules, $R(h)$, then one uses a CAM which searches its memory for the specific

header. Often, however, a *subset* of headers has the same rules associated with each of its elements. In such a case, by using a TCAM and making judicious use of wildcards, it is possible to use many fewer entries to list all the relevant *sets* of headers and their associated rules.

When considering code design and use in TCAMs, we have found it useful to split the task into two sub-tasks: *code design* and *memory organization*. Code design is the design of a code by adding redundancy symbols to headers or sets of headers to form codewords, while memory organization involves taking a (subset of a) code and using it in a way that optimizes the utilization of the memory (and other resources) [12, 13]. In practice, the code designer and the memory organizer are often the same person.

When dealing with code design for TCAMs, we consider two cases: the design of context-oriented and of non-context-oriented codes. In the first case, the code's designer knows something about the set of headers to be stored, and s/he can use this information in designing the codes.

In the second case, the code designer is trying to design a code that will be useful for any possible set of completely- or incompletely-specified headers. The code designer may or may not know something about the general principles the memory organizer will use when storing codewords in the memory, but the code designer *does not know anything at all about the rules associated with the headers the memory organizer will be storing*.

When designing a non-context-oriented code, the code designer must assign a code to each of the 3^k possible *Boolean cubes* that can be used to represent completely- or incompletely-specified headers in the TCAM.

Definition 1 (Boolean cube). A Boolean cube of order w is a vector of length n with w wildcards and $n - w$ elements from $\{0, 1\}$. A Boolean cube of order w represents 2^w distinct binary vectors (and "matches" those vectors).

In what follows, a cube of length k , is referred to as a header; i. e., $h \subseteq \{0, 1, *\}^k$. The term "codeword" is used to refer to a cube, c , of length n , and the input to a TCAM, y , is always a fully specified binary word: $y \in \{0, 1\}^n$. Both headers and codewords are Boolean cubes; in certain contexts the term header is more appropriate, and in others, the term codeword is more suitable.

To fix ideas, first several simple examples of how a memory organizer, who knows which rules are associated with each header, can combine headers are presented and then several code design examples are presented.

Example 1.

Denote by $H \subseteq \{0, 1\}^4$ the set of completely-specified headers of interest without added redundancy. Let

$$H = \left\{ \begin{array}{l} h_1 = (1000), h_2 = (1001), h_3 = (1100) \\ h_4 = (1101), h_5 = (0000) \end{array} \right\}$$

and let the header h_i be associated with the rule $R(h_i)$, $i = 1, \dots, 5$. We consider three scenarios and the ways of combining headers that they lead to.

1. If the rules for h_1, \dots, h_4 are the same and only h_5 has a different rule, then the memory organizer can encode the first four headers as a Boolean cube of order two of the form $h_6 = (1*0*)$ and use a total of two entries.

2. If $R(h_1) = R(h_2)$, $R(h_3) = R(h_4)$, and $R(h_5)$ are three distinct rules, then the memory organizer can encode the headers using two Boolean cubes of order one, $h_7 = (100*)$ and $h_8 = (110*)$, and the single cube $h_5 = (0000)$.

3. If $R(h_1) = R(h_5) \neq R(h_2)$ but $R(h_2) = R(h_3) = R(h_4)$, then the memory organizer can encode the three headers h_2, h_3, h_4 as a cube of order two, $h_6 = (1*0*)$, and the two headers h_1 and h_5 as a cube of order one, $h_9 = (*000)$. S/he must then place h_9 before the cube of order two. As T/CAMs return the *first* address at which a word matches, this *ordering* ensures that the correct rule is associated with h_1 and h_5 .

In order to make the difference between context-oriented codes and non-context-oriented codes clear, the scenarios from Example 1 are now reconsidered from the point of the code designer. As the point of any code is to make different codewords "maximally distant" from one another, it is crucial that the distance between cubes be defined precisely.

Definition 2 (a logic-oriented distance). Let $c_i = (c_{i,n}, \dots, c_{i,2}, c_{i,1})$ and $c_j = (c_{j,n}, \dots, c_{j,2}, c_{j,1})$ be two Boolean cubes. The distance $d(c_i, c_j)$ is defined as

$$d(c_i, c_j) = \left| \left\{ w \mid c_{i,w} \neq c_{j,w}, c_{i,w}, c_{j,w} \neq *, 1 \leq w \leq n \right\} \right|.$$

In what follows: distance will always be taken to mean this logic-oriented distance.

This logic-oriented distance is an extension of the Hamming distance, and when both cubes are of order zero, our logic-oriented distance reverts to the Hamming distance. Two points are particularly worth noting. The logic-oriented definition of distance given here differs from the hardware-oriented distance defined in [11] and is weaker than the distance defined there. For example, using our logic-oriented definition of the distance, $d((000), (***)) = 0$. Using the more hardware-oriented definition in [11], the distance between these elements is 3. (See [7] for more information about this alternate definition of distance.)

Additionally, the logic-oriented extension of the Hamming distance between cubes is not a metric as the triangle inequality does not hold. Consider, for example, $c_1 = (000)$, $c_2 = (110)$, and $c_3 = (**0)$.

In this case $d(c_1, c_2) = 2$, $d(c_1, c_3) = d(c_3, c_2) = 0$, and $d(c_1, c_2) > d(c_1, c_3) + d(c_3, c_2)$.

When a code designer develops a context-oriented code, s/he is assumed to know which codes may be of interest to the memory organizer and which codes are definitely not of interest to the memory organizer. S/he may have additional information as well (such as the order in which the cubes are stored in the memory).

Assume that the code designer (who knows the headers and rules the memory organizer is using) has been told to design a code that can correct a single error. At first glance, it would seem that the code designer must add enough redundancy bits to each header to make the distance between any two codewords at least three, but this is not always necessary.

The intersection of two Boolean cubes c_1 and c_2 , $c_1 \cap c_2$, is defined to be

Definition 3 (intersection of cubes). Let c_1 and c_2 be two cubes of length n , then

$$c_1 \cap c_2 \equiv \{c \in \{0,1\}^n \mid (c \in c_1) \text{ and } (c \in c_2)\}.$$

Cubes are said to be *disjoint* if their intersection is the null set.

It will be shown that when there are non-disjoint codewords it is sometimes possible to design a single-error-correcting code even if the distance between the codewords is less than three. For disjoint codewords, however, we find that:

Theorem 1. *For a (non-context-oriented or context-oriented) code to be t -error correcting, the distance between any two disjoint codewords must be at least $2t + 1$.*

This is so because when two stored entries, h_1 and h_2 , are disjoint, codewords that must match the most corrupted possible version of h_2 — which may have up to t error — must not get “caught” by the corrupted version of h_1 — which may also have up to t errors.

Now the previous scenarios are considered one by one. Redundancy is added to each header in a way that enables the correction of a single error ($t = 1$).

Example 2.

1. In the first case, the redundancy can be added as shown in the following table, and it makes the distance between the cubes ≥ 3 . Here the codewords — composed of the headers with the requisite redundancy symbols — are each six symbols wide.

	header	redundancy
c_6	1*0*	00
c_5	0000	11

2. In the second case, the redundancy can be added as follows.

	header	redundancy
c_7	100*	000
c_8	110*	110
c_5	0000	011

3. In the third and final case, the redundancy can be added as follows.

	header	redundancy
c_9	*000	00
c_6	1*0*	11

Note the difference between the first case and the third case: in the first case $d(c_5, c_6) = 3$, and this allows the input to be correctly located in the presence of a single error. In the third case, however, the distance is only two. In classic coding theory, a distance of two does not provide error correction capability; in a TCAM, it is possible. In the third case, the order of the cubes makes the code single-error correcting. Suppose, for example, that the header and redundancy $y = (100000)$ were entered into a TCAM in which the two entries in the table above had been corrupted to

	header	redundancy
corrupted c_9	*000	10
corrupted c_6	1*0*	10

Because $d(\text{corrupted } c_9, y) \leq 1$, the TCAM returns the rule associated with c_9 . Here, the TCAM works correctly because of the order in which the codewords were stored.

Note that because $(*000) \cap (1*0*) \neq \emptyset$, there are codewords that should be “caught” by both headers. These codewords do not have to be “protected from one another” by the final code, and that is why the final code does not need to maintain a distance of three or greater between different Boolean cubes in order to be single-error correcting.

Because of the fact that the code designer was privy to the headers to be used by the memory organizer, the code designer was able to design relatively efficient codes. As will be seen shortly, when the code designer does not have this foreknowledge, there is no way to achieve such efficiencies.

When building a complete system that employs a TCAM that uses an error-correcting code, it is necessary to implement an encoder that encodes the header being searched for. A simple repetition code turns out to be the optimal code in the non-context-oriented case, and the corresponding encoder has complexity approaching zero. In the case of context-oriented codes, the complexity of the encoder is an open question. If the redundancy were taken to be an arbitrary function of the header, then the encoder would suffer from the “Shannon effect” [14 (Ch. 5), 15], and it would generally require a number of gates that grows exponentially in the number of bits being encoded. At that point, it might be preferable to build a combinatorial circuit that returns the “address” of the header without bothering with a TCAM and error correcting codes.

As our focus is on non-context-oriented codes, we do not consider context-oriented codes further.

Non-Context-Oriented Coding

Constraints Imposed on the Code Designer

When designing a non-context-oriented code, the code designer does not know which headers H contains and certainly does not know the rules to be associated with each header. S/he must, therefore, associate redundancies with each of the 3^k possible headers. S/he must encode the headers in such a way that the distance between any pair of non-intersecting headers h_i and h_j (i. e., cubes for which $h_i \cap h_j = \varnothing$), that the memory organizer might store together in the T/CAM is at least $2t + 1$. In the following, the discussion is restricted to systematic codes — codes for which the information appears in the final coded binary vector in its original form.

The Repetition Code is Optimal

One way of making it possible to correct t errors is to store $2t + 1$ copies of the value to be stored [11] — to use a repetition code. When using a repetition code, if t (or fewer) bits of the stored value get corrupted, the TCAM still recognizes that the stored value matches any value it should have matched, and it does not match any value it should not have matched. Though this scheme looks incredibly wasteful, using our weaker, logic-oriented distance, we provide a proof that there is no error correcting code for a TCAM that is more efficient.

Theorem 2. *The repetition code of length $3k$ and size 3^k is the optimal non-context-oriented single error correcting code for headers of length k .*

Proof: The proof proceeds by producing a sequence that requires at least $3k$ ternary bits to protect against a single error. As the elements of the sequence used generally have many “trailing zeros,” the cube $(abc0\dots0)$ is denoted by C_{abc} .

Because the code designer does not know the elements of H , s/he must encode every possible “header” — whether completely or incompletely specified. In particular, s/he must encode the header $C_0 = (0\dots0)$. Without loss of generality, assume that the header C_0 is associated with a “tail” of redundancy bits that are all zeros. (Any completely specified header will be associated with a completely specified redundancy: using a wildcard as part of the redundancy adds nothing to the distance between the codewords.)

Since the memory organizer may choose to use the header $C_1 = (10\dots0)$ the code has to be able to distinguish between C_0 and C_1 . Thus, C_1 's redundancy must have at least two ones. Denoting the codeword composed of a header $C_{pattern}$ and its associated re-

dundancy by $\tilde{C}_{pattern}$, the distance from \tilde{C}_0 to \tilde{C}_1 is now found to be at least three.

On the other hand, the memory organizer may choose to combine C_0 and C_1 to form $C_* = (*0\dots0)$. As the codeword associated with this header, \tilde{C}_* , must “catch” \tilde{C}_0 and \tilde{C}_1 , \tilde{C}_* 's redundancy must have *s in any place where the redundancies associated with C_0 and C_1 differ. As they differ in at least two places, \tilde{C}_* 's redundancy has at least two *s.

Next consider $C_{01} = (010\dots0)$. Since $C_{01} \cap C_* = \varnothing$, \tilde{C}_{01} has to be distinguishable from \tilde{C}_* . That is, \tilde{C}_{01} 's redundancy must differ by two non-star values from non-star values in each of the preceding redundancies. That brings us to at least four redundancy bits.

The redundancy associated with $C_{**} = (**0\dots0)$ must have *s in any locations where any of the 0-order cubes it contains/covers have redundancy bits that differ from one another. Thus, there must be at least four *s in the redundancy associated with C_{**} . (The redundancies associated with \tilde{C}_0 and \tilde{C}_1 differ from one another in two places, and \tilde{C}_{01} 's redundancy bits differ from each of theirs in two other places.)

Finally, consider $C_{001} = (0010\dots0)$. Since $C_{001} \cap C_{**} = \varnothing$, its associated redundancy bits must differ in at least two non-* locations from those of C_{**} . Thus, its redundancy has at least 6 bits. This pattern continues; for each bit that is added to the header, at least two bits must be added to the redundancy. Thus, any single-error correcting code must have at least $2k$ redundancy bits. It is simple to extend this proof the more general case of a t -error correcting code; at each step rather than adding two bits, one must add $2t$ bits. This proves the following theorem.

Theorem 3. *The repetition code of length $(2t + 1)k$ and size 3^k is the optimal non-context-oriented t -error correcting code for headers of length k .*

A simple consequence of this theorem is that if one needs an ECC for a TCAM, one might as well use the extremely simple repetition code — for no other code can use fewer bits.

Variations on a Theme

We now prove two corollaries of Theorem 3 that concern non-context-oriented codes for which we have some general information about the behaviour of the memory organizer.

From the proof of Theorem 2, it follows that even when the code designer knows that the memory organizer will not use headers that contain one another (for example, that the header (1000) will not be used if the header (**00) is used), the code designer cannot use this knowledge to produce a more efficient non-context-oriented code. Stated more for-

mally, the proof of Theorems 2 and 3 also suffices to prove the following corollary.

Corollary 1. *Knowing that $h_i \not\subset h_j, \forall h_i, h_j, i \neq j$ cannot be used to reduce the number of redundancy bits in a non-context-oriented code.*

Actually, the sequence used in the proofs of Theorem 2 and 3 allows somewhat more to be said.

Corollary 2. *The number of necessary redundancy bits in a non-context-oriented code cannot be reduced by assuming that the memory organizer stores codes in a particularly effective order.*

Proof: Consider the sequence of choices used in the proof of Theorem 2 above and the way they determine the necessary number of redundancy bits. The proof starts by considering \tilde{C}_0 and \tilde{C}_1 . These codes must have a distance of at least three, and as $d(C_0, C_1) \geq 1$, order cannot be used to reduce the necessary distance. Thus \tilde{C}_0 and \tilde{C}_1 's redundancies differ in at least two places. As was seen in the proof, the code associated with C_* must have at least two wildcards in its redundancy. When the code associated with C_{01} is considered, nothing is

known about what the memory organizer is planning. It is possible that the memory organizer will only use \tilde{C}_* and \tilde{C}_{01} . Clearly order cannot help save bits here, and \tilde{C}_{01} 's redundancy must have at least four bits. At each stage in the construction of the sequence in the proof, it is clear that order cannot reduce the number of necessary redundancy bits. Thus, order-related strategies do not allow the number of redundancy bits to be reduced.

Conclusion

Unlike CAMs, TCAMs can store wildcards. This makes them useful when one wants to reduce the size of the memory. Using the new logic-oriented distance, an alternate proof that the $2t + 1$ repetition code is the optimal non-context-oriented code is provided. Moreover, it is shown that knowledge of the order in which codewords will be stored in the TCAM cannot help produce a more efficient non-context-oriented code.

References

- Pagiamtzis K., and Sheikholeslami A. Content-Addressable Memory (CAM) Circuits and Architectures: a Tutorial and Survey. *IEEE Journal of Solid-State Circuits*, March 2006, vol. 41, no. 3, pp. 712–727.
- Arsovski I., Chandler T., and Sheikholeslami A. A Ternary Content-Addressable Memory (TCAM) Based on 4T Static Storage and Including a Current-Race Sensing Scheme. *IEEE Journal of Solid-State Circuits*, Jan. 2003, vol. 38, no. 1, pp. 155–158.
- Seifert N., Gill B., Foley K., Relangi P. Multi-Cell Upset Probabilities of 45 nm High-K Metal Gate SRAM Devices in Terrestrial and Space Environments. *Proc. IEEE Int'l Reliability Physics Symp. (IRPS '08)*, 2008, pp. 181–186.
- Satoh S., Tosaka Y., Wender S. A. Geometric Effect of Multiple-Bit Soft Errors Induced by Cosmic Ray Neutrons on DRAMs. *IEEE Electron Device Letter*, June 2000, vol. 21, no. 6, pp. 310–312.
- Baeg S., Wen S., and Wong R. Minimizing Soft Errors in TCAM Devices: A Probabilistic Approach to Determining Scrubbing Intervals. *IEEE Trans. on Circuits and Systems*, Reg. papers, Apr. 2010, vol. 57, no. 4, pp. 814–822.
- Noda H., Dosaka K., Morishita F., and Arimoto K. A Soft-Error-Immune Maintenance-Free TCAM Architecture with Associated Embedded DRAM. *Proc. IEEE Custom Integrated Circuits Conf.*, 2005, pp. 451–454.
- Pagiamtzis K., Azizi N., and Najm F. N. A Soft-Error Tolerant Content-Addressable Memory (CAM) using an Error-Correcting-Match Scheme. *Proc. IEEE Custom Integrated Circuits Conf. (CICC '06)*, 2006, pp. 301–304.
- Lee H.-J. Immediate Soft Error Detection using Pass Gate Logic for Content Addressable Memory. *Electronics Letters*, 2008, vol. 44, no. 4, pp. 269–270.
- Zhang W. Replication Cache: A Small Fully Associative Cache to Improve Data Cache Reliability. *IEEE Trans. Computers*, Dec. 2005, vol. 54, no. 12, pp. 1547–1555.
- Pontarelli S., Ottavi M., Salsano A. Error Detection and Correction in Content Addressable Memories. *Proc. IEEE 25th Int'l Symp. "Defect and Fault Tolerance in VLSI Systems" (DFT '10)*, Oct. 2010, pp. 420–428.
- Krishnan S. C., Panigrahy R., and Parthasarathy S. Error-Correcting Codes for Ternary Content Addressable Memories. *IEEE Trans. on Comp.*, 2009, vol. 58, no. 2, pp. 275–279.
- Karthik Lakshminarayanan, Anand Rangarajan, and Srinivasan Venkatachary. Algorithms for Advanced Packet Classification with Ternary CAMs. *SIGCOMM Comput. Commun. Rev.*, Aug. 2005, vol. 35, no. 4, pp. 193–204.
- Alex X. Liu, Chad R. Meiners, and Eric Torng. TCAM Razor: A Systematic Approach Towards Minimizing Packet Classifiers in TCAMs. *IEEE/ACM Trans. Netw.*, April 2010, vol. 18, no. 2, pp. 490–500.
- Karpovsky M. G., Stankovic R. S., Astola J. T. *Spectral Logic and its Applications for the Design of Digital Devices*. John Wiley & Sons, 2008. 598 p.
- Wegener I. *The Complexity of Boolean Functions*. New York, John Wiley & Sons, 1987. 458 p.

УДК 621

doi:10.15217/issn1684-8853.2018.1.68

Использование помехоустойчивых кодов в системах троичной ассоциативной памяти: новые перспективыЭнгельберг Ш.^а, PhD, мат., доцент, декан, shlomoe@jct.ac.Керен О.^б, PhD, старший преподаватель, Osnat.Keren@biu.ac.il^аИерусалимский технологический колледж, Хавад Халейми, 21, П.О.Б. 16031, Иерусалим, 91160, Израиль^бУниверситет Бар-Илан, Рамат-Ган, 5290002, Израиль

Введение: в статье «Использование помехоустойчивых кодов в системах троичной ассоциативной памяти» Кришнан с соавторами показал, что при определенных предположениях использование $2t + 1$ копий слов является оптимальной стратегией, гарантирующей реализацию троичных контентных адресных элементов памяти при наличии до t ошибок. **Цель:** вывести новое доказательство результатов Кришнана с соавторами о кодировании для троичных контентно-адресных элементов памяти в целях расширения этих результатов для возможных случаев организации памяти. **Результаты:** представлено новое логически-ориентированное расширение расстояния Хэмминга, благодаря которому предложено альтернативное доказательство того, что кодирование, основанное на повторении, является оптимальным, охватывающим серии не контекстно-ориентированных кодов. Новое доказательство также позволяет расширить результаты Кришнана с соавторами на случаи, когда некая информация об организации памяти имеется в наличии у разработчика кодов. Показано, например, что число необходимых для репродукции битов в не контекстно-ориентированном коде не может быть выкинуто (опущено), если организатор памяти сохраняет коды в частично-эффективном порядке. **Практическая значимость:** результаты данной работы ясно показывают, что повторный код является оптимальным кодом для защиты от ошибок для информации, хранимой в троичных контекстно-адресных элементах памяти. Новое доказательство, предложенное в статье, позволяет расширить результаты Кришнана с соавторами на ряд случаев, когда некоторая информация об организации памяти находится в распоряжении разработчика кодов.

Ключевые слова — контентно-адресные элементы памяти, троичные контекстные адресные элементы памяти, не контекстно-ориентированные коды.

Цитирование: Engelberg S., Keren O. Error-Correcting Codes for Ternary Content Addressable Memories: a New Perspective // Информационно-управляющие системы. 2018. № 1. С. 68–73. doi:10.15217/issn1684-8853.2018.1.68

Citation: Engelberg S., Keren O. Error-Correcting Codes for Ternary Content Addressable Memories: a New Perspective. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 1, pp. 68–73. doi:10.15217/issn1684-8853.2018.1.68

УВАЖАЕМЫЕ АВТОРЫ!

Научная электронная библиотека (НЭБ) продолжает работу по реализации проекта SCIENCE INDEX. После того как Вы регистрируетесь 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.

UDC 621.371

doi:10.15217/issn1684-8853.2018.1.74

IMPACT OF BUILT-UP TERRAIN ON OPERATIONAL PARAMETERS OF SIGNALS IN LAND-SATELLITE COMMUNICATION LINKS

N. S. Blaunstein^a, Dr. Sc., Phys.-Math., Professor, nathan.blaunstein@hotmail.com

M. B. Sergeev^b, Dr. Sc., Tech., Professor, mbse@mail.ru

V. A. Nenashev^b, PhD, Tech., Assistant Professor, nenashev.va@gmail.com

^aBen-Gurion University of the Negev, P.O.B. 653, 1, Ben-Gurion St., Beer-Sheva, 74105, Israel

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

Purpose: Examining the propagation of radio signals in urban environments for different elevations of a moving satellite antenna and stationary or mobile subscribers' antennas located below or above the rooftops, based on a stochastic multi-parametric approach and concentrating on temporal and spatial properties of signal power variations in order to predict the channel pass loss and fading effects. **Results:** In the framework of the stochastic approach proposed in this work, the total signal power, the corresponding probability density functions, and the scales of the temporal and frequency characteristics were analyzed. All the predicted characteristics were examined numerically. The impact of buildings' density and satellite antenna elevation upon the above-mentioned characteristics of the radio signal were examined through the lens of special experiments performed by various research groups worldwide. **Practical relevance:** As the computed and experimentally observed data give a good agreement between the proposed stochastic approach and the approaches proposed by other researchers, it can be used to predict link budget for various built-up terrain environments and for various elevations of moving satellite antennas.

Keywords — Land-Satellite Link, Stochastic Multi-Parametric Model, Buildings' Density and Overlay Profile, Multiplicative Fading Probability, Pass Loss, Link Budget.

Citation: Blaunstein N. S., Sergeev M. B., Nenashev V. A. Impact of Built-up Terrain on Operational Parameters of Signals in Land-Satellite Communication Links. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 1, pp. 74–84. doi:10.15217/issn1684-8853.2018.1.74

Introduction

From the middle of the 20th century worldwide researchers tried to find the adequate models that could predict operational parameters, pass loss, fading and link budget, in the land-satellite communication (LSC) links where the built-up profile, density of buildings and other obstructions, as well as the terrain structure and contours affect signals passing from the moving satellite and ground-located subscriber. Where proposed several approaches that we briefly will present below to understand the matter of the problem under studying. In our description of the matter we follow results discussed in the references [1–27].

Statistical Models

These models correspond to cases for which multipath and line-of-sight are present simultaneously [6–13, 16, 19]. In this section we will describe only two models; Loo's [15, 17] and Lutz [16], which have been used in [10, 11, 22, 23] for designing the unified algorithm for predicting fading phenomena in the land-satellite links. Therefore, we will describe briefly these models and compare their results with those obtained in [10, 11, 22, 23] and with those obtained on the basis of the multi-parametric stochas-

tic approach proposed for the land-satellite links in [20] for the special series of experiments described in [11, 25–27].

We also refer the reader to some other statistical models by Corazza — Vatalaro [19], Xia — Fang [20], Abdi [21], and the three-state model [27], which are lesser used by designers of such land-satellite links.

Loo's Model. This model is a statistical model [15, 17] for a land mobile-satellite link with applications to rural environments. The model assumes that the amplitude of the line-of-sight (LOS) component under foliage attenuation is distributed according to the lognormal probability density function (PDF) and the received multipath component is described by a Rayleigh PDF. The model is statistically described in terms of its PDF or cumulative distribution function (CDF), which were obtained under the hypothesis that foliage not only attenuates but also scatters radio waves. In such assumptions, the total complex fading signal is the sum of a lognormally distributed random signal and a Rayleigh signal [15, 17]:

$$r \exp(j\theta) = z \cdot \exp(j\varphi_0) + w \cdot \exp(j\varphi), \quad z > 0, w > 0, \quad (1)$$

where the phase φ_0 and φ are uniformly distributed between 0 and 2π ; z is lognormal distributed

amplitude, and w is a Rayleigh distributed amplitude. If z is temporally kept constant, it can be assumed that the PDF $p(z)$ is lognormal. The signal random envelope r is lognormal distributed for large values and Rayleigh distributed for small values [15, 17]:

$$p(r) \approx \begin{cases} \frac{1}{r\sqrt{2\pi d_0}} \exp\left[-\frac{(\ln r - \mu)^2}{2d_0}\right] & \text{for } r \gg \sqrt{d_0} \\ \frac{r}{b_0} \exp\left[-\frac{r^2}{2b_0}\right] & \text{for } r \ll \sqrt{b_0} \end{cases} \quad (2)$$

In this equation $\sqrt{d_0}$ and μ are the standard deviation and mean for the lognormal distribution, and b_0 is the variance for the Rayleigh distribution, respectively. The parameter b_0 represents also the average scattered power caused by multipath effects.

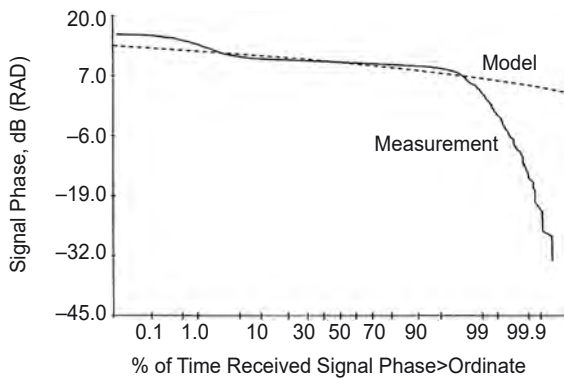
Many calculations with different values for b_0 , d_0 and μ where carried out by Loo with the objective of fitting the results of his model to those derived from measurements made on simulated satellite paths. The measurements site was a rural area with about 35 % tree coverage. The model parameters were obtained by trial and error to fit measured values.

Computational results for the signal envelope, based on Loo's model, we compare to measurements obtained in [15, 17].

As was mentioned in [8], the signal envelope PDF of the model facilitates the calculation of fade margins in the design of communications systems. As for the signal envelope phase distribution, Fig. 1 shows a comparison of the complementary cumulative distribution function (CCDF) for the received signal phase calculated using the well-known equation [8]:

$$CCDF(r) \equiv p(r > R) = 1 - CCD(r) = 1 - \int_0^R p(r) dr, \quad (3)$$

where R is either maximum accepted path loss or noise floor figure of the system.



■ Fig. 1. Loo's phase model and measurement for heavy shadowing at $f = 18.925$ GHz

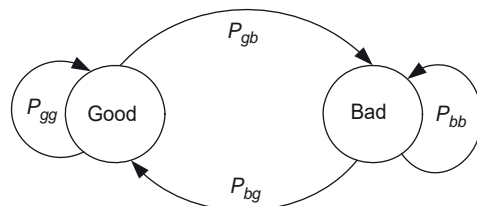
We must note that for the case of infrequent light shadowing, the model shows the best fit around the median region and some deviation near the tails of the distribution. As for the heavy shadowing situation is changed (see Fig. 1), the results indicate that the model shows a correlation between the rate of change of the envelope due to multipath and foliage attenuation for heavy shadowing too excluding only the higher probability of the event. The disadvantage of this model is that the measurements were made up to 30°. Model parameters for higher elevation angles are not available.

Lutz Statistical Model. In this model, which can be considered as a generalization of the Loo's model, the simple statistics of LOS and NLOS (non-line-of-sight) are modeled by two distinct states, *good* and *bad*, called the *Markov's chain*, as shown in Fig. 2. This is appropriate for describing the propagation situation in urban and sub-urban areas where there is a large difference between the shadowed and non-shadowed statistics. The parameters associated with each state and the transition probabilities for evolution between states are empirically derived. The LOS condition is represented by a good state, and the NLOS condition by a bad state. In the good state, the signal is assumed to be Ricean K -factor distributed [8], which depends on the satellite elevation angle and the carrier frequency, so that the PDF of the signal amplitude is given by $P_{good} = P_{Rice}$. In the bad state, the fading statistics of the signal amplitude are assumed to be Rayleigh, with a mean power $S_0 = \sigma^2$, which varies with time.

So the PDF of amplitude is specified as the conditional distribution $p_{Rayleigh}(S|S_0)$, where S_0 varies slowly with a lognormal distribution $p_{LN}(S_0)$, representing the varying effects of shadowing with the NLOS component. For detailed formulas, we refer the reader to the original [8].

Transitions between states are described by a first-order Markov chain. This is a state transition system, in which the transition from one state to another depends only on the current state rather than on any more distant history of the system. The transition probabilities, which summarize all models based on the Markov chain approach, are (see Fig. 2 [8]):

- from good state to good state P_{gg} ;
- from good state to bad state P_{gb} ;



■ Fig. 2. Markov's model of channel states

from bad state to bad state P_{bb} ;
 from bad state to good state P_{bg} .

For a digital communication system, each state transition is taken to represent the transition of one symbol. The transition probabilities can then be found in terms of the mean number of symbol duration spent in each state [8]:

$$P_{gb} = \frac{1}{D_g} \text{ where } D_g \text{ is the mean number of symbol duration in the good state;}$$

$$P_{bg} = \frac{1}{D_b} \text{ where } D_b \text{ is the mean number of symbol duration in the bad state.}$$

The sum of the probabilities leading from any state must equal to the sum of

$$P_{gg} = 1 - P_{gb} \text{ and } P_{bb} = 1 - P_{bg}. \quad (4)$$

The time share of shadowing (the proportion of a symbol in the bad state) is

$$A = \frac{D_b}{D_g + D_b}. \quad (5)$$

Below we will use this formula to find the parameter A , denoted as *the time share of shadowing*, during comparison with physical-statistical and the stochastic multiparametric approaches, where this parameter is derived in other manner. In this comparison we will use the Lutz model as a classical statistical approach.

Physical-Statistical Models

In pure statistical models, the input data and computational effort are quite simple, as the model parameters are fitted to measured data. Such models only apply to hypothetical environments and lack the physical background of the realistic problem. On the other hand, pure deterministic physical models provide high accuracy, but they require actual analytical path profiles and time-consuming computations.

A combination of both approaches has been developed by the authors. The general method relates any channel simulation to the statistical distribution of physical parameters, such as building height, width and spacing, street width or elevation and azimuth angles of the satellite link. This approach is henceforth referred to as the *“Physical-Statistical”* approach [10, 11, 22, 23]. The main concept of such an approach is the following. According to the *physical* approach, input knowledge consists of electromagnetic theory and a full physical understanding of the propagation processes. However, this knowledge is then used to analyze a *statistical*

input data set, yielding a *distribution* of the output predictions. The output predictions are not linked to specific locations. Physical-statistical approach therefore require only simple input data such as distribution parameters (e. g., mean building height and building height variance).

This model describes the geometry of mobile satellite propagation in built-up areas and proposes statistical distributions of building heights, which are used in subsequent analysis. We will consider only two these models:

1) a shadowing model based on the two-state channel Lutz model;

2) a multi-parametric stochastic model.

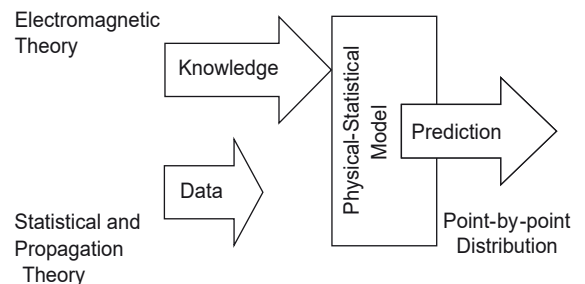
The Model of Shadowing. The geometry of the situation, which was analyzed in [10, 11, 22, 23] by Saunders and Evans with their colleagues, is illustrated in Fig. 3.

It describes a situation where a mobile is situated along a straight street with the direct ray from the satellite impinging on the mobile from an arbitrary direction. The street has buildings, line-up on both sides, with randomly varying heights. In the presented model, the statistics of building height in typical built-up areas is used as input data. A suitable form was sought by comparing with geographical data for the cities of Westminster and Guildford, UK [11, 22, 23]. The PDFs that were selected to fit the data are the log-normal and Rayleigh distributions with unknown parameters the mean value, m , and standard deviation, σ_b . The PDF for the log-normal distribution is [10, 22, 23]

$$p_b(h_b) = \frac{1}{h_b \sqrt{2\pi} \sigma_b} \cdot e^{-\frac{\ln^2(h_b/m)}{2\sigma_b^2}}. \quad (6)$$

The PDF for the Rayleigh distribution is fully described in [8]. We will repeat it using notations made in references [10, 22, 23]:

$$p_b(h_b) = \frac{h_b}{\sigma_b^2} \cdot e^{-\frac{h_b^2}{2\sigma_b^2}}. \quad (7)$$



■ Fig. 3. Algorithm of physical-statistical model

■ **Table 1.** Best-fit parameters for the theoretical PDFs

City	Lognormal PDF		Rayleigh PDF
	Mean m	Standard Deviation	Standard Deviation
Westminster	20.6	0.44	17.6
Guildford	7.1	0.27	6.4

To find the appropriate parameters for these functions in order to fit the data measurements as accurately as possible, the PDF was found by minimizing the maximum difference between the two cumulative distribution functions. The parameters for each PDF are quoted in Table 1 from [10, 22, 23], where all parameters are in meters.

The direct ray is judged to be shadowed when the building height h_b exceeds some threshold height h_T relative to the direct ray height h_s (Fig. 4). The shadowing probability, P_s , can then be expressed in terms of the PDF of the building height, $p_b(h_b)$ as [10, 22, 23]

$$P_s = \Pr(h_b > h_T) = \int_{h_T}^{\infty} p_b(h_b) dh_b. \quad (8)$$

The definition of h_T is obtained by considering shadowing to occur exactly when the direct ray is geometrically blocked by the building face. Using a simple geometry, the following expression is extracted for h_T [10, 22]:

$$h_T = h_r = \begin{cases} h_m + \frac{d_m \tan \varphi}{\sin \theta} & \text{for } 0 < \theta < \pi \\ h_m + \frac{(w - d_m) \tan \varphi}{\sin \theta} & \text{for } -\pi < \theta < 0 \end{cases}. \quad (9)$$

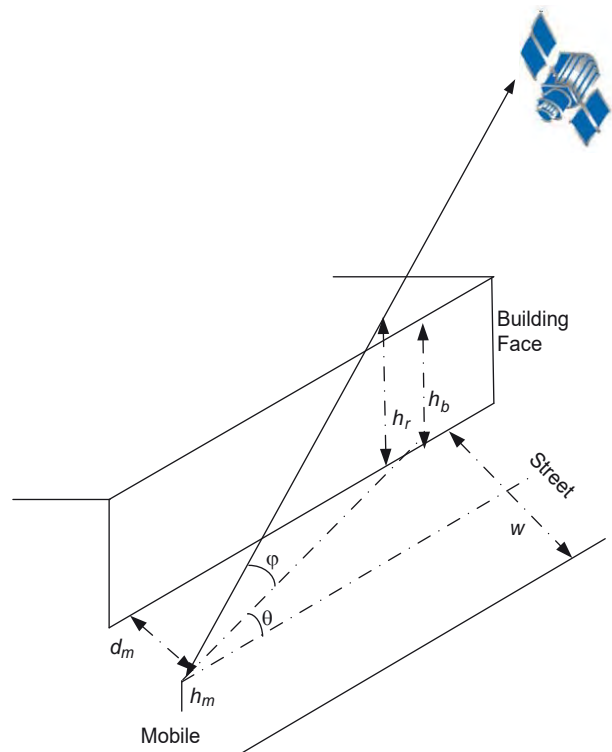
All notations and geometrical parameters in (9) are explained in Fig. 4.

The shadowing model estimates the probability of shadowing for Lutz two-state model [24]. The same Markov chain, shown in Fig. 2, is used, but parameters A , P_{bad} and P_{good} are obtained from actual random distribution of the obstructions above the terrain. Thus,

$$A = \int_{z_1}^{z_2} P_b(h) dh, \quad (10)$$

where h are different heights of obstacles, z_1 and z_2 are the minimum and maximum height of the built-up layer;

$$P_b = \begin{cases} \text{lognormal} + \text{Ricean} \\ \text{lognormal} + \text{Rayleigh} \end{cases}, \quad (11)$$


 ■ **Fig. 4.** Geometry for mobile-satellite communication in built-up areas

where the lognormal PDF is for pure NLOS shadowing. The Ricean PDF describes both the LOS and the multipath component, and the Rayleigh PDF describes the multipath component of the total signal, when the LOS component is absent [8].

Now, using theoretical results obtained by Lutz's pure statistical model and results the Saunders and Evans model, we can combine them into one unified model. This unified model will be compared with the stochastic multiparametric model described in the next section based on the theoretical framework analyzed in [8]. Thus, taking into account the Markov's chain (see Fig. 2), we consider the *bad state* by using the Rayleigh PDF and the *good state* by using the Ricean PDF, as well as shadowing by using the lognormal PDF. By combining all these PDFs in a Markov's chain, we finally can obtain the total PDF that describes the effects of different kinds of fading occurring within the LSC link, caused by terrain obstructions; natural and man-made. As a result we get

$$p(S) = (1 - A) \cdot P_{good} + A \cdot P_{bad} = (1 - A) \cdot P_{Rice}(S) + A \int_0^{\infty} p_{Rayleigh}(S | S_0) \cdot p_{lognorm}(S_0) dS_0. \quad (12)$$

Then, we introduce the corresponding CCDF, which describes the signal stability, being the received signal with amplitude r that prevails upon the maximum accepted path loss, R , in the mul-

tipath channel, caused by fading phenomena. This can be presented in the following form:

$$CCDF = \Pr(r > R) = \int_0^R p(S) dS. \quad (13)$$

All above formulas allow us to present the unified algorithm for fading phenomena estimation in LSC links, both stationary and mobile.

Multi-Parametric Stochastic Approach

As an example of a physical-statistical model, we present the same stochastic approach which was used successfully for land communication channels, rural, sub-urban and urban, and was described in [1–8]. The reason for that is based on the fact that the previous physical-statistical model, as was shown in [10, 11, 22, 23], predicts more strictly the fading effects in different LSC links compared to the pure statistical models [15–21].

At the same time, as was mentioned in [8], the physical-statistical model, which is based on a deterministic distribution of the local built-up geometry introduced in [22, 23], cannot strictly predict any situation when a satellite, moving around the world, has different elevation angles θ_i with respect to a subscriber located at the ground surface, as shown in Fig. 5. As the result, the radio path between the desired subscriber and the satellite crosses different overlay profiles of the buildings because of continuously changing elevation angle of the satellite during its rotation around the Earth with respect to the ground-based subscriber anten-

na. To predict continuously the outage probability of shadowing in real time, a huge amount of data is needed regarding each building, geometry of each radio path between desired user and the satellite during its rotation around the Earth, and finally high-speed powerful computer.

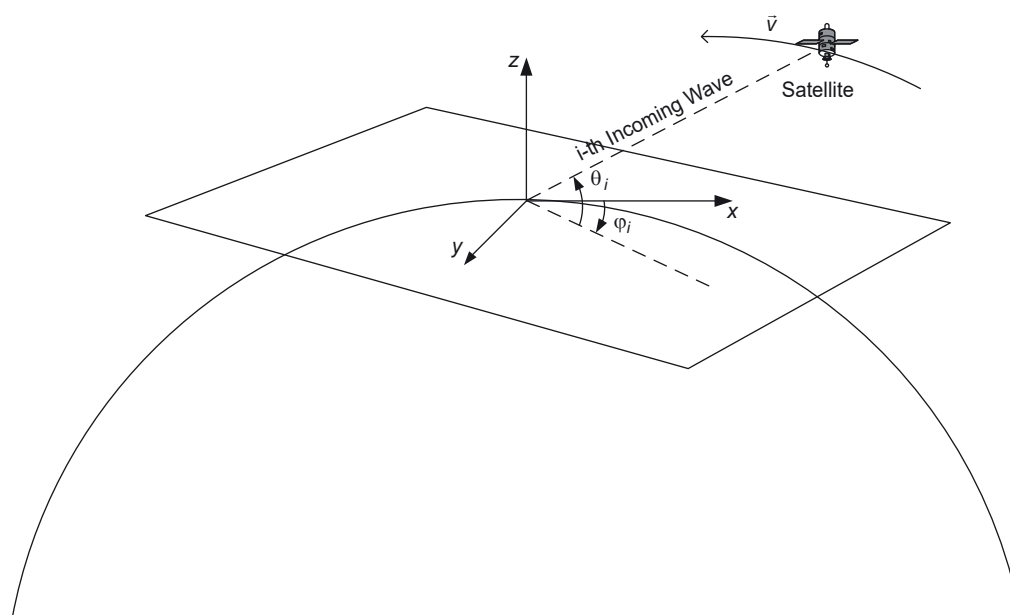
Most of these difficulties can be handled by using the multi-parametric stochastic model [1–8]. Following [8], we take into account both, the buildings distribution at the ground surface and their height profile changing in the vertical plane that is, accounting for the 3D stochastic model of multiple scattering, reflection and diffraction rearranging the corresponding formulas in the case of the LSC link.

Buildings' Overlay Profile. The LSC link is very sensitive to the overlay profile of the buildings, because during its movement around the Earth, depending on the elevation angle φ , the buildings' profile will be continuously changed leading to different effects of shadowing in the current communication link (Fig. 6).

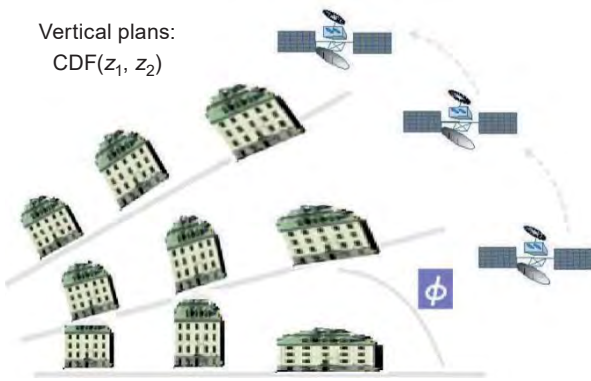
Taking into account the fact that real profiles of urban environment are randomly distributed, the probability function $P_h(z)$, which describes the overlay profile of the buildings, can be presented in the following form [6, 20, 39]:

$$P_b(z) = H(h_1 - z) + H(z - h_1) \cdot H(h_2 - z) \times \left[\frac{(h_2 - z)}{(h_2 - h_1)} \right]^n, \quad n > 0, \quad 0 < z < h_2, \quad (14)$$

where the function $H(x)$ is the Heaviside step function, which is equal 1, if $x > 0$, and is equal 0, if $x < 0$.



■ Fig. 5. General geometry of the land-satellite link



■ Fig. 6. Change of the profile function $F(z_1, z_2)$ in the vertical plane during the movements of a satellite

The graph of this function versus height z of a built-up overlay is presented in Fig. 7. For $n \gg 1$ $P_b(z)$ describes the case where buildings higher than h_1 (minimum level) very rarely exist. The case, where all buildings have heights close to h_2 (maximum level of the built-up layer), is given by $n \ll 1$. For n close to zero, or n approaching infinity, most buildings have approximately the same level that equals h_2 or h_1 , respectively. For $n = 1$, we have the case of building height uniformly distributed in the range of h_1 to h_2 .

The average buildings height \bar{h} , can be found as [8]

$$\bar{h} = h_2 - n \frac{(h_2 - h_1)}{n + 1}. \quad (15)$$

Then, analyzing the built-up layer profile $F(z_1, z_2)$, “illuminating” by the terminal antennas of the heights z_1 and z_2 , we get :

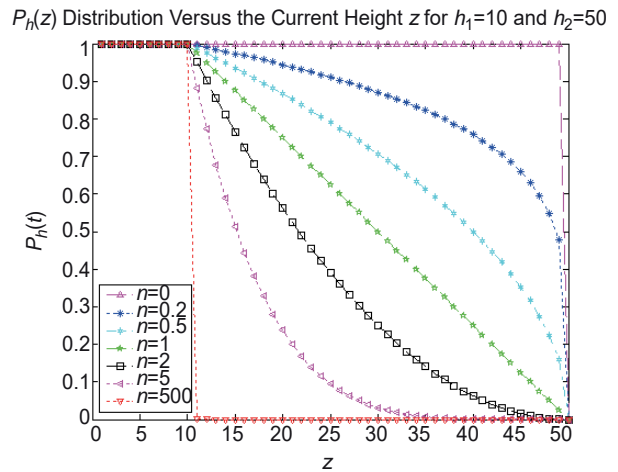
— for the case when the antenna height is above the rooftop level, that is, $z_2 > h_2 > h_1$:

$$F(z_1, z_2) = H(h_1 - z_1) \left[(h_1 - z_1) + \frac{(h_2 - h_1)}{(n + 1)} \right] + H(z_1 - h_1) H(h_2 - z_1) \frac{(h_2 - z_1)^{n+1}}{(n + 1)(h_2 - h_1)^n}; \quad (16a)$$

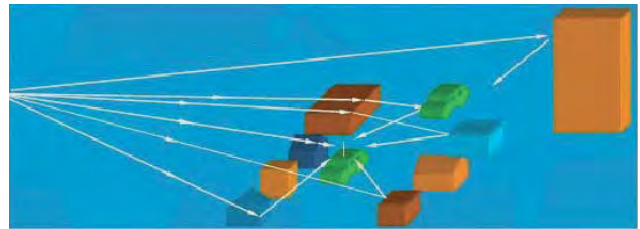
— for the case when the antenna height is below the rooftop level, that is, $z_2 < h_2$:

$$F(z_1, z_2) = H(h_1 - z_1) \times \left[(h_1 - z_1) + \frac{(h_2 - h_1)^{n+1} - (h_2 - z_2)^{n+1}}{(n + 1)(h_2 - h_1)^n} \right] + H(z_1 - h_1) H(h_2 - z_1) \frac{(h_2 - h_1)^{n+1} - (h_2 - z_2)^{n+1}}{(n + 1)(h_2 - h_1)^n}. \quad (16b)$$

Then, the cumulative distribution probability function (CDF) of the event that any subscriber



■ Fig. 7. Buildings’ overlay profile PDF



■ Fig. 8. Horizontal map of multipath phenomenon occurring in the urban scene

located in the built-up layer is affected by obstructions due to shadowing effect can be presented as [1–8]

$$CDF(z_1, z_2, n) = \frac{1}{z_2 - z_1} \int_{z_1}^{z_2} P_h(z) dz \equiv \frac{1}{z_2 - z_1} F(z_1, z_2). \quad (17)$$

The built-up profile function presented by (16) allows us to account for continuous changes of the overlay building profile during the movement of the satellite around the Earth and the corresponding changes of its elevation angle with respect to the position of any subscriber located in areas of service. This CDF will be used in the next section to find the total outage probability of fading phenomena occurring in LSC links.

Multi-ray Pattern Distribution on the Ground Surface. As is shown in [1–7] and mentioned also in [8], at the horizontal plane the array of buildings are randomly distributed according to Poisson’s law. During movement around the Earth, the satellite antenna “illuminates” different land areas with various distributions of obstructions on the ground surface, that is, in the horizontal plane, as shown in Fig. 8.

The proposed multi-parametric model [1–8] allows us to consider the strength of the total field at the receiver as the additive summation of n -time independently scattered waves with independent strengths. The real situation with multipath phenomena occurring in the urban environment is shown in Fig. 8. As was mentioned in [8], in micro-cell land communication links ($d < 1\text{--}3$ km; d is the range from the base station antenna), the singly scattered waves are dominant, whereas the twice- and three-time scattered waves begin to prevail in macro-cell scenarios for $d > 5\text{--}10$ km. Therefore, as was shown in [8], for land-satellite mega-cell links we must additionally consider the three-time reflected and scattered waves. As was shown in [8], the strengths r_i of these waves are distributed according to the Gaussian law with the zero-mean value and dispersion σ_1^2 (for singly scattered waves), σ_2^2 (for twice scattered waves) and σ_3^2 (for three-time scattered waves), which depend strongly on the characteristic features of the terrain. The average number of scattered waves involves also a dependence on the distance from the base station antenna d . Following [8], we can obtain a new multi-ray distribution for the LSC link. Thus:

— for the average number of single scattered waves

$$\bar{N}_1 = \frac{\pi v d^2}{4} K_2(\gamma_0 d); \quad (18a)$$

— for the average number of two-time scattered waves

$$\bar{N}_2 = 9(\pi v d^2)^2 \left[\frac{K_4(\gamma_0 d)}{8!} + \sqrt{\frac{2}{\pi \gamma_0 d}} \frac{K_{7/2}(\gamma_0 d)}{7!} \right]; \quad (18b)$$

— for the average number of three-time scattered waves

$$\bar{N}_3 = 8(\pi v d^2)^3 \times \left[\frac{1}{\gamma_0 d} \frac{K_5(\gamma_0 d)}{10!} + \sqrt{\frac{2}{\pi \gamma_0 d}} \frac{K_{11/2}(\gamma_0 d)}{11!} \right], \quad (18c)$$

where $K_n(\gamma_0 d)$ is the MacDonal'd's function of n -order, γ_0 [km⁻¹] is the density of buildings contours [8].

The probability of receiving once to three-time (i. e., for $n = 1, 2, 3$) scattered waves at the mobile station antenna can be computed according to the following formula [8]:

$$P_n = 1 - \exp[-\bar{N}_n], \quad n = 1, 2, 3. \quad (19)$$

In computations of equation (19) we take into account the effects of independent single (the first term), double (the second term) and triple (the third

term) scattering of rays with the random amplitude r , as well as their mutual influences on each other (the fourth term), that is,

$$\begin{aligned} CDF(r) = & \frac{P_1(1-P_2)(1-P_3)}{P_0} \frac{r}{\sigma_1^2} e^{-\frac{r^2}{2\sigma_1^2}} + \\ & + \frac{P_2(1-P_1)(1-P_3)}{P_0} \frac{r}{\sigma_2^2} e^{-\frac{r^2}{2\sigma_2^2}} + \\ & + \frac{P_3(1-P_1)(1-P_2)}{P_0} \frac{r}{\sigma_3^2} e^{-\frac{r^2}{2\sigma_3^2}} + \\ & + \frac{P_1 P_2 P_3}{P_0} \frac{r}{\sigma_1^2 + \sigma_2^2 + \sigma_3^2} e^{-\frac{r^2}{2(\sigma_1^2 + \sigma_2^2 + \sigma_3^2)}}, \quad (20) \end{aligned}$$

where

$$P_0 = 1 - (1 - P_1)(1 - P_2)(1 - P_3) = 1 - e^{-(\bar{N}_1 + \bar{N}_2 + \bar{N}_3)} \quad (21)$$

is the probability of direct visibility (e. g., LOS component), P_1 , P_2 and P_3 are defined by equation (21) combined with (18a)–(18c), respectively.

Equation (21) is more general form of the cumulative distribution function of the multipath effects occur in the land-satellite links with respect to that describing the same multi-ray effects in the land-to-land communication scenarios (see [1–8]). This CDF will be used in the next section [together with CDF described by (19)] to estimate the total probability of fading for predicting successful communication and service by the satellite antenna of any ground-based subscriber.

Prediction of Fading and Pass Loss in LSC Links via Experimental Data

In this section, we use the unified algorithm, based on use the combination of the physical-statistical (e. g., Saunders and Evans) model and the pure statistical (Lutz) model of the mobile-satellite communication channel, using formulas (12) and (13), and compare it with the proposed stochastic multi-parametric model based on the corresponding $CCDF(r) = 1 - CDF(r)$, of signal random strength envelope (20). This CCDF describes the total probability to achieve a successful communication between any ground-based subscriber and the satellite. It can be presented as [8]

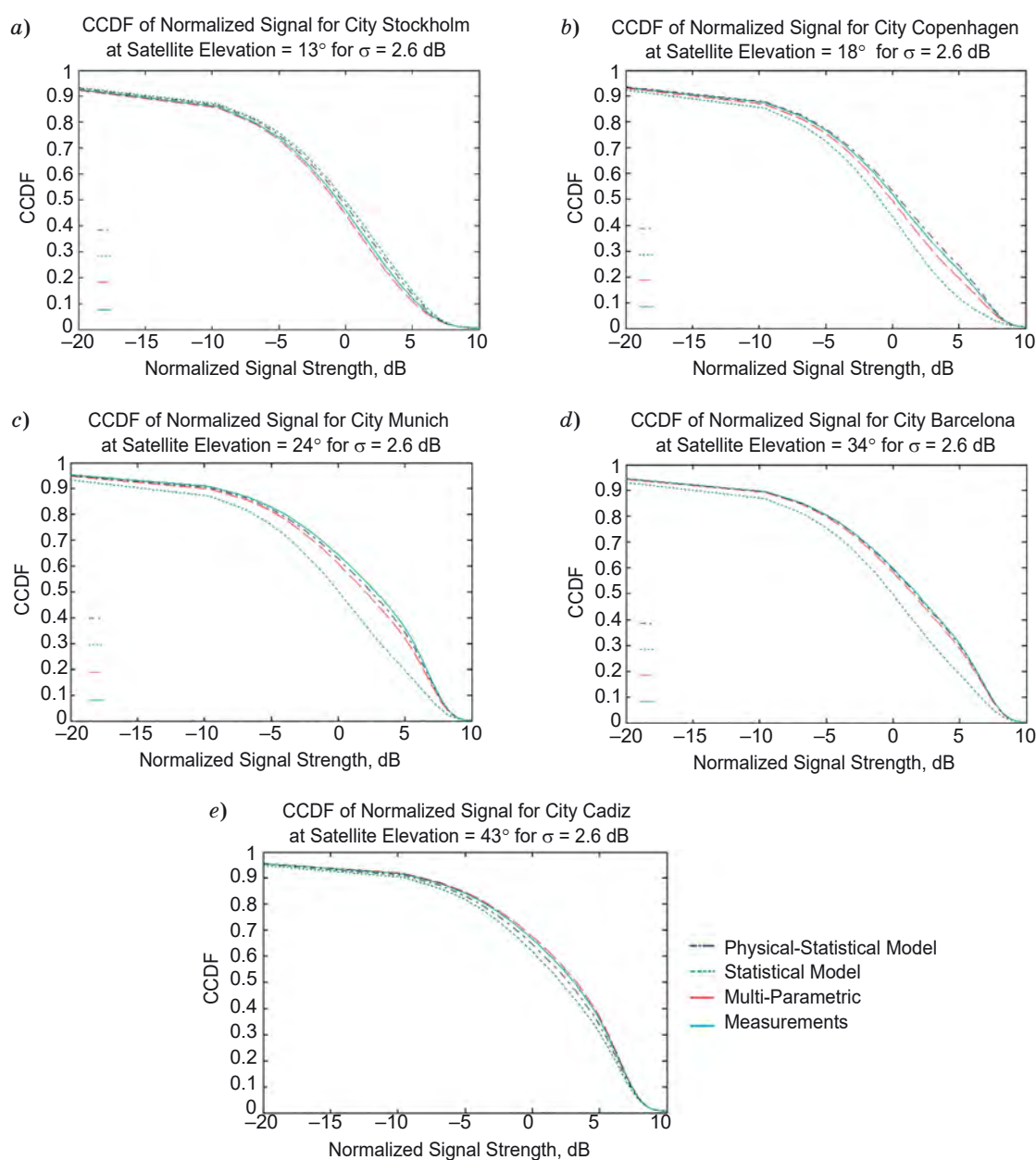
$$CCDF(r) = 1 - CDF(r) \cdot CDF(z_1, z_2, n), \quad (22)$$

where both CDFs are described by (17) and (20), respectively.

In the comparative analysis, we used measured data from references [25, 26], done in several cities in Europe. These tests were narrowband measurements at a single frequency, representing the channel within its coherence bandwidth. The test was transmitted from the ESA ground station in Villafranca, Spain and relayed by the geostationary satellite MARECS at L-band (1.54 GHz). The measurements were conducted in areas with different satellite elevations (Table 2). Using these measurements and using a Rayleigh PDF for building distribution heights and the corresponding equation (22) for CCDF, we constructed a corresponding numerical code to see if there is a good agree-

■ **Table 2.** Parameters of channel model according to [25, 26]

Satellite Elevation	A	$10\log(\epsilon)$, dB	M , dB	σ , dB
13° — Stockholm	0.24	10.2	-8.9	5.1
18° — Copenhagen	0.8	6.4	-11.8	4.0
24° — Munich	0.66	6.0	-10.8	2.8
34° — Barcelona	0.58	6.0	-10.6	2.6
43° — Cadiz	0.54	5.5	-13.6	3.8



■ **Fig. 9.** CCDF of normalized signal for three different models: *a* — for city Stockholm at satellite elevation angle 13°; *b* — for city Copenhagen at satellite elevation angle 18°; *c* — for city Munich at satellite elevation angle 24°; *d* — for city Barcelona at satellite elevation angle 34°; *e* — for city Cadiz at satellite elevation angle 43°

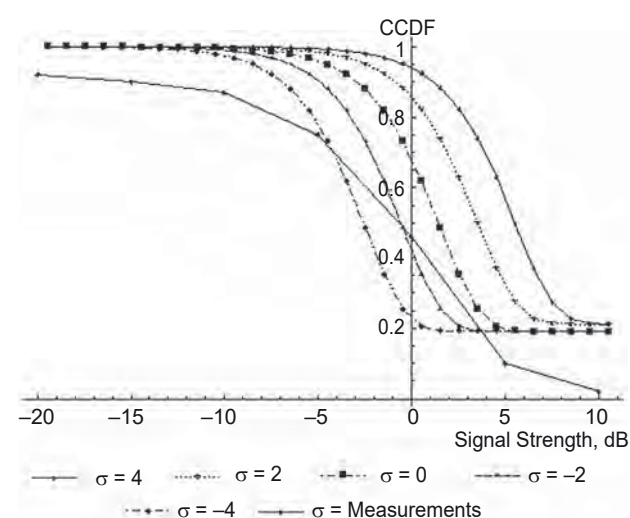
ment with the measured data. It is clear that the CCDF gives us the knowledge of stability of received signal with respect to noise caused by fading phenomena.

The same was done for the multi-parametric model by using equation (22) with the corresponding CDF (17) and (20) for each city. The results of *fading* estimations were compared with those obtained from the pure statistical Lutz model based on Markov's chain.

The main goal of such simulations was to define which one of the three models presented the best fit to the measured data and also the simplest one. In such a comparison we used for the Lutz statistical model simulation $P_{gg} = (0.8, 0.95, 0.85, 0.83, 0.7)$ and $P_{bb} = (0.08, 0.15, 0.25, 0.22, 0.5)$. These results are shown in Figs. 9, *a-e* for five cities presented in Table 2.

The standard deviation, σ , was taken not more than 2.6 dB, obtained from our estimations of each built-up profile. Nevertheless, in reference [16], the authors used $\sigma = 3\div 4$ dB, which is not a realistic case when NLOS regime is very small compared with the LOS component of the total field strength. It is clearly seen from the results of the comparative analysis presented in Fig. 9, *a-e* that the physical-statistical models, Saunders and Evans, and the stochastic multi-parametric, are closer to the experimental data compared to the pure statistical Lutz model. Therefore, in our further analysis, we will compare the proposed physical-statistical models with each other.

Before doing this comparative analysis, let us check the accuracy of the proposed stochastic model in the case when the standard deviation of the CCDF cannot be directly estimated via numerical evaluations. Thus, in Fig. 10, the CCDF of received signal for the city of Stockholm [25–27], at satellite



■ Fig. 10. Comparison between the experimental data obtained in satellite observations over Stockholm (continuous line) and those obtained theoretically from the stochastic approach

elevation angle 13° is presented. The computations were carried out taking the following parameters:

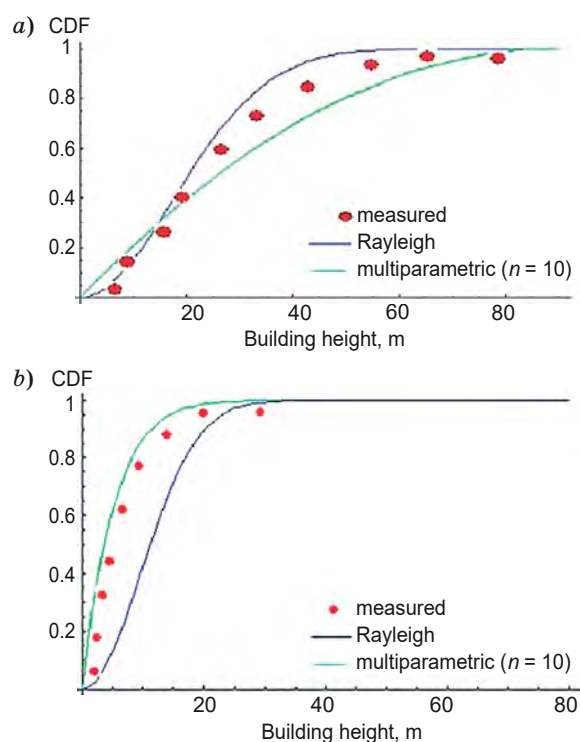
- the building contour density parameter varied from 8 to 12 km^{-1} with the mean value $\bar{\gamma}_0 = 10 \text{ km}^{-1}$;
- the built-up profile parameter n was changed from 0.65 to 1.48;
- the obtained terrain data allow us to estimate the standard deviation of CCDF within the range of $\sigma \in (-4, 4)$ dB.

The results of comparison between the experimental data (continuous curve) and those obtained numerically for the range in the estimated parameters σ shown in Fig. 9 show that the experimental data lie between the two theoretical curves computed for σ changing from -4 to -2 dB, which is close to that parameter estimated by analyzing the topographic map of Stockholm [25–27].

Another land-satellite experiment was compared with the Saunders — Evans physical-statistical model [10, 11, 22, 23], and was carried out in England for two cities, Westmister and Guildford. Here, the CDF, as a total probability of shadowing was investigated using both models. We do not go deeply into the description of these experiments that are fully described in [10, 11, 22, 23]. Here we will only mention that a comparison showed that the best fit was obtained when the CDF of the combined fast/slow fading was described by the corresponding Rayleigh law with the corresponding mean height of buildings $\bar{h} = 20.6$ m and standard deviation $\sigma_b = 17.6$ m (for Westmister) and $\bar{h} = 7.6$ m and $\sigma_b = 6.4$ m (for Guildford). The corresponding comparison analysis is presented in Fig. 11, *a* and *b* for each city, respectively.

Therefore, in this additional comparison with the Saunders — Evans, physical-statistical model and with experimental data, we accounted for these estimations as well as for the average density of buildings' contours within one square kilometer, which was estimated as $\bar{\gamma}_0 \approx 10.6 \text{ km}^{-1}$ (for Westmister) and $\bar{\gamma}_0 \approx 7.5 \text{ km}^{-1}$ (for Guildford). We also estimated the corresponding mean height of buildings for Westmister, $\bar{h} = 18.2$ m with a standard deviation of $\sigma_b = 7.5$ m and for Guildford, $\bar{h} = 5.8$ m and $\sigma_b = 4.6$ m. These estimations have suggested that the two cities have fully different built-up terrain profiles: the parameter of the overlay building profile is $n = 2$ (the amount of small buildings exceeds that of tall buildings) for Westmister, and $n = 10$ (most buildings are small) for Guildford.

We should note that these parameters were taken in a form of average values, but not exactly, as was done by Saunders and Evans with their co-authors in [10, 11, 22, 23] using the local parameters for each position of the moving satellite. Despite this fact, our estimations are within the ranges of estimations obtained by Saunders and Evans, and we can, finally, compare our computations of $CDF = 1 - CCDF$ according to (17), (20) and (22) with those obtained by them using the Rayleigh CDF, as a best fit of measu-



■ **Fig. 11.** Comparison between experimental data (dots) obtained in Westminster (a) and Guildford (b), physical statistical model (blue line) and data obtained from the stochastic approach (green line)

References

1. Blaunstein N., and Christodoulou Ch. G. *Radio Propagation and Adaptive Antennas for Wireless Communication Links — Terrestrial, Atmospheric and Ionospheric*. Hoboken, New Jersey, Wiley Interscience, 2007. 614 p.
2. Blaunstein N. Prediction of Cellular Characteristics for Various Urban Environments. *IEEE Anten. Propag. Magazine*, 1999, vol. 41, no. 6, pp. 135–145.
3. Blaunstein N. D., Katz D., Censor, et al. Prediction of Loss Characteristics in Built-up Areas with Various Buildings' Overlay Profiles. *IEEE Anten. Propag. Magazine*, 2001, vol. 43, no. 6, pp. 181–191.
4. Blaunstein N. Distribution of Angle-of-Arrival and Delay from Array of Buildings Placed on Rough Terrain for Various Elevations of Base Station Antenna. *J. Commun. and Networks*, 2000, vol. 2, no. 4, pp. 305–316.
5. Blaunstein N., Toulch M., Laurila J., Bonek E., et al. Signal Power Distribution in the Azimuth, Elevation and Time Delay Domains in Urban Environments for Various Elevations of Base Station Antenna. *IEEE Trans. Anten. and Propagat.*, 2006, vol. 54, no. 10, pp. 2902–2916.
6. Blaunstein N., Yarkoni N., and Katz D. Spatial and Temporal Distribution of the VHF/UHF Radio Waves in Built-up Land Communication Links. *IEEE Trans. Anten. and Propagat.*, 2006, vol. 54, no. 8, pp. 2345–2355.
7. Blaunstein N., and Ben-Shimol Y. Spectral Properties of Signal Fading and Doppler Spectra Distribution in Urban Mobile Communication Link. *J. Wireless Commun. and Mobile Comput.*, 2006, vol. 6, no. 1, pp. 113–126.
8. Blaunstein N., and Christodoulou Ch. G. *Radio Propagation and Adaptive Antennas for Wireless Communication Networks — Terrestrial, Atmospheric and Ionospheric*. Boca Rocha, Florida, Wiley, 2014. 704 p.
9. Farserotu J., Prasad R. *IP/ATM Mobile Satellite Networks*. Artech House, 2002. 289 p.
10. Saunders S. R. *Antennas and Propagation for Wireless Communication Systems*. John Wiley & Sons, 2001.
11. Evans J. V. Satellite Systems for Personal Communication. *Proc. IEEE*, 1998, vol. 86, no. 7, pp. 1325–1341.
12. Wu W. W. Satellite Communication. *Proc. IEEE*, 1997, vol. 85, no. 6, pp. 998–1010.
13. Fontan F. P., Gonzalez J. P., Ferreiro M. J. S., Castro M. A. V., et al. Complex Envelope Three-state Markov Model Based Simulator for the Narrow-Band LMS Channel. *Int. J. Satellite Communications*, 1997, vol. 15, no. 1, pp. 1–15.

red data. This comparison was shown in Fig. 11 for Westminster and for Guildford, where the corresponding experimental data were plotted by dots.

Conclusion

Even with the data on terrain features using average parameters of the built-up terrain, not the local parameters, we obtained a satisfactory agreement between the theoretical prediction based on both, the Saunders — Evans physical-statistical and multi-parametric stochastic models, and experimental data. This means that the designers of satellite-land links do not need every time to have information on the local built-up terrain parameters as shown in Fig. 6 and 7. It is enough to obtain average parameters of the terrain during satellite movements above the corresponding city, town, village, and so on, and we can for prediction of fading effects to use the stochastic multi-parametric approach.

As can be finally expected, and as seen from Fig. 11, according to the knowledge of the kind of terrain topography from good (open) to bad (urban) scenarios, the main difference is noticeable in the urban environment (as the worst case of strong fading), due to its special propagation features such as multiple diffraction, scattering and reflection from buildings surrounding the subscriber ground-based antenna.

14. Barts R. M., Stutzman W. L. Modeling and Simulation of Mobile Satellite Propagation. *IEEE Trans. Antennas Propagat.*, 1992, vol. 40, no. 4, pp. 375–382.
15. Loo C., Butterworth J. S. Land Mobile Satellite Channel Measurements and Modeling. *Proc. IEEE*, July 1998, vol. 86, no. 7, pp. 1442–1462.
16. Vatalaro F., Mazzenga F. Statistical Channel Modeling and Performance Evaluation in Satellite Personal Communications. *Int. J. Satellite Communications*, 1998, vol. 16, no. 2, pp. 249–255.
17. Loo C. A Statistical Model for Land Mobile Satellite Link. *IEEE Trans. Veh. Technol.*, 1985, vol. VT-34, no. 3, pp. 122–127.
18. Patzold M., Killat U., Laue F. An Extended Suzuki Model for Land Mobile Satellite Channels and its Statistical Properties. *IEEE Trans. Veh. Technol.*, 1998, vol. 47, no. 2, pp. 617–630.
19. Corazza G. E., Vatalaro F. A Statistical Model for Land Mobile Satellite Channels and its Application on Non-Geostationary Orbit Systems. *IEEE Trans. Veh. Technol.*, 1994, vol. 43, no. 3, pp. 738–741.
20. Xia Y., Fang Y. A General Statistical Channel Model for Mobile Satellite Systems. *IEEE Trans. Veh. Technol.*, 2000, vol. 49, no. 3, pp. 744–752.
21. Abdi A., Lau W. C., Alouini M.-S., and Kaveh M. A New Simple Model for Land Mobile Satellite Channels: First- and Second-order Statistics. *IEEE Trans. Wireless Communications.*, 2003, vol. 2, no. 3, pp. 519–528.
22. Saunders S. R., and Evans B. G. A Physical Model of Shadowing Probability for Land Mobile Satellite Propagation. *Electronics Letters*, 1996, vol. 32, no. 17, pp. 1548–1549.
23. Tzaras C., Saunders S. R., and Evans B. G. A Tap-Gain Process for Wideband Mobile Satellite PCN Channels. *Proc. COST 252/259 Joint Workshop*, Bradford, UK, April 21–22, 1998, pp. 156–161.
24. Lutz E., Cygan D., Dippold M., Dolainsky F., Papke W. The Land Mobile Satellite Communication Channel-Recording, Statistics and Channel Model. *IEEE Trans. Veh. Technol.*, 1991, vol. 40, no. 2, pp. 375–385.
25. Butt G., Evans B. G., and Richharia M. Narrowband Channel Statistics from Multiband Propagation Measurements Applicable to High Elevation Angle Land-Mobile Satellite Systems. *IEEE Trans. Select. Areas Commun.*, 1992, vol. 10, no. 8, pp. 1219–1226.
26. Parks M. A. N., Evans B. G., Butt G., and Buonomo S. Simultaneous Wideband Propagation Measurements for Mobile Satellite Communication Systems at L- and S-bands. *Proc. 16th Int. Commun. Systems Conf.*, Washington DC, 1996, pp. 929–936.
27. Karasawa Y., Kimura K., and Minamisono K. Analysis of Availability Improvement in LMSS by Means of Satellite Diversity Based on Three-state Propagation State Model. *IEEE Trans. Veh. Technol.*, 1997, vol. 46, no. 4, pp. 1047–1056.

УДК 621.371

doi:10.15217/issn1684-8853.2018.1.74

Влияние поверхности земли с застройкой на операционные параметры сигналов в каналах связи «Земля — спутник»

Блаунштейн Н. Ш.^а, доктор физ.-мат. наук, профессор, nathan.blaunstein@hotmail.com

Сергеев М. Б.^б, доктор техн. наук, профессор, mbse@mail.ru

Ненашев В. А.^б, канд. техн. наук, доцент, nenashev.va@gmail.com

^аНегевский университет им. Бен-Гуриона, Р.О.В. 653, 1, Бен-Гуриона ул., г. Беэр-Шева, 74105, Израиль

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

Цель: анализ распространения радиосигналов в городской среде для различной направленности (элевации) движущейся спутниковой антенны и стационарных антенн пользователей, расположенных как ниже, так и выше крыш домов. Задача решалась с помощью стохастического многопараметрического подхода при сосредоточенности на временных и пространственных свойствах изменения мощности сигнала для прогноза потерь в канале и эффектов фединга. **Результаты:** на основе стохастического подхода, предложенного в работе, проанализированы полная мощность сигнала, соответствующие функции плотности вероятности, а также масштабы временных и частотных характеристик. Все прогнозируемые характеристики проанализированы численно, влияние плотности застройки домов и элевации спутниковой антенны на вышеуказанные характеристики проанализированы через призму специальных экспериментов, проведенных многими исследователями. **Практическая значимость:** просчитанные и наблюдаемые экспериментально данные показывают хорошее соответствие между предлагаемым стохастическим подходом и подходами, предложенными другими исследователями, поэтому представленный подход можно использовать для прогноза потерь в канале связи для среды с различным профилем застройки домов и для различных элеваций движущейся спутниковой антенны.

Ключевые слова — канал «Земля — спутник», стохастический многопараметрический подход, профиль застройки и плотности застройки домов, вероятность мультипликативного фединга, потери в канале, расчет канала связи.

Цитирование: Blaunstein N. S., Sergeev M. B., Nenashev V. A. Impact of Built-up Terrain on Operational Parameters of Signals in Land-Satellite Communication Links// Информационно-управляющие системы. 2018. № 1. С. 74–84. doi:10.15217/issn1684-8853.2018.1.74

Citation: Blaunstein N. S., Sergeev M. B., Nenashev V. A. Impact of Built-up Terrain on Operational Parameters of Signals in Land-Satellite Communication Links. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 1, pp. 74–84. doi:10.15217/issn1684-8853.2018.1.74

UDC 654.165

doi:10.15217/issn1684-8853.2018.1.85

PREDICTION OF OPERATIONAL PARAMETERS OF RADIO SIGNALS PASSING A LAND-SATELLITE LINK THROUGH STORMTIME IONOSPHERE

N. S. Blaunstein^a, Dr. Sc., Phys.-Math., Professor, nathan.blaunstein@hotmail.com

Y. Ben-Shimol^a, PhD, Electrical Engineering, Senior Lecturer, benshimo@bgu.ac.il

^aBen-Gurion University of the Negev, P.O.B. 653, 1, Ben-Gurion St., Beer-Sheva, 74105, Israel

Introduction: The subject of this research became important during the recent decades due to the increasing demand in globalization of wireless networks by using long-path ionospheric radiowave propagation in land-satellite communication links. **Purpose:** Analyzing the key parameters of VHF/UHF radio waves in a satellite-land link which determine the fading effects occurring in storm-time mid-latitude ionosphere. **Results:** For perturbed plasma parameters experimentally caused by a magnetic storm, the absorption and phase fluctuations of the radio signals were examined. The scintillation index corresponding to signal scintillations in multi-path ionospheric communication links with fading was studied on the base of experimental data. It has been shown that ~10% density irregularities in the storm-time F region yield strong fast fading of radio signals in VHF/UHF frequency band with considerable signal intensity fluctuations (up to 1%) and phase deviations (up to hundreds of radians). **Practical relevance:** The obtained results allow you to predict multiplicative fading in ionospheric sub-channels and its impact on the total link budget in a full land-satellite communication link.

Keywords – Ionospheric Sub-Channel, Ionospheric Plasma Density Irregularities, Signal Intensity Scintillation Index, Amplitude Attenuation, Phase Fluctuation, Root Mean Square, VHF/UHF-Band Radio Waves.

Citation: Blaunstein N., Ben-Shimol Y. Prediction of Operational Parameters of Radio Signals Passing a Land-Satellite Link through Stormtime Ionosphere. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 1, pp. 85–95. doi:10.15217/issn1684-8853.2018.1.85

Introduction

The ionosphere is one of the sub-channel (additional to the terrestrial and atmospheric sub-channels) of the global (e. g., mega-cell) land-satellite and transionospheric communication systems. It is well known that ionospheric irregularities cause strong fluctuations of the amplitude and phase of radio signals, i.e. fast fading. Enhanced irregularities are usually observed during magnetically disturbed periods, especially in the auroral ionosphere (see [1–9] and bibliography there). As the auroral boundary shifts equatorward during magnetic disturbances, it is usually assumed that mid-latitude scintillation events occur inside the expanded auroral zone. As was expected by many researchers dealing with satellite observations of the perturbed stormtime ionosphere, this phenomenon occurs not only in the proximity of the polar and auroral zones, but also in the sub-auroral and middle-latitude ionosphere [10–14]. This event occurs in about 15–17 % of time per year and usually is observed during several-day to several-week periods. It was found that during magnetic storms strong phase and amplitude scintillations (the scintillation index $S_4 \geq 0.5$) of radio signals of the satellite operating at 250 MHz [2] and 1.5 GHz [3] are observed (namely, near Boston, MA, USA and Ithaca, NY, USA) well equatorward of the auroral zone, up to the sub-au-

roral and mid-latitude ionosphere. Satellite and radar observations in the stormtime sub-auroral ionosphere [4–7] revealed irregular structures embedded within fast westward streams, called sub-auroral polarization streams (SAPS) [8]. In particular, strongly irregular plasma density structures are embedded within SAPS wave-like structures (called SAPSWS) [6, 7]. Thus, the Defense Meteorological Satellite Program (DMSP) for satellite observations, which coincided with the [2, 3] scintillation events, revealed SAPSWS-related irregular density troughs. Within the troughs, power spectral densities (PSD) ($\delta n_k = 10\text{--}20\%$ with respect to non-disturbed ionospheric plasma n_0 [9]) of plasma irregularities (electron and ions, since plasma is quasi-neutral) in the range of apparent wave numbers from $k = 2\pi/\lambda = 0.85\text{ km}^{-1}$ to $k = 12\text{ km}^{-1}$, λ is the wavelength, were well-approximated by a plasma density power-law $(\delta n_k/n_0)^2 k^{-p}$ with the spectral index of $5/3 < p \leq 2$ [9].

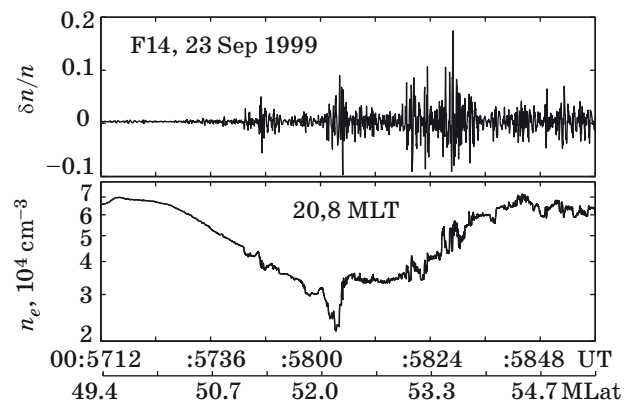
The spectral properties of plasma perturbations allow for predicting spectral characteristics of the radio signal with data (called the bandpass signal [15]), and finally predict the type of signal fading, flat or frequency selective, in such a perturbed ionospheric channel. These aspects are very actual for advanced wireless networks design based on modern techniques of signal processing, namely, orthogonal frequency division multiplexing (OFDM),

where each subscriber obtained very narrow bandwidth. The main goal of such a technique is to eliminate effects of the noise caused by multipath phenomena occurring in the channel with fading and by inter-subscriber interference [16]. Therefore, it is important to predict *a-priori* the type of the channel, frequency or time dispersive (or both), and to estimate the coherent bandwidth and time of coherency of such a channel with fading for the future performance of OFDM access technique [16]. Above example shows the actuality of this paper and research carried out in it.

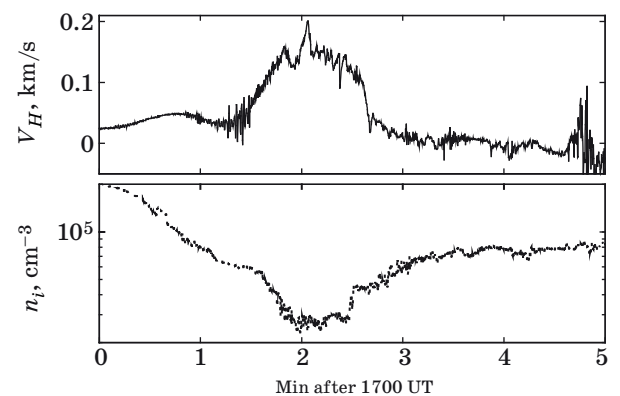
This work describes fading phenomena and their effects on key parameters of radio signals in the land-satellite communication link passing through the structured storm-induced plasma region, filled by the “typical” small- and moderate-scale irregularities. The basic effects of irregular plasma structures, embedded within SAPS, on radio signals propagation via the perturbed stormtime ionosphere are briefly discussed in the first Section. In the Section following it presents the results of computation of signal integral absorption, amplitude and phase “point-to-point” deviations, and variations in the intensity and phase of radio signals propagating through this SAPSWS-related “typical” link. These results help to understand the fading phenomena effects for signals passing multipath ionospheric links during magnetic storms. Then, we conclude main effects of radio signals of various frequency bands occurring in the perturbed ionosphere during magnetic storms.

Satellite Observations of Stormtime SAPS-related Density Irregularities

We start our discussions with analysis of several satellite campaigns carried out during period from the end of 20 century to beginning of 21 century. Thus, Fig. 1 (rearranged from [9]) shows a snapshot of plasma observations from the DMSP F14 satellite near the scintillation event near Boston, MA, USA during the 23 September 1999 magnetic storm [2]. Shown in the bottom panel is the plasma density variation along the satellite track measured by a spherical Langmuir probe sampled at the rate of 24 Hz. A detailed description of the probe and the whole DMSP/SSIES (Special Sensor for Ions, Electrons, and Scintillations) suite can be found in [10–13]. The top panel shows the corresponding waveform of plasma density deviations, $\delta n/n_0$, obtained applying 0.1–9.5 Hz bandpass elliptic filter (see details in [2]). Here δn is the perturbation of the background ionospheric non-disturbed plasma density (denoted by n_0). One can see also that enhanced density irregularities were present well equatorward of the boundary of auroral zone indicated by the vertical dashed



■ Fig. 1. DMSP F14 observations during the 23 September 1999 magnetic storm near Boston, MA, USA: (top) waveform of relative density variations $\delta n/n$ and (bottom) the density variation along the satellite track

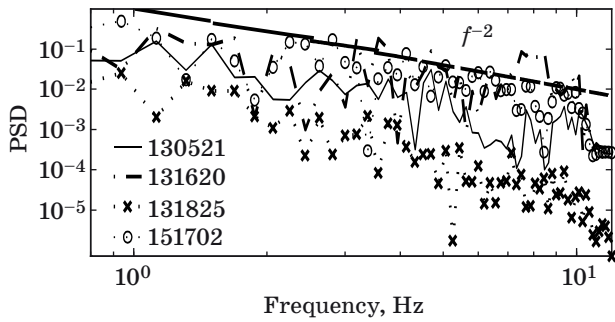


■ Fig. 2. (top) Horizontal (positive westward) component V_H of the convection velocity and (bottom) the plasma density from the DMSP F15 satellite on March 31, 2001 (extracted from [3])

line (i.e., beyond this zone the sub-auroral ino-
sphere is seen to be also perturbed).

We present now another example of stormtime sub-aurora density irregularities observed by the DMSP F13 and F15 satellites during the major magnetic storm of March 31, 2001 [3]. Fig. 2 shows a 5-min snapshot of the horizontal (positive westward) component of the convection velocity V_H (the top panel) and plasma density n_i (bottom) from F15/SSIES obtained by the ion drift meter and Langmuir probe sampled at the rates of 6 and 24 Hz, respectively.

Note that quite similar SAPS/trough patterns were observed throughout the storm also by the DMSP F13 satellite. As in [2, 3, 9, 10], we obtained the density waveforms in the SAPS regions applying 0.1–9.5 Hz bandpass elliptic filter. Their PSD are shown in Fig. 3 as a function of apparent frequencies. Likewise [9, 10], PSD is well represented by the law $\delta n_k/n_0 \sim k^{-p}$, $p = 2$ (or $\delta n_f/n_0 \sim f^{-2}$), where $k = 2\pi f/c$, and p is parameter of PSD [2–9].



■ **Fig. 3.** Power spectral densities of plasma irregularities vs. the apparent frequency from the DMSF satellites and near the times designated in label: e. g., 130521 stand for the satellite F13 at 0521 UT (extracted from [9])

We emphasize that spatial and temporal variations cannot be separated in data from a single spacecraft and thus frequencies (wavelengths) mean apparent frequencies (scale-lengths). Given the satellite speed $v_s = 7.5$ km/s, apparent frequencies 1–10 Hz correspond to wavelengths $\lambda_s = 2\pi/k_s = v_s/f = 7.5\text{--}0.75$ km along the satellite track. Note also that after about 1900 UT short-scale oscillations faded away, whereas “smooth” SAPS/trough plasma structures remained throughout and even after the storm recovery phase. Such behavior is typical of SAPSWS that develop within 10–20 min and decay is ~ 1 hour after substorm expansion onsets [6, 7, 9].

Prediction of Main Parameters of Signals Passing Stormtime Perturbed Ionosphere

Main Parameters of the Radio Channel

In predicting the key parameters of radio signals passing through irregular ionosphere characterized by a spectral index $p = 1 + \langle n \rangle$ ($\langle n \rangle$ is the mean value of the plasma refractive index) or by the PSD index $p' = p - 2$ [15–21] it is important to show that the signal intensity fluctuations at a given frequency can be described by the Ricean K -parameter as a ratio of coherent and incoherent components of the total signal intensity [22]:

$$K = \langle I_{co} \rangle / \langle I_{inc} \rangle. \quad (1)$$

Here $\langle I_{co} \rangle$ is the line-of-sight (LOS) component of the signal (called also the coherent part [14, 15, 22]) and $\langle I_{inc} \rangle$ is the incoherent component of the total signal intensity:

$$\langle I \rangle = \langle I_{co} \rangle + \langle I_{inc} \rangle, \quad (2a)$$

where

$$\langle I_{inc} \rangle = \langle I_{co} \rangle \sqrt{\sigma_I^2}. \quad (2b)$$

The scintillation index

$$\sigma_I^2 = \frac{\langle I^2 \rangle - \langle I \rangle^2}{\langle I \rangle^2} = \frac{\langle I^2 \rangle}{\langle I \rangle^2} - 1 \quad (2c)$$

denoted in [1–9] as S_4 , defines the normalized intensity of the signal fast fading.

For Gaussian zero-mean random signal amplitudes, the relationship between K and σ_I^2 (denoted in experimental works [1–5] as S_4) is simply

$$K^2 = 1 / \sigma_I^2 \equiv 1 / S_4. \quad (3)$$

We note, according to [15] that for Gaussian non-zero-mean random signal amplitude distribution, as clear seen from relations (1)–(2c), formula (3) becomes more complicated. However, as was shown by many researchers investigating stochastic processes occurring in the perturbed ionosphere [15, 16–21, 23–36], the spatial distribution of ionospheric irregularities of natural and artificial nature can be correctly described by *zero-mean Gaussian law*, which according to ergodic theorem leads to the same Gaussian distribution of scattered radio signals in the time domain.

The basic characteristics of signal fading are derived by means of the approach developed in [16–20, 23–32] and summarized in [15, 21]. They are: 1) mean square phase fluctuation $\langle (\Delta\Phi)^2 \rangle$; 2) outer scale L_0 of the disturbed ionospheric plasma structure; 3) inner scale l_0 of the plasma density irregularities.

In addition, to describe radio wave propagation in irregular plasmas, the outer scale L_0 should be compared with the Fresnel scale $d_F = \sqrt{\lambda Z_0}$ [15]. Here λ is the signal wavelength and Z_0 is the distance from the ionospheric layer filled by plasma inhomogeneities to the ground reception plane [15, 25–27]).

The spectrum of the signal intensity fluctuations can be derived via $\sqrt{\langle (\Delta\Phi)^2 \rangle}$ [16–19, 21, 23, 30, 32, 34–36] by assuming that the root mean square (RMS) of phase fluctuations is larger than one radian. Moreover, in [15, 18, 19, 21, 30] have shown theoretically that at $k \gg 1/L_0$ the power spectrum of phase fluctuations $S(k)$ in the perturbed inhomogeneous ionosphere is proportional to k^{-p} (e. g., f^{-p}), where $p = 2$ (see Fig. 3). This condition was proved by numerous observations of perturbed stormtime ionosphere (see [9, 21] and bibliography there).

Main Characteristics of the Radio Signal Fading

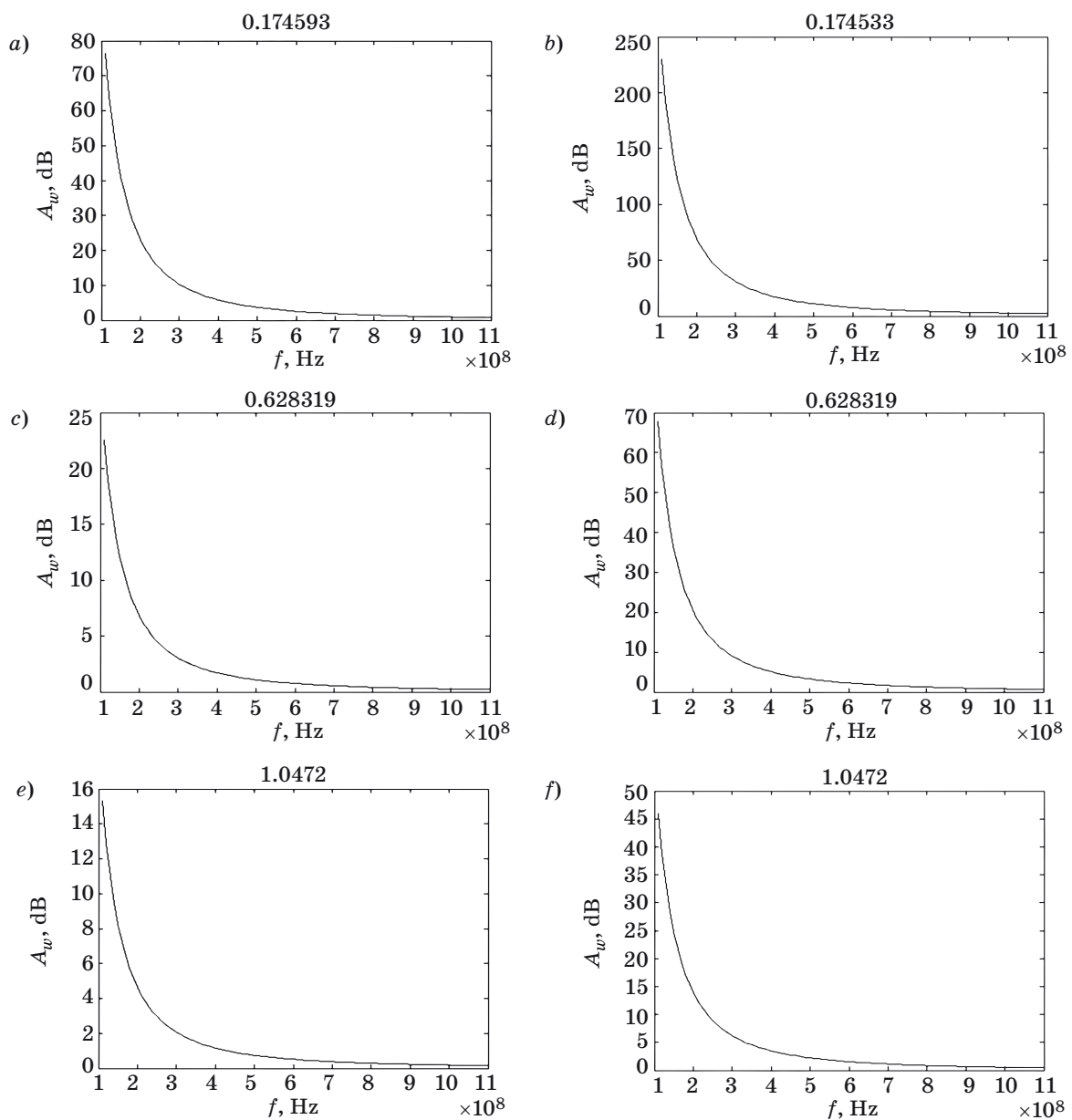
We analyze now propagation characteristics of the radio signal passing disturbed ionospheric region perturbed by geomagnetic storm. As it follows

from experimental data presented above, the background ionospheric plasma can be perturbed essentially in period of magnetic storm preparation. During this period the observed deviations of plasma density around its non-disturbed background can achieve 10–20 % (see Figs. 1, 2). There are several features should be estimated for predicting fading phenomena in the satellite communication links passing through the perturbed ionosphere, such as absorption, point-to point attenuation and phase deviation of radio signal amplitude along the radio path, as well as scattering phenomena caused signal intensity and random phase variations.

Absorption of Radio Signal in the Disturbed Ionosphere

To estimate losses due to absorption, i.e., the part of energy of radio signal, which is absorbed in the perturbed region of the ionosphere, from D to F , we propose here, following [15, 21, 35], to use a special value of absorption in decibels:

$$A_{\omega} = 10 \log \frac{I_0}{I} \approx \approx \frac{4.3}{c} \int \frac{\omega_{pe}^2 (v_{em} + v_{ei} + v_{ee})}{[\omega^2 + (v_{em} + v_{ei} + v_{ee})^2]} ds \text{ [dB]}, \quad (4)$$



■ **Fig. 4.** Absorption of signal vs. frequency for weak magnetic storm (*a, c, e*) and strong magnetic storm (corresponding to Alaska region) (*b, d, f*): *a, b* — grazing angle $\psi \approx 0.175$ rad ($\psi \approx 10^\circ$); *c, d* — $\psi \approx 0.628$ rad ($\psi \approx 36^\circ$); *e, f* — $\psi \approx 1.05$ rad ($\psi \approx 60^\circ$)

where ν_{em} , ν_{ei} and ν_{ee} are the frequencies of electron-neutral, electron-ion and electron-electron collisions, respectively.

The parameter of absorption fully characterizes real ionospheric radio traces and can usually be estimated by measuring radiometric absorption at the fixed frequencies and by knowledge of frequency dependence of coefficient of absorption κ . We should note that in integrand of (4) we account not only for the effects of electron-neutral and electron-ion collisions, but also ion-neutral and electron-electron collisions, based on the elements of kinetic theory [37–40], but not a magneto-ionizing theory usually used in such computations based on hydrodynamic classical approach.

Then, accounting in (4) for the fact that the length of radio path, s , and the height of the ionosphere, h , are related as $h = s \sin \psi$, where ψ is the grazing angle of the satellite antenna with respect to the ground-based antenna, we can estimate the total absorption of radio signal along the radiopath. In Figs. 4, *a–f* this parameter is presented in [dB] vs. the carrier frequency of probing radio signal sent via stormtime ionosphere under different grazing angles, respectively. The corresponding profile of disturbed plasma density $N(h)$ is taken from Fig. 2 and normalized on the non-disturbed plasma density $N_0(h)$ taken from [15, 21] for the mid-latitude regular ionosphere, and accounting in computations of integral (4) for the limits $h_{\min} = 50$ km and $h_{\max} = 500$ km. We consider below two scenarios, weak and the strong magnetic storm, which are shown in Figs. 4, *a, c, e* and 4, *b, d, f* respectively. The corresponding data were taken from observations of magnetic storm, described above for weak storms observed at the middle latitude ionosphere above USA and at the northern regions of USA and Alaska.

As follows from above illustrations, with increase of the magnetic storm strength, from weak to strong, absorption of radio signals increases roughly three times at lower frequencies from 100 to 500 MHz, while at 500–600 MHz, absorption does not depend on frequency; only on the grazing angle. At the same time, with increase of the grazing angle, the tendency of decrease of signal absorption for week storms compared with strong storms is the same, per twice or three times. Results of computations presented above allow us to conclude that magnetic storms occurring in the northern mid-latitude ionosphere can significantly perturb ionospheric plasma and thus cause a dramatic attenuation of radio signals up to 500 MHz. This effect depends on the grazing angle and can be insignificant for grazing angles bigger than $\psi \geq 20^\circ$. Hence with increase of the radiated frequency of radio signal for the same grazing angle or for increase of grazing angle for the same frequency, the tendency of

significant decrease of signal energy absorption is evident.

Signal Point-to-Point Amplitude Attenuation and Phase Deviations

At the same time, we should point out that for radar and radio communication applications, other characteristics of signal fading can give information about radiophysical effects, dealing not with the cumulative intensity integral absorption along the radiopath, but with the point-to-point amplitude and phase deviations of the signal along the radiopath. In this case, it is important to understand of how perturbed parameters of the ionospheric plasma attenuate radio wave amplitude and change its phase. To understand these radiophysical effects, we will consider an arbitrary monochromatic radio wave [15]:

$$U(z, t) = A(z, t) \exp[j(\omega t - kz)]. \quad (5)$$

In this case, the wave number k can be described through the coefficient of wave amplitude attenuation, α , and the coefficient determining changes of wave phase in real time, β , i.e.,

$$k = \alpha + j\beta. \quad (6)$$

Unlike usual description of parameters α and β , we will present them through collision frequencies of plasma particles using elements of kinetic theory (see [15, 21]). For numerical computations, we introduce the additional notations and parameters to describe α and β , that is, $X = \omega_p^2 / \omega^2$, $Z = \nu / \omega$, ω_p is the background plasma density in the perturbed ionosphere [15, 37–39]. In these notations, after straightforward derivations, we obtained that

$$\begin{aligned} k^2 &= \frac{\omega^2}{c^2} \left[1 - \frac{X}{1 - iZ} \right] = \frac{\omega^2}{c^2} \left[1 - \frac{X(1 + iZ)}{1 + Z^2} \right] = \\ &= \frac{\omega^2}{c^2} \left[1 - \frac{X + iZX}{1 + Z^2} \right] \end{aligned} \quad (7a)$$

or

$$k^2 = \frac{\omega^2}{c^2} \left[1 - \frac{X}{1 + Z^2} - i \frac{ZX}{1 + Z^2} \right]. \quad (7b)$$

By introducing now normalized parameters (to the wave number in free space k_0), $\tilde{\alpha} = \frac{\alpha}{k_0}$ and $\tilde{\beta} = \frac{\beta}{k_0}$, we finally get:

$$\tilde{\alpha} = \frac{1 + \left(\frac{\nu}{\omega}\right)^2 - \frac{\omega_p^2}{\omega^2}}{1 + \left(\frac{\nu}{\omega}\right)^2} = \frac{\omega^2 + \nu^2 - \omega_p^2}{\omega^2 + \nu^2}; \quad (8a)$$

$$\tilde{\beta} = \frac{ZX}{1+Z^2} = \frac{\left(\frac{v}{\omega}\right) \frac{\omega_p^2}{\omega^2}}{1 + \left(\frac{v}{\omega}\right)^2} = \frac{v\omega_p^2}{\omega(\omega^2 + v^2)}. \quad (8b)$$

Here $v = v_{em} + v_{ei} + v_{ee} = v_{em}(1 + q + q')$, where $q = v_{ei}/v_{em}$ and $q' = v_{ee}/v_{em}$.

Deviations of radio signal parameters, α and β , in the perturbed ionospheric region, were found by using parameters of the disturbed ionosphere, i. e., variations in the plasma frequency ω_p/ω_{p0} , plasma concentration N/N_0 ($N = N_0 + \delta N$, $\delta N < N_0$), and plasma conductivity σ/σ_0 , where ω_{p0} , N_0 , and σ_0 are the background non-disturbed ionospheric parameters [15, 21]

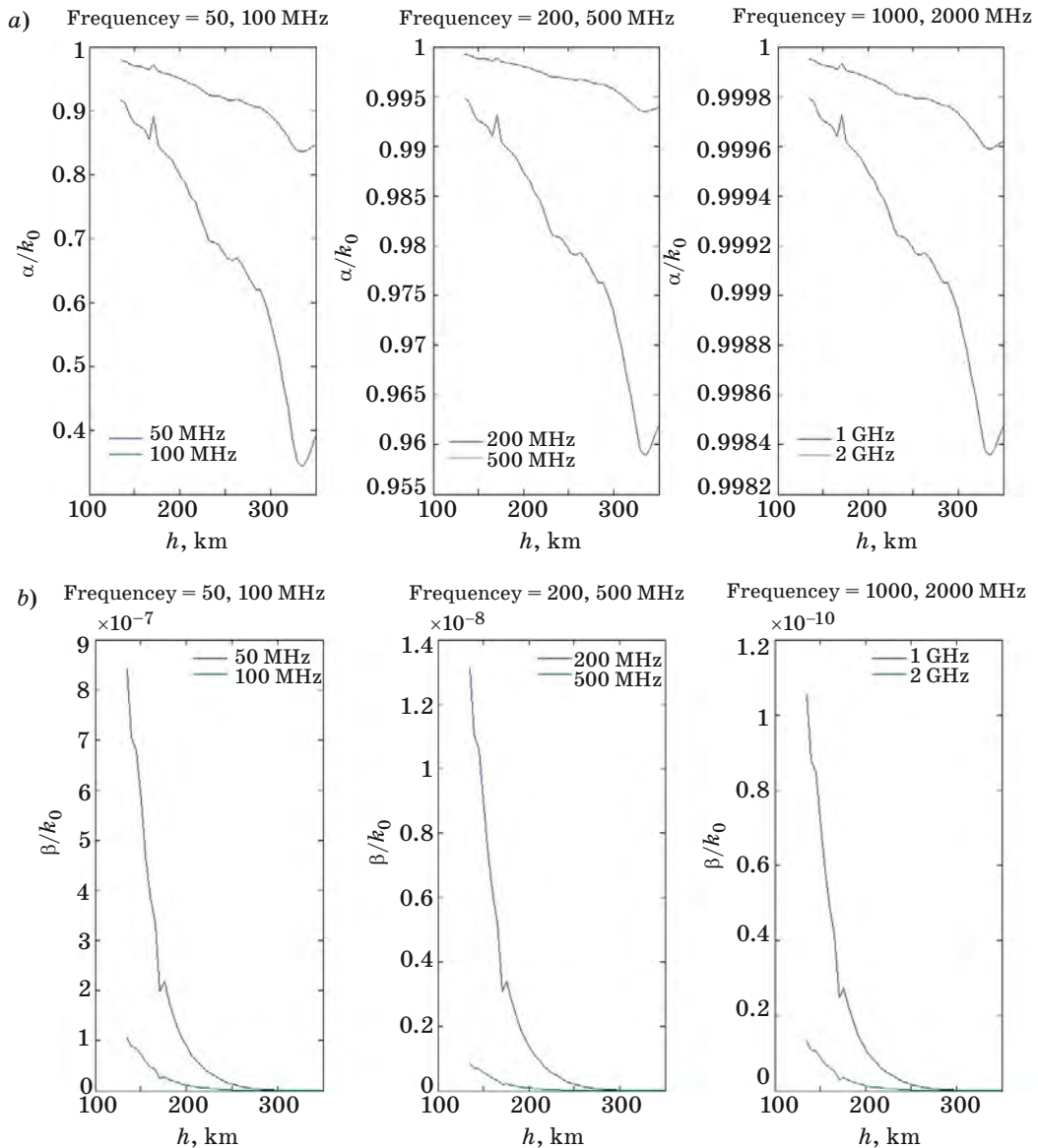
$$\omega_{p0} = (4\pi N_0 e^2 / m_e)^{1/2}; \quad (9)$$

$$\varepsilon_0 = 1 - \left[\omega_0^2 / (\omega^2 + (v_{em} + v_{ei})^2) \right]; \quad (10)$$

$$\sigma_0 = \left[\omega_0^2 (v_{em} + v_{ei}) / 4\pi (\omega^2 + (v_{em} + v_{ei})^2) \right]. \quad (11)$$

Here e and m_e are the charge and mass of plasma electron; other parameters are defined above.

Thus, in Fig. 5, *a, b* we present deviations of radio signal parameters, α and β , compared with those, α_0 and β_0 , obtained for non-disturbed background ionospheric plasma. The computation results are shown in Fig 5, *a* (for normalized attenuation) and Fig. 5, *b* (for normalized phase velocity) versus altitude of the ionosphere for different frequencies of probing waves varied from 50 MHz to 2 GHz (i.e., covering also HF-bandwidth which is important for transionospheric propagation). Thus, we use



■ Fig. 5. Normalized attenuation coefficient (*a*) and phase velocity (*b*) of probing wave vs. ionospheric altitudes for frequencies varied from 50 MHz to 2 GHz

those frequencies which are actual both in land-ionosphere-land navigation ($f < 500\text{--}600$ MHz) and those which usually used in GPS radio monitoring of the ionosphere ($f > 900$ MHz).

As is seen from Fig. 5, *a* with the increase of frequency of the probing wave, the effect of attenuation of the wave energy becomes weaker, and the probing radio wave propagates at the same manner as in the non-disturbed ionosphere. This effect depends strongly on what altitudes the propagation process is observed. Thus, with increase of ionospheric altitude, attenuation affects stronger for all frequencies under consideration.

The same tendency of decreasing phase velocity of radio wave passing the disturbed ionospheric region with increase of frequency of probing wave is clear seen from Fig. 5, *b*. In fact, with increase of frequency from 50 MHz to 2 GHz, the normalized phase velocity, β/β_0 , becomes smaller, i. e., the phase velocity of the probing wave becomes smaller in disturbed plasma compared with that in non-disturbed plasma.

This effect can be easily explained. When the frequency increases (or the wavelength decreases compared to the dimensions of plasma irregularities), the effects of “diffractive scattering” become weaker and, instead of defocusing with strong fading, we observe focusing with weak fading [15, 21].

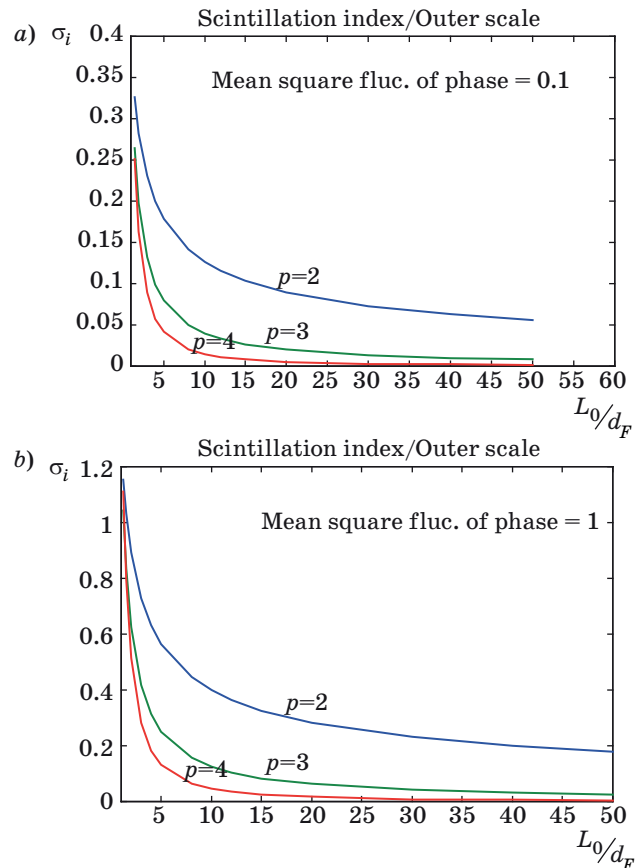
Signal Intensity Fluctuations

To investigate the signal intensity deviations in the perturbed ionosphere and the corresponding scintillation index evaluations, the plane-screen [16–20, 23, 33, 34] and curved-screen ionospheric model [26, 27, 39] were performed and then summarized in [15]. Based on the theoretical self-consistent framework described there, one can estimate the effects of magnetic storm on the above two parameters of the radio signal propagating in perturbed ionospheric communication channel. Therefore, dealing only with situation occurring during magnetic storm described by the following PSD parameter $p = 4$ ($p' = 2$) introduced in first Section (see Fig. 3), we finally get, following [15, 21]

$$I(k) = 32 \langle (\Delta\Phi)^2 \rangle \frac{L_0}{(1 + k^2 L_0^2)^2} \sin^2 \left(\frac{1}{2} k^2 d_F^2 \right). \quad (12)$$

Accounting for self-consistent theoretical framework described in [15, 21] for various scenarios occurring in the strongly perturbed ionosphere, we get for $p' = 2$ ($p = 4$) the following expression of the scintillation index:

$$\sigma_I^2 = \frac{8\sqrt{2}}{3\sqrt{\pi}L_0^3} d_F^3 \langle (\Delta\Phi)^2 \rangle. \quad (13)$$



■ Fig. 6. The RMS of the scintillation index, σ_I , vs. the outer scale L_0 for $1.5d_F \leq L_0 \leq 50d_F$, $\sqrt{\langle (\Delta\Phi)^2 \rangle} = 0.1$ rad (a) and $\sqrt{\langle (\Delta\Phi)^2 \rangle} = 1$ rad (b)

Computations of the scintillation index presented by (13) were compared in [15, 21] with those formulas obtained in [16–20, 23] for weak and moderate plasma perturbations [e. g., for $p = 2$ ($p' = 0$) and $p = 3$ ($p' = 1$), respectively], for an outer scale $L_0 = 10d_F$ and inner scale $l_0 = 10^{-2}d_F$. Figures 6, *a* and 6, *b* present the RMS of the scintillation index, $\sigma_I = \sqrt{\sigma_I^2}$, computed for weak ($\sqrt{\langle (\Delta\Phi)^2 \rangle} = 0.1$ rad) and moderate ($\sqrt{\langle (\Delta\Phi)^2 \rangle} = 1$ rad) signal phase fluctuations versus square mean deviations of signal phase for various PSD parameters $p = 2, 3, 4$ and different scales of ionospheric irregularities.

It is seen that for $p = 2$ the scintillation index with increase of phase fluctuations limits to the unit. In the last case, for more higher spectral index ($p > 2$) σ_I exceeds the unit within the range of $0 < L_0/d_F < 1$, which explain us the focusing properties of the ionospheric layer consisting various irregularities and strong variations of signal phase after passing the perturbed ionosphere.

Signal Phase Deviations

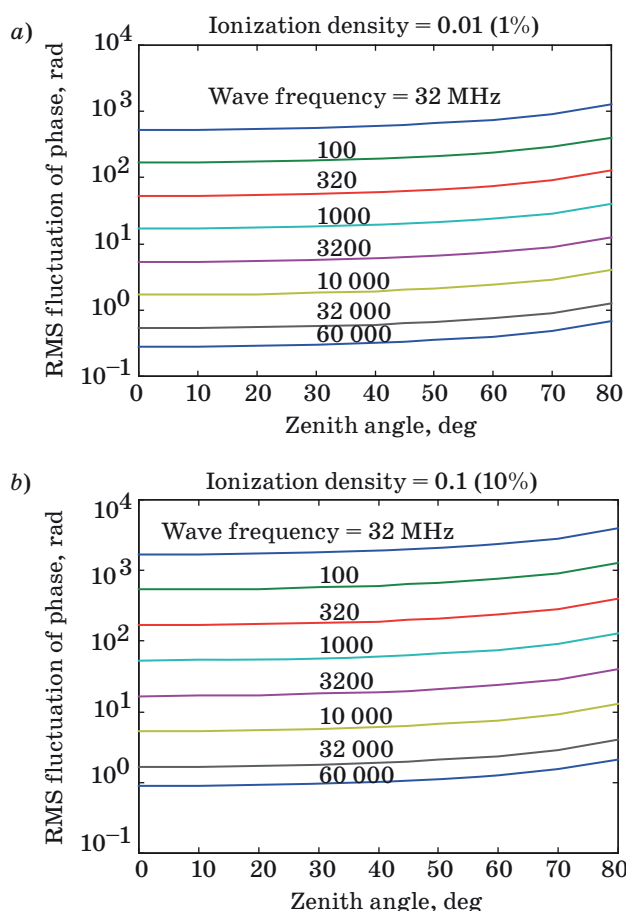
We model the ionospheric F -region, according to [15, 21, 27], as a spherical 3-D layer of mean ionization density N and with the outer scale L_0 (instead of a plane-layer 1-D model, as was performed in [16–20, 24, 29–34]) with the standard fluctuations of ionization density, $(\Delta N/N_0)^2 = (\delta N/N_0)^2$. We take the outer scale L_0 equals H , assuming latter as the thickness of the disturbed ionospheric F -layer. In such notations, the mean square fluctuation of phase of radio signal can be derived with the help of the general model [15, 21, 27], which finally gives:

$$\langle (\Delta\Phi)^2 \rangle = 4r_e^2 N_0^2 \left\langle \left(\frac{\delta N}{N_0} \right)^2 \right\rangle \lambda^2 H^2 \sec \chi. \quad (14)$$

Here r_e is the radius of electron. For numerical computations we take $H = 100$ km and $N_0 = 10^{11}$ m⁻³. In Figs. 7, *a* and *b* the RMS of the signal phase fluctuations, $\sqrt{\langle (\Delta\Phi)^2 \rangle}$, [in radian], are shown versus the zenith angle of a satellite for various frequen-

cies from 32 MHz (HF-band) to 60 GHz (UHF-band) and for different $\delta N/N_0$ in percentages to the total plasma content of $N_0 = 10^{11}$ m⁻³, from 1 % (weak magnetic storm) to 10 % (strong magnetic storm), respectively.

Here, we should mention that the frequency band of signals that we investigate covers whole spectra of useful frequencies operating in existing land-ionosphere-land communication links, GPS satellite links and also in new networks beyond 3-G (third generation), operating at frequencies of 10 GHz to 60 GHz. As is clearly seen from the illustrations presented, the frequency dependence of RMS fluctuations of the signal phase is sufficient only for zenith angles greater 60–65° for mean and large perturbations of the background ionospheric plasma, i. e. 1 to 10 %. With increase of the radiated frequency, the phase fluctuations become strongly depending on $\delta N/N_0$. Thus, for the frequency band from 1 to 10 GHz usually used in satellite communications, for zenith angles of 50–60°, the RMS of phase fluctuations increases from 1–5 rad ($\delta N/N_0 = 0.1$ %) to 20 rad ($\delta N/N_0 = 1$ %) and 100 rad ($\delta N/N_0 = 10$ %).



■ Fig. 7. RMS of signal phase fluctuation vs. zenith angle, for $\delta N/N_0 = 0.01$ (*a*) and for $\delta N/N_0 = 0.1$ (*b*) and for frequencies varied from 32 MHz to 60 GHz

Effects of Fast Fading on the Radio Signal

Based now on the magnetic storm effects experimentally observed during sounding of the disturbed stormtime ionosphere, we will analyze the effects of fading of radio signals determined by the K -factor of fading, defined above in Section “Main Parameters of the Radio Channel”.

For these purposes, we will use relations between the parameter of signal intensity scintillations σ_I^2 and the K -factor described by formula (3), as parameter of fast fading phenomena within the channel [15, 21]. Taking into account variations of scintillation parameter, $S_4 \equiv \sigma_I^2$, observed experimentally in [1, 2], from 0.4 to 0.8 (see also Figs. 1, 2), we get that the corresponding K -parameter varies in the range of 1.1 to 1.6, indicating the existence of direct visibility between the ground-based and satellite antenna. In the above cases the LOS component is at the same order or higher than the NLOS component.

In other words, the coherent component of signal intensity is accompanied by additional effects of multipath phenomena (i. e., by the incoherent component) caused by diffractive scattering of radio signals on small- and moderate-scale plasma density irregularities $\delta N/N_0$, which exist in stormtime ionosphere (see Figs. 1–3). This implies that the coherent (LOS) component exceeds the incoherent (NLOS) component even in stormtime perturbed ionospheric links. As was shown in [15, 21], knowing the K -factor of fast fading, one can predict deviations of the signal data passing through the disturbed ionospheric radio channel.

Conclusions

In this work, we presented theoretical predictions of the signal intensity and phase fluctuations of radio signals propagating in the stormtime subauroral ionosphere. The background ionospheric parameters and density irregularities used in the calculations are taken from the DMSP satellite observations during several magnetic storms that include two events where coincident intense radio-signal scintillations were observed (see first Section and the corresponding bibliography in [1–9]).

Based on experimental observations carried out in [1–9] and summarized in [21], as well as on theoretical framework developed in [15–20, 23–27, 31–34], the following phenomena can be predicted using these results.

1. *Deviation of plasma density*: weak storm — $\delta n / n_0 = 2 \div 5\%$; strong storm — $\delta n / n_0 = 10 \div 20\%$.

2. *Attenuation of signal amplitude*: the change is 3 times from weak to strong magnetic storm (see

Fig. 4); the tendency which does not depend on grazing angle. With decrease of grazing angle, for frequencies up to 600 MHz, the attenuation of radio signal increases.

3. *Signal phase fluctuations*: from tens to hundreds radian from weak to strong magnetic storm (see Fig. 3).

4. *Scintillation index deviations*: from 0.4 to 0.8 (for $p = 2$) with respect to the non-perturbed ionosphere (see Fig. 6).

5. *Effect of fast fading*: $K \in [0.8 \div 1.6]$; I_{co} is at the same order or slightly higher than I_{inco} ; the Ricean law can be used for description of such radio channels.

Finally, the knowledge of the main parameters of land-satellite communication links passing through the perturbed stormtime ionosphere allows estimates the main parameters of signal data, such as the capacity, spectral efficiency, and bit error rate. These aspects needs additional serious theoretical analysis and will be investigated in future.

References

- Kintner P., and B. Ledvina. The Ionosphere, Radio Navigation, and Global Navigation Satellite Systems. *Adv. Space Res.*, 2005, vol. 35, pp. 788–811.
- Basu Su., et al. Ionospheric Effects of Major Magnetic Storms During the International Space Weather Period of September and October 1999: GPS Observations, VHF/UHF Scintillations and in Situ Density Structures at Middle and Equatorial Latitudes. *J. Geophys. Res.*, 2001, vol. 106, pp. 389–399.
- Ledvina B. M., Makela J. J., and Kintner P. M. First Observations of Intense GPS L1 Amplitude Scintillations at Midlatitude. *Geophys. Res. Lett.*, 2002, vol. 29, pp. 1659–1662.
- Erickson P., Foster J., Holt and J. Inferred Electric Field Variability in the Polarization Jet from Millstone Hill E Region Coherent Scatter Radar Observations. *Radio Sci.*, 2002, vol. 37, iss. 2, pp. 1027–1036. doi:10.1029/2000RS002531
- Mishin E., Burke W., Huang C., and Rich F. Electromagnetic Wave Structures Within Subauroral Polarization Streams. *J. Geophys. Res.*, 2003, vol. 108, pp. 1309–1315.
- Mishin E., Burke W., and Viggiano A. Stormtime Subauroral Density Troughs: Ion-Molecule Kinetics Effects. *J. Geophys. Res.*, 2004, vol. 109, A10301. doi:10.1029/2004JA010438
- Mishin E. V., and Burke W. J. Stormtime Coupling of the Ring Current, Plasmasphere and Topside Ionosphere: Electromagnetic and Plasma Disturbances. *J. Geophys. Res.*, 2005, vol. 110, pp. 7209–7216, A07209.
- Foster J., and Burke W. A New Categorization for Subauroral Electric Fields. *EOS Trans. AGU*, 2002, vol. 83, pp. 393–401.
- Mishin E., and Blaunstein N. Irregularities Within Subauroral Polarization Stream-Related Troughs and GPS Radio Interference at Midlatitudes. In: *Midlatitude Ionospheric Dynamics and Disturbances* (eds. P. Kintner, A. Coster, T. Fuller-Rowell, A. Mannucci, M. Mendillo, and R. Heelis). American Geophysical Union, Washington, D. C., 2008. Pp. 291–295. doi:10.1029/181GM26
- Foster J., and Rich F. Prompt Midlatitude Electric Field Effects During Severe Magnetic Storms. *J. Geophys. Res.*, 1998, vol. 103, pp. 26367–26373.
- Hardy D., Schmidt L., Gussenhoven M., Marshall F., Yeh H., Shumaker T., Huber A., and Pantazis J. Precipitating Electron and Ion Detectors (SSJ/4) for Block 5D/Flights 4-10 DMSP Satellites: Calibration and Data Presentation. *Tech. Rep., AFGL-TR-84-0317*, Air Force Geophys. Lab., Hanscom Air Force Base, Mass., 1984.
- Maynard N., Burke W., Basinska E., Erickson G., Hughes W., Singer H., Yahnin A., Hardy D., and Mozer F. Dynamics of the Inner Magnetosphere Near Times of Substorm Onsets. *J. Geophys. Res.*, 1996, vol. 101, pp. 7705–7715.
- Rich F. J., and Hairston M. Large-scale Convection Patterns Observed by DMSP. *J. Geophys. Res.*, 1994, vol. 79, pp. 3827–3835.
- Saunders S. R. *Antennas and Propagation for Wireless Communication Systems*. New York, John Wiley & Sons, 1999. P. 409.
- Blaunstein N., and Christodoulou Ch. G. *Radio Propagation and Adaptive Antennas for Wireless Communication Networks: Terrestrial, Atmospheric and Ionospheric*. Ed. 2. John Wiley & Sons, 2014. 704 p.
- Booker H. G., and Gordon W. E. A Theory of Radio Scattering in the Ionosphere. *Proc. IRE*, 1950, vol. 38, pp. 400–412.

17. Booker H. G., Ratcliffe S. A., and Shinn D. H. Diffraction from an Irregular Screen with Applications to Ionospheric Problems. *Philos. Trans. Roy. Soc. London, Ser. A.*, 1950, vol. 242, pp. 579–607.
18. Booker H. G. A Theory of Scattering by Non-isotropic Irregularities with Application to Radar Reflection from the Aurora. *J. Atmos. Terr. Phys.*, 1956, vol. 8, pp. 204–221.
19. Booker H. G., and Majidi Ahi G. Theory of Refractive Scattering in Scintillation Phenomena. *J. Atmos. Terr. Phys.*, 1981, vol. 43, pp. 1199–1214.
20. Booker H. G. Application of Refractive Scintillation Theory to Radio Transmission through the Ionosphere and the Solar wind and to Reflection from a Rough Ocean. *J. Atmos. Terr. Phys.*, 1981, vol. 43, pp. 1215–1233.
21. Blaunstein N., and Plohotniuc E. *Ionosphere and Applied Aspects of Radio Communication and Radar*. London-New York, CRC Press, 2008. 599 p.
22. Blaunstein N., and Christodoulou Ch. *Radio Propagation and Adaptive Antennas for Wireless Communication Links: Terrestrial, Atmospheric and Ionospheric*. Wiley & Sons, 2007. 614 p.
23. Farley D. T. Incoherent Scatter Radar Probing. In: *Modern Ionospheric Science*. Ed. by H. Kohl, R. Rustner, and K. Schlegel. Katlenburg-Lindau, Copernicus GmbH, 1996. Pp. 415–439.
24. Denisov N. G., and Erukhimov L. M. Statistical Properties of Phase Fluctuations During Total Reflection from a Layer. *Geomagn. and Aeronomy*, 1966, vol. 5, pp. 695–702.
25. Erukhimov L. M., and Rizhkov V. A. Study of Focusing Ionospheric Irregularities by Methods of Radio-Astronomy at Frequencies of 13–54 MHz. *Geomagn. and Aeronomy*, 1971, vol. 5, pp. 693–697.
26. Erukhimov L. M., Komrakov G. P., and Frolov V. L. About the Spectrum of the Artificial Small-Scale Ionospheric Turbulence. *Geomagn. and Aeronomy*, 1980, vol. 20, pp. 1112–1114.
27. Gajlit T. A., Gusev V. D., Erukhimov L. M., and Shpiro P. I. Spectrum of the Phase Fluctuations at the Ionosphere Sounding. *Izvestija vuzov. Radiofizika*, 1983, vol. 26, no. 7, pp. 795–801 (In Russian).
28. Backley R. Diffraction by Random Phase Screen with Very Large RMS Phase Deviation. Two-dimensional screen. *Austral. J. Phys.*, 1971, vol. 24, pp. 373–396.
29. Titheridge J. E. The Diffraction of Satellite Signals by Isolated Ionospheric Irregularities. *J. Atmos. Terr. Phys.*, 1971, vol. 33, pp. 47–69.
30. Crain C. M., Booker H. G., and Fergusson S. A. Use of Refractive Scattering to Explain SHF Scintillations. *Radio Sci.*, 1974, vol. 14, pp. 125–133.
31. Wernik A. W., and Liu C. H. Application of the Scintillation Theory to Ionospheric Irregularities Studies. *J. Artificial Satellites*, 1975, vol. 10, pp. 37–58.
32. Rino C. L., and Fremouw E. J. The Angle Dependence of Single Scattered Wave Fields. *J. Atmos. Terr. Phys.*, 1977, vol. 39, pp. 859–868.
33. Knepp D. L. Multiple Phase-Screen Calculation of the Temporal Behavior of Stochastic Waves. *Proc. IEEE*, 1983, vol. 71, pp. 722–737.
34. Rino C. L. On the Application of Phase Screen Models to the Interpretation of Ionospheric Scintillation Data. *Radio Sci.*, 1982, vol. 17, pp. 855–867.
35. Gurevich A. V., and Tsedilina E. E. *Sverkhdal'nee rasprostranenie korotkikh radiovoln* [Long-Range Propagation of Short Waves]. Moscow, Nauka Publ., 1979. 344 p. (In Russian).
36. Burke W., Rubin A., Maynard M., et al. Ionospheric Disturbances Observed by DMSF at Mid to low Latitudes During Magnetic Storm of June 4–6, 1991. *J. Geophys. Res.*, 2000, vol. 105, pp. 18391–19402. doi:10.1029/1999JA000188
37. Shkarofsky I. P., Johnston T. W., and Bachynski M. P. *The Particle Kinetics of Plasmas*. London, Addison-Wesley, 1966. 518 p.
38. Gurevich A. V. *Nonlinear Phenomena in the Ionosphere*. Berlin, Springer-Verlag, 1978. 465 p.
39. Gel'berg M. G. *Neodnorodnosti vysokoshirotnoi ionosfery* [Inhomogeneities of High-Latitude Ionosphere]. Novosibirsk, Nauka Publ., 1986. 193 p. (In Russian).
40. Silin V. P. *Vvedenie v kineticheskuiu teoriyu gazov* [Introduction in the Kinetic Theory of Gases]. Moscow, Nauka Publ., 1971. 332 p. (In Russian).

УДК 654.165

doi:10.15217/issn1684-8853.2018.1.85

Прогнозирование эксплуатационных параметров прохождения радиосигналов в канале «Земля — спутник» через возмущенную ионосферу

Блаунштейн Н. Ш.^а, доктор физ.-мат. наук, профессор, nathan.blaunstein@hotmail.comБен-Шимол И.^а, PhD, техн., старший преподаватель, benshimo@bgu.ac.il^аНегевский университет им. Бен-Гуриона, П.О.Б. 653, Бен-Гуриона ул., 1, г. Беэр-Шева, 74105, Израиль

Введение: тема исследования стала актуальной в последнее время ввиду глобализации систем беспроводной связи за счет использования дальнего ионосферного распространения радиоволн в каналах связи «Земля — спутник». **Цель:** анализ ключевых параметров ультракоротких радиоволн в канале «Земля — спутник», которые определяют эффекты фединга при магнитных бурях, происходящих в среднеширотной ионосфере. **Результат:** при возмущенных параметрах плазмы, вызванных экспериментально магнитным штормом, проанализированы поглощение и фазовые флуктуации радиосигналов. На основе экспериментальных дан-

ных исследований индекс сцинтилляции σ_f^2 , соответствующий сцинтилляциям сигналов в многолучевых каналах связи с федингом. Показано, что ~10 % нерегулярностей плотности плазмы во время шторма в области F вызывают сильный фединг высокочастотных радиосигналов диапазона ультракоротких радиоволн со значительными флуктуациями интенсивности сигнала (до 1 %) и изменениями фазы сигнала (до сотен радиан). **Практическая значимость:** полученные результаты позволяют прогнозировать мультипликативный фединг в ионосферных субканалах и его вклад в определение потерь в полном канале связи «Земля — спутник».

Ключевые слова — ионосферный субканал, ионосферные нерегулярности плотности плазмы, индекс сцинтилляции интенсивности сигнала, затухание амплитуды, флуктуации фазы, среднеквадратическая величина, радиоволны УКВ-диапазона.

Цитирование: Blaunstein N., Ben-Shimol Y. Prediction of Operational Parameters of Radio Signals Passing a Land-Satellite Link through StormTime Ionosphere // Информационно-управляющие системы. 2018. № 1. С. 85–95. doi:10.15217/issn1684-8853.2018.1.85
Citation: Blaunstein N., Ben-Shimol Y. Prediction of Operational Parameters of Radio Signals Passing a Land-Satellite Link through StormTime Ionosphere. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 1, pp. 85–95. doi:10.15217/issn1684-8853.2018.1.85

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.

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 4, 5, 2002-2003 (*.vsd); Coreldraw (*.cdr); Excel (*.xls); Word (*.doc); AdobeIllustrator (*.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 формате *.ai, *.esp, *.wmf, *.emf, *.svg;

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

Наличие подписуточных 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>) по разным стандартам: Литература — СИБИД РФ, References — один из мировых стандартов.

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 004.932.2

doi:10.15217/issn1684-8853.2018.1.96

ПРЕОДОЛЕНИЕ ДИФРАКЦИОННОГО ПРЕДЕЛА ПРИ ОПТИЧЕСКИХ ИЗМЕРЕНИЯХ ГРАНИЦ ОБЪЕКТОВ

Ф. М. Иночкин^а, ассистент, fedor.inochkin@gmail.com

С. К. Круглов^а, канд. техн. наук, доцент, kruglovsk@gmail.com

И. Г. Бронштейн^б, директор научно-технического центра «Оптико-информационные технологии и системы», kb@jupiter.spb.ru

^аСанкт-Петербургский политехнический университет Петра Великого, Политехническая ул., 29, Санкт-Петербург, 195251, РФ

^бСанкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, Кронверкский пр., 49, Санкт-Петербург, 197101, РФ

Постановка проблемы: дифракционный предел является фундаментальным ограничением точности измерений геометрии объектов в традиционных оптических системах. Вследствие дифракции высокочастотные компоненты изображения полностью затухают, группы малоразмерных элементов границы перестают быть разрешимыми. Оценка границы объекта известными алгоритмами становится сглаженной по отношению к ее истинной форме. **Цель исследования:** разработка метода оценки границы объекта с учетом влияния дифракции на изображение. **Результаты:** для компенсации влияния дифракции предложен метод двумерной аппроксимации изображения алгоритмически вычислимой моделью. Предлагаемая модель аппроксимирует регистрируемое камерой изображение с субпиксельной точностью, учитывает двумерную форму границы, импульсный отклик оптической системы и неравномерность яркости. В основе реализации модели лежит алгоритм синтеза фурье-образа многоугольника, ограничивающего измеряемый объект. В реальных оптических системах метод позволяет скомпенсировать не только влияние дифракции, но и аберраций. Приведены результаты вычислительных экспериментов, а также результаты и методика применения в реальной оптической системе. **Практическая значимость:** результат оценки контура методом двумерной аппроксимации соответствует повышению оптического разрешения приблизительно в 2–3 раза по сравнению с известными одномерными алгоритмами определения границ.

Ключевые слова — определение границ объектов, оптическое сверхразрешение, дифракционный предел, синтезирование изображений, расчет фурье-образа, обратная задача, аппроксимационный метод, геометрические измерения.

Цитирование: Иночкин Ф. М., Круглов С. К., Бронштейн И. Г. Преодоление дифракционного предела при оптических измерениях границ объектов // Информационно-управляющие системы. 2018. № 1. С. 96–105. doi:10.15217/issn1684-8853.2018.1.96

Citation: Inochkin F. M., Kruglov S. K., Bronshtein I. G. Breaking the Diffraction Limit in Optical Edge Estimation Problem. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 1, pp. 96–105 (In Russian). doi:10.15217/issn1684-8853.2018.1.96

Введение

Оптические методы измерения геометрии объектов применяются в широком спектре задач фундаментального и прикладного характера. Совместно с современными электронными средствами считывания изображений на основе твердотельных матричных фотоприемников с миллионами светочувствительных элементов, оптические методы позволяют достичь высокой точности и динамических характеристик измерений. Вместе с тем, несмотря на технологический прогресс, у оптических методов существует фундаментальная проблема — ограничение полосы частот объекта, передаваемых оптической системой, или так называемый дифракционный предел. Следствием этого является размытие изображения объекта, что затрудняет определение его истинной геометрии. Предельная простран-

ственная частота объекта не зависит от фокусного расстояния оптической системы и определяется лишь диаметром ее зрачка и спектральным составом излучения. Так, например, для объектива идеальной оптической схемы с диаметром входного зрачка 30 мм на расстоянии 1 м предельная пространственная частота объекта составит приблизительно 15 мкм^{-1} (в видимой области спектра длин волн). Дальнейшее повышение предельной частоты возможно лишь за счет увеличения диаметра входного зрачка. Это влечет за собой рост массо-габаритных параметров оптической системы, повышение ее стоимости, ужесточение требований к условиям эксплуатации. Кроме того, увеличение диаметра зрачка приводит к пропорциональному снижению глубины резко изображаемого пространства, что в достаточно широком круге задач может приводить к недопустимым потерям информации об объекте.

В связи с этим актуальной является разработка математических и алгоритмических решений, позволяющих восстановить информацию, утраченную вследствие дифракции. Поскольку высокочастотные компоненты изображения полностью затухают, данная задача является некорректной в математическом смысле. Как следствие, для вычисления решения требуется априорная информация об объекте. При определении границ объекта такая информация заключается уже в самой постановке задачи — наблюдаемое изображение есть сигнал ступенчатого перепада яркости на границе объекта и фона. Настоящая статья посвящена методу совместного решения задачи определения границ объектов и преодоления дифракционного предела исключительно вычислительными средствами.

Влияние дифракционного предела на изображение границы объекта

В соответствии с теорией фурье-оптики интенсивность изображения протяженного объекта в некогерентном освещении описывается сверткой

$$I(x, y) = V(x, y) * h(x, y), \quad (1)$$

где (x, y) — координаты точки в плоскости изображения; V — гипотетическое идеальное изображение, представляющее собой масштабированное распределение яркости объекта. Функция h имеет смысл импульсного отклика оптической системы — так называемая функция рассеяния точки (ФРТ). По теореме о свертке выражение (1) можно перевести в частотную область:

$$F(I, f) = F(Q, f) \cdot F(h, f), \quad (2)$$

где F — оператор преобразования Фурье; f — координата в частотной области изображения. Из теории фурье-оптики также известно, что фурье-образ ФРТ ограничен размером зрачка оптической системы [1]. Для оптической системы с круглым зрачком диаметром D : $F(h, f) = 0$, $|f| > f_{\max} = D/(\lambda R)$, где λ — длина волны; R — расстояние от плоскости выходного зрачка до плоскости изображения. С учетом (2) фурье-образ изображения ограничен этой же областью пространственных частот, что и ФРТ. Из этого следуют два важных положения:

1) высокочастотные компоненты фурье-образа изображения полностью затухают;

2) максимальный шаг пространственной дискретизации изображения без дополнительной потери информации может быть определен по теореме отсчетов: $\Delta_D < (2f_{\max})^{-1}$.

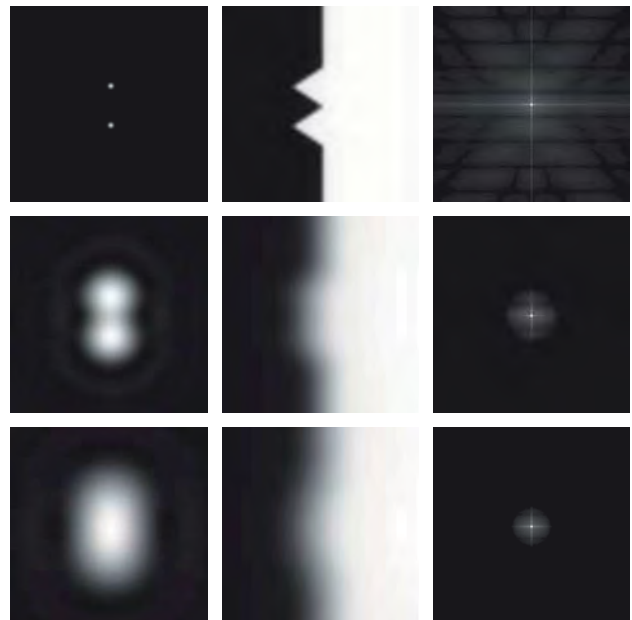
Для идеальной оптической системы без искажений волнового фронта (дифракционно-ограниченная система) ФРТ описывается функцией Эри

$$h_{Airy}(x, y) = \left(\frac{kD^2}{8R} \right)^2 \left[2 \frac{J_1(kDr/2R)}{kDr/2R} \right]^2,$$

где $k = 2\pi/\lambda$, λ — длина волны; J_1 — функция Бесселя первого рода; $r = (x^2 + y^2)^{1/2}$. В таком случае для выполнения условий теоремы отсчетов шаг пространственной дискретизации следует выбрать таким, чтобы диаметр первого темного кольца функции Эри был не менее 4,88 пикселя изображения.

Проанализируем влияние дифракции на изображение границы объекта. Предположим, что граница объекта представляет собой прямую линию за исключением пары рядом расположенных треугольных элементов (рис. 1, верхний ряд). Вследствие дифракции граница между элементами на изображении станет «размытой» (рис. 1, средний ряд). В случае если расстояние между вершинами треугольников будет меньше f_{\max}^{-1} , спад яркости между ними полностью исчезнет — пара элементов становится неразрешимой (рис. 1, нижний ряд).

Очевидно, что, помимо формирования неразрешимых групп элементов границы объекта, дифракция будет искажать результаты оценки



■ **Рис. 1.** Влияние дифракции на изображение: слева — пара точечных источников; по центру — граница объекта; справа — модуль фурье-образа изображения границы

■ **Fig. 1.** Diffraction impact on image: left — a pair of point sources; center — object edge; right — edge image Fourier transform magnitude

геометрии и одиночных элементов, причем чем меньше радиус закругления элемента, тем больше будет это влияние.

Следует отметить, что высокочастотные компоненты объекта могут быть восстановлены по множеству дифракционно-ограниченных изображений с переменными условиями освещения. Для решения этой задачи разрабатываются так называемые методы оптического сверхразрешения. Методы структурированного освещения [2] позволяют получить повышение предельной частоты изображения до двух крат при использовании осветителя той же апертуры, что и система формирования изображения [3]. Широкое семейство методов, основанных на нелинейных эффектах флуоресценции [4–7], позволяет получить теоретически неограниченное увеличение разрешения изображений, на практике достигнуто более чем 10-кратное преодоление дифракционного предела. Известны также методы синтезированной апертуры и методы фурье-птихографии (достигнуто повышение разрешения до четырех крат) [8], однако они применимы лишь для полупрозрачных объектов. Также ведутся исследования по возможности реализации сверхразрешающих оптических систем с элементами из метаматериалов [9], но результаты пока далеки от широкого применения. Из перечисленных лишь методы структурированного освещения обладают достаточной общностью для применения к широкому классу объектов, однако в силу их ограниченного потенциала проблема дифракционного предела остается нерешенной.

Методы оценки границ объектов

Для обнаружения и оценки границ на уровне пикселей изображения наиболее широкое применение нашли подходы на основе дифференцирующих операторов с последующим применением алгоритмов прослеживания границ (далее будем называть такие методы приближенными). В ряде современных приложений требуется погрешность измерения 1/100 пикселя и менее. Для решения этой задачи разрабатываются так называемые субпиксельные методы. Субпиксельные методы обычно применяются как отдельный этап обработки изображения, уточняющие результаты приближенного метода.

Субпиксельные методы можно разделить на три группы: методы интерполяции изображений, методы моментов и аппроксимационные методы. Суть методов интерполяции изображений заключается в применении приближенных методов к интерполированным отсчетам исходного изображения. Следует отметить, что для дифракционно-ограниченных изображений ошибка оценки может составить более одного пикселя уже на

этапе приближенного измерения, и дальнейшая интерполяция отсчетов не позволяет достичь существенного снижения погрешности измерения. Точность методов моментов [10–12] также в значительной степени ограничена точностью приближенного решения [13]. Наиболее точными и независимыми от результата приближенной оценки являются аппроксимационные методы. В их основе лежит идея аппроксимации распределения яркости изображения или его градиента параметризованной моделью. Положение точки границы объекта является одним из параметров модели, вычисляемым в процессе минимизации функционала ошибки аппроксимации. Для аппроксимации непосредственно распределения яркости известны модели, основанные на функции ошибок [14], функции арктангенса [15], логистической функции [16]. Для аппроксимации производной применяют функцию Гаусса [17, 18] и параболическую модель [19], при этом координата точки границы вычисляется как экстремум соответствующей функции.

Главный недостаток перечисленных выше аппроксимационных методов состоит в том, что заложенные в них модели являются одномерными. Как следствие, они не учитывают влияние соседних участков границы объекта на изображение, и полученные таким образом оценки формы границы приобретают сглаженный вид по отношению к истинной. Для компенсации влияния дифракции на результат оценки границы необходимо учитывать ее двумерную структуру. Эта задача может быть решена путем двумерной аппроксимации изображения.

Двумерный аппроксимационный метод определения границ объекта

Идея предлагаемого двумерного аппроксимационного метода заключается в том, чтобы аппроксимировать зарегистрированные камерой данные синтезированным (модельным) изображением. В отличие от одномерных методов, основным параметром модели является не отдельная точка границы, а все множество точек. Обозначим это множество как $\{P_i = (x_i, y_i), i = 1..k\}$. Тогда оценка формы границы заключается в решении обратной задачи:

$$\{P_i\} = \arg \min(E): E = \frac{1}{2} \sum_{(x,y) \in \mu} \Delta I_{x,y}(\{P_i\})^2, \quad (3)$$

где (x, y) — декартовы координаты точки растра изображения; μ — анализируемая область точек растра (в некоторой окрестности границы); ΔI — матрица разности синтезируемого и регистрируемого изображений.

Будем считать прямой задачей вычисление матрицы синтезируемого изображения \mathbf{I}_M , соответствующего заданному множеству точек $\{P_i\}$. Предлагаемая модель описывает изображение объекта с учетом наличия фона:

$$\mathbf{I}_M = \Lambda \left(\left[\hat{I}_{OB} \cdot \hat{M}_{OB} + \hat{I}_{BG} \cdot (\mathbf{I} - \hat{M}_{OB}) \right] * \hat{h} \right);$$

$$\hat{M}_{OB}(\hat{x}, \hat{y}) = \begin{cases} 1, & (\hat{x}, \hat{y}) \in \Omega \\ 0, & (\hat{x}, \hat{y}) \notin \Omega \end{cases}$$

где Λ — оператор двумерной дискретизации; \hat{I}_{OB} — распределение яркости объекта; \hat{M}_{OB} — маска объекта; \hat{I}_{BG} — распределение яркости фона; \hat{h} — импульсный отклик камеры (свертка импульсного отклика оптической системы и фотоприемника); (\hat{x}, \hat{y}) — координаты точки на плоскости изображения; Ω — область точек объекта. Для того чтобы модель могла быть реализована в виде алгоритма, необходимо перейти от непрерывного представления сигналов к дискретному. Будем искать такое представление в виде

$$\mathbf{I}_M = [\mathbf{I}_{OB} \cdot \mathbf{M}_{OB} + \mathbf{I}_{BG} \cdot (\mathbf{I} - \mathbf{M}_{OB})] * \mathbf{h}, \quad (4)$$

где \mathbf{I}_{OB} , \mathbf{M}_{OB} , \mathbf{I}_{BG} , \mathbf{h} — дискретные аналоги \hat{I}_{OB} , \hat{M}_{OB} , \hat{I}_{BG} , \hat{h} соответственно; \mathbf{I} — матрица, все элементы которой равны 1. Произведение матриц — поэлементное.

Основным вопросом, который требуется решить для синтеза модели (4), является вычисление дискретного аналога маски объекта. Поскольку компенсация дифракции связана с оценкой малоразмерных деталей границы объекта, требуются вычисления с субпиксельной точностью. Субпиксельная точность может быть достигнута либо за счет снижения шага дискретизации, либо за счет перехода к вычислениям в частотной (фурье) области. Учитывая, что фурье-образ изображения ограничен дифракцией, а информация о локализации элементов границы кодируется с теоретически неограниченной точностью фазовыми компонентами образа, вычисления в частотной области представляются более перспективным инструментом. Для простых фигур фурье-образ маски может быть вычислен аналитически. Образ произвольных фигур может быть вычислен путем разбиения всей маски на множество треугольников, при этом для каждого треугольника фурье-образ вычисляется аналитически с привлечением лишь элементарных функций [20]. Дискретный аналог маски объекта для синтеза модели (4) формируется путем обратного фурье-преобразования.

Не менее важным вопросом для реализации синтеза является определение матриц распределений яркости \mathbf{I}_{OB} и \mathbf{I}_{BG} . Их задачей является моделирование текстуры объекта и неравномерности освещения, что должно обеспечивать соот-

ветствие модели и данных. Принимая допущение, что в пределах анализируемой области изображения отсутствуют резкие перепады яркости (за исключением самой границы объекта), предложим следующую модель:

$$\begin{aligned} \mathbf{I}_{OB} &= \mathbf{Q}_{OB} * \mathbf{G}(\sigma_{OB}); \\ \mathbf{I}_{BG} &= \mathbf{Q}_{BG} * \mathbf{G}(\sigma_{BG}), \end{aligned} \quad (5)$$

где \mathbf{Q}_{OB} и \mathbf{Q}_{BG} — скрытые параметры яркости в виде матриц коэффициентов той же размерности, что и изображение; \mathbf{G} — сглаживающее ядро (например, функция Гаусса); параметры σ_{OB} и σ_{BG} определяют частотные характеристики неравномерности яркости. Наличие двух параметров позволяет моделировать различия в частотных характеристиках объекта и фона. В частности, при расфокусированном фоне параметр σ_{BG} может значительно превосходить σ_{OB} .

Перейдем к решению обратной задачи (3). В процессе решения требуется найти не только внешние параметры модели $\{P_i\}$, но и скрытые параметры \mathbf{Q}_{OB} , \mathbf{Q}_{BG} . Задача является нелинейной, и ее решение предлагается искать методами численной оптимизации. Свойства функционала позволяют применять градиентные стратегии оптимизации, такие как метод сопряженных градиентов. Для этого требуется определить способ вычисления вектора градиента.

Частные производные по скрытым параметрам распределений яркости объекта могут быть вычислены аналитически, причем сразу для всех элементов матрицы \mathbf{Q}_{OB} . Дифференцируя E с учетом (4) и (5), получим способ вычисления матрицы значений производных:

$$\frac{\partial E}{\partial (\mathbf{Q}_{OB})_{x,y}} = \left([(\Delta \mathbf{I} \cdot \mathbf{C}) \circ \mathbf{h}] \cdot \mathbf{M}_{OB} \right) \circ \mathbf{G};$$

$$\mathbf{C}_{x,y} = \begin{cases} 1, & (x, y) \in \mu \\ 0, & (x, y) \notin \mu \end{cases}$$

где символ « \circ » обозначает корреляцию. Частные производные по параметрам распределения яркости фона вычисляются аналогично, с заменой \mathbf{M}_{OB} на $(\mathbf{I} - \mathbf{M}_{OB})$.

Для частных производных по параметрам (x_i, y_i) простого аналитического решения найти не удается, требуется численное дифференцирование. Общие методы численного дифференцирования потребуют многократного решения прямой задачи, что неэффективно. Однако объем вычислений может быть значительно снижен с учетом пространственной локализации элементов границы. Запишем точное выражение для производной

$$\frac{\partial E}{\partial x_i} = \sum_{(x,y) \in \mu} \Delta \mathbf{I} \cdot \frac{\partial \mathbf{I}_M}{\partial x_i}$$

Дифференцируя (4), получим

$$\frac{\partial \mathbf{I}_M}{\partial x_i} = \left([\mathbf{I}_{OB} - \mathbf{I}_{BG}] \cdot \frac{\partial \mathbf{M}_{OB}}{\partial x_i} \right) * \mathbf{h}.$$

Аппроксимируем производную численно по двум точкам с шагом Δx_i :

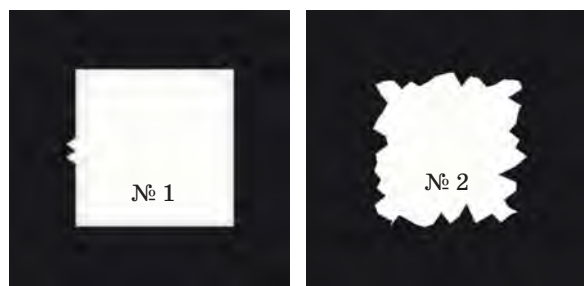
$$\frac{\partial \mathbf{I}_M}{\partial x_i} \approx \frac{1}{\Delta x_i} ([\mathbf{I}_{OB} - \mathbf{I}_{BG}] \cdot \Delta \mathbf{M}_{OB}) * \mathbf{h}.$$

Таким образом, численное дифференцирование по параметру x_i (равно как и y_i) сводится к вычислению $\Delta \mathbf{M}_{OB}$, что эквивалентно решению упрощенной прямой задачи лишь для двух треугольников.

Оценка эффективности двумерного аппроксимационного метода

Рассмотрим эффективность работы метода в вычислительных и физическом экспериментах. Модельные тест-объекты представлены на рис. 2. Объект № 1 служит для демонстрации способности обнаруживать и оценивать геометрию простых элементов, неразрешимых вследствие дифракционного предела. Объект № 2 моделирует сложный контур со множеством неровностей и позволяет оценить работу метода для шероховатых границ объектов. Изображения подобных границ могут быть получены, например, при наблюдении микродефектов механической обработки деталей.

В качестве альтернативного метода для сравнения результатов работы выбран метод одномерной аппроксимации границы функцией ошибок [14]. Начальное решение как предлагаемого, так и альтернативного методов получено алгоритмом прослеживания максимумов модуля градиента изображения. Условием останова процедуры двумерной аппроксимации является снижение величины L_2 -нормы шага до значения 10^{-3} пикселей. Шаг численного дифференцирования



■ **Рис. 2.** Идеальные изображения тест-объектов
 ■ **Fig. 2.** Ideal images of test objects

по координатам точек выбран равным 10 % от длины прилегающих к точке отрезков границы. Анализируемая область μ — полоса шириной 10 пикселей по обе стороны от границы объекта. В качестве сглаживающего ядра компонентов распределений яркости использована функция Гаусса с $\sigma_{BG} = \sigma_{OB} = 10$ пикселям (если не указано другое).

Вычислительные эксперименты

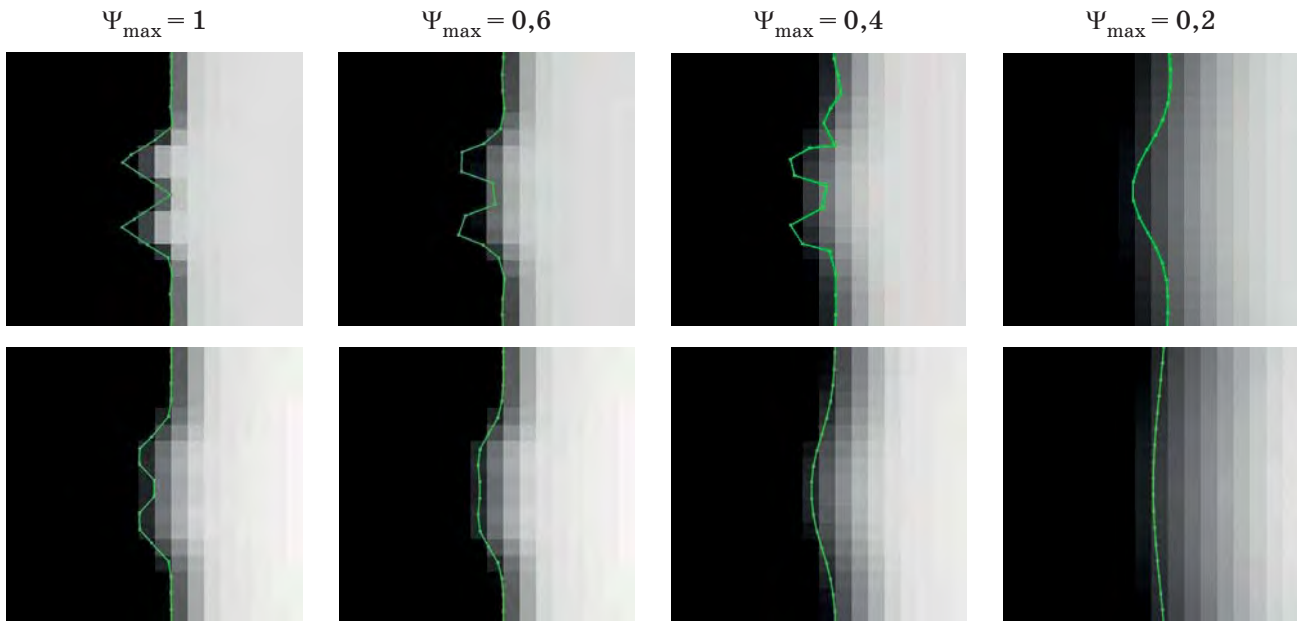
Продемонстрируем способность метода обнаруживать неразрешимые элементы границы. Увеличенные фрагменты смоделированных дифракционно-ограниченных изображений тест-объекта № 1 с результатами работы алгоритмов приведены на рис. 3. Расстояние между вершинами элементов — 4 пикселя. Для описания влияния дифракции введен параметр относительной предельной частоты изображения $\Psi_{\max} = 2\Delta_D f_{\max}$, где Δ_D — шаг раstra изображения. ФРТ оптической системы моделировалась функцией Эри с диаметром первого темного кольца $4,88/\Psi_{\max}$ (пиксель).

Двумерная аппроксимация позволяет обнаружить неразрешимую пару элементов вплоть до предельной частоты $\Psi_{\max} \approx 0,3$. Одномерная аппроксимация разрешает те же элементы лишь для области $\Psi_{\max} > 0,6$, при этом даже в этой области форма границы сглажена по отношению к истинной. Оценим это сглаживание численно с помощью меры ошибки оценки границы (отклонения от ее истинной формы) в следующем виде:

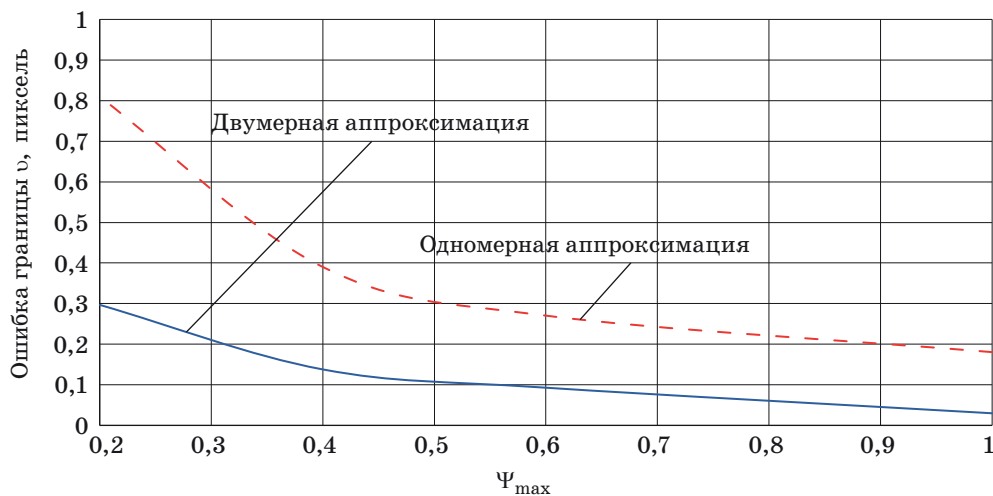
$$v = \sqrt{\frac{1}{N_1 + N_2} \left(\sum_{i=1}^{N_1} \left[\min_{k=1..N_2} \rho(P_{1i}; L_{2k}) \right]^2 + \sum_{k=1}^{N_2} \left[\min_{i=1..N_1} \rho(P_{2k}; L_{1i}) \right]^2 \right)},$$

где N_1 и N_2 — количество точек; P_{1i} и P_{2k} — множества точек границ; L_{1i} и L_{2k} — множества отрезков границ (все — для первой и второй границ соответственно); ρ — оператор вычисления расстояния между точкой и отрезком. Данная мера инвариантна к смещению точек вдоль отрезков, лежащих на одной прямой, поскольку это не изменяет форму границы. Зависимость ошибки оценки границы от относительной предельной частоты показана на рис. 4 (приведены данные для объекта № 1). Ошибка оценки границы соответствует 2–3-кратному увеличению оптического разрешения по отношению к той же оценке, полученной одномерным методом.

Эффективность методов, основанных на моделях, в значительной степени определяется адекватностью модели экспериментальным данным.



■ **Рис. 3.** Способность метода обнаруживать неразрешимые элементы границы: верхний ряд — двумерная аппроксимация; нижний ряд — одномерная аппроксимация
 ■ **Fig. 3.** Method ability to detect unresolvable edge details: top row — two-dimensional approximation (proposed); bottom row — one-dimensional approximation

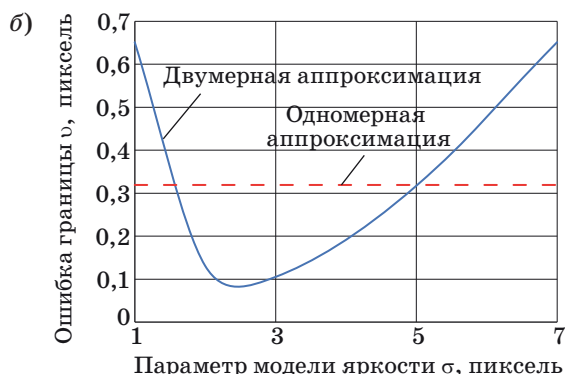
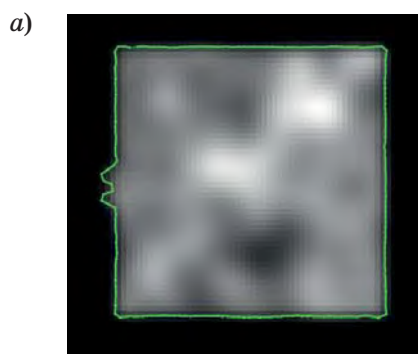


■ **Рис. 4.** Зависимость ошибки оценки границы от относительной предельной частоты
 ■ **Fig. 4.** Edge estimation error as a function of relative cutoff frequency

Приведенные выше оценки получены в условиях равномерного распределения яркости объекта и фона. Однако реальные объекты могут быть подвержены неоднородному освещению и наличию текстуры. Наибольшую трудность при этом представляют изображения объектов с текстурами. В таких случаях адекватность модели обеспечивается корректным выбором параметров σ_{OB} и σ_{BG} в компонентах (5). Пример работы метода для изображения с текстурой приведен на рис. 5 (для

значения $\Psi_{max} = 0,6$). Текстура на рис. 5, а смоделирована белым шумом, сглаженным гауссовым фильтром с $\sigma = 3$ пикселям, величина относительной модуляции текстуры 20 %.

При значениях $\sigma < 2$ (рис. 5, б) соответствие модели и данных обеспечивается за счет высокочастотных составляющих распределений яркости модели, при этом параметры границы объекта неустойчивы. При значениях $\sigma > 5$ модель не может достаточно хорошо аппроксимировать



■ **Рис. 5.** Пример работы метода для изображения с текстурой (а) и зависимость ошибки оценки границы от параметра модели (б)

■ **Fig. 5.** Example of method application to a textured image (а), edge estimation error as a function of the model parameter (б)

данные, и точки границы приобретают смещение по отношению к их истинному положению. При оптимальном выборе параметров метод обеспечивает такое же качество решения, как и для объекта без текстуры.

Физический эксперимент

В реальных оптических системах на изображение влияет не только дифракция, но и аберрации, вызванные недостатками оптической схемы или технологии изготовления. В таком случае ФРТ оптической системы может принимать вид, существенно отличающийся от функции Эри. При достаточно больших аберрациях модель изображения (4) не обеспечивает соответствие регистрируемому сигналу, что снижает качество решения и сходимость метода. Решением проблемы является измерение реальной ФРТ камеры, например, путем регистрации изображения точечного источника.

Основные параметры эксперимента приведены в табл. 1, а результаты — в табл. 2. Параметры подобраны таким образом, что пара элементов границы объекта № 1 оказывается неразрешимой. Результаты работы метода для реальной оптической системы с аберрациями представлены на рис. 6, а-ж. Изображение точечного источника, примененное в процедуре синтеза, показано на рис. 6, б в том же масштабе, что и изображение объекта. В качестве точечного источника использовано отверстие диаметром 1 мм.

Как и в вычислительном эксперименте, метод позволяет обнаружить и оценить неразрешимые элементы границы (рис. 6, е). На рис. 6, в приведены синтезированные в результате работы метода изображения (I_M), а также разность зарегистрированного и синтезированного изображений (ΔI , разностный сигнал усилен в 100 раз). На разностных изображениях преобладает шум,

■ **Таблица 1.** Параметры физического эксперимента

■ **Table 1.** Physical experiment parameters

Параметр	Значение
Относительное отверстие	f/16
Расчетный диаметр диска Эри: изображение/объект	10 мкм/5 мм
Ψ_{\max}	0,5
Фотоприемник	КМОП МТ9Р031, монохромный
Шаг пикселя	2,2 мкм
Расстояние до объекта	12 м
Масштаб изображений	1 мм/пиксель
Размеры изображений	100 × 100 пикселей

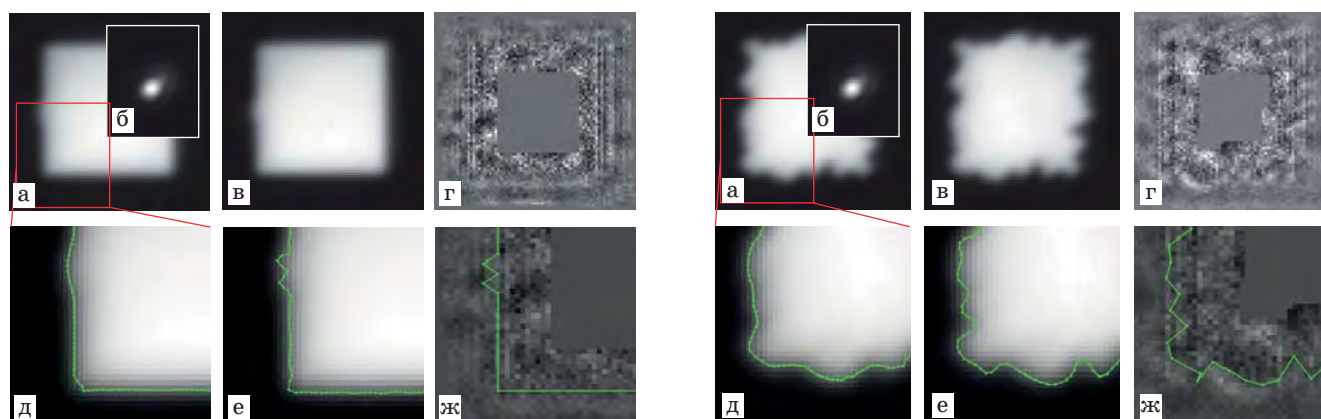
■ **Таблица 2.** Ошибка оценки границы в физическом эксперименте

■ **Table 2.** Physical experiment edge estimation error

Метод	Объект № 1	Объект № 2	
Одномерная аппроксимация	0,41	0,67	
Двумерная аппроксимация	С измеренной ФРТ	0,13	0,26
	Функция Эри	0,35	0,62

демонстрируя хорошее соответствие модели (4) реальным экспериментальным данным.

По сравнению с альтернативным методом, ошибка оценки границы снижается в 2,9 и 2,4 раза для объектов № 1 (слева) и 2 (справа), что соответствует результатам вычислительного экс-



■ **Рис. 6.** Пример работы метода с реальной оптической системой: *a* — изображение объекта; *б* — изображение точечного источника; *в* — синтезированное изображение; *г* — разностное изображение; *д* — одномерная аппроксимация; *е* — двумерная аппроксимация; *ж* — истинная граница

■ **Fig. 6.** Example of method application to real optical system: *a* — acquired image; *б* — point source image; *в* — reconstructed image; *г* — difference image; *д* — one-dimensional approximation; *е* — two-dimensional approximation; *ж* — ground-truth

перимента. При этом использование измеренной, а не теоретической ФРТ является существенным фактором эффективного применения метода.

Заключение

Предложенный метод двумерной аппроксимации изображения является эффективным алгоритмическим инструментом преодоления дифракционного предела оптических систем в задачах измерения геометрии объектов. За счет аппроксимации двумерного распределения яркости, а также учета влияния фактической ФРТ оптической системы метод позволяет скомпенсировать вызванное дифракцией и аберрациями сглаживание изображения.

Метод является итерационным, и качество решения зависит от стратегии численной минимизации функционала ошибки аппроксимации. Свойства функционала позволяют реализовать вычислительно-эффективные градиентные стратегии. Эффект применения метода соответствует приблизительно 2–3-кратному повышению оптического разрешения для одномерных методов и, в частности, позволяет оценить форму неразрешимых элементов границы.

Результатом работы метода двумерной аппроксимации является не только оценка формы границы, но и синтезированное изображение объекта. Это дает возможность верификации результатов работы алгоритма путем анализа соответствия синтезированного и зарегистрированного изображений. Также таким образом можно проводить косвенные оценки состояния оптической системы, объекта наблюдения, условий его освещения.

Недостатком метода является увеличение ошибки оценки в случае несоответствия модели и экспериментальных данных. В частности, наличие текстуры объекта требует адекватного выбора модели распределения яркости, а аберрации в оптической системе требуют измерения ее фактической ФРТ. Другой недостаток — большой объем вычислений, связанный со сложностью решения прямой задачи. Поиск решения на типичном современном процессоре в один поток исполнения занимает несколько десятков секунд. Отметим, что предложенное решение прямой задачи на основе прямого расчета фурье-образа эффективно поддается параллелизации для вычислений на графических процессорах, в таком случае время поиска решения может быть сокращено на один-два порядка.

Литература

1. Goodman J. W. Introduction to Fourier Optics. — Roberts and Company Publishers, 2005. — 491 p.
2. Lukosz W. Optical Systems with Resolving Powers Exceeding the Classical Limit II // Journal of the Op-

tical Society of America. 1967. Vol. 57. N 7. P. 932–941. doi:10.1364/josa.57.000932

3. Gustafsson M. G. L. Surpassing the Lateral Resolution Limit by a Factor of Two using Structured Illumination Microscopy // Journal of Microscopy. 2000. Vol. 198. N 2. P. 82–87. doi:10.1046/j.1365-2818.2000.00710.x

4. Hell S. W., Wichmann J. Breaking the Diffraction Resolution Limit by Stimulated Emission: Stimulated-Emission-Depletion Fluorescence Microscopy // *Optics Letters*. 1994. Vol. 19. N 11. P. 780. doi:10.1364/ol.19.000780
5. Betzig E., Patterson G. H., Sougrat R., Lindwasser O. W., Olenych S., Bonifacino J. S., Davidson M. W., Lippincott-Schwartz J., Hess H. F. Imaging Intracellular Fluorescent Proteins at Nanometer Resolution // *Science*. 2006. Vol. 313. N 5793. P. 1642–1645. doi:10.1126/science.1127344
6. Rust M. J., Bates M., Zhuang X. Sub-Diffraction-Limit Imaging by Stochastic Optical Reconstruction Microscopy (STORM) // *Nature Methods*. 2006. Vol. 3. N 10. P. 793–796. doi:10.1038/nmeth929
7. Hess S. T., Girirajan T. P. K., Mason M. D. Ultra-High Resolution Imaging by Fluorescence Photoactivation Localization Microscopy // *Biophysical Journal*. 2006. Vol. 91. N 11. P. 4258–4272. doi:10.1529/biophysj.106.091116
8. Zheng G., Horstmeyer R., Yang C. Wide-Field, High-Resolution Fourier Ptychographic Microscopy // *Nature Photonics*. 2013. Vol. 7. N 9. P. 739–745. doi:10.1038/nphoton.2013.187
9. Dong X. H., Wong A. M. H., Kim M., Eleftheriades G. V. Superresolution Far-Field Imaging of Complex Objects using Reduced Superoscillating Ripples // *Optica*. 2017. Vol. 4. N 9. P. 1126. doi:10.1364/optica.4.001126
10. Lyvers E. P., Mitchell O. R., Akey M. L., Reeves A. P. Subpixel Measurements using a Moment-Based Edge Operator // *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 1989. Vol. 11. N 12. P. 1293–1309. doi:10.1109/34.41367
11. Ghosal S., Mehrotra R. Orthogonal Moment Operators for Subpixel Edge Detection // *Pattern Recognition*. 1993. Vol. 26. N 2. P. 295–306. doi:10.1016/0031-3203(93)90038-x
12. Xiang F., Wang Z., Yuan X. Subpixel Edge Detection: An Improved Zernike Orthogonal Moments Method // 2013 5th Intern. Conf. on Intelligent Human-Machine Systems and Cybernetics. Hangzhou, 2013. P. 19–22. doi:10.1109/ihmsc.2013.11
13. Fabijańska A. Subpixel Edge Detection in Blurry and Noisy Images // *International Journal of Computer Science and Applications*. 2015. N 12(2). P. 1–19.
14. Hagara M., Kulla P. Edge Detection with Sub-Pixel Accuracy Based on Approximation of Edge with Erf Function // *Radioengineering*. 2011. N 20(2). P. 516–524.
15. Sun Q., Hou Y., Tan Q. A Subpixel Edge Detection Method Based on an Arctangent Edge Model // *Optik — International Journal for Light and Electron Optics*. 2016. Vol. 127. N 14. P. 5702–5710. doi:10.1016/j.ijleo.2016.03.058
16. Chen P., Chen F., Han Y., Zhang Z. Sub-Pixel Dimensional Measurement with Logistic Edge Model // *Optik — International Journal for Light and Electron Optics*. 2014. Vol. 125. N 9. P. 2076–2080. doi:10.1016/j.ijleo.2013.10.020
17. Shang Y., Chen J., Tian J. Sub-Pixel Edge Detection Algorithm Based on Gauss Fitting // *Journal of Computer Applications*. 2011. Vol. 31. N 1. P. 179–181. doi:10.3724/sp.j.1087.2011.00179
18. Fabijańska A. Gaussian-Based Approach to Subpixel Detection of Blurred and Unsharp Edges // *Proc. of the 2014 Federated Conf. on Computer Science and Information Systems*. Warsaw, 2014. P. 641–650. doi:10.15439/2014F136
19. Bailey D. G. Sub-Pixel Profiling // 2005 5th Intern. Conf. on Information Communications & Signal Processing. Bangkok, 2005. P. 1311–1315. doi:10.1109/ICICS.2005.1689268
20. Inochkin F. M., Kruglov S. K., Bronshtein I. G., Kompan T. A., Kondratjev S. V., Korenev A. S., Pukhov N. F. Subpixel Edge Estimation with Lens Aberrations Compensation Based on the Iterative Image Approximation for High-Precision Thermal Expansion Measurements of Solids // *Optical Measurement Systems for Industrial Inspection X: Proc. SPIE*. Munich, 2017. Vol. 10329. doi:10.1117/12.2270204

UDC 004.932.2

doi:10.15217/issn1684-8853.2018.1.96

Breaking the Diffraction Limit in Optical Edge Estimation ProblemInochkin F. M.^a, Assistant Professor, fedor.inochkin@gmail.comKruglov S. K.^a, PhD, Tech., Associate Professor, kruglovsk@gmail.comBronshtein I. G.^b, Head of a Research Center, kb@jupiter.spb.ru^aPeter the Great St. Petersburg Polytechnic University, 29, Polytechnicheskaya St., 195251, Saint-Petersburg, Russian Federation^bSaint-Petersburg National Research University of Information Technologies, Mechanics and Optics, 49, Kronverkskii Pr., 197101, Saint-Petersburg, Russian Federation

Introduction: Diffraction is a fundamental limit for object geometry measurement accuracy in conventional optical systems. Due to diffraction, high-frequency components of an image are fully attenuated, and groups of small-sized details become non-resolvable. Edge estimation results obtained by known algorithms become smoothed compared against actual edge geometry. **Purpose:** Development of a method to estimate an object edge with a compensation of the diffraction impact on the image. **Results:** To compensate for the

diffraction, a method of 2D image approximation by a numerical model is proposed. The proposed model reconstructs a camera-acquired image with subpixel accuracy, taking into account the optical system's 2D point spread function along with the non-uniformity of the object intensity. The model implementation is based on the Fourier image synthesis algorithm for the measured object's bounding polygon. In real optical systems, the method provides the opportunity to compensate not only for the diffraction but also for lens aberrations by means of point spread function measurement. Numerical simulation results are presented, along with the results and technique for a real system application. **Practical relevance:** The edge estimation error for 2D image approximation corresponds to 2- or 3-fold improvement in optical resolution as compared to the conventional 1D edge estimation algorithms.

Keywords — Edge Detection, Optical Superresolution, Diffraction Limit, Image Reconstruction, Fourier Image Calculation, Inverse Problem, Approximation Method, Object Geometry Measurements.

Citation: Inochkin F. M., Bronshtein I. G. Breaking the Diffraction Limit in Optical Edge Estimation Problem. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 1, pp. 96–105 (In Russian). doi:10.15217/issn1684-8853.2018.1.96

References

- Joseph W. Goodman. *Introduction to Fourier Optics*. Roberts and Company Publishers, 2005. 491 p.
- Lukosz W. Optical Systems with Resolving Powers Exceeding the Classical Limit II. *Journal of the Optical Society of America*, 1967, vol. 57, no. 7, pp. 932–941. doi:10.1364/josa.57.000932
- Gustafsson M. G. L. Surpassing the Lateral Resolution Limit by a Factor of Two using Structured Illumination Microscopy. *Journal of Microscopy*, 2000, vol. 198, no. 2, pp. 82–87. doi:10.1046/j.1365-2818.2000.00710.x
- Hell S. W., Wichmann J. Breaking the Diffraction Resolution Limit by Stimulated Emission: Stimulated-Emission-Depletion Fluorescence Microscopy. *Optics Letters*, 1994, vol. 19, no. 11, p. 780. doi:10.1364/ol.19.000780
- Betzig E., Patterson G. H., Sougrat R., Lindwasser O. W., Olenych S., Bonifacino J. S., Davidson M. W., Lippincott-Schwartz J., Hess H. F. Imaging Intracellular Fluorescent Proteins at Nanometer Resolution. *Science*, 2006, vol. 313, no. 5793, pp. 1642–1645. doi:10.1126/science.1127344
- Rust M. J., Bates M., Zhuang X. Sub-Diffraction-Limit Imaging by Stochastic Optical Reconstruction Microscopy (STORM). *Nature Methods*, 2006, vol. 3, no. 10, pp. 793–796. doi:10.1038/nmeth929
- Hess S. T., Girirajan T. P. K., Mason M. D. Ultra-High Resolution Imaging by Fluorescence Photoactivation Localization Microscopy. *Biophysical Journal*, 2006, vol. 91, no. 11, pp. 4258–4272. doi:10.1529/biophysj.106.091116
- Zheng G., Horstmeyer R., Yang C. Wide-Field, High-Resolution Fourier Ptychographic Microscopy. *Nature Photonics*, 2013, vol. 7, no. 9, pp. 739–745. doi:10.1038/nphoton.2013.187
- Dong X. H., Wong A. M. H., Kim M., Eleftheriades G. V. Superresolution Far-Field Imaging of Complex Objects using Reduced Superoscillating Ripples. *Optica*, 2017, vol. 4, no. 9, p. 1126. doi:10.1364/optica.4.001126
- Lyvers E. P., Mitchell O. R., Akey M. L., Reeves A. P. Subpixel Measurements using a Moment-Based Edge Operator. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1989, vol. 11, no. 12, pp. 1293–1309. doi:10.1109/34.41367
- Ghosal S., Mehrotra R. Orthogonal Moment Operators for Subpixel Edge Detection. *Pattern Recognition*, 1993, vol. 26, no. 2, pp. 295–306. doi:10.1016/0031-3203(93)90038-x
- Xiang F., Wang Z., Yuan X. Subpixel Edge Detection: An Improved Zernike Orthogonal Moments Method. *2013 5th Intern. Conf. on Intelligent Human-Machine Systems and Cybernetics*, Hangzhou, 2013, pp. 19–22. doi:10.1109/ihmsc.2013.11
- Fabijańska A. Subpixel Edge Detection in Blurry and Noisy Images. *International Journal of Computer Science and Applications*, 2015, no. 12(2), pp. 1–19.
- Hagara M., Kulla P. Edge Detection with Sub-Pixel Accuracy Based on Approximation of Edge with Erf Function. *Radioengineering*, 2011, no. 20(2), pp. 516–524.
- Sun Q., Hou Y., Tan Q. A Subpixel Edge Detection Method Based on an Arctangent Edge Model. *Optik — International Journal for Light and Electron Optics*, 2016, vol. 127, no. 14, pp. 5702–5710. doi:10.1016/j.ijleo.2016.03.058
- Chen P., Chen F., Han Y., Zhang Z. Sub-Pixel Dimensional Measurement with Logistic Edge Model. *Optik — International Journal for Light and Electron Optics*, 2014, vol. 125, no. 9, pp. 2076–2080. doi:10.1016/j.ijleo.2013.10.020
- Shang Y., Chen J., Tian J. Sub-Pixel Edge Detection Algorithm Based on Gauss Fitting. *Journal of Computer Applications*, 2011, vol. 31, no. 1, pp. 179–181. doi:10.3724/sp.j.1087.2011.00179
- Fabijańska A. Gaussian-Based Approach to Subpixel Detection of Blurred and Unsharp Edges. *Proc. of the 2014 Federated Conf. on Computer Science and Information Systems*, Warsaw, 2014, pp. 641–650. doi:10.15439/2014F136
- Bailey D. G. Sub-Pixel Profiling. *2005 5th Intern. Conf. on Information Communications & Signal Processing*, Bangkok, 2005, pp. 1311–1315. doi:10.1109/ICICS.2005.1689268
- Inochkin F. M., Kruglov S. K., Bronshtein I. G., Kompan T. A., Kondratjev S. V., Korenev A. S., Pukhov N. F. Subpixel Edge Estimation with Lens Aberrations Compensation Based on the Iterative Image Approximation for High-Precision Thermal Expansion Measurements of Solids. *Proc. SPIE "Optical Measurement Systems for Industrial Inspection X"*, Munich, 2017, vol. 10329. doi:10.1117/12.2270204

МЕТОД ВЕЙВЛЕТ-АНАЛИЗА ПАТТЕРНОВ ДВИГАТЕЛЬНОЙ АКТИВНОСТИ НА ЭКСПЕРИМЕНТАЛЬНЫХ ДАННЫХ МНОГОКАНАЛЬНОЙ ЭЛЕКТРОЭНЦЕФАЛОГРАФИИ ЧЕЛОВЕКА ДЛЯ УПРАВЛЕНИЯ ВНЕШНИМИ УСТРОЙСТВАМИ

А. Е. Руннова^а, канд. физ.-мат. наук, доцент, anefila@gmail.com

В. А. Максименко^а, канд. физ.-мат. наук, доцент, maximenkovl@gmail.com

С. В. Пчелинцева^а, канд. техн. наук, доцент, pchelintseva@inbox.ru

Р. А. Куланин^а, научный сотрудник, rkulanin2010@yandex.ru

А. Е. Храмов^а, докт. физ.-мат. наук, профессор, hramovae@gmail.com

^аСаратовский государственный технический университет им. Гагарина Ю. А., Политехническая ул., 77, Саратов, 410054, РФ

Постановка проблемы: большой интерес современной фундаментальной и технической науки к разработке нейроинтерфейсов связан с созданием устройств как для повышения качества жизни и преодоления ограничений возможностей людей, так и в рамках исследования функционирования различных отделов головного мозга и сознания человека. Среди основных проблем современных методов расшифровки паттернов регистрируемой активности головного мозга, лежащих в основе систем управления подобными устройствами (нейроинтерфейсами), выделим низкую эффективность методов распознавания паттернов и большие затраты времени тренировки операторов нейроинтерфейсов. **Цель исследования:** разработка методов для распознавания паттернов, ассоциированных с разнообразной двигательной активностью нетренированных операторов. **Результаты:** на основе вейвлетных технологий предложены новые способы изучения разнотипных паттернов активности головного мозга, коррелируемых с различными типами двигательной активности: модифицирован в целях сокращения расчетного времени морле-базис, описаны подходы оценки энергетической мощности процессов на электроэнцефалограмме на основе расчета скелетонных характеристик. Приведены примеры обработки экспериментальных данных, зарегистрированных в группе добровольцев без тренировочного периода. Впервые выявлены характерные особенности паттернов различной двигательной активности (воображаемой и действительной, движений ног/рук). Показана универсальность обнаруженных паттернов на электроэнцефалографических данных для добровольцев, принявших участие в экспериментальных работах. **Практическая значимость:** выявленные особенности «двигательных» паттернов, а также методы обработки электроэнцефалографических сигналов могут быть использованы для построения устройств — нейроинтерфейсов, расшифровывающих электрическую активность головного мозга практически без тренировки оператора.

Ключевые слова — электроэнцефалография, паттерн, непрерывное вейвлетное преобразование, частотно-временная динамика.

Цитирование: Руннова А. Е., Максименко В. А., Пчелинцева С. В., Куланин Р. А., Храмов А. Е. Метод вейвлет-анализа паттернов двигательной активности на экспериментальных данных многоканальной электроэнцефалографии человека для управления внешними устройствами // Информационно-управляющие системы. 2018. № 1. С. 106–115. doi:10.15217/issn1684-8853.2018.1.106

Citation: Runnova A. E., Maksimenko V. A., Pchelintseva S. V., Kulanin R. A., Hramov A. E. Wavelet Method for Motor Activity Pattern Analysis on Experimental Data from Human Multichannel Electroencephalography for External Device Control. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 1, pp. 106–115 (In Russian). doi:10.15217/issn1684-8853.2018.1.106

Введение

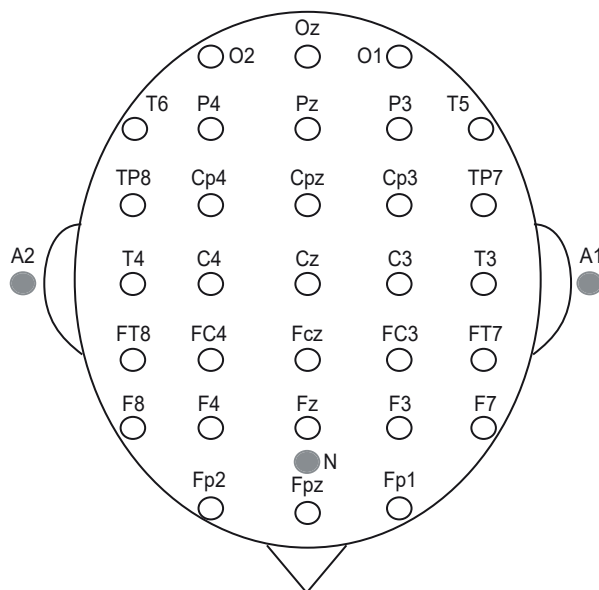
Представленная научная работа лежит в области нейронаучных исследований. На сегодня нейронаука — это активно развивающееся междисциплинарное научное направление, сочетающее в себе физико-математические и технические подходы к обработке биологических и медицинских данных. Подобные комбинации методов и современных регистрирующих устройств как позволяют продвинуться в понимании принципов функционирования головного мозга и сознания человека, так и способствуют решению ряда прикладных задач, связанных с лечением и (или) улучшением качества жизни людей, в част-

ности разработке нейроинтерфейсных устройств. Особо важную роль нейроинтерфейсы играют в создании различных протезов и экзоскелетов, реализующих возможности восстановления двигательной активности человека при ее нарушениях и управляемых на основе расшифровки регистрируемых сигналов с головного мозга человека. Наиболее часто используются при создании подобных устройств электроэнцефалографические регистрации активности головного мозга, что связано с неинвазивностью, удобством применения, безопасностью и мобильностью современных электроэнцефалографов [1–4]. Кроме того, современные регистрирующие приборы позволяют сочетать высокое качество сигналов электро-

энцефалограммы (ЭЭГ) с относительно низкой ценой их эксплуатации. В настоящее время обработка и расшифровка подобных регистрируемых сигналов привлекает к себе обширное внимание исследователей [5–7]. Однако существенным недостатком многочисленных подходов остается низкая эффективность, сильная индивидуализация каждого оператора и длительный срок обучения оператора для достижения стабильно высокого качества работы алгоритма. В данной статье предлагаются методы обработки коротких паттернов на ЭЭГ-данных на основе вейвлетного преобразования и приводятся результаты их применения на экспериментальных сигналах. Усилия по созданию подходов к диагностике особенностей ЭЭГ-паттернов направлены на распознавание сигналов, соответствующих различной двигательной активности человека, — движений как в реальности, так и только в воображении человека. Отметим, что именно опыт «воображаемой» двигательной активности является основным объектом работы при функционировании нейроинтерфейса, управляющего протезными конструкциями. Целью работы является создание объективного и, в идеале, универсального инструмента работы с ЭЭГ-данными, не требующего длительного обучения оператора нейроинтерфейса для управления внешними устройствами.

Экспериментальные данные

Проведена серия экспериментальных работ в группе из 25 неоплачиваемых условно здоровых добровольцев (15 женщин, 10 мужчин, возраст варьировался от 20 до 34 лет, физическое состояние добровольцев соответствовало характеристикам нормального индекса массы тела и среднего уровня физической активности). Все добровольцы были проинструктированы о необходимости перед проведением исследований в течение трех дней соблюдать режим полноценного ночного отдыха. Исследования велись в утренний и дневной периоды (9–13 часов) через 2 часа после небольшого приема пищи с ограничением потребления кофеина и (или) иных стимулирующих добавок к пище. В рамках проводимых исследований использовался монополярный способ регистрации и расширенная расстановка электродов ЭЭГ «10–10», позволяющая оценить активность коры головного мозга по проекциям всех основных зон [1, 3]. На рис. 1 приведена схема расстановки 31 активного регистрирующего электрода с указанием их принятых обозначений. Дополнительно отмечены три электрода (A1, A2 и N), обеспечивающие необходимые, так называемые индифферентные, точки, в которых электрические процессы минимальны и относительно которых



■ **Рис. 1.** Схема расстановки скальповых электродов «10–10»

■ **Fig. 1.** The scheme of the scalp electrodes arrangement “10–10”

ведется регистрация ЭЭГ в активных каналах. Регистрация ЭЭГ-данных велась на электроэнцефалографическом оборудовании «Энцефалан-ЭЭГР-19/26» (ООО НПКФ «Медиком МТД», РФ). Частота дискретизации ЭЭГ составляла 250 Гц, частотный диапазон данных составлял от 0,016 до 70 Гц с применением режекторного фильтра на 50 Гц. Дизайн и оборудование экспериментальной работы одобрены комиссией по этике СГТУ им. Гагарина Ю. А.

После установки всех необходимых датчиков специалист-психолог проводил короткую личную беседу с добровольцем, подготавливая его к концентрации на своих двигательных ощущениях. Активная фаза эксперимента предполагала чередование периодов реальной и воображаемой двигательной активности. Испытуемый сидел в удобном положении, получая аудиальные команды, которые дублировались для его удобства в виде текста на экран перед ним. В эксперименте использовались реальные и воображаемые движения левой/правой рукой (взмахи и сжатия в кулак), а также движения в бедренном суставе левой/правой ноги. После каждой команды в течение 2–2,5 мин следовали серии ее выполнения оператором. Между сериями двигательной активности для расслабления, отдыха, а также восстановления способности к концентрации испытуемого вводились паузы, во время которых проигрывалась легкая музыка. Для сравнения с состоянием пассивного бодрствования в конце и начале каждого эксперимента были дополнительно зарегистрированы короткие фрагменты фоновой активности (7–10 мин). Особо

выделим тот факт, что все добровольцы не имели возможности «тренировки» своих способностей к воображению или осознанию двигательной активности, и, таким образом, исследование велось на неподготовленных операторах.

Обработка данных

В основу создания методики обнаружения коротких паттернов на ЭЭГ положен анализ энергетических характеристик регистрируемых сигналов на базе непрерывного вейвлетного преобразования [8, 9]:

$$W(s, t_0) = \frac{1}{\sqrt{s}} \int_{-\infty}^{+\infty} x(t) \psi^* \left(\frac{t-t_0}{s} \right) dt, \quad (1)$$

где $x(t)$ — временная реализация экспериментального сигнала; $\psi_{s,t_0}(t)$ — материнский вейвлет; s — временной масштаб, определяющий ширину вейвлета, однозначно связанный с классическими частотами преобразования Фурье; символ «*» обозначает комплексное сопряжение. Вейвлетный анализ является на сегодня одним из наиболее мощных математических инструментов для обработки сложных нестационарных зашумленных данных, ограниченных короткими временными рядами [8, 10].

Важным фактором успешности применения вейвлетного преобразования к различным экспериментальным данным является выбор функции материнского вейвлета (вейвлет-базиса). В подобных задачах, связанных с анализом ЭЭГ-сигналов, обычно используется широко известный вейвлет Морле [8]

$$\psi(\eta) = (1/\sqrt[4]{\pi}) \exp(i\Omega_0 \eta) \exp(-\eta^2/2) \quad (2)$$

с параметром $\Omega_0 = 2\pi$, что обеспечивает однозначную взаимосвязь между временным масштабом s вейвлетного преобразования и частотой f фурье-преобразования, а именно $f = 1/s$. Вейвлет Морле имеет очевидные преимущества по сравнению с вещественными вейвлет-базисами: высокую частотно-временную локализацию, возможность проводить оценку мгновенной частоты и фазы ритмических процессов, наглядное представление результатов спектрального анализа в виде частотно-временных спектров, простоту интерпретации результатов расчетов из-за сходства усредненных энергетических спектров (скалограмм) с классическим спектром мощности, основанным на преобразовании Фурье. Материнский вейвлет Морле хорошо зарекомендовал себя при использовании в исследованиях динамики мозга на ЭЭГ-данных у людей и животных, в частности, при изучении эпилептических

явлений [10, 11], когнитивных процессов [12–14] и т. д. [7–9]. Однако поставленная задача, связанная с исследованием «двигательных паттернов» на ЭЭГ, требует анализа крайне коротких временных рядов. В связи с этим была проведена адаптация параметров классического вейвлета Морле. Целью данной модификации являлось уменьшение длины участка ЭЭГ, необходимого для расчета вейвлетного спектра в низкочастотной области [15]. На начальном этапе проводилась оценка длины данного участка, необходимой для использования базиса Морле. В результате была построена зависимость длины временного интервала от частоты в диапазоне 1–30 Гц (рис. 2, а). Полученные результаты сопоставлены с временным интервалом, ассоциирующимся с изменениями частотно-временных свойств сигнала при выполнении или воображении движений. С учетом результатов сопоставления сформулированы требования к оптимальной ширине временного окна, необходимого для построения вейвлетного спектра в дельта-области (1–5 Гц). В результате в классический базис Морле (2) был введен множитель, обеспечивающий ускорение снижения амплитуды огибающей вейвлетной функции ψ :

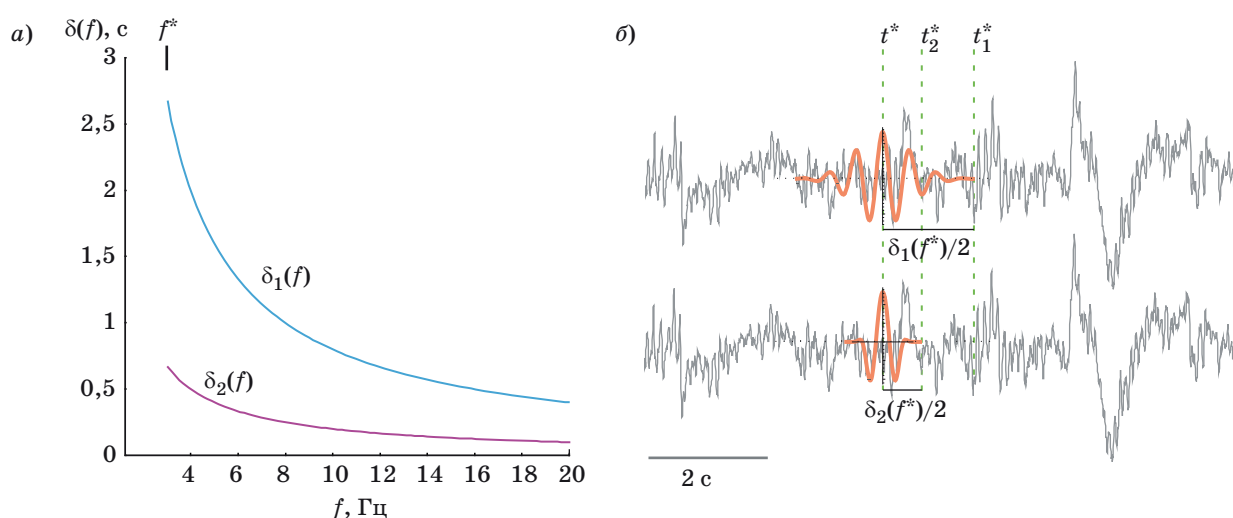
$$\psi_m(\eta) = \frac{1}{\sqrt[4]{\pi}} \exp(i\Omega_0 \eta) \exp(-5\eta^4). \quad (3)$$

Оценка ширины временного окна, необходимого для выделения низкочастотных спектральных компонент на основе классического (2) и модифицированного (3) вейвлетов Морле, приведена на рис. 2, б, где t_1^* , t_2^* обозначают максимальные временные интервалы, необходимые для расчета энергии спектра на основе классической и модифицированной вейвлетной функции Морле соответственно. Таким образом, для вычисления энергии на частоте 1 Гц модифицированный вейвлет (3) использовал интервал времени 2 с, в то время как стандартный вейвлет Морле (2) — 8 с. Полученное преимущество позволило эффективно выявлять быстрые изменения частотно-временной структуры ЭЭГ, ассоциирующиеся с движениями.

Для обработки ЭЭГ-данных применялось вейвлетное преобразование (1) с базисом (3) в частотном диапазоне 1–40 Гц. Для каждого момента времени t_0 был рассчитан моментальный спектральный срез $E_{t_0}(s)$ сигнала $W(s, t)$:

$$E_{t_0}(s) = |W(s, t_0)|^2. \quad (4)$$

Соотношение (4) рассчитывается по всему диапазону масштабов s вейвлетного преобразования. Затем полученные вейвлетные спектры усреднялись на временных интервалах, соответствую-



■ **Рис. 2.** Сравнение длительностей временных окон, необходимых для расчета классического и модифицированного вейвлетного преобразования: *a* — зависимость длины временного интервала, необходимого для вычисления энергии вейвлетного спектра на различных частотах при помощи вейвлетной функции Морле ($\delta_1(f)$) и модифицированного вейвлета ($\delta_2(f)$); *б* — пример использования вейвлетной функции Морле и модифицированного вейвлета для анализа отрезка ЭЭГ

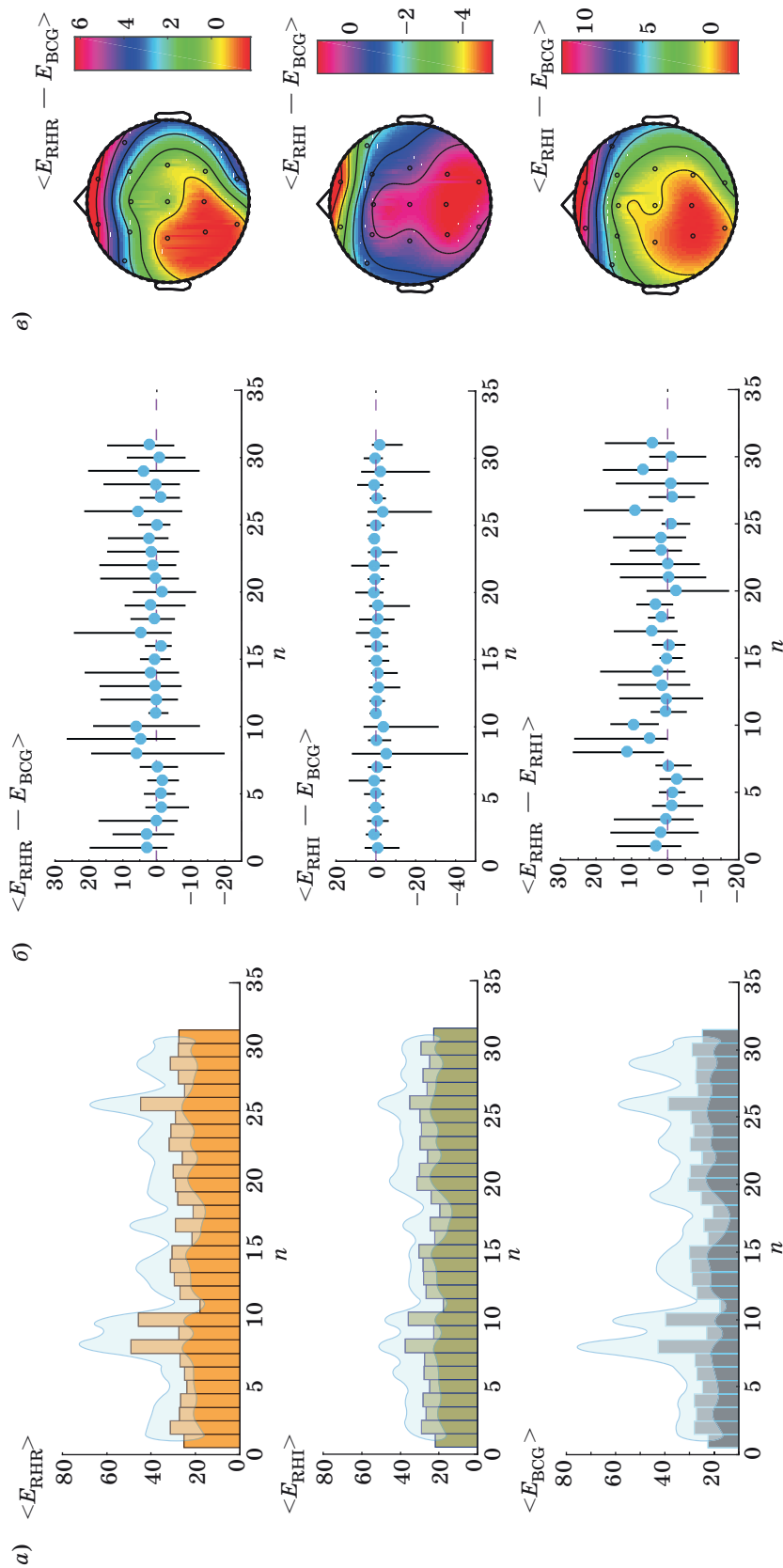
■ **Fig. 2.** Comparison of the durations of the time windows necessary for the calculation of the classical and modified wavelet transforms: *a* — the dependence of the length of the time interval necessary to calculate the energy of the wavelet spectrum at various frequencies using the Morlet wavelet function ($\delta_1(f)$) and the modified wavelet ($\delta_2(f)$); *б* — an example of using the Morlet wavelet function and a modified wavelet for analyzing an EEG segment

ющих различным типам фоновой активности и периодам различных воображаемых и реальных движений. Длительность рассматриваемого временного интервала была выбрана 4 с. В результате для каждого типа событий для каждого испытуемого был получен набор $N \times M$ вейвлетных спектров (M — число разных событий, $N = 31$ — число каналов ЭЭГ). На основе полученных спектров были рассчитаны средние вейвлетные энергии. На рис. 3, *a* представлены рассчитанные распределения средней энергии ЭЭГ-сигналов по различным отведениям, усредненные по группе испытуемых для соответствующих сессий: реальное движение правой рукой (обозначено как RHR), воображаемое движение правой рукой (RHI) и фоновой активности (BCG). На рисунке также показан разброс значений в пределах группы испытуемых. На рис. 3, *б* показаны рассчитанные изменения энергии, соответствующие различным отведениям ЭЭГ, ассоциирующиеся с переходом от фоновой активности — к реальному движению ($\langle E_{RHR} - E_{BCG} \rangle$), от фоновой активности — к воображаемому движению ($\langle E_{RHI} - E_{BCG} \rangle$), от воображаемого движения — к реальному ($\langle E_{RHR} - E_{RHI} \rangle$). Рис. 3, *в* демонстрирует пространственные локализации наиболее существенных изменений распределения величин, характеризующих изменения энергии при переходе от одного типа активности к другому, по различным отведениям ЭЭГ на двумерной плоскости. На рис. 3 хорошо прослеживается от-

личие паттернов, соответствующих двигательной активности, от фоновой пассивной участка ЭЭГ. Кроме того, необходимо отметить разную структуру и при сравнении энергетической характеристики фоновой ЭЭГ (BCG) с моментами реальной (RHR) и воображаемой (RHI) активности. Внизу рис. 3 приведена горизонтальная таблица обозначений всех активных каналов, наименования каналов являются международными и соотношены со схемой на рис. 1

Таким образом, спектральная структура временных интервалов регистрации событий, связанных с реальной или воображаемой двигательной активностью, существенно различается по интенсивности колебательной активности на различных каналах регистрации. Обратимся теперь к более детальному с точки зрения спектральной структуры анализу конкретных особенностей, связанных с реальной или воображаемой двигательной активностью определенного типа. Для этих целей также исследованы интервалы времени, соответствующие отдельным событиям, но общий частотный диапазон был разделен на классически выделяемые в электрофизиологии спектральные интервалы, и затем каждый частотный интервал анализировался отдельно.

Подобная методика предоставила возможности и детального изучения свойств каждого паттерна, и описания обнаруженных особенностей в традиционных для нейронауки терминах. Нами были рассмотрены три классических частотных диа-



1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
O1	P3	CP3	C3	T5	TP7	T3	FC3	F3	Fp1	FT7	F7	Oz	Pz	CPz	Cz	FCz	Fz	Fpz	F8	FT8	Fp2	F4	FC4	T4	TP8	T6	C4	CP4	P4	O2

■ Рис. 3. Энергетическая характеристика активности головного мозга на ЭЭГ: а — распределения средней энергии ЭЭГ-сигналов по различным отведениям n , усредненные по группе испытуемых для соответствующих сессий; б — изменения энергии, соответствующие различным отведениям ЭЭГ; в — распределения величин, характеризующих изменения энергии при переходе от одного типа активности к другому, по различным отведениям ЭЭГ на двумерной плоскости

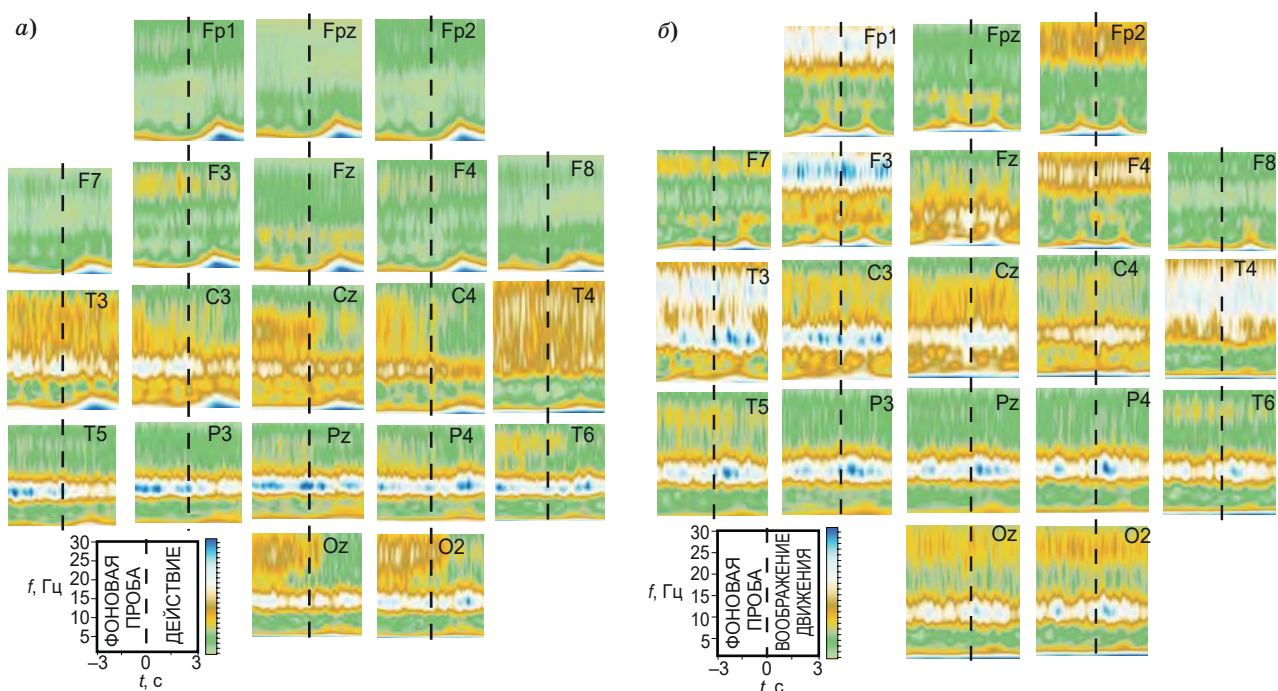
■ Fig. 3. Energy characteristic of brain activity on the EEG: a — the average energy distributions of the EEG signals for different leads n , averaged over the group of subjects for the respective sessions; б — energy changes corresponding to different EEG leads; в — distributions of the quantities characterizing the changes in energy during the transition from one type of activity to another, along different EEG leads on a two-dimensional plane

пазона, выбранных согласно традиционным представлениям о структуре ЭЭГ-сигналов [1–5, 16]: дельта-диапазон (1–5 Гц), альфа-диапазон (8–12 Гц) и бета-диапазон (20–30 Гц), — каждый из которых рассматривался отдельно. Для каждого уникального события рассмотрен интервал времени длительностью 6 с, содержащий интервалы (по 3 с) до и после начала выполнения действия. Для каждого интервала получена частотно-временная зависимость спектральной энергии для каждого из $N = 31$ отведений ЭЭГ. В результате для каждого из испытуемых получен набор из N зависимостей. На рис. 4 данные зависимости, являющиеся результатом обработки данных отдельных «обычных» эпизодов двигательной активности одного из испытуемых, расположены на листе в соответствии с положением отведений ЭЭГ. Рис. 4, а соответствует выполнению реального движения, рис. 4, б — воображаемого движения. Штриховая линия обозначает момент времени, соответствующий началу выполнения действия, каждая схема содержит интервал времени (3 с), предшествующий выполнению движения, и интервал времени (3 с), соответствующий выполнению движения. Легенда в левом нижнем углу показывает границы рассматриваемого пространственно-временного окна.

Кроме того, для анализа общей картины изменения спектральных характеристик нами проведено усреднение полученных значений энергии по боль-

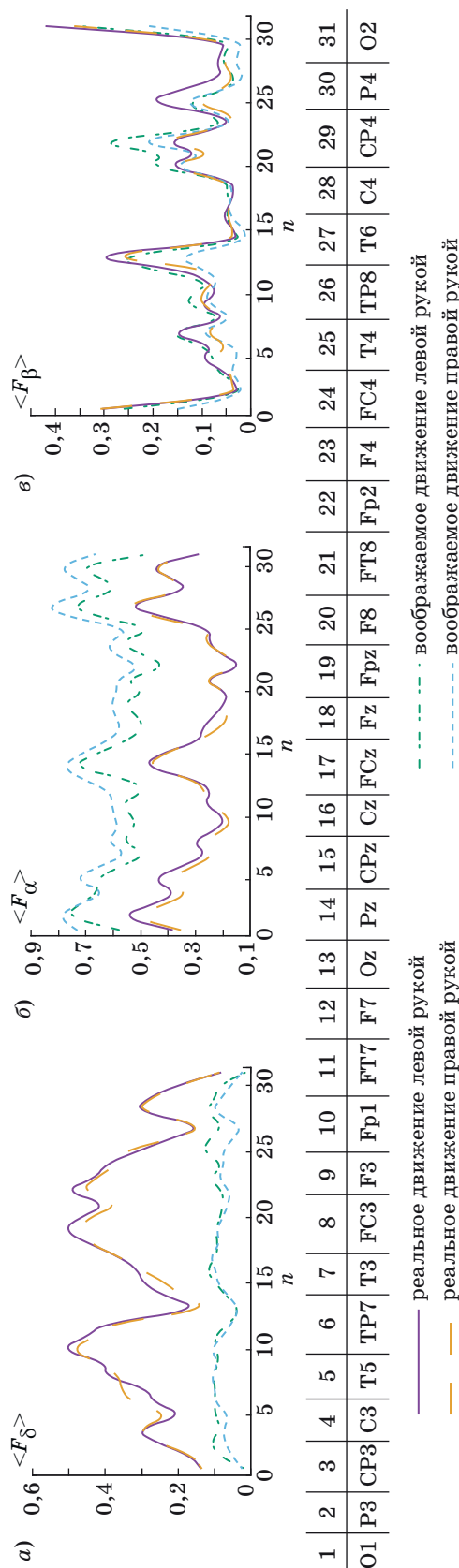
шому (~100) числу однотипных событий отдельно для каждого из испытуемых. Для вышеуказанных частотных диапазонов альфа, бета и гамма были рассчитаны значения вейвлетной энергии, усредненные по интервалу времени 4 с после подачи звукового сигнала (сигнала к началу осуществления или воображения движения), отдельно для каждого из 31 отведения ЭЭГ. Полученные значения энергии были затем усреднены по большому (~100) числу однотипных событий отдельно для каждого испытуемого. На рис. 5 представлены результаты на примере четырех различных типов событий (двух реальных — движение правой и левой рукой и двух воображаемых — воображение движения правой и левой рукой соответственно).

Для уточнения характера качественных отличий между частотной динамикой паттернов, соответствующих различным типам движений, проведена детализация анализа энергетических соотношений по всему частотному спектру. Для этих целей анализируемая спектральная область (1–40 Гц) была разбита на узкие (~2 Гц) интервалы частот, в которых проанализирована вейвлетная энергия при тех же типах движения. На рис. 6 демонстрируются результаты энергетического анализа с расширенной частотной детализацией в пространственных распределениях по скальповому расположению электродов. Плавающее значение длительности интервала



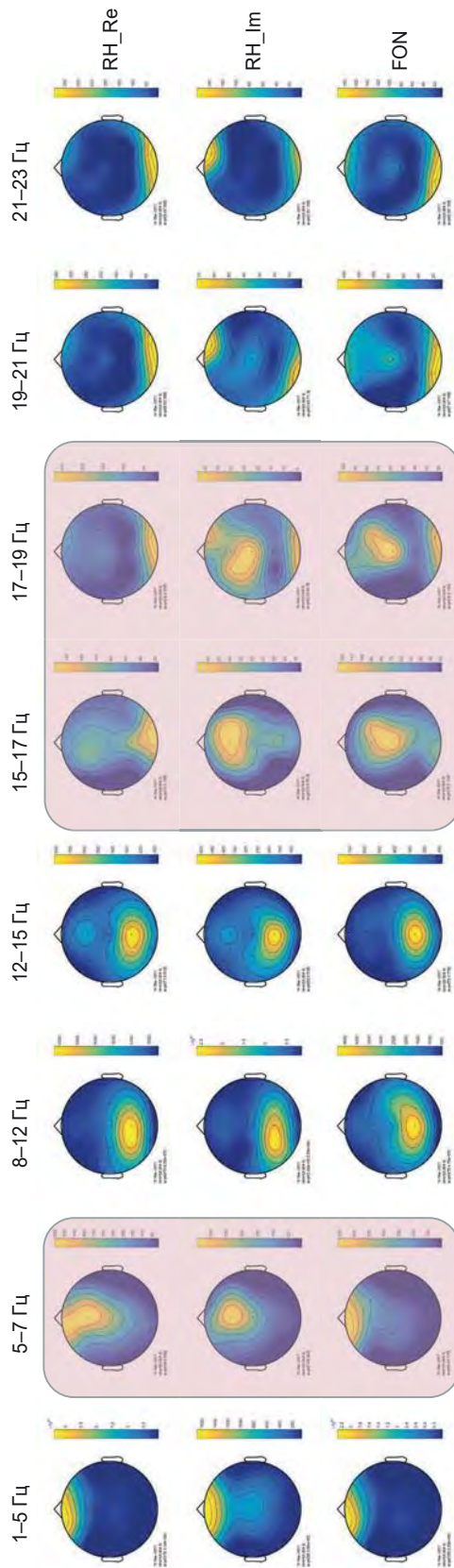
■ **Рис. 4.** Пространственно-временные зависимости, характеризующие эпизод реального (а) и воображаемого (б) движения правой рукой

■ **Fig. 4.** Spatio-temporal dependence for an episode of the real (a) and of an imaginary (b) movement with the right hand



■ Рис. 5. Усредненные по частотным диапазонам дельта [1–5 Гц] (а), альфа [8–12 Гц] (б) и бета [20–30 Гц] (в) зависимости спектральной энергии сигналов ЭЭГ, соответствующие выполнению реальных и воображаемых движений

■ Fig. 5. The dependences of the spectral energy of the EEG signals, averaged over the frequency ranges of alpha delta [1–5 Hz] (a), alpha [8–12 Hz] (b) and beta [20–30 Hz] (c), corresponding real and imaginary movements



■ Рис. 6. Значения спектральной энергии сигналов ЭЭГ, усредненной по набору частотных интервалов, для реальных (RH_Re), воображаемых (RH_Im) движений правой рукой и для фоновой активности (FON)

■ Fig. 6. The values of the spectral energy of EEG signals averaged over a set of frequency intervals for real (RH_Re), imaginary (RH_Im) motions with the right hand and for background activity (FON)

рассматриваемых частот выбрано с позиции наибольшей наглядности. В рамках — наиболее отличающиеся по своим энергетическим характеристикам спектральные диапазоны.

Заключение

В результате расчета и анализа вейвлетных характеристик (1), (3), (4) были выявлены некоторые характерные особенности структуры сигналов ЭЭГ, ассоциирующихся с различными типами движений (воображаемыми и реальными). В первую очередь, анализ изменения энергии электрических колебаний, регистрируемых посредством ЭЭГ, показывает, что во время последовательного выполнения реальных движений наблюдается увеличение энергии сигналов ЭЭГ, снимаемых с лобной области. При этом наблюдается одновременное (менее значительное) уменьшение энергии в затылочной области. Наименьшие изменения рассматриваемой энергетической характеристики наблюдаются в теменной области. В отличие от реального движения, воображаемое движение характеризуется уменьшением энергии в лобных долях и увеличением энергии в затылочной и теменной областях (см. рис. 3).

В то же время проведенный анализ изменений энергетических характеристик во времени показывает, что при реальных и воображаемых движениях наблюдается противофазное изменение энергии ЭЭГ в лобной и теменной областях: в случае реального движения наблюдается локальное увеличение энергии в течение первой секунды после предъявления звукового стимула (сигнала к началу выполнения движения), а в случае воображаемого — локальное уменьшение. При этом средняя во времени энергия колебаний ЭЭГ в лобной области оказывается большей при реальном движении, а в теменной — при воображаемом (см. рис. 3).

Анализ энергетических изменений в частотно-временных доменах показал, что в случае воображаемого движения большая часть энергии сосредоточена в диапазоне 8–12 Гц (альфа-ритм) в затылочной и теменной областях. При воображении движения происходит увеличение энергии в данной области спектра. В то же время в лобных долях увеличение энергии в диапазоне 8–12 Гц (мю-ритм) сопровождается более значительным уменьшением энергии в низкочастотной области (1–5 Гц). В случае реального движения наблюдается уменьшение энергии в области частот бета-диапазона (15–30 Гц) в затылочной части, уменьшение энергии в диапазоне 8–12 Гц — в затылочной и теменной областях и увеличение энергии на низких частотах — в лобных долях (см. рис. 4).

Отметим, что проведение статистического усреднения по набору событий демонстрирует сохранность выявленных особенностей в частотных диапазонах гамма, альфа и бета по «двигательным» паттернам на ЭЭГ (см. рис. 6). Кроме того, проведенная частотная детализация спектрального состава ЭЭГ-паттернов позволяет уточнить рамки наблюдающихся отличий между паттернами. Максимальные различия, способные достоверно определяться в численной обработке, лежат в диапазонах 5–7 и 15–19 Гц.

Таким образом, в результате проведенного вейвлетного анализа были установлены достоверные различия между случаями реального и воображаемого движения. Полученные результаты носят междисциплинарный характер и являются важными как для понимания физиологических механизмов электрической активности головного мозга, так и для проектирования нейроинтерфейсов «мозг — компьютер», управляемых воображением тех или иных движений.

Работа поддержана РФФИ (проект № 16-29-08221_офи_м).

Литература

- Schomer D. H., Lopes da Silva F. H. Nidermeyer's. Electroencephalography. — Philadelphia: Lippincott Williams & Wilkins, 2011. — 1296 p.
- Гусельников В. И. Электрофизиология головного мозга. — М.: Высш. шк., 1976. — 423 с.
- Жирмунская Е. А. Клиническая электроэнцефалография. — М.: МЭЙБИ, 1991. — 118 с.
- Зенков Л. Р. Клиническая электроэнцефалография (с элементами эпилептологии). — М.: МЕДпресс-информ, 2004. — 368 с.
- Freeman W. J., Quiroga R. Q. Imaging Brain Function with EEG: Advanced Temporal and Spatial Analysis of Electroencephalographic. — N. Y.: Springer Science+Business Media, 2013. — 245 p.
- Kropotov J. D. Quantitative EEG, Event-Related Potentials and Neurotherapy. — Elsevier Science Publishing, 2008. — 600 p.
- Shanbao Tong, Nitish V. Quantitative EEG Analysis Methods and Clinical Applications. — Boston: Artech House, 2009. — 421 p.
- Hramov A. E., Koronovskii A. A., Makarov V. A., Pavlov A. N., Sitnikova E. Wavelets in Neuroscience. — Springer Heidelberg New York Dordrecht London, 2015. — 314 p. doi:10.1007/978-3-662-43850-3
- Короновский А. А., Храмов А. Е. Непрерывный вейвлетный анализ и его приложения. — М.: Физматлит, 2003. — 176 с.
- Sitnikova E. Yu., Hramov A. E., Grubov V. V., Koronovskii A. A. Rhythmic Activity in EEG and Sleep

- in Rats with Absence Epilepsy // Brain Research Bulletin. 2016. Vol. 120. P. 106–116.
11. Hramov A. E., Koronovskii A. A., Midzyanovskaya I. S., Sitnikova E. Yu., van Rijn C. M. On-off Intermittency in Time Series of Spontaneous Paroxysmal Activity in Rats with Genetic Absence Epilepsy // Chaos. 2006. Vol. 16. N 4. P. 043111.
 12. Runnova A. E., Zhuravlev M. O., Pisarchik A. N., Khranova M. V., Grubov V. V. The Study of Cognitive Processes in the Brain EEG during the Perception of Bistable Images using Wavelet Skeleton // Proc. SPIE. 2017. Vol. 10063. P. 1006319. doi:10.1117/12.2250403
 13. Руннова А. Е., Журавлев М. О., Лопатин Д. В. Математическое моделирование некоторых аспектов когнитивного распознавания сложных объектов с пространственной перспективой // Вестник ТГУ. 2017. № 1(22). С. 33–38. doi:10.20310/1810-0198-2017-22-1-33-38
 14. Runnova A. E., Zhuravlev M. O., Khranova A. E., Pisarchik A. N. The Study of Evolution and Depression of the Alpha-rhythm in the Human Brain EEG by Means of Wavelet-based Methods // Proc. SPIE. 2017. Vol. 10337. P. 1033713. doi:10.1117/12.2267699
 15. van Luijtelaaar G., Lüttjohann A., Makarov V. V., Maksimenko V. A., Koronovskii A. A., Hramov A. E. Methods of Automated Absence Seizure Detection, Interference by Stimulation, and Possibilities for Prediction in Genetic Absence Models // Journal of Neuroscience Methods. 2016. Vol. 260. P. 144–158.
 16. Jensen O., Goel P., Kopell N., Pohja M., Hari R., Ermentrout B. On the Human Sensorimotor-cortex Beta Rhythm: Sources and Modeling // Neuroimage. 2005. N 2(26). P. 347–355. doi:10.1016/j.neuroimage.2005.02.008

UDC 616.8, 57.087, 51-7

doi:10.15217/issn1684-8853.2018.1.106

Wavelet Method for Motor Activity Pattern Analysis on Experimental Data from Human Multichannel Electroencephalography for External Device Control

Runnova A. E.^a, PhD, Phys.-Math, Associate Professor, anefila@gmail.comMaksimenko V. A.^a, PhD, Phys.-Math, Associate Professor, maximenkovl@gmail.comPchelintseva S. V.^a, PhD, Tech., Associate Professor, pchelintseva@inbox.ruKulanin R. A.^a, Research Fellow, rkulanin2010@yandex.ruHramov A. E.^a, Dr. Sc., Phys.-Math., Professor, hramovae@gmail.com^aYuri Gagarin State Technical University of Saratov, 77, Politechnicheskaja St., 410054, Saratov, Russian Federation

Introduction: The modern fundamental and technical science has a great interest in neural interface development, as it aims at both improving life quality and studying the functioning of a human brain. Neural interface devices are based on recognizing the patterns of recorded brain activity, and their main problems are low efficiency of pattern recognition techniques and long time necessary to train the operators. **Purpose:** Developing methods for recognizing patterns associated with various motor activity of untrained operators. **Results:** Based on wavelet technologies, new methods are proposed for studying various patterns of brain activity which correlate with various types of motor activity. Morle basis is modified in order to speed up the calculation. Approaches to assessing energy capacity of various electroencephalography processes are discussed, based on the calculation of skeleton characteristics. We give examples of processing experimental data obtained from untrained volunteers. Characteristic features of patterns for various motor activity (imaginary or real, leg or hand movement) have been discovered. We demonstrate that the electroencephalography patterns are universal for everyone from the volunteering group. **Practical relevance:** The revealed features of "motor" patterns and the methods of electroencephalography processing can be used in constructing neural interfaces which decode brain electrical activity almost without operator's training.

Keywords — Electroencephalography, Pattern, Continuous Wavelet Transformation, Time-Frequency Dynamics.

Citation: Runnova A. E., Maksimenko V. A., Pchelintseva S. V., Kulanin R. A., Hramov A. E. Wavelet Method for Motor Activity Pattern Analysis on Experimental Data from Human Multichannel Electroencephalography for External Device Control. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 1, pp. 106–115 (In Russian). doi:10.15217/issn1684-8853.2018.1.106

References

1. Schomer D. H., Lopes da Silva F. H. Nidermeyer's. *Electroencephalography*. Philadelphia, Lippincott Williams & Wilkins, 2011. 1296 p.
2. Gusel'nikov V. I. *Elektrofiziologija golovnogo mozga* [Brain Electrophysiology]. Moscow, Vysshaya shkola Publ., 1976. 423 p. (In Russian).
3. Zhirmunskaja E. A. *Klinicheskaja elektroentsefalografija* [Clinical Electroencephalography]. Moscow, MEIBI Publ., 1991. 118 p. (In Russian).
4. Zenkov L. R. *Klinicheskaja elektroentsefalografija (s elementami epileptologii)* [Clinical Electroencephalography (with Elements of Epileptology)]. Moscow, MEDpress-inform Publ., 2004. 368 p. (In Russian).
5. Freeman W. J., Quiroga R. Q. *Imaging Brain Function with EEG: Advanced Temporal and Spatial Analysis of Electroencephalographic*. Springer Science+Business Media New York, 2013. 245 p.
6. Kropotov J. D. *Quantitative EEG, Event-Related Potentials and Neurotherapy*. Elsevier Science Publishing, 2008. 600 p.
7. Shanbao Tong, Nitish V. *Quantitative EEG Analysis Methods and Clinical Applications*. Boston, Artech House, 2009. 421 p.
8. Hramov A. E., Koronovskii A. A., Makarov V. A., Pavlov A. N., Sitnikova E. *Wavelets in Neuroscience*. Springer Heidelberg New York Dordrecht London, 2015. 314 p. doi: 10.1007/978-3-662-43850-3

9. Koronovskii A. A., Khramov A. E. *Nepřeryvnyi veivletnyi analiz i ego prilozheniia* [Continuous Wavelet Analysis and Applications]. Moscow, Fizmatlit Publ., 2003. 176 p. (In Russian).
10. Sitnikova E. Yu., Hramov A. E., Grubov V. V., Koronovskii A. A. Rhythmic Activity in EEG and Sleep in Rats with Absence Epilepsy. *Brain Research Bulletin*, 2016, vol. 120, pp. 106–116.
11. Hramov A. E., Koronovskii A. A., Midzyanovskaya I. S., Sitnikova E. Yu., van Rijn C. M. On-off Intermittency in Time Series of Spontaneous Paroxysmal Activity in Rats with Genetic Absence Epilepsy. *Chaos*, 2006, vol. 16, no. 4, pp. 043111.
12. Runnova A. E., Zhuravlev M. O., Pisarchik A. N., Khramova M. V., Grubov V. V. The Study of Cognitive Processes in the Brain EEG during the Perception of Bistable Images using Wavelet Skeleton. *Proc. SPIE*, 2017, vol. 10063, pp. 1006319. doi:10.1117/12.2250403
13. Runnova A. E., Zhuravlev M. O., Lopatin D. V. The Mathematical Modeling of Some Aspects of Cognitive Recognition of Complex Objects with the Spatial Perspective. *Vestnik TGU*, 2017, no. 1(22), pp. 33–38 (In Russian). doi:10.20310/1810-0198-2017-22-1-33-38
14. Runnova A. E., Zhuravlev M. O., Khramova A. E., Pisarchik A. N. The Study of Evolution and Depression of the Alpha-rhythm in the Human Brain EEG by Means of Wavelet-based Methods. *Proc. SPIE*, 2017, vol. 10337, pp. 1033713. doi:10.1117/12.2267699
15. van Luijtelaaar G., Lüttjohann A., Makarov V. V., Maksimenko V. A., Koronovskii A. A., Hramov A. E. Methods of Automated Absence Seizure Detection, Interference by Stimulation, and Possibilities for Prediction in Genetic Absence Models. *Journal of Neuroscience Methods*, 2016, vol. 260, pp. 144–158.
16. Jensen O., Goel P., Kopell N., Pohja M., Hari R., Ermentrout B. On the Human Sensorimotor-Cortex Beta Rhythm: Sources and Modeling. *Neuroimage*, 2005, no. 2(26), pp. 347–355. doi:10.1016/j.neuroimage.2005.02.008

УДК 519.21: 537.86

ББК 22.17

X46

Хименко В. И.

Случайные данные: структура и анализ / В. И. Хименко. — М.: Техносфера, 2017. — 424 с. (Серия «Мир фотоники»)

ISBN 978-5-94836-497-1

Книга посвящена одной из наиболее общих проблем физики и техники, биологии и естествознания — проблеме извлечения информации из случайных данных (наблюдений, измерений, экспериментальных исследований). Сам процесс извлечения информации включает в себя этапы сбора данных, построения моделей реальных процессов и систем, анализа и интерпретации полученных результатов.

Даны описание и детальный анализ структуры наиболее важных с точки зрения приложений моделей временных рядов, непрерывных случайных процессов, случайных потоков событий, случайных полей и изображений. Представлено большое количество новых результатов по вероятностному анализу неоднородных данных, отображениям случайных процессов на фазовой плоскости, характеристикам выбросов и характеристикам превышений заданных уровней. Показано широкое разнообразие практических задач, которые решаются (или могут решаться) на основе рассмотренных моделей случайных функций.

Книга ориентирована на широкий круг читателей — на тех специалистов, аспирантов и студентов, которые изучают, исследуют и применяют на практике модели и методы анализа различных по своей физической природе случайных данных.



УДК 004.891 + 311.2

doi:10.15217/issn1684-8853.2018.1.116

СИНТЕЗ СТРУКТУР БАЙЕСОВСКОЙ СЕТИ ДОВЕРИЯ ДЛЯ ОЦЕНКИ ХАРАКТЕРИСТИК РИСКОВАННОГО ПОВЕДЕНИЯ

А. В. Суворова^а, канд. физ.-мат. наук, старший научный сотрудник, suvalv@gmail.com

А. Л. Тулупьев^{а, б}, доктор физ.-мат. наук, доцент, alexander.tulupyev@gmail.com

^аСанкт-Петербургский институт информатики и автоматизации РАН, 14-я линия В. О., 39, Санкт-Петербург, 199178, РФ

^бСанкт-Петербургский государственный университет, Университетская наб., 7–9, Санкт-Петербург, 199034, РФ

Постановка проблемы: необходимость оценивания параметров поведения (как индивидуального, так и на уровне популяции) возникает в различных областях социологических, психологических, эпидемиологических, маркетинговых исследований, исследований безопасности. Однако прямая оценка интенсивности рискованного поведения не всегда доступна, как следствие, требуется развитие косвенных методов оценки. Ранее был предложен подход к моделированию рискованного поведения на основе байесовской сети доверия по данным о нескольких последних эпизодах такого поведения, но для практического применения необходимы изменения этой модели в целях снижения ее зависимости от первоначальных предположений экспертов о взаимосвязях между элементами модели. **Цель:** предложить модификацию модели, которая не требует задания структуры экспертами; провести сравнение этой модели с первоначальной. **Методы:** для проверки модели разработана программа, генерирующая тестовые данные в соответствии с теоретическими предположениями модели. Для построения структуры байесовской сети доверия по сгенерированным данным использован алгоритм оптимизации меры качества сети Hill-Climbing с мерой качества Bayesian Information Criterion. **Результаты:** предложено развитие подхода к построению модели рискованного поведения на основе байесовской сети доверия по совокупности наблюдений, включающей сведения об эпизодах такого поведения. Проведено сравнение двух структур такой модели: предложенной экспертами и построенной автоматически по данным. В то время как формальные меры качества показывают преимущество автоматически обученной структуры, качество предсказания лучше у модели с экспертно заданной структурой. Таким образом, для решения практических задач можно использовать любую из предложенных моделей; выбор может быть обусловлен условием конкретной задачи.

Ключевые слова — моделирование поведения, байесовские сети доверия, синтез структур, машинное обучение, рискованное поведение.

Цитирование: Суворова А. В., Тулупьев А. Л. Синтез структур байесовской сети доверия для оценки характеристик рискованного поведения // Информационно-управляющие системы. 2018. № 1. С. 116–122. doi:10.15217/issn1684-8853.2018.1.116

Citation: Suvorova A. V., Tulupyev A. L. Bayesian Belief Network Structure Synthesis for Risky Behavior Rate Estimation. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 1, pp. 116–122 (In Russian). doi:10.15217/issn1684-8853.2018.1.116

Введение

Решение многих задач в различных областях исследования основано на изучении поведения человека или группы, в частности, на изучении численных параметров исследуемого поведения, сравнении их с нормативными значениями или отслеживании динамики показателей. Примерами подобных задач являются оценивание частоты индивидуального поведения, создающего угрозу безопасности, например информационной системы [1]; моделирование поведения в группе [2]; оценивание интенсивности рискованного поведения на уровне популяции [3].

Ряд задач эпидемиологии и охраны общественного здоровья тесно связан с оценкой вреда, который может быть причинен индивидом обществу, самому себе и (или) другому индивиду. В таком случае с риском связывают эпизоды определен-

ного поведения индивида, а численной характеристикой такого риска выступает интенсивность поведения [4]. Одной из таких задач является оценивание риска передачи или получения ВИЧ-инфекции, а в качестве интересующего исследователя типа поведения — употребление инъекционных наркотиков и рискованное сексуальное поведение [5]. Однако прямая оценка интенсивности (т. е. непосредственное измерение числа эпизодов) рискованного поведения не всегда доступна в силу экономических причин (подобные исследования для сбора данных сложно, долго и дорого организовывать) [6, 7]. Ретроспективное измерение (опросы самих респондентов о числе эпизодов поведения определенного типа) сталкивается с проблемой точности ответов из-за особенностей воспоминания и самооценки, а также с проблемой целенаправленного занижения интенсивности из-за социальной нежелательности многих видов

рискованного поведения [7]. Вот почему требуется развитие косвенных методов оценки.

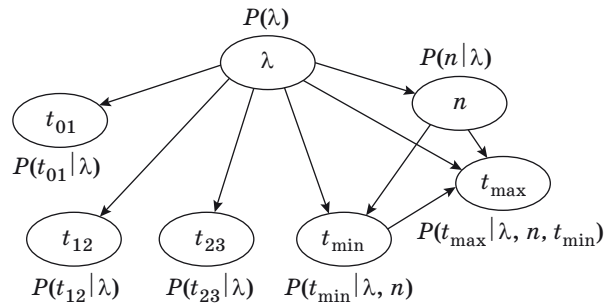
В работах [8, 9] предложен подход к построению байесовской сети доверия (БСД) для моделирования рискованного поведения на основе данных о нескольких последних эпизодах такого поведения. БСД позволяет представить сложные взаимосвязи между элементами, входящими в модель, в виде разложения на более простые элементы, что упрощает как описание всей системы, так и ее интерпретацию [9–11], в то время как выводы делаются с учетом полной системы. Кроме того, аппарат БСД позволяет учитывать как имеющиеся статистические данные, так и экспертную информацию об интересующей исследователя области [9–11]. Причем уровень вовлечения экспертов может быть очень разным: от выбора методов, используемых для построения модели, до полного задания модели, включая ее структуру и параметры [12–15].

В данной работе исследовано влияние экспертных предположений о структуре модели на ее качество. Для этого предложена модификация модели, при построении структуры которой использован алгоритм автоматического обучения по данным, а затем проведено сравнение моделей с исходной и модифицированной структурой.

Описание исходной модели

В качестве исходных данных для вычисления оценки интенсивности рискованного поведения [16] используются сведения о трех последних эпизодах поведения t_{01} , t_{12} , t_{23} (точнее, длины интервалов между этими эпизодами) и сведения о минимальном и максимальном интервале (t_{\min} и t_{\max}) между эпизодами за исследуемый промежуток времени. Моделью поведения выступает пуассоновский случайный процесс, т. е. длины интервалов между эпизодами распределены экспоненциально. Кроме того, модель содержит также оцениваемую величину λ , соответствующую интенсивности поведения, и скрытую переменную n — число эпизодов, произошедших за исследуемый промежуток времени.

Для построения байесовской сети возможные значения всех непрерывных величин (λ , t_{01} , t_{12} , t_{23} , t_{\min} , t_{\max}) разбиваются на дискретные интервалы; таким образом, распределения являются мультиномиальными. Во всех примерах, рассмотренных в данной работе, используется дискретизация вида: для случайной величины, соответствующей интенсивности поведения, $\lambda^{(1)} = [0; 0,01]$, $\lambda^{(2)} = [0,01; 0,03]$, $\lambda^{(3)} = [0,03; 0,05]$, $\lambda^{(4)} = [0,05; 0,1]$, $\lambda^{(5)} = [0,1; 0,2]$, $\lambda^{(6)} = [0,2; 0,3]$, $\lambda^{(7)} = [0,3; 0,5]$, $\lambda^{(8)} = [0,5; 0,7]$, $\lambda^{(9)} = [0,7; 1]$, $\lambda^{(10)} = [1; \infty)$; для случайных величин $t_{j,j+1}$, t_{\min} , t_{\max} , характеризующих длины интервалов между эпизодами,



■ **Рис. 1.** Структура БСД для моделирования рискованного поведения, заданная экспертно

■ **Fig. 1.** Expert-based structure of the BBN for risky behavior modelling

$t^{(1)} = [0; 0,1)$, $t^{(2)} = [0,1; 1)$, $t^{(3)} = [1; 7)$, $t^{(4)} = [7; 30)$, $t^{(5)} = [30; 180)$, $t^{(6)} = [180; \infty)$.

В первоначальной модели [8] структура задана экспертно (рис. 1).

Тензоры \mathbf{P} условной вероятности, характеризующие переходы между узлами сети: $\mathbf{P} = \{P(t_{j,j+1}|\lambda), P(t_{01}|\lambda), P(t_{\min}|n, \lambda), P(t_{\max}|n, \lambda, t_{\min}), P(n|\lambda), P(\lambda)\}$, — определяются аналитически, согласно предположению о пуассоновском процессе в качестве модели поведения ($l_s = 1, \dots, k_s$, где k_s — число дизъюнктивных промежутков при дискретизации случайных величин; $s = 0, \dots, 4$; $j = 1, 2$; $i = 1, \dots, m$, где m — число дизъюнктивных промежутков при дискретизации величины λ) [8]:

$$p\left(t_{j,j+1}^{(l_j)} \mid \lambda^{(i)}\right) = e^{-a\lambda^{(i)}} - e^{-b\lambda^{(i)}}, \quad j = 0, 1, 2,$$

$$t_{j,j+1}^{(l_j)} = [a; b);$$

$$p\left(t_{\min}^{(l_3)} \mid n, \lambda^{(i)}\right) = e^{-an\lambda^{(i)}} - e^{-bn\lambda^{(i)}}, \quad t_{\min}^{(l_3)} = [a; b);$$

$$p\left(n \mid \lambda^{(i)}\right) = \frac{\left(\lambda^{(i)}T\right)^n}{n!} e^{-\lambda^{(i)}T};$$

$$p\left(t_{\max}^{(l_4)} \mid n, \lambda^{(i)}, t_{\min}^{(l_3)}\right) = e^{(n-1)\lambda^{(i)}t_{\min}^{(l_3)}} \times \left(\left(e^{-\lambda^{(i)}t_{\min}^{(l_3)}} - e^{-\lambda^{(i)}b} \right)^{n-1} - \left(e^{-\lambda^{(i)}t_{\min}^{(l_3)}} - e^{-\lambda^{(i)}a} \right)^{n-1} \right),$$

$$t_{\max}^{(l_4)} = [a; b).$$

Предложенная модель была апробирована на тестовых данных [17]; показано, что модель имеет хорошие показатели качества при сравнении с фактическими значениями интенсивности поведения, т. е. сведения о последних эпизодах поведения позволяют характеризовать поведение в целом.

Обучение структуры байесовской сети доверия

Для оценивания влияния экспертных предположений о структуре на качество модели в данной работе также исследуется структура, полученная автоматически по данным. Таблицы условных вероятностей также вычисляются на основе данных. Из-за того, что данные о числе эпизодов реального поведения за заданный промежуток времени получить организационно сложно, а иногда и невозможно (например, при исследовании рискованного сексуального поведения), то для тестирования модели была разработана программа, генерирующая «эпизоды поведения» в соответствии с теоретическими предположениями модели, т. е. в соответствии с пуассоновской моделью поведения. Все вычисления и анализ выполнены с помощью языка R [18], в частности, для работы с байесовскими сетями использовался пакет bnlearn [19].

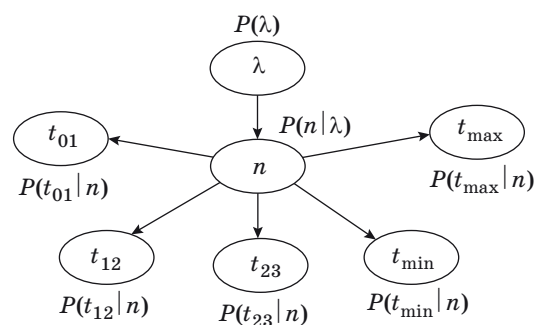
Сначала были сгенерированы 300 значений интенсивности, соответствующие значениям случайной величины, распределенной по гамма-распределению с параметрами $k = 1,1$; $\theta = 0,3$. С одной стороны, большая часть значений меньше 0,5, что соотносится со многими примерами реального поведения, с другой — в данных есть значения для всех интервалов, на которые разбито значение λ при дискретизации.

Далее для каждого значения интенсивности генерируется 20 «респондентов» — последовательностей точек, расстояния между которыми подчиняются экспоненциальному распределению с соответствующим значением интенсивности. Из каждой такой последовательности выделяются исходные данные для оценки: длины интервалов между тремя последними точками, минимальный и максимальный интервал за промежуток длиной 365 «дней», удаляются последовательности, у которых нет хотя бы двух точек за этот промежуток. Таким образом, конечный обучающий набор включает 5936 «респондентов», причем для каждого дополнительно известно исходное значение интенсивности, что позволяет сравнить его с итоговой оценкой.

Аналогичным образом были сгенерированы тестовые данные для дальнейшей оценки построенных моделей (50 значений интенсивности, 15 последовательностей для каждого значения, всего 730 «респондентов»).

Для построения структуры БСД по сгенерированным данным был использован алгоритм оптимизации меры качества сети Hill-Climbing [19], мерой качества являлся BIC (Bayesian Information Criterion).

Полученная в результате обучения структура БСД представлена на рис. 2.



■ Рис. 2. Структура БСД для моделирования рискованного поведения, обученная на данных

■ Fig. 2. Data-based structure of the BBN for risky behavior modelling

Следует отметить, что данная структура имеет достаточно простую интерпретацию: интенсивность поведения (другими словами, частота поведения) определяется через число эпизодов поведения, произошедших за рассматриваемый промежуток времени (в текущем примере 365 дн.). Однако, как уже отмечалось, получить в явном виде число эпизодов для большинства примеров поведения невозможно. Таким образом, n является скрытой переменной, определяемой через исходные данные и определяющей в свою очередь искомое значение λ .

Для полного определения моделей (см. рис. 1, 2) далее было проведено автоматическое обучение параметров представленных байесовских сетей на имеющихся данных, т. е. вычисление условных вероятностей для всех пар переменных, соединенных в разработанных структурах ребром.

Сравним полученные модели. Согласно алгоритму построения, на обучающей выборке мера качества структуры, представленной на рис. 2, выше, чем первоначальной, заданной экспертно (см. рис. 1), как для BIC (−40165 и −54991 соответственно), так и для меры максимального правдоподобия (−38649 и −38704). На тестовой выборке меры качества структуры, обученной по данным, также выше, хотя и незначительно в случае меры максимального правдоподобия (BIC: −5990 vs −16836; максимальное правдоподобие: −4474 vs −4477).

Однако, так как основное назначение предложенных моделей — оценивание интенсивности, то следующий этап сравнения моделей состоит в оценке качества не самих структур, а предсказаний согласно каждой из предложенных моделей. Соответствие предсказанного и исходного значений интенсивности для предложенных моделей представлено в табл. 1 и 2. Отметим, что при указанной дискретизации переменной λ задача оценивания интенсивности индивидуального поведения является задачей классификации по 10 непересекающимся классам.

■ **Таблица 1.** Предсказание по экспертно заданной структуре

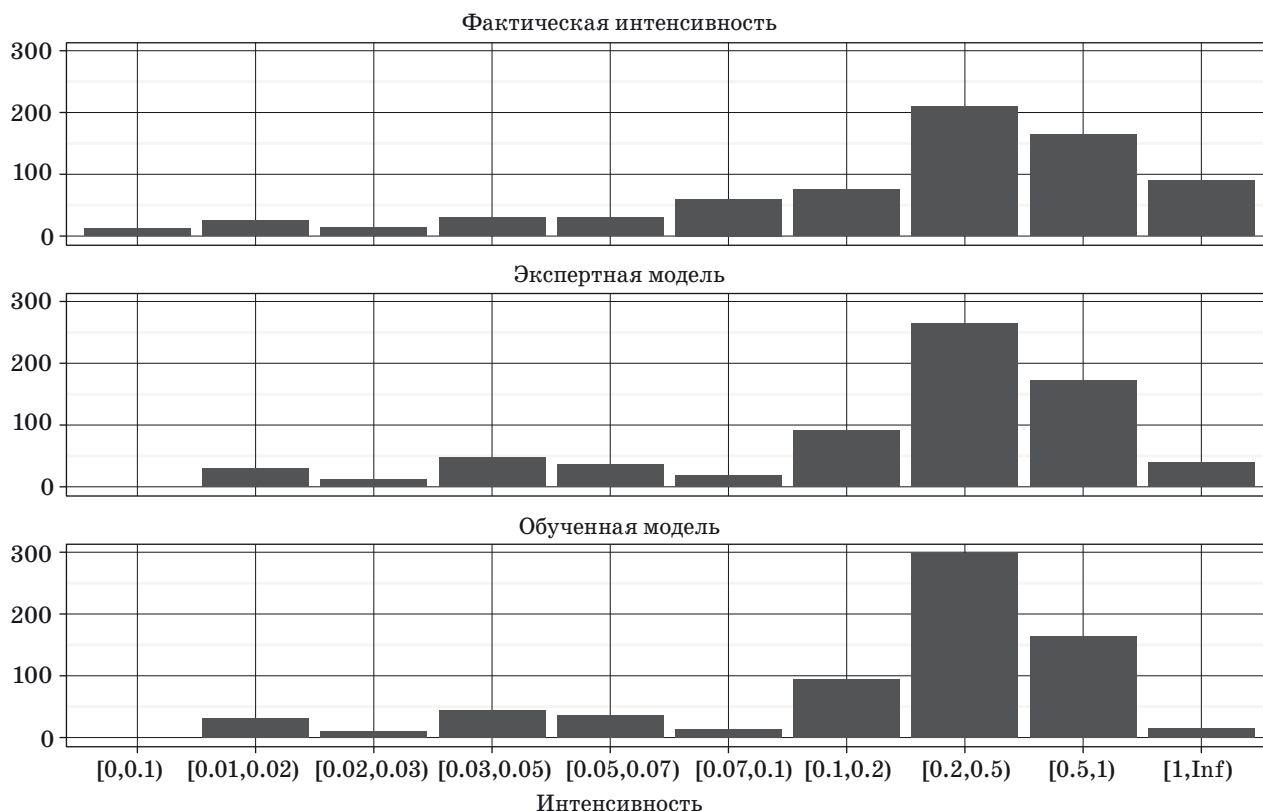
■ **Table 1.** Confusion matrix for prediction on model with expert-based structure

		Оценка интенсивности									
		$\lambda^{(1)}$	$\lambda^{(2)}$	$\lambda^{(3)}$	$\lambda^{(4)}$	$\lambda^{(5)}$	$\lambda^{(6)}$	$\lambda^{(7)}$	$\lambda^{(8)}$	$\lambda^{(9)}$	$\lambda^{(10)}$
Исходное значение	$\lambda^{(1)}$	0	10	1	1	0	0	0	0	0	0
	$\lambda^{(2)}$	0	14	4	6	1	0	0	0	0	0
	$\lambda^{(3)}$	0	4	3	4	3	0	0	0	0	0
	$\lambda^{(4)}$	0	2	3	17	4	3	1	0	0	0
	$\lambda^{(5)}$	0	1	1	9	11	4	5	0	0	0
	$\lambda^{(6)}$	0	0	0	10	16	11	23	0	0	0
	$\lambda^{(7)}$	0	0	0	1	1	0	48	24	1	0
	$\lambda^{(8)}$	0	0	0	0	0	0	14	162	34	0
	$\lambda^{(9)}$	0	0	0	0	0	0	0	77	81	7
	$\lambda^{(10)}$	0	0	0	0	0	0	0	2	56	32

■ **Таблица 2.** Предсказание по обученной на данных структуре

■ **Table 2.** Confusion matrix for prediction on model with data-based structure

		Оценка интенсивности									
		$\lambda^{(1)}$	$\lambda^{(2)}$	$\lambda^{(3)}$	$\lambda^{(4)}$	$\lambda^{(5)}$	$\lambda^{(6)}$	$\lambda^{(7)}$	$\lambda^{(8)}$	$\lambda^{(9)}$	$\lambda^{(10)}$
Исходное значение	$\lambda^{(1)}$	0	24	2	0	0	0	0	0	0	0
	$\lambda^{(2)}$	0	18	6	3	2	0	0	0	0	0
	$\lambda^{(3)}$	0	4	2	8	1	0	0	0	0	0
	$\lambda^{(4)}$	0	5	0	18	3	2	2	0	0	0
	$\lambda^{(5)}$	0	0	1	9	10	3	7	0	0	0
	$\lambda^{(6)}$	0	0	0	7	20	8	25	0	0	0
	$\lambda^{(7)}$	0	0	0	0	1	0	47	27	0	0
	$\lambda^{(8)}$	0	0	0	0	0	0	14	172	24	0
	$\lambda^{(9)}$	0	0	0	0	0	0	0	94	70	1
	$\lambda^{(10)}$	0	0	0	0	0	0	0	5	71	14



■ **Рис. 3.** Распределения интенсивности, полученные на разных моделях

■ **Fig. 3.** Predicted rate distributions for different models

Средняя доля правильных предсказаний (ассурагу) для 10-классовой классификации согласно модели со структурой, заданной экспертно, немного выше, чем точность оценивания согласно модели с автоматически обученной структурой (90,7% vs 89,8%). Другие показатели качества предсказания также немного выше для экспертной модели (точность (precision): 0,53 vs

0,49; полнота (recall): 0,5 vs 0,48). Отметим, что и в той и в другой модели неправильно классифицируемые значения в большинстве случаев отнесены к интервалам, смежным исходному значению, т. е. даже неправильное предсказание не дает сильного смещения оценки.

Обобщенную оценку по группе в целом можно получить за счет свойств распределения Дирихле,

являющегося сопряженным априорным распределением к мультиномиальному распределению [20]. Апостериорное распределение интенсивности поведения в группе вычисляется путем сложения векторов вероятностей индивидуальных распределений и последующей их нормировки [21]. На рис. 3 представлено фактическое распределение значений интенсивности и распределения, полученные с помощью двух рассматриваемых моделей.

Для рассматриваемых тестовых данных не выявлено различий между распределением интенсивности, посчитанной по первоначальной (экспертной) модели, и распределением интенсивности, вычисленной по модели с автоматически обученной структурой ($\chi^2 = 13,86$, $df = 8$, $p\text{-value} = 0,09$). Другими словами, обе модели дают значимо не различающиеся предсказания интенсивности поведения в группе.

Заключение

В работе предложено развитие подхода к построению модели рискованного поведения на основе байесовской сети доверия по совокупности наблюдений, включающей сведения об эпизодах такого поведения. Первоначальная структура модели основана на экспертных знаниях, что может влиять на качество модели. Для снижения подобного влияния рассмотрена другая структура модели, построенная с помощью алгоритма машинного обучения, выявляющего статистические взаимосвязи между элементами из данных и не использующего предположения экспертов о таких взаимосвязях. Результаты сравнения этих структур на автоматически генерируемых данных показывают,

что качество предсказания немного выше у модели с экспертно заданной структурой. Формальные меры качества структуры (BIC и мера максимального правдоподобия) ожидаемо выше у структуры, обученной автоматически, так как используемые алгоритмы структурного обучения ориентированы именно на оптимизацию меры качества BIC.

Отметим, однако, что в целом качество предсказания обеих моделей достаточно высокое и не значительно отличается друг от друга. Кроме того, содержательная интерпретация взаимосвязей является обоснованной как для исходной, так и для обученной структуры. Одним из дальнейших направлений исследования является сравнение приведенных в статье моделей не на автоматически генерируемых данных, а на реальных данных о поведении (для такого тестирования, однако, необходимо получить сведения не только о последних эпизодах поведения, но и о фактической интенсивности, что часто невозможно для реального поведения).

Таким образом, на данном этапе для решения практических задач можно использовать любую из предложенных моделей. Выбор может быть обусловлен условием конкретной задачи: например, при наличии некоторой обучающей выборки для оценивания параметров можно использовать структуру, приведенную на рис. 2, а при отсутствии таких данных — первоначальную структуру с аналитически вычисленными условными вероятностями. Перечисленные возможные действия аналитика учитываются при разработке инструментария, автоматизирующего решение задач оценивания параметров поведения.

Статья содержит материалы исследований, частично поддержанных грантом РФФИ 16-31-60063-мол_а_дк.

Литература

1. Азаров А. А., Тулупьева Т. В., Фильченков А. А., Тулупьев А. Л. Вероятностно-реляционный подход к представлению модели комплекса «информационная система — персонал — критичные документы» // Тр. СПИИРАН. 2012. Вып. 1(20). С. 57–71. doi:10.15622/sp.20.3
2. Афанасьев И. В. Возможности математического моделирования поведения аудитории с помощью динамических математических моделей // Актуальные проблемы современной науки. 2006. № 4. С. 212–218.
3. Leigh B. C., Stall R. Substance use and Risky Sexual Behavior for Exposure to HIV: Issues in Methodology, Interpretation, and Prevention // *American Psychologist*. 1993. Vol. 48(10). P. 1035.
4. Varghese B., Maher J. E., Peterman T. A., Branson B. M., Steketee R. W. Reducing the Risk of Sexual HIV Transmission: Quantifying the Per-act Risk for HIV on the Basis of Choice of Partner, Sex Act, and Condom use // *Sexually Transmitted Diseases*. 2002. Vol. 29(1). P. 38–43.
5. Lemelin C., Lussier Y., Sabourin S., Brassard A., Naud C. Risky Sexual Behaviours: The Role of Substance use, Psychopathic Traits, and Attachment Insecurity Among Adolescents and Young Adults in Quebec // *The Canadian Journal of Human Sexuality*. 2014. Vol. 23(3). P. 189–199. doi:0.3138/cjhs.2625
6. Bolger N., Davis A., Rafaeli E. Diary Methods: Capturing Life as it is Lived // *Annual Review of Psychology*. 2003. Vol. 54(1). P. 579–616. doi:10.1146/annurev.psych.54.101601.145030
7. Graham C. A., Catania J. A., Brand R., Duong T., Cancchola J. A. Recalling Sexual Behavior: A Methodological Analysis of Memory Recall Bias via Interview using the Diary as the Gold Standard // *Journal of Sex Research*. 2003. Vol. 40(4). P. 325–332. doi:10.1080/00224490209552198

8. Суворова А. В. Моделирование социально-значимого поведения по сверхмалой неполной совокупности наблюдений // Информационно-измерительные и управляющие системы. 2013. Т. 11. № 9. С. 34–37.
9. Тулупьев А. Л., Николенко С. И., Сироткин А. В. Байесовские сети: логико-вероятностный подход. — СПб.: Наука, 2006. — 607 с.
10. Neapolitan R. E. Learning Bayesian Networks. — Pearson Prentice Hall, 2003. — 674 p.
11. Pearl J. Causality: Models, Reasoning, and Inference. — Cambridge: Cambridge University Press, 2000. — 400 p.
12. Constantinou A. C., Fenton N., Marsh W., Radlinski L. From Complex Questionnaire and Interviewing Data to Intelligent Bayesian Network Models for Medical Decision Support // Artificial Intelligence in Medicine. 2016. P. 75–93. doi:10.1016/j.artmed.2016.01.002
13. Du Y., Guo Y. Evidence Reasoning Method for Constructing Conditional Probability Tables in a Bayesian Network of Multimorbidity // Technology and Health Care. 2015. Vol. 23(s1). P. S161–S167. doi:10.3233/thc-150950
14. Mkrtchyan L., Podofillini L., Dang V. N. Bayesian Belief Networks for Human Reliability Analysis: A Review of Applications and Gaps // Reliability Engineering & System Safety. 2015. Vol. 139. P. 1–16. doi:10.1016/j.res.2015.02.006
15. Trucco P., Cango E., Ruggeri F., Grande O. A Bayesian Belief Network Modelling of Organisational Factors in Risk Analysis: A Case Study in Maritime Transportation // Engineering and System Safety. 2008. Vol. 93. P. 845–856. doi: 10.1016/j.res.2007.03.035
16. Тулупьева Т. В., Пащенко А. Е., Тулупьев А. Л., Красносельских Т. В., Казакова О. С. Модели ВИЧ-рискованного поведения в контексте психологической защиты и других адаптивных стилей. — СПб.: Наука, 2008. — 140 с.
17. Suvorova A., Tulupyeva T. Bayesian Belief Networks in Risky Behavior Modelling // Advances in Intelligent Systems and Computing. 2016. Vol. 451. P. 95–102. doi:10.1007/978-3-319-33816-3_10
18. R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, Austria. <http://www.R-project.org/> (дата обращения: 20.10.2016).
19. Scutari M. Learning Bayesian Networks with the Bnlearn R Package // arXiv preprint. arXiv:0908.3817. 2009.
20. Frigiyik B. A., Kapila A., Gupta M. R. Introduction to the Dirichlet Distribution and Related Processes. UWEE Tech. Rep. UWEETR-2010-0006. — Washington: UWEE, 2010. — 27 p.
21. Суворова А. В., Тулупьев А. Л., Сироткин А. В. Байесовские сети доверия в задачах оценивания интенсивности рискованного поведения // Нечеткие системы и мягкие вычисления. 2014. № 2. С. 115–129.

UDC 004.891 + 311.2

doi:10.15217/issn1684-8853.2018.1.116

Bayesian Belief Network Structure Synthesis for Risky Behavior Rate EstimationSuvorova A. V.^a, PhD, Phys.-Math., suvalv@gmail.comTulupyev A. L.^{a,b}, Dr. Sc., Phys.-Math., Associate Professor, alexander.tulupyev@gmail.com^aSaint-Petersburg Institute for Informatics and Automation of the RAS, 39, 14 Line, V. O., 199178, Saint-Petersburg, Russian Federation^bSaint-Petersburg State University, 7–9, Universitetskaya Emb., 199034, Saint-Petersburg, Russian Federation

Introduction: Studies in sociology, psychology, epidemiology, marketing or information security often face the issue of estimating the behavior rate (either individually or at the level of a population). Direct methods of behavior rate estimation are sometimes not available; hence, it is important to develop indirect methods. Earlier studies proposed an approach to risky behavior modeling based on a Bayesian belief network using the data about several last behavior episodes as its initial data. However, to apply this model in practice, we need to reduce its dependency from the initial experts' assumptions about the relations between the elements of the model. **Purpose:** To propose a model modification which would not require an expert-based model structure, and to compare the modified model with the initial one. **Methods:** To test the model, we used an automatically generated dataset which followed some initial assumptions about the data. To form the structure of a Bayesian belief network, we used a score-based hill-climbing algorithm with Bayesian information criterion score. **Results:** We proposed to modify the approach to risky behavior modeling in terms of Bayesian belief network based on the data about several last behavior episodes. The initial expert-based model and the model with a data-based structure were compared. Formal scores were better for the data-based structure, while the prediction quality was slightly better for the expert-based model. Hence, we can use both these models for practical applications; the choice depends on the assumptions and limitations of a particular task.

Keywords — Behavior Modelling, Bayesian Belief Network, Structure Synthesis, Machine Learning, Risky Behavior.

Citation: Suvorova A. V., Tulupyev A. L. Bayesian Belief Network Structure Synthesis for Risky Behavior Rate Estimation. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 1, pp. 116–122 (In Russian). doi:10.15217/issn1684-8853.2018.1.116

References

1. Azarov A. A., Tulupyeva T. V., Filchenkov A. A., Tulupyev A. L. Probabilistic Relational Approach to Representing "Informational System — Personnel — Critical Documents" Complex Model. *Trudy SPIIRAN* [SPIIRAS Proceedings], 2012, no. 1(20), pp. 57–71 (In Russian). doi:10.15622/sp.20.3
2. Afanasyev I. V. Capabilities of Dynamic Mathematical Models for Audience Behavior Modeling. *Aktual'nye problemy sovremennoi nauki*, 2006, no. 4, pp. 212–218 (In Russian).
3. Leigh B. C., Stall R. Substance use and Risky Sexual Behavior for Exposure to HIV: Issues in Methodology, Interpretation, and Prevention. *American Psychologist*, 1993, no. 48(10), pp. 1035.
4. Varghese B., Maher J. E., Peterman T. A., Branson B. M., Steketee R. W. Reducing the Risk of Sexual HIV Transmission: Quantifying the Per-act Risk for HIV on the Basis of Choice of Partner, Sex Act, and Condom use. *Sexually Transmitted Diseases*, 2002, no. 29(1), pp. 38–43.
5. Lemelin C., Lussier Y., Sabourin S., Brassard A., Naud C. Risky Sexual Behaviours: The Role of Substance use, Psychopathic Traits, and Attachment Insecurity Among Adolescents and Young Adults in Quebec. *The Canadian Journal of Human Sexuality*, 2014, no. 23(3), pp. 189–199. doi:0.3138/cjhs.2625
6. Bolger N., Davis A., Rafaeli E. Diary Methods: Capturing Life as it is Lived. *Annual Review of Psychology*, 2003, no. 54(1), pp. 579–616. doi:10.1146/annurev.psych.54.101601.145030
7. Graham C. A., Catania J. A., Brand R., Duong T., Canchola J. A. Recalling Sexual Behavior: A Methodological Analysis of Memory Recall Bias via Interview using the Diary as the Gold Standard. *Journal of Sex Research*, 2003, no. 40(4), pp. 325–332. doi:10.1080/00224490209552198
8. Suvorova A. V. Socially Significant Behavior Modeling on the Base of Super-Short Incomplete Set of Observations. *Informatsionno-izmeritel'nye i upravliaiushchie sistemy*, 2013, no. 9(11), pp. 34–37 (In Russian).
9. Tulupyev A. L., Nikolenko S. I., Sirotkin A. V. *Baiesovskie seti: logiko-veroiatnostnyi podkhod* [Bayesian Networks: A Probabilistic Logic Approach]. Saint-Petersburg, Nauka Publ., 2006. 607 p. (In Russian).
10. Neapolitan R. E. *Learning Bayesian Networks*. Pearson Prentice Hall, 2003. 674 p.
11. Pearl J. *Causality: Models, Reasoning, and Inference*. Cambridge, Cambridge University Press, 2000. 400 p.
12. Constantinou A. C., Fenton N., Marsh W., Radlinski L. From Complex Questionnaire and Interviewing Data to Intelligent Bayesian Network Models for Medical Decision Support. *Artificial Intelligence in Medicine*, 2016, pp. 75–93. doi:10.1016/j.artmed.2016.01.002
13. Du Y., Guo Y. Evidence Reasoning Method for Constructing Conditional Probability Tables in a Bayesian Network of Multimorbidity. *Technology and Health Care*, 2015, no. 23(s1), pp. S161–S167. doi:10.3233/thc-150950
14. Mkrtchyan L., Podofilini L., Dang V. N. Bayesian Belief Networks for Human Reliability Analysis: A Review of Applications and Gaps. *Reliability Engineering & System Safety*, 2015, no. 139, pp. 1–16. doi:10.1016/j.res.2015.02.006
15. Trucco P., Cango E., Ruggeri F., Grande O. A Bayesian Belief Network Modelling of Organisational Factors in Risk Analysis: A Case Study in Maritime Transportation. *Engineering and System Safety*, 2008, no. 93, pp. 845–856. doi:10.1016/j.res.2007.03.035
16. Tulupyeva T. V., Pashhenko A. E., Tulupyev A. L., Krasnoselskih T. V., Kazakova O. S. *Modeli VICH-riskovannogo povedeniia v kontekste psikhologicheskoi zashchity i drugikh adaptivnykh stilei* [HIV Risky Behavior Models in the Context of Psychological Defense and Other Adaptive Styles]. Saint-Petersburg, Nauka Publ., 2008. 140 p. (In Russian).
17. Suvorova A., Tulupyeva T. Bayesian Belief Networks in Risky Behavior Modelling. *Advances in Intelligent Systems and Computing*, 2016, no. 451, pp. 95–102. doi:10.1007/978-3-319-33816-3_10
18. *A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. Available at: <http://www.R-project.org/> (accessed 20 October 2016).
19. Scutari M. Learning Bayesian Networks with the Bnlearn R Package. *arXiv preprint*. arXiv:0908.3817, 2009.
20. Frigyi B. A., Kapila A., Gupta M. R. *Introduction to the Dirichlet Distribution and Related Processes*. UWEE Tech. Rep. UWEE-TR-2010-0006. Washington, UWEE, 2010. 27 p.
21. Suvorova A. V., Tulupyev A. L., Sirotkin A. V. Bayesian Belief Networks for Risky Behavior Rate Estimates. *Nechetkie sistemy i miagkie vychisleniia* [Fuzzy Systems and Soft Computing], 2014, no. 2, pp. 115–129 (In Russian).

UDC 621.371

doi:10.15217/issn1684-8853.2018.1.123

MODEL OF DEEP FADING

V. Lyandres^a, PhD, Professor, lyandres@ee.bgu.ac.il

^aBen-Gurion University of the Negev, P.O.B. 653, 1, Ben-Gurion St., Beer-Sheva, 84105, Israel

Introduction: Fading, i. e. random changes in the level of a radio signal, is one of the main problems in signal processing, as the result strongly depends on the fading depth. The distribution of the received signal envelope is sometimes left-shifted relative to the Rayleigh law. A radio propagation channel with such a sub-Rayleigh fading may be considered a "critical" one. **Purpose:** Synthesizing a model of a narrow-band random process with an envelope distribution left-shifted relative to the Rayleigh law. **Results:** The synthesis of a model is based on representing the process as a reaction of a stable dynamic system to white Gaussian noise excitation. We have obtained nonlinear stochastic second-order differential equations to simulate the fading of a radio signal having an envelope with Nakagami or Weibull distribution. The envelope is considered a Markov continuous process. An analytical expression for the envelope correlation function is obtained. It is shown that, at least, for Nakagami fading, the correlation interval of the envelope almost does not depend on its depth. **Practical relevance:** In various applications, including indoor radio communication, the level of the received signal can be critically low during long time intervals. The proposed model used as a simulator core for such a propagation channel provides the opportunity to evaluate the performance quality of a communication system at the stage of its development.

Keywords — Sub-Rayleigh Fading, Markovian Diffusion Processes, Stochastic Differential Equations, Nakagami Distribution Envelope, Weibull Distribution Envelope, Radio Signal Propagation Model.

Citation: Lyandres V. Model of Deep Fading. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 1, pp. 123–127. doi:10.15217/issn1684-8853.2018.1.123

Introduction

A radio channel is a rather complex medium. It is not repeatable, so the problem of the radio propagation modeling is a part of the system design, and there is a challenge to assure the required system's quality during all its stages. Any channel model is based on generation of the so-called fading carrier, i.e. its reaction on the pure sinusoid excitation. The channel output, in turn, represents a narrow band random process with certain statistical characteristics of its envelope and phase. The dominant approach to simulation of fading carrier is based on the supposition that it represents a sum of diffuse and line-of-sight components. The diffuse stationary component results from a large number of independent and identically distributed scatters. It might be considered as a product of multiplying the in-phase and quadrature components of the transmitted carrier by two mutually non-correlated band-limited white Gaussian noises. If the both have the same dispersion and zero mean then the fading carrier envelope is Rayleigh distributed and does not depend on the carrier phase, which, in turn, has a uniform probability density function (PDF) [1].

If the line-of-sight component of the received signal is not equal to zero, the fading carrier becomes nonstationary, and the PDF of its envelope is right shifted with respect to the Rayleigh law. This, so called generalized Gaussian fading model usually continues to be valid, but this assertion is not true in the case of a fading, which is deeper than the Rayleigh one. Such phenomenon takes

place if only a limited number of scatters actually contribute to the received signal. It is observed in the High-Frequency (HF) channel with significant variability [2], in Very-High-Frequency (VHF) indoor communication channel [3]. It describes the unwanted phenomenon which usually accompanies target detection by sea microwave radar [4]. In all these cases the measured envelope PDF exhibits large deviations from the Rayleigh distribution. The fading carrier quadrature components continue to be non-correlated but now due their non-Gaussian nature they become mutually dependent [1]. Disregarding of this fact leads to misinterpretation of experimental data and to the system performance decreasing as now the receivers designed under the Gaussian assumption are not optimal [5]. As well as channels with such so-called sub-Rayleigh fading have relative lack of reliability, namely they must be considered in the attempts of obtaining estimation of the system lower bound performance.

The Model of the Fading Carrier

Throughout the paper, we assume that the simulated propagation channel is not frequency-selective, and so in the case of a sinusoidal transmitted signal the received signal (fading carrier) may be written as

$$x(t) = r(t) \cos(\omega_0 t + \theta), \quad (1)$$

where $x(t)$ is a stationary process and the PDF of its envelope $P_r(r)$ is described by [6] the Nakagami law

$$P_r(r) = \begin{cases} \frac{2}{\Gamma(m)} \left(\frac{m}{\sigma^2}\right) r^{2m-1} \exp\left(-\frac{mx^2}{\sigma^2}\right), & r \geq 0 \\ 0, & r < 0 \end{cases}, \quad (2)$$

or by the Weibull law

$$P_r(r) = \begin{cases} \alpha \beta r^{\alpha-1} \exp(-\beta r^\alpha), & r \geq 0 \\ 0, & r < 0 \end{cases}. \quad (3)$$

In (2) $\Gamma(\cdot)$ is the gamma function, $m \geq 0.5$ is the fading parameter (the shape factor) and $2\sigma^2 = \langle r^2 \rangle$. The case of $m = 1$ corresponds to the Rayleigh PDF and if $0.5 \leq m < 1$ the fading is deeper (a sub-Rayleigh one). If $m = 0.5$ PDF (2) is the one-sided Gaussian. If in (3) parameter $\alpha < 2$ the corresponding fading is deeper than Rayleigh. If $\alpha = 2$ the PDF is Rayleigh. If $\alpha = 1$ it becomes one-sided exponential. Consideration of some experimental results [5] shows that in indoor channel m and α appear to be very close to those limiting values.

It is evident fact that a strong connection exists between the PDF of $x(t)$ and $P_r(r)$ principally gives an opportunity to generate $x(t)$ as a product of a synthesized process $r(t)$ and a Gaussian band pass process, but we propose a different approach consisting of the direct generation of the fading carrier by a second-order dynamic system excited by a white Gaussian noise (WGN). It is clear that generally this system must be nonlinear.

If the generating stochastic differential equation (SDE) is a stochastic modification of the Duffing equation with the operator

$$\ddot{x} + \mu \dot{x} + \omega_0^2 f(x) x = \sqrt{\frac{N_0}{2}} \xi, \quad (4)$$

where $\mu > 0$; $f(x)$ is a positively defined function; $\xi(t)$ is WGN with unit power spectral density; ω_0 is the natural frequency of the oscillator (5), the envelope PDF $P_r(r)$ may be the Rayleigh one (in the linear case) or over-Rayleigh, but never sub-Rayleigh [7], so it is not appropriate for our purpose. Instead (4) we consider the following generating SDE

$$\dot{x} + f(r) \dot{x} + \omega_0^2 x = \sqrt{\frac{N_0}{2}} \xi, \quad (5)$$

where $f(r)$ once more is a positively defined damping function.

The Hilbert transform $\hat{x}(t)$ of the process $x(t)$ may be approximately written as

$$\hat{x}(t) \approx -\frac{\dot{x}(t)}{\omega_0} \quad (6)$$

and its envelope in turn as

$$r(t) = \sqrt{x^2(t) + \hat{x}^2(t)} \approx \sqrt{x^2(t) + \frac{\dot{x}^2(t)}{\omega_0^2}}. \quad (7)$$

From mutual consideration of (6) and (7) we may come to the following system of two first order differential equations

$$\dot{x} \approx -\omega_0 \sqrt{r^2 - x^2}; \quad (8)$$

$$\dot{r} = \frac{\dot{x}}{r} \left(x + \frac{\dot{x}}{\omega_0^2} \right). \quad (9)$$

With the help of (5) the brackets in (9) may be written as

$$x + \frac{\dot{x}}{\omega_0^2} = \frac{\sqrt{K} \xi - f(r) \dot{x}}{\omega_0^2} \quad (10)$$

so, for the envelope $r(t)$ of the process $x(t)$ generated by SDE (5) we obtain the following differential equation

$$\dot{r} = -\frac{\sqrt{r^2 - x^2}}{r \omega_0} \left[\sqrt{K} \xi + \omega_0 \sqrt{r^2 - x^2} f(r) \right]. \quad (11)$$

The system of differential equations (8) and (9) generates a vector Markovian process (a, r) defined by its partial drift $a_x(x, r)$, $a_r(x, r)$ and diffusion $b_x(x, r)$, $b_r(x, r)$ functions [8]

$$a_x(x, r) = -\omega_0 \sqrt{r^2 - x^2}; \quad (12)$$

$$a_r(x, r) = -\frac{r^2 - x^2}{r} f(r) + \frac{N_0}{4\omega_0^2} \frac{x^2}{r^3}; \quad (13)$$

$$b_x(x, r) = 0; \quad (14)$$

$$b_r(x, r) = \frac{N_0}{2\omega_0^2} \frac{r^2 - x^2}{r^2}. \quad (15)$$

The Fokker — Planck equation (FPE) for the transition probabilities $P_{x,r}(x, r, t)$ is written as

$$\begin{aligned} \frac{\partial}{\partial t} P_{x,r}(x, r, t) = & -\frac{\partial}{\partial x} \left[a_x(x, r) P_{x,r}(x, r, t) \right] - \\ & -\frac{\partial}{\partial r} \left[a_r(x, r) P_{x,r}(x, r, t) \right] + \\ & + \frac{1}{2} \frac{\partial^2}{\partial r^2} \left[b_r(x, r) P_{x,r}(x, r, t) \right]. \end{aligned} \quad (16)$$

While the operator of the generating SDE (5) does not depend on time, the stationary PDFs $P_{x,r}(x, r) = \lim_{t \rightarrow \infty} P_{x,r}(x, r, t)$ exists and is written as [1]

$$P_{x,r}(x, r) = \frac{P_r(r)}{\pi\sqrt{r^2 - x^2}}, \quad (17)$$

where $P_r(r)$ is the stationary PDF of the r -component of the vector process (x, r) . Substituting expressions (12)–(15) in (16) we derive the following stationary FPE:

$$\begin{aligned} & \frac{\partial}{\partial r} \left[\frac{\sqrt{r^2 - x^2}}{r^2} P_r(r) \right] = \\ & = -\frac{4\omega_0^2 \sqrt{r^2 - x^2}}{N_0 r} f_1(r) P_r(r) + \frac{P_r(r)}{\omega_0^2 r^3 \sqrt{r^2 - x^2}}. \end{aligned} \quad (18)$$

Integrating (18) respect to $x \in (-r, r)$ we obtain the expression for $f(r)$ as a function of the generating SDE parameters (N_0, ω_0) and the required envelope PDF $P_r(r)$ is

$$f(r) = \frac{N_0}{4\omega_0^2 r} \left[\frac{1}{r} - \frac{d}{dr} \ln P_r(r) \right]. \quad (19)$$

For the Nakagami PDF of the envelope this function is written as

$$f(r) = \frac{N_0}{2\omega_0^2} \left(\frac{m}{\sigma^2} + \frac{1-m}{r^2} \right) \quad (20)$$

and the SDE (5) itself, which generates the corresponding stationary pass band process takes the form

$$\ddot{x} + \frac{N_0}{2\omega_0^2} \left(\frac{m}{\sigma^2 r^{2(1-m)}} + \frac{1-m}{r^2} \right) \dot{x} + \omega_0^2 x = \sqrt{\frac{N_0}{2}} \xi. \quad (21)$$

It is clear now that a process with a sub-Rayleigh fading ($0.5 \leq m < 1$) is represented as a stationary reaction to WGN excitation of the passive oscillator with energy dependent damping. In the Rayleigh case the oscillator is linear. Evidently, the following equivalent equation, is more convenient for simulation application then (20) and may be written as

$$r^2 \ddot{x} + \frac{N_0}{2\omega_0^2} \left(\frac{m}{\sigma^2} r^{2m} + 1 - m \right) \dot{x} + \omega_0^2 r^2 x = \sqrt{\frac{N_0}{2}} r^2 \xi. \quad (22)$$

In the case of Weibull PDF of the envelope, the nonlinear function in (4) is written as

$$f(r) = \frac{N_0}{4\omega_0^2} \left(\frac{2-\alpha}{r^2} + \alpha\beta r^{\alpha-2} \right) \quad (23)$$

and the generating SDE is

$$r^2 \ddot{x} + \frac{N_0}{4\omega_0^2} \left(2 - \alpha + \alpha\beta r^\alpha \right) \dot{x} + \omega_0^2 r^2 x = \sqrt{\frac{N_0}{2}} r^2 \xi. \quad (24)$$

In the case $1 \leq \alpha < 2$ (a sub-Rayleigh case) the oscillator corresponding to (24) is, once more, a passive one with the energy dependent damping.

Correlation Function of the Envelope of Sub-Rayleigh Fading Carrier

Let us return to (5) and consider instead of the vector process (x, r) another one (r, φ) , where r is defined by (7) and

$$\varphi(t) = -\omega_0 t - \text{arctg} \left(\frac{\dot{x}}{\omega_0 x} \right). \quad (25)$$

Differentiation of (25) with respect to time gives

$$\dot{\varphi} = -\frac{x(\ddot{x} + \omega_0^2 x)}{\omega_0^2 r^2}. \quad (26)$$

Substituting in the system of equations (7) and (26), defining together the process (r, φ) , expression for \ddot{x} from (5) and

$$\dot{x} = -\omega_0 r \sin(\omega_0 t + \varphi) \quad (27)$$

we easily obtain equations generating $r(t)$ and $\varphi(t)$. Their simplification is based on the assumption of the effective filtering of high harmonics in the generating SDE solution. The corresponding procedure for the envelope $r(t)$ which interests us leads to the first-order SDE

$$\dot{r} = -\frac{C}{2} r + \frac{\omega_0^2 N_0}{8r} + \zeta(t). \quad (28)$$

In (28) $\zeta(t)$ is a WGN with unit spectral density. From (28) it is clear that the envelope $r(t)$ of the Gaussian narrow band process is Rayleigh distributed, i. e.

$$P_r(r) = \frac{r}{\sigma^2} \exp \left(-\frac{r^2}{2\sigma^2} \right), \quad (29)$$

where

$$\sigma^2 = \frac{\omega_0^2 N_0}{4C}. \quad (30)$$

Since the results of eliminating the vibrations are not related to the specific form of $f(r)$ in (5), the envelope of the corresponding non-Gaussian process $x(t)$, just as in the linear case may be considered as a one-dimensional Markovian process with the diffusion function $b_r(r)$ independent of the PDF.

As the stationary PDF of the Markovian one-dimensional process with the diffusion function constant and equal to unity is defined as [8]

$$P_r(r) = \text{const} \exp\left[-\int a_r(r) dr\right], \quad (31)$$

its drift function is

$$a_r(r) = \frac{d}{dr} \ln P_r(r), \quad (32)$$

and so in the case of Nakagami $P_r(r)$ the corresponding SDE may be written as

$$\dot{r} = -\frac{2m-1}{r} + \frac{8Cmr}{\omega_0^2 N_0} + \zeta_1(t). \quad (33)$$

The SDE (33) generates one of the few Markovian processes for which the exact expression for the correlation function can be obtained. It is written as [8]

$$B_r(\tau) = \frac{\omega_0^2 N_0 \Gamma^2(m+0.5)}{16C\pi m \Gamma(m)} \times \sum_{i=1}^{\infty} \frac{\Gamma^2(i-0.5)}{\Gamma(i+m)i!} \exp\left(-i \frac{8Cm}{\omega_0^2 N_0} |\tau|\right). \quad (34)$$

For the sub-Rayleigh envelope PDF ($m < 1$) this expression is well approximated by the first member of the expansion:

$$B_r(\tau) \approx \frac{\omega_0^2 N_0 \Gamma^2(m+0.5)}{16Cm^2 \Gamma^2(m)} \exp\left(-\frac{8Cm}{\omega_0^2 N_0} |\tau|\right), \quad (35)$$

which with the help of (30) may be rewritten as

$$B_r(\tau) \approx \frac{\sigma^2 \Gamma^2(m+0.5)}{4m^2 \Gamma^2(m)} \exp\left(-\frac{2m}{\sigma^2} |\tau|\right). \quad (36)$$

Thus the correlation function of the fading carrier envelope is rather close to the exponent (a widely used model of the fading correlation). Formally, from (36) it appears as if the correlation interval of a sub-Rayleigh fading exceeds the same for the Rayleigh one, i. e. sub-Rayleigh fading is slower than the latter. This interpretation, nevertheless, ignores the fact that an increase of the channel fading

depth practically always leads to a decrease of the envelope mean value

$$\bar{r} = \frac{\sigma \Gamma(m+0.5)}{\sqrt{m} \Gamma(m)}, \quad (37)$$

which depends slightly on m if $0.5 < m \leq 1$. So the required decrease of the envelope PDF parameter σ may be obtained by the decline of N_0 in (30) and (5) which, in turn, leads to decreasing of the sub-Rayleigh envelope correlation interval, and as a result to its approximate independence on the parameter m value.

Unfortunately, in the case of the Weibull envelope PDF the corresponding Fokker — Planck equation cannot be solved analytically and due this it is impossible to obtain closed expression for the Weibull fading correlation function.

Conclusions

This paper has addressed the problem of modeling fading in situations where the Gaussian assumption no longer applies. The received fading signal is represented by the corresponding nonlinear oscillator (generating system) excited by WGN. The approach is quite different from the one, which generates the fading process as a sum of sinusoids [9] as well as from its representation as a product of a certain baseband and complex Gaussian processes [3]. Its main advantage is that the marginal PDF of the signal envelope and its correlation function are easy controlled. In particular, the former can be accommodated by a suitable choice of the spectral density of the WGN and of the energy dependent oscillator damping. With regard to simulation of non-Rayleigh fading the procedure of the corresponding oscillator synthesis is presented. In the sub-Rayleigh case (Nakagami- $m < 1$, and Weibull- $\alpha < 2$) the oscillator appears to be a passive one. The correlation function of the presented model envelope is analyzed. It is rather close to an exponent, but depends on the fading depth.

References

1. Rytov S., Kravtsov Y., Tatarskii V. *Principles of Statistical Radiophysics. Vol. 2. Correlation Theory of Random Processes*. Springer-Verlag, 1988. 234 p.
2. Furman W. N., Nieto J. W. The Effects of Channel Variability on High Data Rate HF Communications. *Proc. 10th Int. Conf. on Ionospheric Radio Systems and Techniques*, 2006, Beijing, China, 2006, pp. 115–119.
3. Polydorou D. S., Capsalis C. N. A New Theoretical Model for the Prediction of Rapid Fading Variations in an Indoor Environment. *IEEE Trans. Veh. Technology*, 1997, vol. 46, no. 3, pp. 748–755.
4. Beckman P., Spizzichino A. *The Scattering of Electromagnetic Waves from Rough Surfaces*. Pergamon Press, 1963.
5. Hashemi H., et al. Measurements and Modeling of Temporal Variations of the Indoor Radio Propagation Channel. *IEEE Trans. Veh. Technology*, 1994, vol. 43, no. 3, pp. 733–737.
6. Zayezdny A., Tabak D., Wulich D. *Engineering Applications of Stochastic Processes. Theory, Problems and Solutions*. John Wiley&Sons, 1989. 509 p.

7. Lyandres V., Shahaf M. Envelope Correlation Function of Narrow-band non-Gaussian Process. *Int. J. Nonlinear Mechanics*, 1995, vol. 30, no. 3, pp. 359–369.
8. Primak S., Kontorovich V., Lyandres V. *Stochastic Methods and their Applications to Communications. Stochastic Differential Equations Approach*. John Wiley&Sons, 2004. 434 p.
9. Patzold M., Killat U., Lane F., et al. On the Statistical Properties of Deterministic Simulation Models for Mobile Fading Channel. *IEEE Trans. Veh. Technology*, 1998, vol. 47, no. 1, pp. 254–269.

UDC 621.371

doi:10.15217/issn1684-8853.2018.1.123

Модель глубокого федингаЛяндрес В.^а, PhD, профессор, lyandres@ee.bgu.ac.il^аНегевский университет им. Бен-Гуриона, ПОБ 653, Бен-Гуриона ул., 1, Беэр-Шева, 84105, Израиль

Введение: качество приема радиосигнала в существенной степени зависит от диапазона случайных изменений его уровня, то есть от глубины фединга. Иногда распределение огибающей принимаемого сигнала оказывается сдвинутым влево относительно закона Рэлея. Радиоканал с таким распределением фединга может рассматриваться как «критический». **Цель:** синтез модели узкополосного случайного процесса с распределением огибающей, сдвинутого влево относительно закона Рэлея. **Результаты:** процедура синтеза модели основана на представлении процесса как реакции устойчивой динамической системы, возбуждаемой белым гауссовым шумом. Получены нелинейные устойчивые стохастические дифференциальные уравнения второго порядка для моделирования замирающей несущей радиосигнала с огибающей, распределенной по законам Накагами или Вейбулла. Огибающая рассматривается как марковский непрерывный процесс. Получено аналитическое выражение для корреляционной функции. Показано, что по крайней мере в случае замираний по Накагами интервал корреляции огибающей почти не зависит от глубины замираний. **Практическая значимость:** в различных приложениях, включая радиосвязь в закрытых помещениях, уровень замирающего сигнала в течение значительных временных интервалов критически низок. Предложенная модель, используемая в качестве основного блока имитатора такого радиоканала, позволяет осуществить оценку качества системы связи на уровне ее проектирования.

Ключевые слова — рэлеевские замирания, марковские диффузионные процессы, стохастические дифференциальные уравнения, распределение Накагами огибающей, распределение Вэйбулла огибающей, модель распространения радиосигнала.

Цитирование: Lyandres V. Model of Deep Fading// Информационно-управляющие системы. 2018. № 1. С. 123–127. doi:10.15217/issn1684-8853.2018.1.123

Citation: Lyandres V. Model of Deep Fading. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 1, pp. 123–127. doi:10.15217/issn1684-8853.2018.1.123

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.

БЕЗРУЧКО
Борис
Петрович



Профессор кафедры динамического моделирования и биомедицинской инженерии Саратовского государственного университета им. Н. Г. Чернышевского. В 1969 году окончил физический факультет Саратовского государственного университета им. Н. Г. Чернышевского по специальности «Радиофизика». В 1995 году защитил диссертацию на соискание ученой степени доктора физико-математических наук. Является автором более 250 научных публикаций. Область научных интересов — стохастическая динамика и хаос, нелинейные колебания и волны, радиофизика, анализ временных рядов. Эл. адрес: bezruchkobp@gmail.com

БЛАУНШТЕЙН
Натан
Шаевич



Гражданин Израиля. Профессор Иерусалимского технологического института, профессор-эмитус 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

БУТЕНКО
Анна
Анатольевна



Старший преподаватель кафедры романо-германской филологии и переводоведения Института филологии и журналистики Саратовского национального исследовательского государственного университета им. Н. Г. Чернышевского. В 2000 году окончила Саратовский государственный университет по специальности «Филолог». Преподаватель английского языка и литературы. Переводчик в сфере профессиональной коммуникации. Является автором десяти научных публикаций. Область научных интересов — теория перевода, когнитивная лингвистика. Эл. адрес: butenkoaan@yandex.ru

БЕН-ШИМОЛ
Иегуда



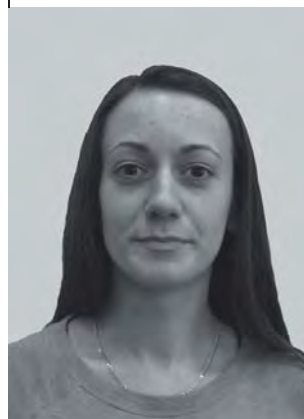
Гражданин Израиля. Профессор кафедры систем связи инженерного факультета Негевского университета им. Бен-Гуриона, Беэр-Шева, Израиль. В 1986 году окончил Негевский университет им. Бен-Гуриона по специальности «Электротехника и вычислительная техника». В 1997 году защитил диссертацию на соискание ученой степени доктора наук (PhD). Является автором 53 научных публикаций. Область научных интересов — проводные и беспроводные системы связи, комбинаторная оптимизация, параллельные, распределенные и символьные вычисления. Эл. адрес: benshimo@bgu.ac.il

БРОНШТЕЙН
Игорь
Григорьевич



Генеральный директор ЗАО «КБ Юпитер», директор Центра оптико-информационных технологий и систем Санкт-Петербургского национального исследовательского университета информационных технологий, механики и оптики, лауреат премии Правительства РФ в области науки и техники за 2005 год. В 1973 году окончил Ленинградский институт точной механики и оптики по специальности «Специальные оптические приборы». Является автором 60 научных публикаций. Область научных интересов — оптико-информационные системы. Эл. адрес: kb@jupiter.spb.ru

ВАТАМАНЮК
Ирина
Валерьевна



Аспирантка лаборатории речевых и мультимедийных интерфейсов Санкт-Петербургского института информатики и автоматизации РАН. В 2014 году окончила Санкт-Петербургский государственный университет аэрокосмического приборостроения по специальности «Роботы и робототехнические системы». Является автором 27 научных публикаций. Область научных интересов — системы компьютерного зрения, киберфизические системы. Эл. адрес: vatamaniuk@iias.spb.su

**ГОРОДЕЦКИЙ
Андрей
Емельянович**



Профессор, заведующий лабораторией интеллектуальных электромеханических систем Института проблем машиноведения РАН, Санкт-Петербург, заслуженный деятель науки и техники.

В 1965 году окончил Ленинградский политехнический институт им. М. И. Калинина по специальности «Автоматика и телемеханика».

В 1993 году защитил диссертацию на соискание ученой степени доктора технических наук.

Является автором более 250 научных публикаций и 70 патентов на изобретения.

Область научных интересов — математическое моделирование, оптимальное управление, идентификация и диагностика.

Эл. адрес:
gorodetsky@mail23.ipme.ru

**ИНОЧКИН
Федор
Михайлович**



Ассистент высшей школы программной инженерии Санкт-Петербургского политехнического университета Петра Великого.

В 2012 году окончил факультет технической кибернетики Санкт-Петербургского политехнического университета по специальности «Информатика и вычислительная техника».

Является автором 25 научных публикаций и одного патента на изобретение.

Область научных интересов — алгоритмы формирования и обработки изображений, системы технического зрения, оптико-информационные системы.

Эл. адрес:
fedor.inochkin@gmail.com

**ИШБУЛАТОВ
Юрий
Михайлович**



Аспирант кафедры динамического моделирования и биомедицинской инженерии Саратовского национального исследовательского государственного университета им. Н. Г. Чернышевского. В 2017 году окончил магистратуру Саратовского национального исследовательского государственного университета им. Н. Г. Чернышевского по специальности «Биотехнические системы и технологии».

Является автором 24 научных публикаций и 11 свидетельств о регистрации программного обеспечения.

Область научных интересов — теория разрядных вычислений, методы проектирования спецпроцессоров для систем контроля и управления, оптико-информационные системы.

Эл. адрес: ishbulatov95@mail.ru

**КАРАВАЕВ
Анатолий
Сергеевич**



Доцент кафедры динамического моделирования и биомедицинской инженерии Саратовского национального исследовательского государственного университета им. Н. Г. Чернышевского. В 2004 году окончил Саратовский государственный университет им. Н. Г. Чернышевского по специальности «Физика».

В 2007 году защитил диссертацию на соискание ученой степени кандидата физико-математических наук.

Является автором 76 научных публикаций, в том числе шести монографий, а также 45 объектов авторского права.

Область научных интересов — математическое моделирование, электронные системы контроля и управления, методы нелинейной динамики для анализа экспериментальных данных.

Эл. адрес: karavaevas@gmail.com

**КАРАСЁВ
Евгений
Юрьевич**



Программист лаборатории речевых и многомодальных интерфейсов Санкт-Петербургского института информатики и автоматизации РАН.

В 2017 году окончил Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики по специальности «Компьютерная фотоника».

Является автором одной научной публикации и трех патентов на изобретения и полезные модели. Область научных интересов — сетевое взаимодействие, реестр, передача мультимедийных данных, киберфизические системы.

Эл. адрес:
gnomskg@gmail.com

**КЕРЕН
Оснат**



Гражданка Израиля. Старший преподаватель инженерного факультета Университета Бар-Илан, Рамат-Ган, Израиль.

В 1988 году окончила магистратуру Израильского технологического института «Технион» в Хайфе по специальности «Электротехника».

В 1999 году защитила диссертацию на соискание ученой степени доктора наук (PhD) в Университете Тель-Авива.

Является автором 85 научных публикаций.

Область научных интересов — спектральные методы в логическом проектировании, схемы принятия решений, отказоустойчивость, онлайн-тестирование и кодирование компьютерных систем.

Эл. адрес: osnat.keren@biu.ac.il

**КИСЕЛЕВ
Антон
Робертович**



Ведущий научный сотрудник НИИ кардиологии Саратовского государственного медицинского университета им. В. И. Разумовского, научный сотрудник отделения хирургического лечения интерактивной патологии Национального научно-практического центра сердечно-сосудистой хирургии им. А. Н. Бакулева Минздрава России, Москва. В 2004 году окончил Саратовский государственный медицинский университет по специальности «Лечебное дело». В 2011 году защитил диссертацию на соискание ученой степени доктора медицинских наук. Является автором более 180 научных публикаций и одного патента на изобретение. Область научных интересов — кардиология, физиология кровообращения. Эл. адрес: kiselev@cardio-it.ru

**КРАСИЛЬНИКОВ
Николай
Николаевич**



Профессор кафедры информационно-сетевых технологий Санкт-Петербургского государственного университета аэрокосмического приборостроения, заслуженный деятель науки и техники РФ. В 1950 году окончил Ленинградский политехнический институт по специальности «Техническая физика». В 1963 году защитил диссертацию на соискание ученой степени доктора технических наук. Является автором более 250 научных публикаций, в том числе пяти монографий. Область научных интересов — цифровая обработка изображений, статистическая теория передачи и восприятия изображений, математическое моделирование процессов обработки информации зрительной системой человека. Эл. адрес: NNKrasilnikov@yandex.ru

**КРУГЛОВ
Сергей
Константинович**



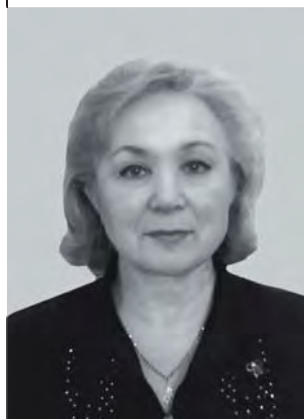
Доцент высшей школы программной инженерии Санкт-Петербургского политехнического университета Петра Великого. В 1981 году окончил Ленинградский политехнический институт им. М. И. Калинина по специальности «Автоматизированные системы обработки информации и управления». В 1996 году защитил диссертацию на соискание ученой степени кандидата технических наук. Является автором более 60 научных публикаций. Область научных интересов — методы обработки сигналов и изображений, спектральный анализ, оптоэлектронные системы. Эл. адрес: kruglovsk@gmail.com

**КОТЕНКО
Игорь
Витальевич**



Профессор, заведующий лабораторией проблем компьютерной безопасности Санкт-Петербургского института информатики и автоматизации РАН. В 1983 году окончил Военно-космическую академию им. А. Ф. Можайского по специальности «Математическое обеспечение автоматизированных систем управления», в 1987 году — Военную академию связи по специальности «Инженерная автоматизированных систем управления». В 1999 году защитил диссертацию на соискание ученой степени доктора технических наук. Является автором более 450 научных публикаций. Область научных интересов — безопасность компьютерных сетей, обнаружение компьютерных атак, межсетевые экраны, защита от вирусов и сетевых червей и др. Эл. адрес: ivkote@comsec.spb.ru

**КРАСИЛЬНИКОВА
Ольга
Ивановна**



Доцент кафедры информационно-сетевых технологий Санкт-Петербургского государственного университета аэрокосмического приборостроения. В 1978 году окончила Ленинградский институт авиационного приборостроения по специальности «Радиотехника». В 1991 году защитила диссертацию на соискание ученой степени кандидата технических наук. Является автором более 90 научных публикаций и семи учебных пособий. Область научных интересов — цифровая обработка изображений. Эл. адрес: OIKrasilnikova@yandex.ru

**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: rkulandin2010@yandex.ru

КУРБАНОВ
Вугар
Гариб оглы



Старший научный сотрудник лаборатории интеллектуальных электромеханических систем Института проблем машиноведения РАН, Санкт-Петербург, доцент кафедры управления в технических системах Санкт-Петербургского государственного университета аэрокосмического приборостроения. В 1976 году окончил Азербайджанский государственный университет им. С. М. Кирова по специальности «Прикладная математика». В 1983 году защитил диссертацию на соискание ученой степени кандидата физ.-мат. наук. Является автором более 90 научных публикаций. Область научных интересов — математическое моделирование процессов управления и др. Эл. адрес: vugar_borchali@yahoo.com

ЛЯНДРЕС
Владимир
Залманович



Гражданин Израиля. Профессор кафедры электронных и компьютерных систем Негевского университета им. Бен-Гуриона, Беэр-Шева, Израиль. В 1966 году окончил Ленинградский электротехнический институт связи им. проф. Бонч-Бруевича по специальности «Радиотехника». В 1972 году защитил диссертацию на соискание ученой степени кандидата технических наук. Является автором более 60 научных публикаций, в том числе двух монографий и 17 патентов на изобретения. Область научных интересов — модели помех в каналах связи, марковские модели случайных процессов, алгоритмы дискретной оптимизации, алгоритмы частотного планирования. Эл. адрес: lyandres@ee.bgu.ac.il

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

ЛАПШЕВА
Елена
Евгеньевна



Руководитель образовательного центра непрерывной подготовки IT-специалистов, старший преподаватель кафедры информатики и программирования Саратовского национального исследовательского государственного университета им. Н. Г. Чернышевского. В 1992 году окончила Саратовский государственный университет по специальности «Прикладная математика», в 2017 году — магистратуру по специальности «Биотехнические системы и технологии». Является автором более 30 научных публикаций и трех свидетельств о регистрации программ. Область научных интересов — обработка сигналов биомедицинской природы, большие данные. Эл. адрес: lapsheva@yandex.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

МАРТЫНОВА
Любовь
Александровна



Старший научный сотрудник, ведущий научный сотрудник научно-исследовательского центра «Системы освещения обстановки» АО «Концерн «ЦНИИ «Электронприбор», Санкт-Петербург. В 1985 году окончила Ленинградский кораблестроительный институт по специальности «Прикладная математика». В 2013 году защитила диссертацию на соискание ученой степени доктора технических наук. Является автором 90 научных публикаций. Область научных интересов — системный анализ, математическое моделирование, оценка эффективности, обработка разнородной информации. Эл. адрес: martynowa999@bk.ru

НЕНАШЕВ
Вадим
Александрович



Доцент кафедры вычислительных систем и сетей Санкт-Петербургского государственного университета аэрокосмического приборостроения. В 2012 году окончил Санкт-Петербургский государственный университет аэрокосмического приборостроения по специальности «Информатика и вычислительная техника». В 2017 году защитил диссертацию на соискание ученой степени кандидата технических наук. Является автором 22 научных публикаций. Область научных интересов — компьютерное моделирование, системы навигации и радиолокации, статистический анализ, электродинамика. Эл. адрес: granat89@mail.ru

ПОНОМАРЕНКО
Владимир
Иванович



Доцент, ведущий научный сотрудник Саратовского филиала Института радиотехники и электроники им. В. А. Котельникова РАН, профессор Саратовского национального исследовательского государственного университета им. Н. Г. Чернышевского. В 1982 году окончил Саратовский государственный университет им. Н. Г. Чернышевского по специальности «Радиофизика и электроника». В 2008 году защитил диссертацию на соискание ученой степени доктора физ.-мат. наук. Является автором более 300 научных публикаций, пяти патентов на изобретения и авторских свидетельств. Область научных интересов — хаотическое поведение нелинейных систем и др. Эл. адрес: ponomarenkovi@gmail.com

ПЧЕЛИНЦЕВА
Светлана
Вячеславовна



Доцент кафедры автоматизации, управления, мехатроники Саратовского государственного технического университета им. Гагарина Ю. А., старший научный сотрудник научно-образовательного центра «Нелинейная динамика сложных систем». В 2002 году окончила факультет электронной техники и приборостроения Саратовского государственного технического университета по специальности «Роботы и робототехнические комплексы». В 2005 году защитила диссертацию на соискание ученой степени кандидата технических наук. Является автором более 40 научных публикаций. Область научных интересов — методы математического моделирования манипуляционных и робототехнических систем. Эл. адрес: pchelintseva@inbox.ru

РОНЖИН
Андрей
Леонидович



Профессор кафедры электромеханики и робототехники Санкт-Петербургского государственного университета аэрокосмического приборостроения, заместитель директора по научной работе Санкт-Петербургского института информатики и автоматизации РАН. В 1999 году окончил Санкт-Петербургский государственный университет аэрокосмического приборостроения по специальности «Автоматизированные системы обработки информации и управления». В 2010 году защитил диссертацию на соискание ученой степени доктора технических наук. Является автором 250 научных публикаций. Область научных интересов — разработка речевых и многомодальных интерфейсов, киберфизические системы. Эл. адрес: ronzhin@iias.spb.su

РУННОВА
Анастасия
Евгеньевна



Доцент кафедры автоматизации, управления, мехатроники Саратовского государственного технического университета им. Гагарина Ю. А., старший научный сотрудник Научно-образовательного центра «Нелинейная динамика сложных систем», докторант Саратовского государственного технического университета им. Гагарина Ю. А. 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

САВЕЛЬЕВ
Антон
Игоревич



Старший научный сотрудник лаборатории речевых и многомодальных интерфейсов Санкт-Петербургского института информатики и автоматизации РАН. В 2012 году окончил Санкт-Петербургский государственный университет аэрокосмического приборостроения по специальности «Роботы и робототехнические системы». В 2016 году защитил диссертацию на соискание ученой степени кандидата технических наук. Является автором 22 научных публикаций и пяти патентов на изобретения и полезные модели. Область научных интересов — человеко-машинное взаимодействие, робототехника, инфокоммуникационные системы. Эл. адрес: saveliev@iias.spb.su

СЕРГЕЕВ
Михаил
Борисович



Профессор, директор Института вычислительных систем и программирования, заведующий кафедрой вычислительных систем и сетей Санкт-Петербургского государственного университета аэрокосмического приборостроения, почетный работник высшего профессионального образования РФ.

В 1980 году окончил ЛЭТИ по специальности «Электронные вычислительные машины».

В 2001 году защитил диссертацию на соискание ученой степени доктора технических наук.

Является автором более 150 научных публикаций и 14 патентов на изобретения.

Область научных интересов — теория разрядных вычислений, методы проектирования спецпроцессоров для систем контроля и управления и др.

Эл. адрес: mbse@mail.ru

ТАРАСОВА
Ирина
Леонидовна



Старший научный сотрудник лаборатории интеллектуальных электромеханических систем Института проблем машиноведения РАН, Санкт-Петербург, доцент Санкт-Петербургского политехнического университета.

В 1978 году окончила Ленинградский политехнический институт им. М. И. Калинина по специальности «Автоматические системы управления».

В 1998 году защитила диссертацию на соискание ученой степени кандидата технических наук. Является автором более 60 научных публикаций.

Область научных интересов — математическое моделирование, оптимальное управление, идентификация и диагностика.

Эл. адрес: til@msa2.ipme.ru

ФЕДОРЧЕНКО
Андрей
Владимирович



Аспирант, младший научный сотрудник лаборатории проблем компьютерной безопасности Санкт-Петербургского института информатики и автоматизации РАН.

В 2014 году окончил Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» по специальности «Компьютерная безопасность».

Является автором 33 научных публикаций.

Область научных интересов — безопасность компьютерных сетей, обнаружение вторжений, вредоносные программы.

Эл. адрес: fedorchenko@comsec.spb.ru

СУВОРОВА
Алёна
Владимировна



Старший научный сотрудник лаборатории теоретических и междисциплинарных проблем информатики Санкт-Петербургского института информатики и автоматизации РАН.

В 2010 году окончила Санкт-Петербургский государственный университет по специальности «Прикладная информатика в социологии».

В 2013 году защитила диссертацию на соискание ученой степени кандидата физико-математических наук.

Является автором 65 научных публикаций и 12 свидетельств о регистрации программ для ЭВМ и баз данных.

Область научных интересов — машинное обучение, байесовские сети, анализ данных, моделирование поведения, рискованное поведение.

Эл. адрес: suvalv@gmail.com

ТУЛУПЬЕВ
Александр
Львович



Заведующий лабораторией теоретических и междисциплинарных проблем информатики Санкт-Петербургского института информатики и автоматизации РАН.

В 1992 году окончил Санкт-Петербургский государственный университет по специальности «Прикладная математика», в 1997 году — по специальности «Социология», в 2003 году — Йельский университет, США.

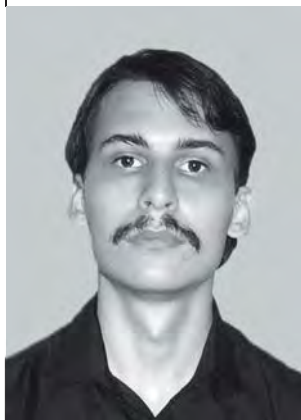
В 2010 году защитил диссертацию на соискание ученой степени доктора физико-математических наук.

Является автором 300 научных публикаций.

Область научных интересов — представление и обработка данных и знаний с неопределенностью и др.

Эл. адрес: alexander.tulup'ev@gmail.com

ХОРЕВ
Владимир
Сергеевич



Доцент кафедры динамического моделирования и биомедицинской инженерии Саратовского государственного университета им. Н. Г. Чернышевского, член общества IEEE.

В 2011 году окончил магистратуру Саратовского государственного университета им. Н. Г. Чернышевского по специальности «Биомедицинская инженерия».

В 2015 году защитил диссертацию на соискание ученой степени кандидата физико-математических наук.

Является автором 20 научных публикаций.

Область научных интересов — математическое моделирование, динамика систем с запаздыванием, методы нелинейной динамики для анализа экспериментальных данных.

Эл. адрес: khorevvs@gmail.com

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

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 40 научных публикаций. Область научных интересов — моделирование сложных систем, методы обработки телеметрической информации, техническая диагностика, методы обработки изображений. Эл. адрес: yakim78@yandex.ru

ЭНГЕЛЬБЕРГ
Шломо
Цви



Гражданин Израиля и США. Доцент кафедры электротехники, декан школы технических и компьютерных наук Иерусалимского технологического колледжа, Израиль. В 1991 году окончил магистратуру Нью-Йоркского университета по специальности «Математика». В 1994 году защитил диссертацию на соискание ученой степени доктора математических наук (PhD) в Нью-Йоркском университете, США. Является автором десятков научных публикаций и четырех монографий. Область научных интересов — прикладная математика, кодирование и теория информатики, обработка сигналов, теория управления. Эл. адрес: shlomoe@jct.ac.il